

Syracuse University

SURFACE

Dissertations - ALL

SURFACE

August 2017

Wireless Throughput and Energy Efficiency under QoS Constraints

Mustafa Ozmen
Syracuse University

Follow this and additional works at: <https://surface.syr.edu/etd>



Part of the [Engineering Commons](#)

Recommended Citation

Ozmen, Mustafa, "Wireless Throughput and Energy Efficiency under QoS Constraints" (2017).

Dissertations - ALL. 763.

<https://surface.syr.edu/etd/763>

This Dissertation is brought to you for free and open access by the SURFACE at SURFACE. It has been accepted for inclusion in Dissertations - ALL by an authorized administrator of SURFACE. For more information, please contact surface@syr.edu.

ABSTRACT

Mobile data traffic has experienced unprecedented growth recently and is predicted to grow even further over the coming years. As one of the main driving forces behind this growth, wireless transmission of multimedia content has significantly increased in volume and is expected to be the dominant traffic in data communications. Such wireless multimedia traffic requires certain quality-of-service (QoS) guarantees.

With these motivations, in the first part of the thesis, throughput and energy efficiency in fading channels are studied in the presence of randomly arriving data and statistical queueing constraints. In particular, Markovian arrival models including discrete-time Markov, Markov fluid, and Markov-modulated Poisson sources are considered, and maximum average arrival rates in the presence of statistical queueing constraints are characterized. Furthermore, energy efficiency is analyzed by determining the minimum energy per bit and wideband slope in the low signal-to-noise ratio (SNR) regime.

Following this analysis, energy-efficient power adaptation policies in fading channels are studied when data arrivals are modeled as Markovian processes and statistical QoS constraints are imposed. After formulating energy efficiency (EE) as maximum throughput normalized by the total power consumption, optimal power control policies that maximize EE are obtained for different source models.

Next, throughput and energy efficiency of secure wireless transmission of delay sensitive data generated by random sources are investigated. A fading broadcast model in which the transmitter sends confidential and common messages to two re-

ceivers is considered. It is assumed that the common and confidential data, generated from Markovian sources, is stored in buffers prior to transmission, and the transmitter operates under constraints on buffer/delay violation probability. Under such statistical QoS constraints, the throughput is determined. In particular, secrecy capacity is used to describe the service rate of buffers containing confidential messages. Moreover, energy efficiency is studied in the low signal-to-noise (SNR) regime.

In the final part of the thesis, throughput and energy efficiency are addressed considering the multiuser channel models. Five different channel models, namely, multiple access, broadcast, interference, relay and cognitive radio channels, are considered. In particular, throughput regions of multiple-access fading channels are characterized when multiple users, experiencing random data arrivals, transmit to a common receiver under statistical QoS constraints. Throughput regions of fading broadcast channels with random data arrivals in the presence of QoS requirements are studied when power control is employed at the transmitter. It is assumed that superposition coding with power control is performed at the transmitter with interference cancellation at the receivers. Optimal power control policies that maximize the weighted combination of the average arrival rates are investigated in the two-user case. Energy efficiency in two-user fading interference channels is studied when the transmitters are operating subject to QoS constraints. Specifically, energy efficiency is characterized by determining the corresponding minimum energy per bit requirements and wideband slope regions. Furthermore, transmission over a half-duplex relay channel with secrecy and QoS constraints is studied. Secrecy throughput is derived for the half duplex two-hop fading relay system operating in the presence of an eavesdropper. Fundamental limits on the energy efficiency of cognitive radio transmissions are analyzed in the presence of statistical quality of service (QoS) constraints. Minimum energy per bit and wideband slope expressions are obtained in order to identify the performance limits in terms of energy efficiency.

WIRELESS THROUGHPUT AND ENERGY
EFFICIENCY UNDER QOS CONSTRAINTS

By

Mustafa Ozmen

B.S., Bilkent University, Ankara, Turkey, 2011

DISSERTATION

Submitted in partial fulfillment of the requirements for the degree of
Doctor of Philosophy in Electrical and Computer Engineering

Syracuse University
August 2017

Copyright © 2017 Mustafa Ozmen

All rights reserved

For all the loving support and motivation:

This thesis is dedicated to my wonderful family Günay Özmen,
Yusuf Özmen.

TABLE OF CONTENTS

List of Figures	x
1 Introduction	1
1.1 Literature Review	2
1.1.1 Wireless Throughput and Energy Efficiency under Statistical Queueing Constraints	2
1.1.2 Secure Transmission over Wireless Fading Channels	5
1.1.3 Throughput and Energy Efficiency in Multiuser Channels	6
1.2 Main Contributions	8
1.3 Bibliographic Note	13
2 Preliminaries on Statistical Queueing Constraints	16
2.1 Queueing Constraints	16
2.2 Effective Bandwidths of Different Source Models	19
3 Wireless Throughput and Energy Efficiency with Random Arrivals and Statistical Queueing Constraints	25
3.1 Throughput with Markovian Source Models	25
3.1.1 Channel Model	27
3.1.2 Discrete-Time Markov Sources	28
3.1.3 Markov Fluid Sources	36
3.1.4 Continuous-Time Markov Modulated Poisson Sources	40

3.1.5	Discrete-Time Markov Modulated Poisson Process	43
3.1.6	Comparative View of Source Models and Performance Levels	45
3.2	Energy Efficiency Analysis	46
3.2.1	Energy Efficiency Metrics	46
3.2.2	Energy Efficiency with Constant Arrival Rate	48
3.2.3	Energy Efficiency with Discrete-Time Markov Sources	49
3.2.4	Energy Efficiency with Markov Fluid Sources	54
3.2.5	Energy Efficiency with Markov-Modulated Poisson Process	59
3.3	Energy Efficiency in Multiple-Antenna Channels	62
3.3.1	Channel Model	62
3.3.2	Effective Capacity of Wireless Transmissions with MIMO Channels	63
3.3.3	Energy Efficiency of MIMO Channels with Discrete Markov Sources	64
3.3.4	Energy Efficiency of MIMO Channels with Markov Fluid Sources	68
4	Energy-Efficient Power Control in Fading Channels with Markovian Sources and QoS Constraints	71
4.1	Channel Model	71
4.2	Energy-Efficient Power Control	72
4.2.1	Discrete Markov Source	73
4.2.2	Markov Fluid Source	83
4.2.3	Markov-Modulated Poisson Processes	86
4.3	Optimal Power Control with EE Constraints	86
4.3.1	Discrete Markov Source	87
4.3.2	Markov Fluid Source	89
4.3.3	Markov Modulated Poisson Processes	90
4.4	Optimal Power Control with Average Power Constraints	90
4.5	Optimal Power Control in Multichannel Systems	94
4.5.1	Discrete Markov Source	95

4.5.2	Markov Fluid Source	100
5	Secure Transmission of Delay-Sensitive Data over Wireless Fading Channels	104
5.1	Channel Model	104
5.1.1	Instantaneous Secrecy Capacity of Confidential Messages and Capacity of Common Message Transmissions	105
5.2	Throughput of Secure Transmissions with Random Data Arrivals Under QoS Constraints	108
5.3	Energy Efficiency Of Secure Transmissions with Random Data Arrivals Under QoS Constraints	113
5.3.1	Minimum Energy per Bit	113
5.3.2	Wideband Slope	117
5.4	Throughput and Energy Efficiency with no Channel Knowledge at the Transmitter	125
5.4.1	Effective Capacity with no Channel Knowledge at the Transmitter .	126
5.4.2	Energy Efficiency with Discrete Markov Sources	126
5.4.3	Energy Efficiency with Markov Fluid Sources	128
6	Energy Efficiency of Fixed-Rate Transmissions with Markov Arrivals under Queueing Constraints	130
6.1	Channel Model and Fixed-Rate Transmissions	130
6.2	Energy Efficiency of Fixed-Rate Transmission of ON-OFF Markov Sources	132
6.2.1	Markov Fluid Sources	132
6.2.2	Discrete-Time Markov Sources	136
7	Analysis of Multiuser Channels with Markov Arrivals under QoS Constraints	138
7.1	Throughput Regions of Multiple-Access Fading Channels with Markov Arrivals and QoS Constraints	138
7.1.1	Channel Model	138

7.1.2	MAC Throughput Region	139
7.1.3	Numerical Results	141
7.2	Power Control in Fading Broadcast Channels with Random Arrivals and QoS Constraints	143
7.2.1	Channel Model	143
7.2.2	Throughput Regions of Fading Broadcast Channels with Power Control	146
7.2.3	Numerical Results and Discussion	152
7.3	Energy Efficiency in Fading Interference Channels under QoS Constraints .	157
7.3.1	Channel Model	157
7.3.2	Energy Efficiency in Fading Interference Channels	158
7.3.3	Numerical Results and Discussion	165
7.4	Energy Efficiency in Fading Relay Channels under Secrecy and QoS Con- straints	167
7.4.1	Channel Model	167
7.4.2	Preliminaries	170
7.4.3	Energy Efficiency Of Two-Hop Wireless Communication	172
7.4.4	Numerical Results and Discussions	175
7.5	Energy Efficiency in Cognitive Radio Channels with Markov Arrivals . . .	180
7.5.1	Channel Sensing	180
7.5.2	Cognitive Radio Channel Model	182
7.5.3	Effective Capacity of Cognitive Radio Channels	183
7.5.4	Energy Efficiency Metrics	186
7.5.5	Energy Efficiency in Cognitive Radio Channels with Markov Arrivals	187
8	Future Research Directions	194
8.1	QoS-Driven Energy-Efficient Power Control in Cognitive Radio Channels with Markov Arrivals	194

8.2	Throughput and Optimal Resource Allocation in the Finite Blocklength Regime under QoS Constraints	195
A	Appendix	196
A.1	Proof of Theorem 3.1.1:	196
A.2	Proof of Theorem 3.1.2:	197
A.3	Proof of Theorem 3.1.3:	200
A.4	Proof of Theorem 3.1.4:	201
A.5	Proof of Theorem 3.1.5:	202
A.6	Proof of Theorem 3.1.7:	203
A.7	Proof of Theorem 3.1.8:	204
A.8	Proof of Theorem 3.2.1:	205
A.9	Proof of Theorem 7.5.1:	207
A.10	Proof of Theorem 3.3.3:	208
A.11	Proof of Theorem 3.2.4:	209
A.12	Proof of Proposition 5.3.1	210
A.13	Proof of Proposition 5.3.2	211
A.14	Proof of Proposition 5.3.3	212
A.15	Proof of Proposition 5.3.4	212
A.16	Proof of Proposition 5.3.5	213
A.17	Proof of Proposition 5.3.6	214
A.18	Proof of Proposition 5.4.1	215
A.19	Proof of Proposition 5.4.2	217
	References	219

LIST OF FIGURES

3.1	System Model.	27
3.2	Maximum average arrival rate r_{avg}^* vs. effective capacity $C_E(\text{SNR})$ for different source statistics. No fading correlation, i.e., $\rho = 0$. $\theta = 1$	29
3.3	Maximum average arrival rate r_{avg}^* vs. signal-to-noise ratio SNR for different values of θ and source statistics. No fading correlation, i.e., $\rho = 0$	33
3.4	Required SNR vs. ON probability, P_{ON} , for a given fixed average arrival rate. No fading correlation, i.e., $\rho = 0$	33
3.5	Maximum average arrival rate r_{avg}^* vs. QoS exponent θ for different values of P_{ON} . No fading correlation, i.e., $\rho = 0$	34
3.6	Maximum average arrival rate r_{avg}^* vs. QoS exponent θ for different fading correlations. $P_{\text{ON}} = 0.5$	35
3.7	Buffer overflow probability $\Pr\{Q > q\}$ vs. buffer threshold q for different values of θ . $p_{11} = p_{22} = 0.8$, SNR = 0 dB	36
3.8	Delay violation probability $\Pr\{D > d\}$ vs. delay threshold d for different values of θ . $p_{11} = p_{22} = 0.8$, SNR = 0 dB	37
3.9	Maximum average arrival rate r_{avg}^* vs. signal-to-noise ratio SNR for different values of θ and source statistics. No fading correlation, i.e., $\rho = 0$	39
3.10	Required SNR vs. ON probability, P_{ON} , for a given average arrival rate. $\theta = 0.5$. No fading correlation, i.e., $\rho = 0$	40
3.11	Maximum average arrival rate r_{avg}^* vs. QoS exponent θ for different values of α and β . $P_{\text{ON}} = 0.5$, $\rho = 0$, and SNR = 0 dB.	41

3.12	Maximum average arrival rate r_{avg}^* vs. signal-to-noise ratio SNR for different values of θ and different source statistics. No fading correlation, i.e., $\rho = 0$	43
3.13	Maximum average arrival rate r_{avg}^* vs. QoS exponent θ for different source statistics. $\rho = 0$	44
3.14	Normalized maximum average arrival rate $\frac{1}{m}r_{\text{avg}}^*$ vs. energy per bit $\frac{E_b}{N_0}$ for different fading correlations. $\theta = 1, m = 10$	50
3.15	Maximum average arrival rate $\frac{1}{m}r_{\text{avg}}^*$ vs. energy per bit $\frac{E_b}{N_0}$ for different values of $P_{\text{ON}} = s$ when $\rho = 0.75, \theta = 1$, and $m = 10$	52
3.16	Maximum average arrival rate $\frac{1}{m}r_{\text{avg}}^*$ vs. energy per bit $\frac{E_b}{N_0}$ for different values of s when channel blocks are uncorrelated and $\theta = 1$. Number of states for the arrival process is $n = 10$	53
3.17	Maximum average arrival rate $\frac{1}{m}r_{\text{avg}}^*$ vs. energy per bit $\frac{E_b}{N_0}$ for different θ values when channel blocks are uncorrelated and $s = 0.5$. Number of states for arrival process is $n = 10$	54
3.18	Maximum average arrival rate r_{avg}^* vs. signal-to-noise ratio SNR when channel blocks are uncorrelated. Number of states of the arrival process is $n = 10$	55
3.19	Maximum average arrival rate $\frac{1}{m}r_{\text{avg}}^*$ vs. energy per bit $\frac{E_b}{N_0}$ when $\rho = 0.75, \theta = 1$, and $m = 10$	56
3.20	Maximum average arrival rate $\frac{1}{m}r_{\text{avg}}^*$ vs. energy per bit $\frac{E_b}{N_0}$ when channel blocks are uncorrelated. $\theta = 1, \beta = 100$ and the number of states of the arrival process is $n = 10$	58
3.21	Maximum average arrival rate r_{avg}^* vs. energy per bit $\frac{E_b}{N_0}$ when channel blocks are uncorrelated and $\alpha = \beta = 50$. Number of states of the arrival model is 10.	58
3.22	Maximum average arrival rate $\frac{1}{m}r_{\text{avg}}^*$ vs. energy per bit $\frac{E_b}{N_0}$ when $\rho = 0.75, \theta = 1$, and $m = 10$	60

3.23	Maximum average arrival rate $\frac{1}{m}r_{\text{avg}}^*$ vs. energy per bit $\frac{E_b}{N_0}$ when channel blocks are uncorrelated. $\theta = 1, \beta = 100$ and the number of Markov states for the arrival process is $n = 10$	61
3.24	Maximum average arrival rate $\frac{1}{m}r_{\text{avg}}^*$ vs. energy per bit $\frac{E_b}{N_0}$ when channel blocks are uncorrelated and $\alpha = \beta = 50$. Number of Markov states for the arrival process is $n = 10$	62
3.25	Maximum average arrival rate r_{avg}^* vs. energy per bit $\frac{E_b}{N_0}$ when $\theta = 0.1$ and $n_T = 5$	67
3.26	Maximum average arrival rate r_{avg}^* vs. energy per bit $\frac{E_b}{N_0}$ when $\theta = 0.1$ and $n_R = 5$	67
3.27	Maximum average arrival rate r_{avg}^* vs. energy per bit $\frac{E_b}{N_0}$ when $P_{\text{ON}} = 0.2$	68
3.28	Maximum average arrival rate r_{avg}^* vs. energy per bit $\frac{E_b}{N_0}$ when $\theta = 1$ and $n_T = 5$	70
3.29	Maximum average arrival rate r_{avg}^* vs. energy per bit $\frac{E_b}{N_0}$ when $\theta = 1$ and $n_R = 5$	70
4.1	Energy efficiency EE vs. maximum average arrival rate r_{avg}^* when $\varrho = 1, \mathbf{P}_c = 1, N_0 = 1, \epsilon = 1$	78
4.2	Power control function $\mu(\theta, z)$ vs. z when $\varrho = 1, \mathbf{P}_c = 1, N_0 = 1$ and $\epsilon = 1$	79
4.3	Energy efficiency EE vs. maximum average arrival rate r_{avg}^* for different values of the circuit power \mathbf{P}_c when $\varrho = 1, p_{11} = p_{22} = 0.5, N_0 = 1, \epsilon = 1$	79
4.4	Maximum EE vs. QoS exponent θ when $\mathbf{P}_c = 1, N_0 = 1, \epsilon = 1$	80
4.5	Buffer overflow probability $\Pr\{Q \geq q\}$ vs. buffer threshold q for cases with optimal EE power control and fixed power when $\mathbf{P}_c = 1, N_0 = 1, \epsilon = 1$. Discrete Markov source with $p_{11} = p_{22} = 0.75$	81
4.6	Delay violation probability $\Pr\{D \geq d\}$ vs. delay threshold d for cases with optimal EE power control and fixed power when $\mathbf{P}_c = 1, N_0 = 1, \epsilon = 1$. Discrete Markov source with $p_{11} = p_{22} = 0.75$	81

4.7	Energy efficiency EE vs. maximum average arrival rate r_{avg}^* when $\rho = 1$, $\mathbf{P}_c = 1, N_0 = 1, \epsilon = 1$	85
4.8	Energy efficiency EE vs. maximum average arrival rate r_{avg}^* with different power control schemes when $\rho = 0.5, \alpha = 2, \beta = 8, \mathbf{P}_c = N_0 = 1 = \epsilon = 1$	85
4.9	Throughput r_{avg}^* gain % vs. EE gain %. $\theta = 1$. (Discrete Markov Source)	89
4.10	Throughput r_{avg}^* gain % vs. EE gain %. $\theta = 1$. (Markov Fluid Source)	90
4.11	Maximum average arrival rate r_{avg}^* vs. average power $\bar{\mathbf{P}}$ when $\theta = 1$. (Discrete Markov Source)	92
4.12	Maximum average arrival rate r_{avg}^* vs. average power $\bar{\mathbf{P}}$ when $\theta = 1$. (Markov Fluid Source)	93
4.13	Energy efficiency EE vs. maximum average arrival rate r_{avg}^* when $N_0 = 1$, $\rho = 1$	98
4.14	Throughput r_{avg}^* gain % vs. EE gain % when $N_0 = 1, \rho = 1$	98
4.15	Energy efficiency EE vs. maximum average arrival rate r_{avg}^* when $N_0 = 1$, $K = 4$	99
4.16	Energy efficiency EE vs. maximum average arrival rate r_{avg}^* when $N_0 = 1$, $\rho = 1$	101
4.17	Throughput r_{avg}^* gain % vs. EE gain % when $N_0 = 1, \rho = 1$	102
4.18	Energy efficiency EE vs. maximum average arrival rate r_{avg}^* when $N_0 = 1$, $\alpha = \beta = 5$	103
5.1	Two-receiver broadcast channel model.	104
5.2	Maximum average arrival rate of the confidential message of the first user $r_{\text{avg},1}^*$ vs. average signal-to-noise ratio SNR when $\theta_1 = 1$ and $\delta_1 = 0.5$	110
5.3	Buffer overflow probability $\Pr\{Q > q\}$ vs. buffer threshold q for both confidential and common messages when $\theta_1 = 0.5, \theta_2 = 2, \theta_0 = 1$, $\delta_1 = \delta_2 = 0.7, p_{11} = p_{22} = 0.8$ and SNR = 1.	111

5.4	Maximum average arrival rate of the confidential message of the second user $r_{\text{avg},2}^*$ vs. average signal-to-noise ratio SNR when $\theta_2 = 1$, $\rho = 0.05$ and $\delta_2 = 0.5$	112
5.5	Maximum average arrival rate of first user's confidential message $r_{\text{avg}1}^*$ vs. energy per bit $\frac{E_b}{N_0}$ in dB when $\theta_1 = 1$, $\rho = 0.05$ and $\delta_1 = 0.5$	120
5.6	Maximum average arrival rate of common message $r_{\text{avg}0}^*$ vs. energy per bit $\frac{E_b}{N_0}$ in dB when $\theta_0 = 1$, $\rho = 0.05$ and $\delta_1 = \delta_2 = 0.5$	121
5.7	Maximum average arrival rate of first user's confidential message $r_{\text{avg}1}^*$ vs. energy per bit $\frac{E_b}{N_0}$ when $\theta_1 = 1$, $\gamma = 1$ and $\delta_1 = 0.5$	123
5.8	Maximum average arrival rate of common message $r_{\text{avg}0}^*$ vs. energy per bit $\frac{E_b}{N_0}$ in dB when $\rho = 0.8$, $\gamma = 1$, $p_{11} = 0.1$ and $p_{22} = 0.9$	124
5.9	Buffer overflow probability $\Pr\{Q > q\}$ vs. buffer threshold q for different values of θ . $p_{11} = p_{22} = 0.8$, SNR = 0.05	128
5.10	Maximum average arrival rate r_{avg}^* vs. energy per bit $\frac{E_b}{N_0}$ with various source statistics when $\theta = 0.5$	129
6.1	Maximum average arrival rate r_{avg}^* vs. energy per bit $\frac{E_b}{N_0}$ when $\theta = 1, 10$	136
7.1	Throughput regions for discrete Markov sources when $\theta = 1$, SNR = 10 and $\rho = 0$	142
7.2	Throughput regions for discrete Markov sources when $q = 0.5$, SNR = 10 and $\rho = 0$	143
7.3	Throughput regions for discrete Markov sources when $q = 0.5$, $\theta = 0.7$ and $\rho = 0$	144
7.4	Throughput regions for Markov fluid sources when $\theta = 1$, SNR = 10 and $\rho = 0.1$	145
7.5	Fading broadcast channel with random arrivals.	145

7.6	Maximum average arrival rate regions $(r_{\text{avg}1}^*, r_{\text{avg}2}^*)$ of the two users when $\theta = 2$, SNR = 0.1, $\gamma = 1$ and source has discrete Markov property when $P_{\text{ON}} = 0.5$	152
7.7	Sum rate vs. SNR when $\theta = 2$ and $\gamma = 1$ with discrete Markov source.	155
7.8	Maximum average arrival rate regions $(r_{\text{avg}1}^*, r_{\text{avg}2}^*)$ of the two users when $\theta = 2$, SNR = 0.1, $\gamma = 1$ and source has Markov fluid property.	155
7.9	Sum rate vs. SNR when $\theta = 2$ and $\gamma = 1$ with Markov fluid source.	156
7.10	The system model.	157
7.11	$\mathcal{S}_{0,1}$ vs. $\mathcal{S}_{0,2}$ for different interference levels for time division with power control (curved line) and treating interference as noise (rectangular regions). $\theta_1 = \theta_2 = 1$	166
7.12	$\mathcal{S}_{0,1}$ vs. $\mathcal{S}_{0,2}$ for different QoS exponents and different interference levels. for time division with power control (curved line) and treating interference as noise (rectangular regions).	167
7.13	Minimum energy per bit achieved by transmitter 1 vs. time-sharing parameters α_1 and α_2	168
7.14	Minimum energy per bit achieved by transmitter 1 vs. time-sharing parameters α_1 and α_2	168
7.15	Channel Model	168
7.16	Locations of source (S), eavesdropper (E), relay (R) and destination (D).	175
7.17	Maximum arrival rate $R_E(\theta_1, \theta_2)$ vs. energy per bit E_b/N_0 without eavesdropper.	176
7.18	Maximum arrival rate $R_E(\theta_1, \theta_2)$ vs. energy per bit E_b/N_0 in the presence of an eavesdropper when $\theta_1 = 0.1, \theta_2 = 0.01$	177
7.19	Maximum arrival rate $R_E(\theta_1, \theta_2)$ vs. distance between source and relay d_0 when $\theta_1 = \theta_2 = 0.1$	178

7.20	Maximum arrival rate $R_E(\theta_1, \theta_2)$ vs. distance between source and relay d_0 when $\theta_2 = 0.1$ and SNR = 1.	179
7.21	Minimum energy per bit E_b/N_{0min} vs. distance between source and relay d_0 when $\theta_2 = 0.1$ and SNR = 1.	180
7.22	The state-transition model for the cognitive radio channel with four states.	184
7.23	Maximum average arrival rate r_{avg}^*/m vs. energy per bit $\frac{E_b}{N_0}$ with different source burstiness when $\theta = 1$	191
7.24	Minimum energy per bit and P_d-P_f vs sensing duration n	192
7.25	Minimum energy per bit and P_d-P_f vs sensing threshold λ	193

CHAPTER 1

INTRODUCTION

MOBILE data traffic has experienced unprecedented growth recently and is predicted to grow even further over the coming years. For instance, it is projected that global mobile data traffic will increase 7-fold between 2016 and 2021, reaching 48.3 exabytes per month by 2021 [5]. As one of the main driving forces behind this growth, wireless transmission of multimedia content has significantly increased in volume and is expected to be the dominant traffic in data communications. Indeed, mobile video traffic already accounted for 60 percent of the total mobile data traffic in 2016 and is predicted to become more than three-fourths of the world's mobile data traffic by 2021 [5].

This exponential growth in the flow of mobile data and multimedia content has significant implications on wireless networks. For one, wireless multimedia traffic requires certain quality-of-service (QoS) guarantees. For instance, in voice over IP (VoIP), multimedia streaming, interactive video, and online gaming applications, constraints on delay, packet loss, or buffer overflow probabilities need to be imposed so that acceptable performance and quality levels can be met for the end-users. Another consequence is heterogeneity in network traffic. Wireless networks now carry heterogeneous traffic in diverse environments, and successful design of networks, efficient use of resources, and effective QoS provisioning for multimedia communications critically depend on the appropriate choice

of source traffic models. For instance, while voice traffic can be accurately modeled as an ON/OFF process with fixed-rate data arrivals in the ON state, data traffic can be bursty and video traffic, which exhibits correlations, can be modeled statistically using autoregressive, Markovian, or Markov-modulated processes [4].

Finally, it is important to note that this increased traffic together with the given QoS requirements need to be supported by wireless systems equipped with only limited bandwidth and power resources. Especially, due to limited energy available for mobile units and rising energy costs and environmental concerns, energy efficiency in wireless communications is a key concern (see e.g., [2] and [33]). Therefore, it is crucial to identify the fundamental performance limits (e.g., in terms of maximum achievable throughput and minimum energy per bit) in order to determine how to most effectively utilize the scarce resources. With this motivation, in this thesis we investigate the throughput and energy efficiency of wireless systems when data arrivals are in general random, and QoS constraints in the form of limitations on the asymptotic buffer overflow probabilities are imposed.

1.1 Literature Review

1.1.1 Wireless Throughput and Energy Efficiency under Statistical Queueing Constraints

Satisfying QoS requirements is critical for most communication networks, and how to satisfy QoS constraints for various source traffic models has been one of the key considerations in the networking literature. In particular, besides conventional queueing theory, network calculus has been introduced by Cruz in early 1990s as a theory to address the delay and other deterministic service guarantees in networks by dealing with queueing systems [34] – [36]. Subsequently, Chang in [6] developed the stochastic version of the network calculus. More specifically, the theory of effective bandwidth of a time-varying source has been formulated to identify the minimum amount of transmission rate that is needed to satisfy

the statistical QoS requirements (see also [7] – [38]). This theory is based on the logarithmic moment generating function of the arrival process and is related to the large deviation principle. Moreover, statistical QoS constraints are imposed as limitations on buffer/delay violation probabilities. Effective bandwidths of various source models have been investigated extensively in the literature. For instance, Elwalid and Mitra studied the effective bandwidth of Markovian traffic sources (including Markov-modulated fluid and Markov-modulated Poisson sources) in [11] under constraints on the buffer overflow probability. It is shown that effective bandwidth is given by the maximum eigenvalue of a matrix derived from source parameters and service requirements. In [12], effective bandwidth formulations were provided for multi-class Markov fluids as well as memoryless (Poisson) and discrete-time Markov sources. In [39], the authors studied the effective bandwidths of general stationary sources and derived a first order approximation of the effective bandwidth in terms of the mean arrival rate and index of dispersion.

In wireless communications, the instantaneous channel capacity varies randomly depending on the channel conditions. Hence, in addition to the source characteristics, transmission rates for reliable communication are also time-varying. In such cases, randomly time-varying servers can be considered in the queueing system model. Indeed, motivated by the wireless channel, Stolyar in [40], Venkataramanan and Lin in [41], and Sadiq and de Veciana in [42] employed tools from the theory of large deviations and investigated scheduling rules (e.g., MaxWeight, Exponential, and Radial Sum-Rate Monotonic scheduling) while controlling the large deviations of queues. Following another method, the time-varying channel capacity can be incorporated into the theory of effective bandwidth by regarding the channel service process as a time-varying source with negative rate and using the source multiplexing rule ([9, Example 9.2.2]). Using a similar approach, as a dual concept to effective bandwidth, Wu and Negi defined in [10] the effective capacity, which describes the maximum constant arrival rate that a given time-varying service process can support while satisfying the statistical QoS requirements. Indeed, work in [10] revitalized the consid-

eration of statistical queueing constraints in the context of wireless communications, and the effective capacity of wireless transmissions has been investigated intensively in various settings (see e.g., [43]–[21]). For instance, Tang and Zhang in [44] considered the effective capacity when both the receiver and transmitter know the instantaneous channel gains, and derived the optimal power policy that maximizes the system throughput under QoS constraints. Liu *et al.* in [27] considered fixed-rate transmission schemes and analyzed the effective capacity and related resource requirements for Markov wireless channel models and Markov fluid sources. In [47] and [48], effective capacity of cognitive radio channels was studied. In [49], multi-antenna communication in the presence of queueing limitations was investigated. Soret *et al.* in [51] addressed correlated Rayleigh fading channels and studied the effective capacity under different adaptive rate policies. In this study, performance in the presence of probabilistic delay constraints and variable rate sources was also analyzed by considering a Gaussian autoregressive source model.

Energy efficiency in the presence of statistical QoS constraints has also been addressed recently. For instance, the fundamental limits of energy efficiency in the low signal-to-noise ratio (SNR) regime in fading channels were determined under QoS constraints in [18]. Musavian and Le-Ngoc in [19] incorporated the circuit power consumption into their analysis of energy efficiency. Ru *et al.* in [20] investigated the minimum energy per bit and wideband slope in a hybrid cellular system. Liu in [21] considered the optimal power control to achieve the maximum energy efficiency. Helmy and Musavian in [22] considered a multichannel scenario in which they obtained the optimal power allocation for each channel to achieve the maximum global energy efficiency. Authors in [23] studied the energy-efficient design in downlink OFDMA systems. In a recent study in [26], the authors analyzed energy-efficient resource allocation strategies in MIMO-OFDM systems in the presence of random arrivals and statistical QoS requirements. In particular, they characterized the optimal energy-efficient queue-length based resource allocation policy that minimizes the total power consumption while satisfying the QoS requirements. Further-

more, in [28]–[30], power control policies were examined under QoS constraints.

1.1.2 Secure Transmission over Wireless Fading Channels

Addressing security considerations is essential in wireless communication networks due to the ease in eavesdropping of wireless transmissions. With this motivation, information-theoretic security has been extensively investigated. For instance, in [68] and [69] wiretap channels with fading have been studied whereas authors in [70] and [71] incorporated the multiple antenna settings to wiretap channels. Furthermore, the energy efficiency of secure and reliable communication schemes have been addressed in several recent studies. The work in [72] addressed secure communication in the low signal-to-noise ratio (SNR) regime and identified the minimum energy *per secret bit* and the wideband slope (which are two key performance metrics in the low SNR regime [57]). Motivated similarly by energy efficiency requirements, Comaniciu and Poor in [73] investigated the security-energy tradeoff from an information theoretic perspective. Zhang *et al.* in [74] studied three-node MIMO wiretap channels in order to design an energy efficient precoder. Ng *et al.* in [75] considered secure OFDMA systems and addressed the energy efficient resource allocation problem. Kalantari *et al.* in [76] investigated the power control in wiretap interference channels where users either work together or act as selfish nodes. Similar to our motivation, Chen and Lei in [77] took energy efficiency, security and QoS guarantees into account jointly and worked on maximizing the secrecy energy efficiency while having constraints on delay. In [78] and [79], Zhu *et al.* investigated the cross layer scheduling of OFDMA networks with both open and private data transmissions. In [80], two medium-access protocols were proposed and the mean service rate, the source's data queue and the secret keys queue was analyzed. Shafie and Al-Dhahir [81] proposed a network scheme that consists of a source node and a destination in the presence of buffer aided relay node and an eavesdropper, while taking the data burstiness of source and energy recycling process at the relay into account. In [82], secure and stable throughput region is investigated by employ-

ing beamforming based cooperative jamming that depends on the channel side information available at the transmitter. In [83], authors assumed that only the distribution of eavesdropper is known at the transmitter and studied the problem that maximizes the long-term data admission rate while having constraints on the secrecy outage and stability of the data queue. Khalil *et al.* in [84] derived upper and lower bounds on the secrecy capacity of the flat fading channel with limitations on delay. For more details regarding the advances in this rich field of physical-layer security in wireless communications, we refer to surveys and overviews provided in [85]–[91].

1.1.3 Throughput and Energy Efficiency in Multiuser Channels

Multiple-access channel (MAC) model, in which multiple users share a communication medium to send their messages to a common receiver, is one of the main building blocks of multiple-user communication scenarios, modeling, for instance, uplink in cellular and satellite communications and wireless LANs. It is well-known that Gaussian MAC capacity region is achieved by having simultaneous transmissions from the users (i.e., superposition coding) with successive cancelation decoding at the receiver [58]. Similar transmission and reception strategies are optimal in multiple-access fading channels as well [59]. In [60], effective capacity framework was employed to study the throughput regions of multiple-access fading channel subject to statistical queueing constraints under the assumption that arrival rates to all users are constant.

Broadcast channel (BC) model, in which we have one sender transmitting to multiple receivers, is one of the main building blocks of multiuser wireless networks. For instance, downlink in cellular, broadcasting, and satellite communication systems is modeled as a broadcast channel. Due to the importance and common usage of these channel models, they have been extensively studied from an information theoretic perspective in order to design and analyze efficient transmission strategies (see e.g., [58] and [93] and references therein). For instance, Li and Goldsmith in [94] characterized the ergodic capacity region

of fading broadcast channels and determined the optimal resource allocation policies. They showed that the capacity region, which is achieved by superposition coding and successive decoding whose order is determined by the effective noise levels, is convex. As noted before, providing QoS guarantees has become a very important consideration in wireless networks due to the fact that mobile multimedia data traffic (consisting of e.g. voice over IP, streaming and interactive video) has surged in recent years with the widespread use of social networking tools, video-sharing sites, and online gaming applications. With this motivation, in [95] authors have studied the throughput regions of fading broadcast channels in the presence of QoS constraints. However, only constant-rate arrivals were addressed in [95].

Due to the broadcast nature of wireless transmissions, interference is a common form of distortion experienced in especially densely-deployed wireless networks. Interference channel models explicitly take into account this type of distortion and have been extensively studied from an information-theoretic perspective (see e.g., [96], [97], and references therein).

Secure communication of confidential messages is a key concern in wireless networks due to the broadcast nature of wireless transmissions. Recently, information-theoretic physical-layer security for wireless communications has drawn considerable attention. In particular, based on the seminal work in [65], secrecy capacity of wireless links have been extensively studied. In [65], Wyner addressed the security problem in a wiretap channel by using information theoretic methods. In this model, the wiretapper receives a noisier version of the signal received at the intended user. Thus, the secrecy capacity is defined as the supremum of the achievable communication rates from the transmitter to the intended receiver while keeping the eavesdropper ignorant of the message. Recently, studies have also been conducted on cooperation for secrecy [98]. For instance, in [99] and [100], the authors considered decode-and-forward (DF) strategies in relay networks.

Energy and bandwidth are two critical resources in wireless communications. Due to

unprecedented growth in mobile applications and wireless networks and the fact that much of the prime radio spectrum has already been allocated for specific applications, the scarcity in the spectrum has become a serious concern. On the other hand, recent measurements have shown that the licensed spectrum is considerably underutilized across many time and frequency slots. This has led to much interest in dynamic spectrum access strategies and cognitive radio systems [101] which can more effectively harness the available bandwidth by, for instance, utilizing the spectrum holes.

1.2 Main Contributions

In Chapter 3, throughput and energy efficiency in fading channels are studied in the presence of randomly arriving data and statistical queueing constraints. In particular, we consider Markovian arrival models including discrete-time Markov, Markov fluid, and Markov-modulated Poisson sources. Employing the effective bandwidth of time-varying sources and effective capacity of time-varying wireless transmissions, maximum average arrival rates in the presence of statistical queueing constraints are characterized. For the two-state (ON/OFF) source models, throughput is determined in closed-form as a function of the source statistics, channel characteristics, and quality of service (QoS) constraints. Throughput is further studied in certain asymptotic regimes. Furthermore, we analyze energy efficiency by determining the minimum energy per bit and wideband slope in the low signal-to-noise ratio (SNR) regime. Overall, the impact of source characteristics, QoS requirements, and channel fading correlations on the throughput and energy efficiency of wireless systems is identified.

We note that the studies on the effective capacity of wireless channels have primarily concentrated on constant arrival rates in the analysis of the throughput and energy efficiency. Departing from this approach, we in this chapter explicitly take into account the randomness and burstiness of the source traffic. In particular, we address Markovian

source models including discrete-time Markov, Markov fluid, and Markov modulated Poisson sources, and conduct a performance analysis. More specifically, our contributions can be listed as follows:

- A framework with which source randomness can be incorporated in the throughput analysis of wireless transmissions is provided.
- For two-state (ON/OFF) source models, closed-form expressions are obtained for the maximum average arrival rate in terms of the source statistics, effective capacity of wireless transmissions, and the QoS exponent θ , which quantifies how strict the QoS constraints are.
- Throughput is characterized in the low- θ and high-SNR regimes.
- An energy efficiency analysis is conducted and minimum energy per bit and wide-band slope expressions are determined for both constant and random arrival models.
- Via both analytical and numerical results, the impact of source randomness, fading correlations, and queueing constraints on the wireless throughput and energy efficiency is identified.
- Throughput of multiple-input multiple-output (MIMO) wireless communication channels is studied an energy efficiency analysis is conducted under statistical queueing constraints.

In Chapter 4, we again explicitly model the information flows stochastically using Markovian models and study energy-efficient wireless transmission strategies in the presence of statistical QoS constraints imposed on buffer overflow probabilities. In particular, we identify energy-efficient power control policies in fading channels for different source arrival models. We further investigate the tradeoff between throughput and energy efficiency (EE).

The studies on energy efficiency and power control conducted with effective capacity formulations have mainly centered around the assumption that sources have constant arrival rates. In this chapter, we take into account the stochastic nature of information flows and investigate the effect of the randomness and burstiness of the source traffic on the energy-efficient design of wireless systems. Specifically, we consider Markovian source models (namely discrete-time Markov, Markov fluid, and both discrete and fluid Markov modulated Poisson processes (MMPP)) and determine the optimal energy-efficient power allocation policies. The contributions of this chapter can be further detailed as follows:

- Considering two-state (ON/OFF) source models, throughput expressions are provided and subsequently energy efficiency metric is identified for discrete-time Markov, Markov fluid, and MMPP arrival models. Overall, an analytical framework is provided to study the energy efficiency of wireless transmissions in the presence of random data arrivals and statistical queueing constraints.
- After taking both the circuit and transmission power into account, optimal power allocation policies that maximize the energy efficiency are determined for different source models.
- Power control policies that maximize the throughput under either energy efficiency or power constraints are also obtained.
- In addition to single-channel systems, power allocation and control strategies that maximize the throughput in multichannel systems under energy efficiency constraints are determined.
- Via both analytical and numerical results, the impact of source randomness, channel fading, queueing constraints, and power control strategies on the energy efficiency is identified. Tradeoff between energy efficiency and throughput is explored.

In Chapter 5, we study the secure communication of delay-sensitive data traffic generated by Markovian sources (e.g., discrete-time Markov, Markov fluid, discrete-time and continuous-time Markov modulated Poisson sources) and investigate the fundamental performance limits of secure throughput and energy efficiency under statistical buffer/delay violation constraints. In particular, we can list the contributions of this chapter as follows:

- Considering two-state (ON/OFF) Markovian source models, throughput expressions for common and confidential messages in terms of source statistics, effective capacity of wireless transmissions of common and confidential messages, and QoS exponent θ are provided.
- Energy efficiency metrics, namely the minimum energy per bit and wideband slope, are identified for discrete-time Markov, Markov fluid, and Markov-modulated Poisson arrival models again in terms of important system, channel, and source parameters.
- The effect of source randomness, channel correlation, secrecy requirements, buffer/delay QoS constraints on the performance metrics are identified for both common and confidential messages from both analytical characterizations and numerical results.
- Throughput and energy efficiency metrics are obtained when the transmitter knows the channel statistics but not the realizations of the channel fading, and therefore sends the confidential data at a fixed rate.

In Chapter 6, we assume that the time-varying channel conditions are not known at the transmitter and consequently, transmission rate is fixed. Fixed-rate transmission over the Rayleigh fading channel is modeled as an ON-OFF Markov fluid process. Under these modeling assumptions, our main contributions are the introduction of a general framework for performance analysis in the low-power regime, determination of closed-form expressions for the minimum energy per bit and wideband slope, and characterization of the impact of source and channel parameters and queueing constraints on the energy efficiency.

In Chapter 7, we conduct throughput and energy efficiency analysis for multiuser scenarios under QoS constraints. More specifically, we have the following contributions.

- In Section 7.1, we investigate the throughput regions of multiple-access fading channels when the users experience random arrivals and operate in the presence of quality-of-service (QoS) constraints. Consideration of QoS guarantees is motivated by the recent exponential growth of wireless transmissions of multimedia content. In this study, a more general scenario is considered in which users experience random Markov arrivals. In particular, we combine the theory of effective bandwidth of time-varying random arrivals and the theory of effective capacity of time-varying wireless transmissions in order to characterize the throughput regions in multiple-access fading channels.
- In Section 7.2, we characterize the throughput regions in fading broadcast channels when superposition coding with successive decoding is employed together with power control. We further propose an optimal power control algorithm. We determine the throughput region and sum-throughput for the two-user case and compare different strategies including time division multiplexing with power control and superposition coding without power adaptation.
- In Section 7.3, we investigate the energy efficiency in interference channels when users operate under QoS limitations. More specifically, we consider three strategies for communication in the two-user case, which are time division with power control, treating interference as noise, and simultaneous decoding. As metrics of energy efficiency, we determine the corresponding minimum energy per bit and wideband slope regions for these strategies. We compare the performances in the presence of different levels of interference and different QoS constraints.
- In Section 7.4, we consider a two-hop wireless channel setting in which the relay node, employing the DF strategy, helps the communication between the source and

destination. We assume that the wireless communication is half-duplex and hence the relay can not transmit and receive simultaneously. We adopt the energy per bit as the metric of energy efficiency. We impose constraints on the buffer overflow probabilities. We study secure cooperative communications under QoS constraints and investigate the energy efficiency by determining the minimum energy per bit.

- In Section 7.5, we address the efficient use of both bandwidth and energy resources by investigating the energy efficiency of cognitive radio systems. Motivated by the recent rapid growth in mobile multimedia applications which may exhibit bursty traffic and require certain QoS guarantees for acceptable performance levels at the end-users, we consider a setting in which data arrivals are modeled as a two-state Markov process and statistical buffer constraints are imposed at the cognitive transmitter.

1.3 Bibliographic Note

Our work has appeared in the following journals and conferences:

- The results in Section 3.1 and 3.2 appeared in journal paper:
 - M. Ozmen and M. C. Gursoy, “Wireless throughput and energy efficiency with random arrivals and statistical queueing constraints,” *IEEE Trans. Inform. Theory*, vol. 62, no. 3, pp. 1375-1395, March 2016.

and in the conference paper:

- M. Ozmen and M. C. Gursoy, “Impact of channel and source variations on the energy efficiency under QoS constraints,” Proc. of the IEEE International Symposium on Information Theory (ISIT), 2012.
- The results in Section 3.3 appeared in conference paper:

- M. Ozmen and M. C. Gursoy, “Energy efficiency in multiple-antenna channels with Markov arrivals and queueing constraints,” Proc. of the IEEE International Symposium on Information Theory (ISIT), 2015.

- The results in Chapter 4 appeared in journal paper:

- M. Ozmen and M. C. Gursoy, “Energy-efficient power control in fading channels with markovian sources and QoS constraints,” *IEEE Trans. Communications*, vol. 64, no. 12, pp. 5349-5364, Dec. 2016.

and in the conference paper:

- M. Ozmen and M. C. Gursoy, “Energy-efficient power control policies in fading channels with Markov arrivals and QoS constraints,” Proc. of IEEE Global Conference on Signal and Information Processing (GlobalSIP), 2013.

- The results in Chapter 5 appeared in journal paper:

- M. Ozmen and M. C. Gursoy, “Secure transmission of delay-sensitive data over wireless fading channels,” *IEEE Trans. Inform. Foren. & Security.*, vol. 12, no. 9, pp. 2036-2051, Sept. 2017.

and in the conference paper:

- M. Ozmen and M. C. Gursoy, “Throughput and energy efficiency under queueing and secrecy constraints,” Proc. of IEEE Conference Record of the Forty Sixth Asilomar Conference on Signals, Systems and Computers (ASILOMAR), 2012.

- The results in Chapter 6 appeared in letter:

- M. Ozmen and M. C. Gursoy, “Energy efficiency of fixed-rate transmissions with markov arrivals under queueing constraints,” *IEEE Commun. Letters*, vol. 18, no. 4, pp. 608-611, Apr. 2014.

- The results in Section 7.1 appeared in letter:
 - M. Ozmen and M. C. Gursoy, “Throughput regions of multiple-access fading channels with markov arrivals and QoS constraints,” *IEEE Wireless Commun. Letters*, vol. 2, no. 5, pp. 499-502, Oct. 2013.

- The results in Section 7.2 appeared in conference paper:
 - M. Ozmen and M. C. Gursoy, “Power control in fading broadcast channels with random arrivals and QoS constraints,” Proc. of the IEEE Global Communications Conference (GlobeCom), 2014.

- The results in Section 7.3 appeared in conference paper:
 - M. Ozmen, M. C. Gursoy and P. Varshney, “Energy efficiency in fading interference channels under QoS constraints,” Proc. of IEEE International Symposium on Information Theory and its Applications (ISITA), 2012.

- The results in Section 7.4 appeared in conference paper:
 - M. Ozmen, C. Ye, M. C. Gursoy and S. Velipasalar, “Energy efficiency in fading relay channels under secrecy and QoS constraints,” Proc. of the IEEE 25th Annual International Symposium on Personal, Indoor, and Mobile Radio Communication (PIMRC), 2014.

- The results in Section 7.5 appeared in conference paper:
 - M. Ozmen, G. Ozcan and M. C. Gursoy, “Energy efficiency in cognitive radio channels with Markov arrivals,” Proc. of IEEE First International Black Sea Conference on Communications and Networking (BlackSeaCom), 2013.

CHAPTER 2

PRELIMINARIES ON STATISTICAL QUEUEING CONSTRAINTS

2.1 Queueing Constraints

We assume that the data to be transmitted is generated from random sources and is first stored in a buffer before transmission. Statistical constraints are imposed on the queue length. In particular, we assume that the buffer violation/overflow probability satisfies

$$\lim_{q \rightarrow \infty} \frac{\log \Pr\{Q \geq q\}}{q} = -\theta \quad (2.1)$$

where Q denotes the stationary queue length, and θ is the decay rate of the tail distribution of the queue length. The above limiting formula implies that for large q , we have

$$\Pr\{Q \geq q\} \approx e^{-\theta q}. \quad (2.2)$$

Indeed, a closer approximation is [10]

$$\Pr\{Q \geq q\} \approx \varsigma e^{-\theta q} \quad (2.3)$$

where $\varsigma = \Pr\{Q > 0\}$ is the probability of non-empty buffer¹. From (2.3), we notice that, for a sufficiently large threshold, the buffer overflow probability should decay exponentially with rate controlled by the QoS exponent θ . Note that as θ increases, stricter queueing or QoS constraints are imposed, while looser queueing constraints are implied by smaller values of θ . Conversely, for a given buffer threshold q and overflow probability limit $\epsilon = \Pr\{Q \geq q\}$, the desired value of θ can be determined as

$$\theta = \frac{1}{q} \log_e \frac{\varsigma}{\epsilon}. \quad (2.4)$$

In the given setting, the delay violation probability is also characterized to decay exponentially and is approximated by [52]

$$\Pr\{D \geq d\} \approx \varsigma e^{-\theta a^*(\theta)d} \quad (2.5)$$

where D is the queueing delay in the buffer at steady state, d is the delay threshold, and $a^*(\theta)$ is the effective bandwidth of the arrival process, described below.

Next, we introduce the notions of effective bandwidth and effective capacity which we subsequently employ to formulate the wireless throughput in fading channels in the presence of random arrivals and statistical queueing constraints.

Effective Bandwidth

Effective bandwidth characterizes the minimum constant transmission (or service) rate required to support the given random data arrival process while the buffer overflow proba-

¹Probability of non-empty buffer can be approximated from the ratio of average arrival rate to average service rate [22].

bility is limited or more explicitly the statistical queueing constraint described by (2.1) is satisfied. Let $\{a(k), k = 1, 2, \dots\}$ be a sequence of nonnegative random variables, describing the random arrival rates. Let also the time-accumulated arrival process be denoted by $A(t) = \sum_{k=1}^t a(k)$. Then, the effective bandwidth is given by the asymptotic logarithmic moment generating function of $A(t)$ [6], i.e.,

$$a^*(\theta) = \lim_{t \rightarrow \infty} \frac{1}{\theta t} \log \mathbb{E} \{e^{\theta A(t)}\}. \quad (2.6)$$

In Section 2.2, we describe the effective bandwidth of different source arrival models in detail.

Effective Capacity

Effective capacity, as a dual concept to effective bandwidth, identifies the maximum constant arrival rate that can be supported by a given time-varying service process while satisfying (2.1). Let $\{\nu[k], k = 1, 2, \dots\}$ denote the discrete-time stationary and ergodic stochastic service process and $S[t] \triangleq \sum_{k=1}^t \nu[k]$ be the time-accumulated service process. Then, the effective capacity is given by [10]

$$C_E(\text{SNR}, \theta) = - \lim_{t \rightarrow \infty} \frac{1}{\theta t} \log_e \mathbb{E} \{e^{-\theta S[t]}\}. \quad (2.7)$$

Note that we have assumed that the fading coefficients $\{h_i\}$ change independently from one block of m symbols to another. Under this assumption, effective capacity simplifies to

$$C_E(\text{SNR}, \theta) = -\frac{1}{\theta} \log_e \mathbb{E} \{e^{-\theta \nu}\} \quad (2.8)$$

where ν is the instantaneous service (or equivalently transmission) rate in one block. If the channel input sequence $\{x_i\}$ is an independent and identically distributed (i.i.d.) sequence of Gaussian random variables with zero mean and variance \mathcal{E} , then the service rate can be

written as

$$\nu = \sum_{i=1}^m \log_2(1 + \text{SNR}z_i) \quad (2.9)$$

where we have defined $z_i = |h_i|^2$. Hence, the effective capacity in the units of bits/block is

$$C_E(\text{SNR}, \theta) = -\frac{1}{\theta} \log_e \mathbb{E} \left\{ e^{-\theta \sum_{i=1}^m \log_2(1 + \text{SNR}z_i)} \right\}. \quad (2.10)$$

Remark 1. *In the special case of independent channel coefficients in each block and Rayleigh fading, we can express the effective capacity in closed-form as*

$$C_E(\text{SNR}, \theta) = -\frac{m}{\theta} \log_e \left[\text{SNR}^{-\frac{\theta}{\log_e 2}} e^{\frac{1}{\text{SNR}}} \Gamma \left(1 - \frac{\theta}{\log_e 2}, \frac{1}{\text{SNR}} \right) \right] \quad (2.11)$$

$$= m \log_2(\text{SNR}) - \frac{m}{\theta \text{SNR}} - \frac{m}{\theta} \log_e \Gamma \left(1 - \frac{\theta}{\log_e 2}, \frac{1}{\text{SNR}} \right) \quad (2.12)$$

where $\Gamma(s, w) = \int_w^\infty \tau^{s-1} e^{-\tau} d\tau$ is the upper incomplete gamma function.

2.2 Effective Bandwidths of Different Source Models

Discrete-Time Markov Sources

In this subsection, we consider discrete-time Markov source models. Assume that the transition probability matrix of the n -state irreducible and aperiodic Markov source process is denoted by \mathbf{J} , and λ_i is the arrival rate in state i . Moreover, $\mathbf{\Lambda} = \text{diag} \{ \lambda_1, \lambda_2, \dots, \lambda_n \}$ is the diagonal matrix of arrival rates. Then, the effective bandwidth of this discrete Markov source is given by [9]

$$a(\theta) = \frac{1}{\theta} \log_e [\text{sp} (e^{\theta \mathbf{\Lambda}} \mathbf{J})] \quad (2.13)$$

where $\text{sp}(\cdot)$ is the spectral radius of the input matrix. Note that the stationary distribution $\boldsymbol{\pi}$ can be found from the solution of

$$\begin{aligned}\boldsymbol{\pi}\mathbf{1} &= 1, \\ \boldsymbol{\pi}\mathbf{J} &= \boldsymbol{\pi}\end{aligned}\tag{2.14}$$

where $\boldsymbol{\pi} = [\pi_1, \pi_2, \dots, \pi_n]$ and $\mathbf{1} = [1, \dots, 1]^T$.

In order to unveil the key relationships and tradeoffs, we consider a particularly simple two-state model. We assume that data arrival is either in the ON or OFF state in each block duration of m symbols. When the state is ON, λ bits arrive (i.e., the arrival rate is λ bits/block), while there are no arrivals in the OFF state. For this two-state model, the state transition probability matrix is given as

$$\mathbf{J} = \begin{bmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{bmatrix}.\tag{2.15}$$

Given the above transition matrix \mathbf{J} , the effective bandwidth for this ON-OFF Markov model can be derived as [9]

$$a^*(\theta, \lambda) = \frac{1}{\theta} \log_e \left(\frac{p_{11} + p_{22} e^{\lambda\theta} + \sqrt{(p_{11} + p_{22} e^{\lambda\theta})^2 - 4(p_{11} + p_{22} - 1)e^{\lambda\theta}}}{2} \right)\tag{2.16}$$

where p_{11} denotes the probability of staying in the OFF state from one block to another. Similarly, p_{22} denotes the probability of staying in the ON state. The probabilities of transitioning from one state to a different one are therefore denoted by $p_{21} = 1 - p_{22}$ and $p_{12} = 1 - p_{11}$. For these transition probabilities, we can easily see that the probability of the ON state in the steady state is

$$P_{\text{ON}} = \frac{1 - p_{11}}{2 - p_{11} - p_{22}}.\tag{2.17}$$

Therefore, the average arrival rate is

$$r_{\text{avg}} = \lambda P_{\text{ON}} = \lambda \frac{1 - p_{11}}{2 - p_{11} - p_{22}} \quad (2.18)$$

which is equal to the average departure rate when the queue is in steady state [8].

Markov Fluid Sources

In this subsection, we address Markov fluid sources where the source arrival process is modeled as a continuous-time Markov chain. Assume that \mathbf{G} is the irreducible transition rate matrix of the Markov chain, λ_i is the arrival rate in the i^{th} state, and $\mathbf{\Lambda} = \text{diag}\{\lambda_1, \lambda_2, \dots, \lambda_n\}$. Then, the effective bandwidth of this source is given by [11], [12]

$$a^*(\theta) = \mu\left(\mathbf{\Lambda} + \frac{1}{\theta}\mathbf{G}\right) \quad (2.19)$$

where $\mu(\cdot)$ denotes the maximum real eigenvalue of the input matrix. We also note that the stationary distribution $\boldsymbol{\pi}$ of the continuous-time Markov chain can be found by solving

$$\begin{aligned} \boldsymbol{\pi}\mathbf{1} &= 1, \\ \boldsymbol{\pi}\mathbf{G} &= \mathbf{0} \end{aligned} \quad (2.20)$$

where $\boldsymbol{\pi} = [\pi_1, \pi_2, \dots, \pi_n]$, $\mathbf{0} = [0, \dots, 0]^T$ and $\mathbf{1} = [1, \dots, 1]^T$.

In order to derive closed-form expressions in our analysis, we again consider two states (ON/OFF). When there is no arrival, the state is OFF. When the state is ON, the arrival rate is λ bits/block. The transition rate matrix for a two-state Markov fluid is in the form of

$$\mathbf{G} = \begin{bmatrix} -\alpha & \alpha \\ \beta & -\beta \end{bmatrix}, \quad (2.21)$$

where α is the transition rate from OFF state to ON state whereas β is the transition rate

from ON state to OFF state. Using (2.19), we can express the effective bandwidth as

$$a^*(\theta) = \frac{1}{2\theta} \left[\theta\lambda - (\alpha + \beta) + \sqrt{(\theta\lambda - (\alpha + \beta))^2 + 4\alpha\theta\lambda} \right]. \quad (2.22)$$

The probability of ON state, π_2 , is required to define the average rate. Inserting the generator matrix \mathbf{G} in (2.21) into (2.20), we obtain the ON state probability as

$$\pi_2 = P_{\text{ON}} = \frac{\alpha}{\alpha + \beta}. \quad (2.23)$$

Therefore, the average arrival rate of the two-state Markov fluid process is

$$r_{\text{avg}} = \lambda P_{\text{ON}} = \lambda \frac{\alpha}{\alpha + \beta}. \quad (2.24)$$

Continuous-Time Markov Modulated Poisson Sources

In this subsection, we assume that the data arrival to the buffer is a Poisson process whose intensity is controlled by a continuous-time Markov chain. For instance, the intensity of the Poisson arrival process is λ_i in the i^{th} state of the Markov chain. Therefore, the source arrival is modeled as a Markov-modulated Poisson process (MMPP). Assuming that the \mathbf{G} is the irreducible transition rate matrix of the Markov chain and $\mathbf{\Lambda} = \text{diag} \{ \lambda_1, \lambda_2, \dots, \lambda_n \}$ is the diagonal matrix of the intensities of the Poisson arrivals in different states, the effective bandwidth is given by [11], [12]

$$a^*(\theta) = \frac{1}{\theta\mu} \left((e^\theta - 1) \mathbf{\Lambda} + \mathbf{G} \right). \quad (2.25)$$

As in previous sections, we consider a two-state (ON/OFF) model in which there are no arrivals in the OFF state (i.e., the intensity is 0) and the intensity of the Poisson arrival process is λ bits/block in the ON state. Assuming the same generator matrix \mathbf{G} as in (2.21),

we can express the effective bandwidth as

$$a^*(\theta) = \frac{1}{2\theta} [(e^\theta - 1)\lambda - (\alpha + \beta)] + \frac{1}{2\theta} \sqrt{[(e^\theta - 1)\lambda - (\alpha + \beta)]^2 + 4\alpha(e^\theta - 1)\lambda}. \quad (2.26)$$

Note that the average arrival rate in bits/block is again given by

$$r_{\text{avg}} = \lambda P_{\text{ON}} = \lambda \frac{\alpha}{\alpha + \beta}. \quad (2.27)$$

We further note that if the transition rate $\beta = 0$, then we have $P_{\text{ON}} = 1$. In this case, MMPP model specializes to a pure Poisson source with intensity λ bits/block, and the effective bandwidth of this source is given by

$$a^*(\theta) = \frac{1}{\theta} (e^\theta - 1)\lambda. \quad (2.28)$$

Discrete-Time Markov Modulated Poisson Sources

In this source model, the data arrival to the buffer is a Poisson process whose intensity is controlled by a discrete-time Markov chain. Again, the intensity of the Poisson arrival process is λ_i in the i^{th} state of the Markov chain and the source arrival is modeled as a Markov-modulated Poisson process (MMPP). Assume that the transition probability matrix of the n -state irreducible and aperiodic Markov source process is denoted by \mathbf{J} and $\mathbf{\Lambda} = \text{diag}\{\lambda_1, \lambda_2, \dots, \lambda_n\}$ is the diagonal matrix of the intensities of the Poisson arrivals in different states, the effective bandwidth is given by [11], [12]

$$a(\theta) = \frac{1}{\theta} \log_e \left[\text{sp} \left(e^{(e^\theta - 1)\mathbf{\Lambda}} \mathbf{J} \right) \right] \quad (2.29)$$

We again consider a two-state model in which the intensity of the Poisson arrival process is λ and 0 in the ON and OFF states of the Markov chain, respectively. Therefore, the source

arrival is modeled as a Markov-modulated Poisson process (MMPP). Assuming that the matrix \mathbf{J} in (2.15) is the transition probability matrix of the Markov chain, the effective bandwidth is given by [6].

$$a(\theta) = \frac{1}{\theta} \log_e \left(\frac{p_{11} + p_{22}e^{\lambda(e^\theta - 1)}}{2} + \frac{\sqrt{(p_{11} + p_{22}e^{\lambda(e^\theta - 1)})^2 - 4(p_{11} + p_{22} - 1)e^{\lambda(e^\theta - 1)}}}{2} \right) \quad (2.30)$$

CHAPTER 3

WIRELESS THROUGHPUT AND ENERGY EFFICIENCY WITH RANDOM ARRIVALS AND STATISTICAL QUEUEING CONSTRAINTS

3.1 Throughput with Markovian Source Models

In this section, we formulate the throughput of wireless fading channels when the data arrivals are random and statistical queueing constraints are imposed. More specifically, we consider Markovian arrival models introduced in Section 2.2, namely discrete-time Markov sources, Markov fluids and Markov-modulated Poisson arrivals. The states in these Markov processes are differentiated by the corresponding arrival rates in these states, e.g., the arrival rate in the i^{th} state is λ_i . If the stationary distribution of the Markov process is denoted by

π , the average arrival rate in an n -state Markov source model simply becomes

$$r_{\text{avg}} = \sum_{i=1}^n \pi_i \lambda_i \quad (3.1)$$

which is equal to the average departure rate when the queue is in steady state [8].

We seek to determine the throughput by identifying the maximum average arrival rate that can be supported by the fading channel described in Section 3.1.1 while satisfying the statistical QoS limitations given in the form in (2.1). As shown in [8, Theorem 2.1], (2.1) is satisfied, i.e., buffer violation probability decays exponentially fast with rate controlled by the QoS exponent θ , if the effective bandwidth of the arrival process is equal to the effective capacity of the service process, i.e.,

$$a^*(\theta) = C_E(\text{SNR}, \theta). \quad (3.2)$$

Hence by solving (3.2), we can determine the maximum average arrival rate $r_{\text{avg}}^*(\text{SNR}, \theta)$. By specifying the effective bandwidth of different source models and incorporating the effective capacity of time-varying wireless transmissions in (2.10), the maximum average arrival rate can be determined for general n -state Markovian source models. Indeed, several n -state source models are addressed in Section 3.2. However, in our analysis in this section, to illustrate the impact of the arrival and system parameters in a lucid setting, we concentrate on the two-state (ON-OFF) arrival models and provide closed-form expressions for the maximum average arrival rates in terms of the source parameters and the effective capacity of the wireless transmissions. We also identify the characteristics of the throughput in the low- θ and high-SNR regimes. We note that the analysis throughout this section is applicable to any arbitrary fading correlation within each fading block, with the exception of high-SNR characterizations which are obtained under the assumption of i.i.d. fading.

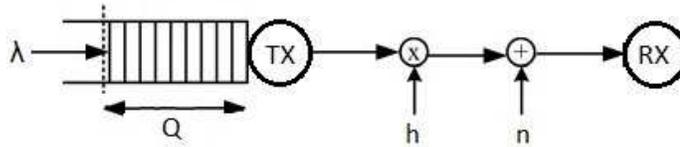


Fig. 3.1: System Model.

3.1.1 Channel Model

As depicted in Fig. 3.1, we consider a point-to-point link with a single transmitter and single receiver. In this system, the data generated by the source is initially stored in a buffer at the transmitter before it is transmitted over a wireless channel. We consider a flat-fading channel between the transmitter and receiver, and assume a block-fading model with a block duration of m symbols. Hence, fading varies independently from one block to another. On the other hand, we further assume that in each block duration of m symbols, fading can be arbitrarily correlated. The channel input-output relation within each block can be expressed as

$$y_i = h_i x_i + n_i \text{ for } i = 1, 2, \dots, m \quad (3.3)$$

where x_i and y_i are the channel input and output, respectively. The average energy of the input is \mathcal{E} , i.e.,

$$\mathbb{E}\{|x_i|^2\} = \mathcal{E}. \quad (3.4)$$

n_i denotes the zero-mean, circularly-symmetric, complex Gaussian noise with variance $\mathbb{E}\{|n_i|^2\} = N_0$. Hence, the signal-to-noise ratio is

$$\text{SNR} = \frac{\mathbb{E}\{|x|^2\}}{\mathbb{E}\{|n|^2\}} = \frac{\mathcal{E}}{N_0}. \quad (3.5)$$

Above in (3.3), h_i denotes the fading coefficient. Fading coefficients are assumed to be identically distributed, and the fading distribution can be arbitrary with finite variance.

While the ensuing analysis is applicable to a general class of fading distributions, we use a Gauss-Markov fading model in the numerical results and assume that the Gaussian fading coefficients in each block of m symbols follow the correlation pattern $h_i = \rho h_{i-1} + w_i$ where w_i is an independent, zero-mean Gaussian random variable with variance $\mathbb{E}\{|w_i|^2\} = (1 - \rho^2)\sigma_h^2$, $\rho \in [0, 1]$, and σ_h^2 is the common variance of the fading coefficients $\{h_i\}$. Note that when $\rho = 1$, we have full correlation, whereas $\rho = 0$ models the case of independent fading.

3.1.2 Discrete-Time Markov Sources

In this section, we consider two-state (ON/OFF) discrete Markov sources described in Section 2.2, and initially characterize the maximum average arrival rate r_{avg}^* that can be supported by the fading channel while satisfying the statistical QoS limitations given in the form in (2.1).

Theorem 3.1.1. *For the two-state (ON/OFF) discrete Markov source, the maximum average arrival rate (in bits/block) as a function of the QoS exponent θ , effective capacity of the fading channel $C_E(\text{SNR}, \theta)$, and the state transition probabilities is expressed as*

$$r_{\text{avg}}^*(\text{SNR}, \theta) = \frac{P_{\text{ON}}}{\theta} \log_e \left(\frac{e^{2\theta C_E(\text{SNR}, \theta)} - p_{11} e^{\theta C_E(\text{SNR}, \theta)}}{1 - p_{11} - p_{22} + p_{22} e^{\theta C_E(\text{SNR}, \theta)}} \right). \quad (3.6)$$

Proof: See Appendix A.1.

Note that r_{avg}^* above is formulated in terms of the effective capacity, C_E , of wireless transmissions. In Fig. 3.2, we plot the the maximum average arrival rate as a function of the effective capacity for different source characteristics when $\theta = 1$. It is easy to verify that when $P_{\text{ON}} = \frac{1-p_{11}}{2-p_{11}-p_{22}} = 1$ or equivalently $p_{22} = 1$, (3.6) simplifies to $r_{\text{avg}}^*(\text{SNR}, \theta) = C_E(\text{SNR}, \theta)$. Hence, when the source is always ON and therefore the arrivals are at a constant

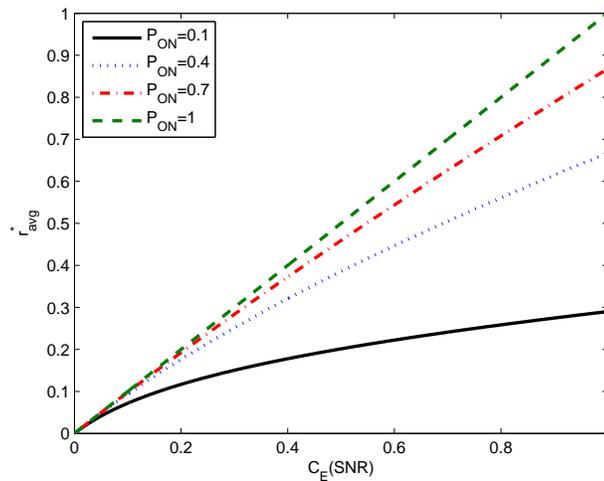


Fig. 3.2: Maximum average arrival rate r_{avg}^* vs. effective capacity $C_E(\text{SNR})$ for different source statistics. No fading correlation, i.e., $\rho = 0$. $\theta = 1$.

rate, maximum average arrival rate is equal to the effective capacity, as also observed in Fig. 3.2. On the other hand, we notice in this figure that as P_{ON} diminishes and the source becomes more bursty, throughput diminishes as well and smaller average arrival rates are supported for given effective capacity.

As also indicated in the above discussion and seen in (3.6), $r_{\text{avg}}^*(\text{SNR}, \theta)$ is in general a function of the state transition probabilities of the Markov arrival process in the presence of buffer constraints. On the other hand, as shown in the following result, this dependence disappears if no buffer constraints are imposed, i.e., when $\theta = 0$.

Theorem 3.1.2. *As the statistical queueing constraints are relaxed by letting the QoS exponent θ approach zero, the maximum average arrival rate converges to*

$$\lim_{\theta \rightarrow 0} r_{\text{avg}}^*(\text{SNR}, \theta) = \sum_{i=1}^m \mathbb{E} \{ \log_2(1 + \text{SNR}z_i) \} \text{ bits/block.} \quad (3.7)$$

Moreover, the first derivative of r_{avg}^* with respect to θ at $\theta = 0$ is

$$\begin{aligned} \left. \frac{\partial r_{\text{avg}}^*(\text{SNR}, \theta)}{\partial \theta} \right|_{\theta=0} &= -\frac{1}{2} \sum_{i,j=1}^m \text{cov}\{\log_2(1 + \text{SNR}z_i), \log_2(1 + \text{SNR}z_j)\} \\ &\quad - \frac{\eta}{2} \left(\sum_{i=1}^m \mathbb{E}\{\log_2(1 + \text{SNR}z_i)\} \right)^2 \end{aligned} \quad (3.8)$$

where we define η as

$$\eta = \frac{(1 - p_{22})(p_{11} + p_{22})}{(1 - p_{11})(2 - p_{11} - p_{22})}. \quad (3.9)$$

Proof: See Appendix A.2.

We see from (3.7) that if no statistical buffer constraints are imposed i.e., if $\theta = 0$, then the maximum average arrival rate is equal to the Shannon capacity of the block-fading channel, and therefore is independent of the statistical characteristics of the discrete Markov arrival model. Moreover, the dependence of the maximum arrival rate in this regime on the channel statistics is only through the marginal distributions of the fading coefficients. Hence, channel correlation in each fading block does not play any role. However, this radically changes when $\theta > 0$. For instance, we notice from (3.8) that even with a small increase in θ , r_{avg}^* starts varying with the source and channel statistics, as exemplified by the dependence of the first derivative on η and the covariance function.

Having discussed the low- θ regime above, we next provide a characterization of $r_{\text{avg}}^*(\text{SNR}, \theta)$ at high SNR values for i.i.d. Rayleigh fading.

Theorem 3.1.3. *Assume that the channel fading coefficients are i.i.d. in each block and fading power $z = |h|^2$ is exponentially distributed with unit mean (i.e., Rayleigh fading is experienced). Then, we have*

$$\frac{1}{m} r_{\text{avg}}^*(\text{SNR}, \theta) = \begin{cases} \frac{P_{ON}}{\theta \log_2 e} \log_2 \text{SNR} + \mathcal{O}(1) & \text{if } \theta > \frac{1}{\log_2 e} \\ P_{ON} \log_2 \text{SNR} + \mathcal{O}(1) & \text{if } 0 < \theta < \frac{1}{\log_2 e} \\ \log_2 \text{SNR} + \mathcal{O}(1) & \text{if } \theta = 0 \end{cases} \quad (3.10)$$

as $\text{SNR} \rightarrow \infty$.

Proof: See Appendix A.3.

Note that the high-SNR slope is defined as [55]

$$\mathcal{S}_\infty = \lim_{\text{SNR} \rightarrow \infty} \frac{\frac{1}{m} r_{\text{avg}}^*(\text{SNR}, \theta)}{\log_2 \text{SNR}}. \quad (3.11)$$

Theorem 3.1.3 shows that the high-SNR slope of the maximum arrival rate for the two-state discrete Markov source that can be supported in the i.i.d Rayleigh fading channel is

$$\mathcal{S}_\infty = \begin{cases} \frac{P_{\text{ON}}}{\theta \log_2 e} & \text{if } \theta > \frac{1}{\log_2 e} \\ P_{\text{ON}} & \text{if } 0 < \theta < \frac{1}{\log_2 e} \\ 1 & \text{if } \theta = 0 \end{cases}. \quad (3.12)$$

It is interesting to observe from Theorem 3.1.2 that when no buffer constraints are imposed i.e., when $\theta = 0$, the high-SNR slope is $\mathcal{S}_\infty = 1$, again independent of the source statistics. On the other hand, when $\theta > 0$, \mathcal{S}_∞ becomes proportional to the ON probability and is now less than one unless the arrival rate is constant. Furthermore, for θ values greater than $\frac{1}{\log_2 e}$, \mathcal{S}_∞ starts decreasing with increasing θ . Hence, the result in Theorem 3.1.3 quantifies the performance degradation experienced at high SNR levels due to source randomness and statistical buffer constraints.

Let us further simplify the source model and set $p_{11} = 1 - s$ and $p_{22} = s$. The source is now described by the single parameter s . Notice that with this choice we have $P_{\text{ON}} = s$ and hence s becomes a measure of the burstiness of the source. The smaller the s , the less frequently the data arrives and the more bursty the source becomes. At the other extreme, if $s = 1$, source is ON all the time and we have constant arrival rate. Furthermore, with the above choice of p_{11} and p_{22} , the expression for the maximum average arrival rate simplifies

to

$$r_{\text{avg}}^*(\text{SNR}, \theta) = \frac{s}{\theta} \log_e \left(\frac{e^{\theta C_E(\text{SNR}, \theta)} - (1 - s)}{s} \right), \quad (3.13)$$

which can readily be seen to be a diminishing function as s decreases. Therefore, source burstiness generally hurts the throughput if we keep all other variables fixed.

We can further observe this in Fig. 3.3, where we plot the maximum average arrival rate (or equivalently the throughput) as a function of SNR for different values of s and the QoS exponent θ . Numerical analysis verifies that as the source becomes more bursty with lower values of s , throughput diminishes. Conversely, throughput is maximized when $s = 1$ i.e., when we have constant arrival rates. It is also interesting to notice from (3.13) that the arrival rate in the ON state, which is given by $\lambda^* = \frac{r_{\text{avg}}^*(\text{SNR}, \theta)}{s}$, increases as s diminishes. Hence, smaller s implies that data arrives less frequently but with bursts of increased rates. We also observe in Fig. 3.3 that the throughput reduction due to burstiness is more severe at high SNRs. This is indeed a consequence of the fact that high-SNR slope gets smaller as $P_{\text{ON}} = s$ decreases, as discussed above. Finally, we see in Fig. 3.3 that performance degradation is experienced as θ increases and hence stricter buffer constraints are imposed.

In Fig. 3.4, we plot the SNR levels required to support a given average arrival rate as a function of the ON-state probability for different values of the QoS exponent θ . We observe that as P_{ON} decreases and hence the source becomes more bursty, required SNR level increases in general. Interestingly, a sharper increase is experienced under stricter buffer constraints (e.g., when $\theta = 0.5$ rather than $\theta = 0.1$), indicating higher power/energy costs in these cases.

The low- θ regime is investigated in Fig. 3.5 where we plot the maximum average arrival rate r_{avg}^* vs. QoS exponent θ for different P_{ON} values. We set $\text{SNR} = 1$. We notice that all three curves converge to the same throughput value $r_{\text{avg}}^*(0)$ as $\theta \rightarrow 0$, confirming the result

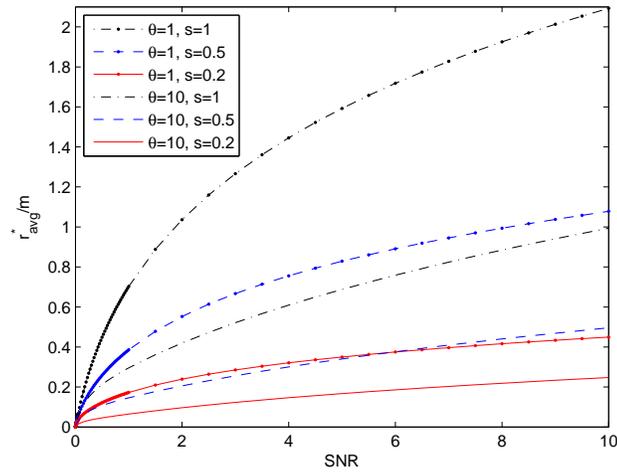


Fig. 3.3: Maximum average arrival rate r_{avg}^* vs. signal-to-noise ratio SNR for different values of θ and source statistics. No fading correlation, i.e., $\rho = 0$.

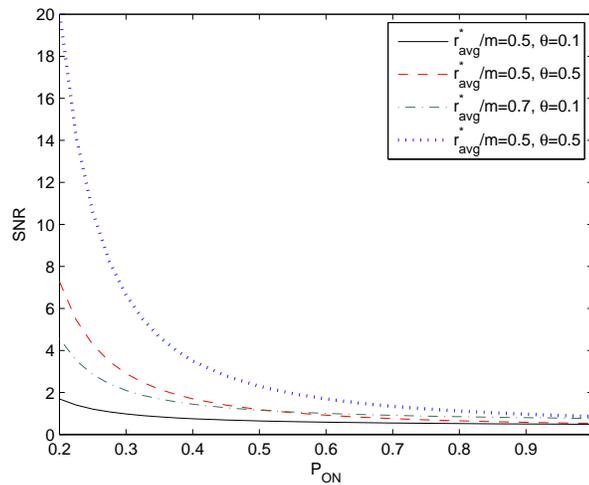


Fig. 3.4: Required SNR vs. ON probability, P_{ON} , for a given fixed average arrival rate. No fading correlation, i.e., $\rho = 0$.

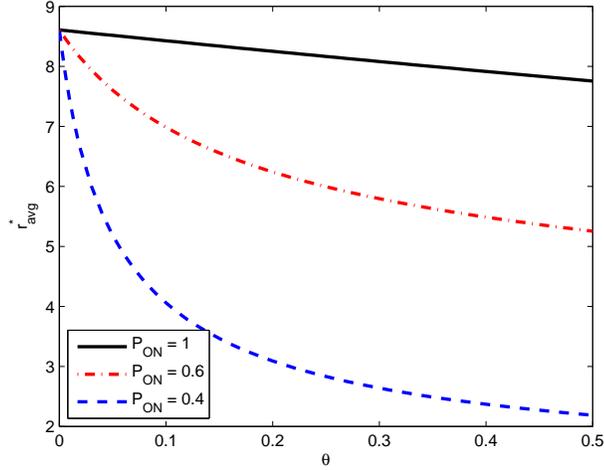


Fig. 3.5: Maximum average arrival rate r_{avg}^* vs. QoS exponent θ for different values of P_{ON} . No fading correlation, i.e., $\rho = 0$.

in (3.7). Hence, source characteristics do not affect the throughput if no queuing constraints are imposed. As θ increases, throughput diminishes and the reduction in r_{avg}^* is more severe for more bursty sources (e.g., when $P_{\text{ON}} = 0.4$). We notice that, as predicted by (3.8), this is already reflected by the different slopes of r_{avg}^* in the vicinity of $\theta = 0$. Hence, overall the system for more bursty sources becomes more cautious and supports smaller average arrival rates in order to avoid buffer overflows.

In Fig. 3.6, we again plot the throughput as a function θ but for different values of ρ , which quantifies the correlation between fading coefficients in each fading block. We fix $\text{SNR} = 1$ and set $P_{\text{ON}} = 0.5$. Similar to burstiness, fading correlation does not have any effect on the throughput when $\theta = 0$. When $\theta > 0$, higher correlation (i.e., larger ρ) results in lower supported throughput under the same QoS constraints.

Finally, we have conducted simulations to further verify the theoretical analysis and results. In particular, in the simulations, for fixed QoS exponent θ , SNR, and state transition probabilities p_{11} and p_{22} of the ON/OFF discrete Markov source, we initially determine the maximum average arrival rate from (3.6) and the corresponding maximum arrival rate in the ON state. Then, using the given statistical characterizations and the maximum ar-

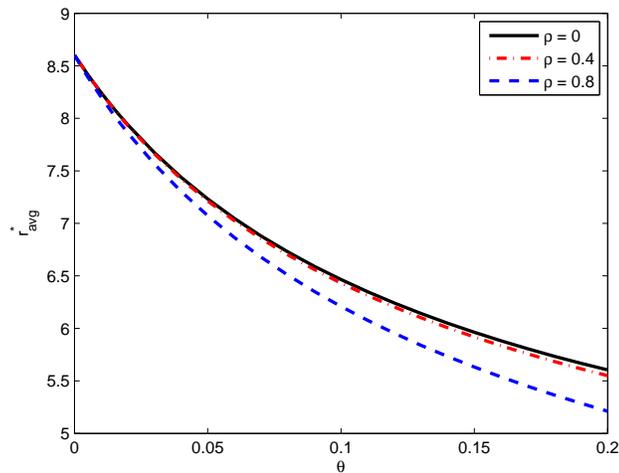


Fig. 3.6: Maximum average arrival rate r_{avg}^* vs. QoS exponent θ for different fading correlations. $P_{\text{ON}} = 0.5$.

rival rate, we generate random Markov arrivals and assume that the arriving data is initially stored in the buffer before being transmitted. Transmission rates are simulated by generating realizations of i.i.d. Gaussian fading coefficients. Throughout this process, we track the queue evolution and the buffer state (i.e., the queue length) as the Markov arrivals occur (and hence more data gets stored) and transmissions at varying rates according to the generated fading coefficients are performed, clearing some data off the buffer. In Figs. 3.7 and 3.8, we plot the simulated buffer overflow probability $\Pr\{Q \geq q\}$ and delay violation probability $\Pr\{D > d\}$, respectively, as functions of the corresponding thresholds, following 10^7 runs of the simulation. We notice that while the theoretical analysis makes use of results from the theory of large deviations and is generally applicable for large thresholds, the simulation results are interestingly in excellent agreement with the theoretical predictions even at small values of the thresholds. For instance, we note from (2.3) that $\log_e \Pr\{Q \geq q\} \approx \log_e \varsigma - \theta q$ and hence is expected to decay linearly in q with slope θ . We indeed observe this linear decay in Fig. 3.7 (where the overflow probabilities are plotted in logarithmic scale) for even small to moderate values of q . Moreover, the slopes of the simulated curves, denoted by θ_{sim} , are very close to the originally selected value of

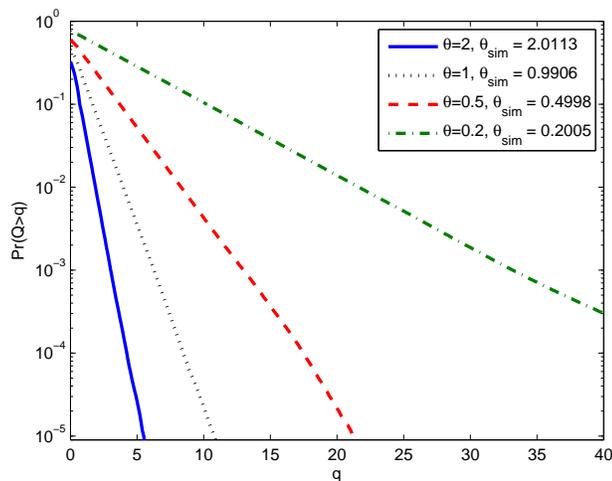


Fig. 3.7: Buffer overflow probability $\Pr\{Q > q\}$ vs. buffer threshold q for different values of θ . $p_{11} = p_{22} = 0.8$, SNR = 0 dB

θ . Similar conclusions apply to Fig. 3.8 as well. In this figure, delay violation probabilities are determined by keeping track of the delay experienced by the data stored in the buffer until transmission. We again notice that the logarithm of the delay violation probability decays linearly with threshold d (or equivalently the delay violation probability diminishes exponentially with d). Note that the slope of the linear decay is predicted from (2.5) to be $\theta a^*(\theta)$ where $a^*(\theta)$ is the effective bandwidth of the source. Again, the slope of the simulated curves are almost the same as this theoretical slope value, as indicated in the legend on the figure.

3.1.3 Markov Fluid Sources

In this section, we consider Markov fluid sources. In the following, we go through similar steps as in the previous subsection and initially determine the maximum average arrival rates of ON/OFF Markov fluid sources that can be supported by the wireless channel as a function of the source transition rates and the effective capacity of wireless transmissions. Subsequently, we give characterizations of the maximum average arrival rates in the low- θ and high-SNR regimes.

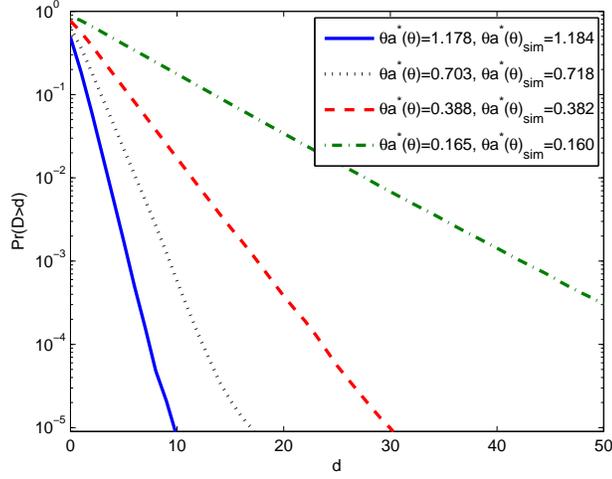


Fig. 3.8: Delay violation probability $\Pr\{D > d\}$ vs. delay threshold d for different values of θ . $p_{11} = p_{22} = 0.8$, $\text{SNR} = 0$ dB

Theorem 3.1.4. *For the two-state (ON/OFF) Markov fluid source, the maximum average arrival rate is given as*

$$r_{avg}^*(\text{SNR}, \theta) = P_{ON} \frac{\theta C_E(\text{SNR}, \theta) + \alpha + \beta}{\theta C_E(\text{SNR}, \theta) + \alpha} C_E(\text{SNR}, \theta). \quad (3.14)$$

Proof: See Appendix A.4.

Note that maximum average arrival rate generally depends on the transition rate matrix of the Markov fluid source. At the same time, similar to the discrete case, when there are no QoS constraints, source characteristics do not have any impact on the throughput. However, this changes drastically when $\theta > 0$ even if θ is vanishingly small. These properties are demonstrated analytically in the result below.

Theorem 3.1.5. *As the statistical queueing constraints are relaxed by letting the QoS exponent θ approach zero, we have*

$$\lim_{\theta \rightarrow 0} r_{avg}^*(\text{SNR}, \theta) = \sum_{i=1}^m \mathbb{E} \{ \log_2(1 + \text{SNR}z_i) \} \text{ bits/block}, \quad (3.15)$$

and

$$\begin{aligned} \left. \frac{\partial r_{avg}^*(SNR, \theta)}{\partial \theta} \right|_{\theta=0} &= -\frac{1}{2} \sum_{i,j=1}^m \text{cov}\{\log_2(1 + SNRz_i), \log_2(1 + SNRz_j)\} \\ &\quad - \frac{\zeta}{2} \left(\sum_{i=1}^m \mathbb{E}\{\log_2(1 + SNRz_i)\} \right)^2 \end{aligned} \quad (3.16)$$

where ζ is defined as

$$\zeta = \frac{2\beta}{\alpha(\alpha + \beta)}. \quad (3.17)$$

Proof: See Appendix A.5.

We note that when $\theta > 0$, r_{avg}^* depends on the source and channel statistics. In (3.16), we observe the dependence of even the first derivative on channel correlations and source statistics via the covariance function and the parameter ζ , respectively.

Next we present a high-SNR characterization of the throughput for Rayleigh fading.

Theorem 3.1.6. *Assume that the channel fading coefficients are i.i.d. in each block and fading power $z = |h|^2$ is exponentially distributed with unit mean (i.e., Rayleigh fading is experienced). Then, we have*

$$\frac{1}{m} r_{avg}^*(SNR, \theta) = \begin{cases} \frac{P_{ON}}{\theta \log_2 e} \log_2 SNR + \mathcal{O}(1) & \text{if } \theta > \frac{1}{\log_2 e} \\ P_{ON} \log_2 SNR + \mathcal{O}(1) & \text{if } 0 < \theta < \frac{1}{\log_2 e} \\ \log_2 SNR + \mathcal{O}(1) & \text{if } \theta = 0 \end{cases} \quad (3.18)$$

as $SNR \rightarrow \infty$.

The proof of Theorem 3.1.6 is omitted due to its similarity to the proof of Theorem 3.1.3 in Appendix A.3. Similar conclusions as in Section 2.2 immediately apply.

Note that the throughput expression in (3.14) suggests that for sufficiently high SNR

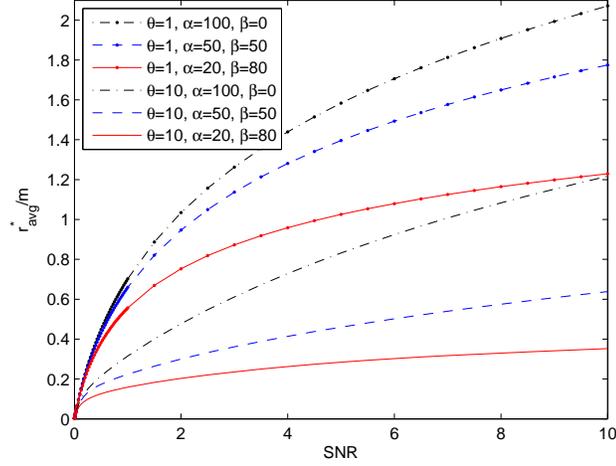


Fig. 3.9: Maximum average arrival rate r_{avg}^* vs. signal-to-noise ratio SNR for different values of θ and source statistics. No fading correlation, i.e., $\rho = 0$.

levels leading to $\theta C_E(\text{SNR}, \theta) \gg \alpha + \beta$, we have

$$r_{\text{avg}}^*(\text{SNR}, \theta) \approx P_{\text{ON}} C_E(\text{SNR}, \theta). \quad (3.19)$$

Hence, at high SNRs, the maximum average arrival rate depends on the source statistics only through the ON probability. This is noted in the high-SNR behavior in (3.18) as well.

In Fig. 3.9, we plot r_{avg}^* vs. SNR curves for different α , β , and θ values. We immediately observe that throughput diminishes with increasing θ and decreasing P_{ON} . In Fig. 3.10, we analyze the effect of r_{avg} , P_{ON} , and $\alpha + \beta$ on the required SNR levels. For Markov fluid sources, ON state probability is not the sole indicator of burstiness. Having low α and β values also indicates that source is more bursty as the transition between ON and OFF states becomes less frequent. Hence, OFF state can be more persistent. When α and β are large, state transitions occur more rapidly, leading to lower required SNR levels. Again, we notice that the burstiness is harmful for the system.

In Fig. 3.11, we plot the maximum average arrival rate r_{avg}^* as a function of θ for different values of α and β . Notice that by keeping $\alpha = \beta$, the ON probability P_{ON} is

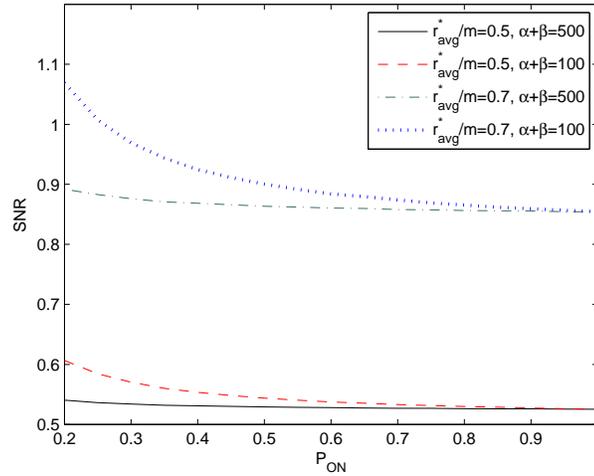


Fig. 3.10: Required SNR vs. ON probability, P_{ON} , for a given average arrival rate. $\theta = 0.5$. No fading correlation, i.e., $\rho = 0$.

fixed at 0.5, while average durations of ON and OFF states vary as the values of $\alpha = \beta$ change. For example, higher α and β values lead to shorter periods for ON and OFF states on average. As an outcome of this fact, we observe in the figure that higher throughput is achieved with sources having higher $\alpha + \beta$.

3.1.4 Continuous-Time Markov Modulated Poisson Sources

Now, we address two-state (ON/OFF) MMPP sources. Similarly as for the previous source models, we determine the maximum average arrival rate of the MMPP source, which can be supported by the fading channel in the presence of QoS constraints, and investigate the throughput in the low- θ and high-SNR regimes. The results can be immediately specialized to pure Poisson sources by setting $\beta = 0$.

Theorem 3.1.7. *For the two-state (ON/OFF) MMPP source model, the maximum average arrival rate is*

$$r_{avg}^*(SNR, \theta) = P_{ON} \frac{\theta [C_E(SNR, \theta) + \alpha + \beta]}{(e^\theta - 1) [\theta C_E(SNR, \theta) + \alpha]} C_E(SNR, \theta). \quad (3.20)$$

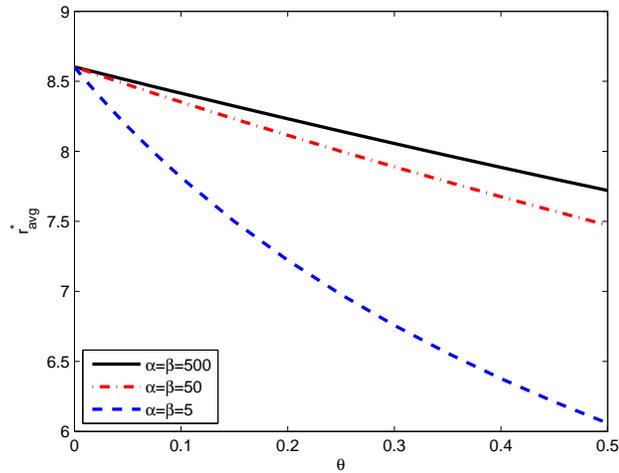


Fig. 3.11: Maximum average arrival rate r_{avg}^* vs. QoS exponent θ for different values of α and β . $P_{\text{ON}} = 0.5$, $\rho = 0$, and $\text{SNR} = 0$ dB.

Proof: See Appendix A.6.

It is interesting to observe that the throughput with the MMPP source is almost identical to that with the Markov fluid source model, save only for the multiplicative factor $\frac{\theta}{e^\theta - 1}$ in (3.20). Note that $\frac{\theta}{e^\theta - 1} < 1$ for $\theta > 0$ and diminishes exponentially fast with increasing θ . Hence, the throughput is generally smaller with MMPP sources and decreases fast with θ . This can be attributed to the much more randomness/burstiness we experience with an MMPP source with respect to the previous Markov models. Note that the arrival rate in the ON state, rather than being a constant as in the previous cases, is determined by a Poisson process. Hence, the presence of the term $\frac{\theta}{e^\theta - 1}$ is due to this Poisson property. Indeed, if we have a pure Poisson source, the maximum average arrival rate is $r_{\text{avg}}^*(\text{SNR}, \theta) = \frac{\theta}{(e^\theta - 1)} C_E(\text{SNR}, \theta)$ obtained by setting $\beta = 0$. The cost of this additional randomness is reflected in the following results as well.

Theorem 3.1.8. *As the statistical queueing constraints are relaxed by letting the QoS ex-*

ponent θ approach zero, we have

$$\lim_{\theta \rightarrow 0} r_{avg}^*(SNR, \theta) = \sum_{i=1}^m \mathbb{E} \{ \log_2(1 + SNRz_i) \} \text{ bits/block}, \quad (3.21)$$

and

$$\begin{aligned} \left. \frac{\partial r_{avg}^*(SNR, \theta)}{\partial \theta} \right|_{\theta=0} &= -\frac{1}{2} \sum_{i,j=1}^m \text{cov} \{ \log_2(1 + SNRz_i), \log_2(1 + SNRz_j) \} \\ &\quad - \frac{\zeta}{2} \left(\sum_{i=1}^m \mathbb{E} \{ \log_2(1 + SNRz_i) \} \right)^2 - \frac{1}{2} \sum_{i=1}^m \mathbb{E} \{ \log_2(1 + SNRz_i) \} \end{aligned} \quad (3.22)$$

where

$$\zeta = \frac{2\beta}{\alpha(\alpha + \beta)}. \quad (3.23)$$

Proof: See Appendix A.7.

When the system is free of QoS limitations, the maximum average arrival rate for the MMPP source again turns out to be equal to the Shannon capacity. However, the throughput has a steeper decline in the low- θ regime due to the third term on the right-hand side of (3.22).

Theorem 3.1.9. *Assume that the channel fading coefficients are i.i.d. in each block and fading power $z = |h|^2$ is exponentially distributed with unit mean (i.e., Rayleigh fading is experienced). Then, we have*

$$\frac{1}{m} r_{avg}^*(SNR, \theta) = \begin{cases} \frac{P_{ON}}{(e^\theta - 1) \log_2 e} \log_2 SNR + \mathcal{O}(1) & \text{if } \theta > \frac{1}{\log_2 e} \\ \frac{\theta}{e^\theta - 1} P_{ON} \log_2 SNR + \mathcal{O}(1) & \text{if } 0 < \theta < \frac{1}{\log_2 e} \\ \log_2 SNR + \mathcal{O}(1) & \text{if } \theta = 0 \end{cases} \quad (3.24)$$

as $SNR \rightarrow \infty$.

Since the ratio between the MMPP throughput and Markov fluid throughput always

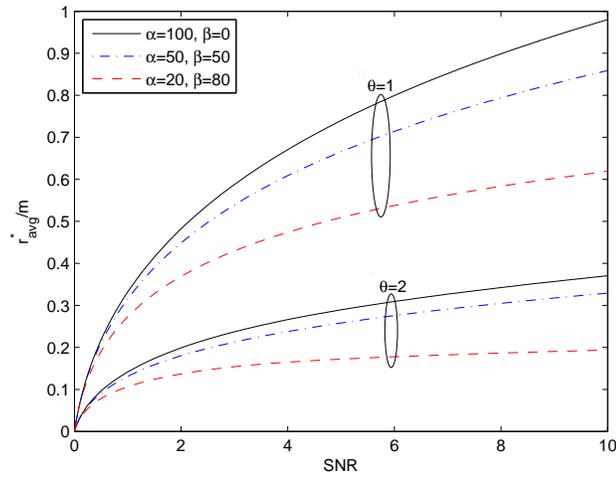


Fig. 3.12: Maximum average arrival rate r_{avg}^* vs. signal-to-noise ratio SNR for different values of θ and different source statistics. No fading correlation, i.e., $\rho = 0$.

stays at $\frac{\theta}{e^\theta - 1}$, we can immediately obtain the above high-SNR characterization, using the formulations in (3.18).

In the numerical results, we have similar conclusions as in the Markov fluid case. The primary difference is the reduced throughput for given θ , which, for instance, is readily seen when we compare Figs. 3.9 and 3.12, where we have throughput vs. SNR curves for Markov fluid and MMPP sources, respectively. In Fig. 3.13, we display the maximum average arrival rate r_{avg}^* as a function of θ . We set $\alpha + \beta = 100$ and $\text{SNR} = 1$, and vary α and β and hence the ON probability. We note that as P_{ON} decreases, the performance degrades faster with increasing θ , as indicated by the steeper slopes.

3.1.5 Discrete-Time Markov Modulated Poisson Process

In order to determine the maximum average arrival rate in terms of $C_E(\theta)$, we insert the effective bandwidth expression in (2.30) into (3.2) and obtain

$$\left(p_{11} + p_{22}e^{r(e^\theta - 1)} - 2e^{\theta C_E} \right)^2 = (p_{11} + p_{22}e^{r(e^\theta - 1)})^2 - 4(p_{11} + p_{22} - 1)e^{r(e^\theta - 1)}. \quad (3.25)$$

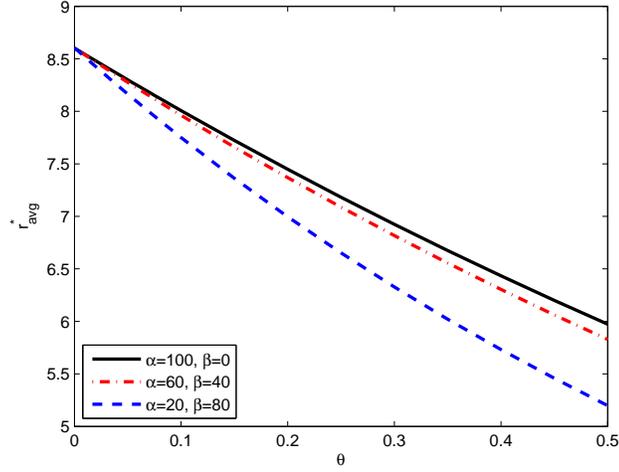


Fig. 3.13: Maximum average arrival rate r_{avg}^* vs. QoS exponent θ for different source statistics. $\rho = 0$.

After solving the above equation for r , we obtain the maximum average arrival rate as

$$r_{\text{avg}}^*(\theta) = \frac{P_{\text{ON}}}{(e^\theta - 1)} \left[\log_e \left(\frac{e^{2\theta C_E(\theta)} - p_{11} e^{\theta C_E(\theta)}}{(1 - p_{11} - p_{22}) + p_{22} e^{\theta C_E(\theta)}} \right) \right]. \quad (3.26)$$

Having formulated the maximum average arrival rates in terms of the effective capacity and source statistics, we next identify the optimal power control policies, maximizing the energy efficiency. In order to have convex optimization problems below, we need to show that throughput $r_{\text{avg}}^*(\theta)$ is concave in $\text{SNR} = \mathbb{E}\{\mu(\theta, z)\}$. In [18, Lemma 1], it is proven that effective capacity is a concave function of SNR. In [11], it is shown that effective bandwidth of the source is strictly monotonically increasing and is also convex in source arrival rates. Therefore, inverse function of the effective bandwidth $a^{*-1}(C_E(\theta))$ exists and is a nondecreasing concave function of the effective capacity, which is concave in SNR. Using the composition properties of concave functions [31], we immediately conclude that the maximum average arrival rate

$$r_{\text{avg}}^*(\theta) = P_{\text{ON}} a^{*-1}(C_E(\theta)) \quad (3.27)$$

is also a concave function of SNR.

3.1.6 Comparative View of Source Models and Performance Levels

In our analysis, we have considered discrete-time Markov, Markov fluid, and MMPP arrival models. All models possess the Markovian property in the sense that the evolution of the Markov chains and hence the state transitions satisfy the Markov condition and are described by the transition probability matrix in the case of discrete-time models and by the transition rate matrix in the case of fluid (or equivalently continuous-time) models. Also, state holding times are geometrically distributed in discrete-time models and exponentially distributed in continuous-time models, and hence exhibit the memoryless property.

At the same time, there are distinct differences between different source models. For instance, transitions between states occur in discrete time steps in discrete-time Markov models while the Markov chain can spend a continuous amount of time in any state in Markov fluid models (i.e., the length of time spent in any state is a continuous random variable or more explicitly holding times are exponentially distributed as also noted above). MMPP models are further differentiated. In the discrete-time Markov and Markov fluid models, arrival rates are assumed to be constant in any given state. On the other hand, when the arrivals are modeled as MMPP, arrival rate is Poisson distributed in each state with a different intensity. Hence, MMPP sources exhibit a higher level of variation in this sense and can be regarded as a more bursty source.

We also remark that ON/OFF discrete-time Markov and Markov fluid source models can be easily specialized to the source with a constant arrival rate by letting ON state probability $P_{ON} = 1$. On the other hand, when $P_{ON} = 1$ in the ON/OFF MMPP source, we have a pure Poisson arrival source.

Finally, we note that although there is a certain degree of similarity in the analysis of discrete-time Markov and Markov fluid models and their throughput performances (e.g.,

high-SNR characterizations are the same in Theorems 3.1.3 and 3.1.6), the set of results for one model do not immediately follow from those for the other model as seen in the throughput formulations in (3.6) and (3.14) and the definitions of η and ζ in (3.9) and (3.17), respectively. However, there is a clear distinction when MMPP sources are considered. As also discussed in Section 3.1.4, higher level of burstiness of MMPP sources penalizes the performance, and lower throughput levels are achieved in general with these sources.

3.2 Energy Efficiency Analysis

In this section, we conduct a low-SNR analysis and investigate the energy efficiency in fading channels when data arrivals are random and statistical queueing constraints are imposed. We first identify the energy efficiency metrics. Subsequently, we consider different source arrival models and provide closed-form expressions for the energy efficiency metrics when the arrival rate is constant or follows a two-state Markovian model. We also numerically analyze specific n -state Markovian sources. Similarly as in the previous section, arbitrary fading correlation within each fading block is considered in the analysis.

3.2.1 Energy Efficiency Metrics

Before defining the energy efficiency metrics, we briefly describe the concavity of the maximum average arrival rate as a function of SNR in the two-state (ON/OFF) arrival models (or if the arrival rates in an n -state model can be expressed as multiples of a certain single rate). In [18, Lemma 1], it was proven that effective capacity is a concave function of SNR. Elwalid and Mitra [11] showed that the effective bandwidth of a source is monotonically increasing when any arrival rate λ_i increases and is convex in the arrival rates $\{\lambda_1, \lambda_2, \dots, \lambda_N\}$. In the ON/OFF arrival models, we have a single arrival rate λ . Since effective bandwidth is a monotonically increasing and convex function of λ , the inverse function of the effective bandwidth a^{*-1} exists and is a nondecreasing concave function. More

specifically, the maximum arrival rate can be expressed as $\lambda^*(\text{SNR}, \theta) = a^{*-1}(C_E(\text{SNR}, \theta))$, which is a nondecreasing concave function of the effective capacity, which is concave in SNR. Using the composition properties of concave functions [31], we realize that the maximum arrival rate is concave in SNR. Thus, the maximum average arrival rate $r_{\text{avg}}^*(\text{SNR}, \theta)$ is also concave in SNR.

In our analysis, following the approach in [57], we study the minimum energy per bit and the wideband slope, which is defined as the slope of the spectral efficiency curve at zero spectral efficiency, as the performance metrics of energy efficiency. While minimum bit energy is a performance measure in the limit as $\text{SNR} \rightarrow 0$ (due to the concavity of the throughput), wideband slope has emerged as a tool that enables us to analyze the energy efficiency at low but nonzero SNR levels. In our setup, we define energy per bit as

$$\frac{E_b}{N_0} = \frac{\text{SNR}}{r_{\text{avg}}^*(\text{SNR}, \theta)/m} \quad (3.28)$$

where the normalization with m is due to our assumption that r_{avg}^* is in the units of bits per m symbols (or equivalently per block).

The minimum energy per bit $\frac{E_b}{N_{0 \min}}$ under QoS constraints can be obtained from

$$\frac{E_b}{N_{0 \min}} = \lim_{\text{SNR} \rightarrow 0} \frac{\text{SNR}}{r_{\text{avg}}^*(\text{SNR}, \theta)/m} = \frac{1}{\dot{r}_{\text{avg}}^*(0)/m}. \quad (3.29)$$

At $\frac{E_b}{N_{0 \min}}$, the slope \mathcal{S}_0 of the throughput versus E_b/N_0 (in dB) curve is defined as [57]

$$\mathcal{S}_0 = \lim_{\frac{E_b}{N_0} \downarrow \frac{E_b}{N_{0 \min}}} \frac{r_{\text{avg}}^*(\text{SNR}, \theta)/m}{10 \log_{10} \frac{E_b}{N_0} - 10 \log_{10} \frac{E_b}{N_{0 \min}}} 10 \log_{10} 2. \quad (3.30)$$

The wideband slope can also be found from

$$\mathcal{S}_0 = -\frac{2(\dot{r}_{\text{avg}}^*(0)/m)^2}{\dot{r}_{\text{avg}}^*(0)/m} \log_e 2 \quad (3.31)$$

where $\dot{r}_{\text{avg}}^*(0)$ and $\ddot{r}_{\text{avg}}^*(0)$ are the first and second derivatives, respectively, of the function $r_{\text{avg}}^*(0)$ with respect to SNR at zero SNR. $\frac{E_b}{N_{0 \min}}$ and \mathcal{S}_0 essentially provide a linear approximation of the throughput curve at low SNR levels.

3.2.2 Energy Efficiency with Constant Arrival Rate

In this section, we assume that the source arrival rate is fixed. Hence, we investigate the energy efficiency in the absence of source randomness and examine the impact of fading correlation and queueing constraints. As discussed in the previous section, effective capacity, $C_E(\text{SNR}, \theta)$, characterizes the maximum constant arrival rate in the presence of QoS constraints described by the QoS exponent θ . Hence, we in this case have $r_{\text{avg}}^*(\text{SNR}, \theta) = C_E(\text{SNR}, \theta)$. In the following result, we provide the minimum bit energy and wideband slope expressions under these assumptions.

Theorem 3.2.1. *Assume that the source arrival rate is constant. Then, the minimum energy per bit and wideband slope expressions as a function of the QoS exponent θ are given, respectively, by*

$$\frac{E_b}{N_{0 \min}} = \frac{\log_e 2}{\mathbb{E}\{z\}} \quad (3.32)$$

and

$$\mathcal{S}_0 = \frac{2(\mathbb{E}\{z\})^2}{\frac{\theta}{m \log_e 2} \sum_{i,j=1}^m \text{cov}\{z_i, z_j\} + \mathbb{E}\{z^2\}} \quad (3.33)$$

where $\text{cov}(z_i, z_j) = \mathbb{E}\{z_i z_j\} - \mathbb{E}\{z_i\}\mathbb{E}\{z_j\}$ is the covariance of z_i and z_j .

Proof: See Appendix A.8.

Remark 2. *As can be seen in (3.32), the minimum energy per bit, which is achieved in the asymptotic regime in which SNR vanishes, does not depend on the QoS exponent θ , hence*

is not affected by the presence of the buffer limitations. Indeed, this is the fundamental limit in Gaussian channels [57]. Wideband slope \mathcal{S}_0 , on the other hand, depends on the QoS constraints via the QoS exponent θ . It can be easily seen that higher the value of θ , the stricter the QoS constraints are and the smaller the value of the wideband slope is, indicating the increased energy requirements. Furthermore, it can be readily verified that wideband slope decreases with increased fading correlation. Or conversely, variations in the channel conditions are favorable for improved energy efficiency.

In Fig. 3.14, we plot the normalized maximum average arrival rate $\frac{1}{m}r_{\text{avg}}^*$ as a function of the energy per bit $\frac{E_b}{N_0}$ for different correlation factors ρ when $\theta = 1$ and $\mathbb{E}\{z\} = \mathbb{E}\{|h|^2\} = \sigma_h^2 = 1$. As predicted by Theorem 3.2.1, all curves converge to the same minimum energy per bit of $\frac{E_b}{N_0 \min} = \frac{\log_e 2}{\mathbb{E}\{z\}} = \log_e 2 = -1.59$ dB as SNR and hence r_{avg}^* vanish. On the other hand, wideband slopes are different for different values of ρ . As discussed above, as ρ and hence correlation diminishes from 1 to 0, slopes increase progressively. It is also interesting to note that in the absence of QoS constraints, i.e., when $\theta = 0$, such a distinction disappears. The wideband slope becomes $\mathcal{S}_0 = \frac{2(\mathbb{E}\{z\})^2}{\mathbb{E}\{z^2\}}$, which clearly does not depend on the fading correlation.

3.2.3 Energy Efficiency with Discrete-Time Markov Sources

Starting with this subsection, we incorporate random arrivals into our energy efficiency analysis and determine how source randomness affects the performance.

ON-OFF Discrete-Time Markov Sources

We assume that data arrival is either in the ON or OFF state in each block duration of m symbols. As we have previously stated, in the ON state, λ bits arrive (i.e., the arrival rate is λ bits/block) while there are no arrivals in the OFF state. Below, we provide our results on energy efficiency.

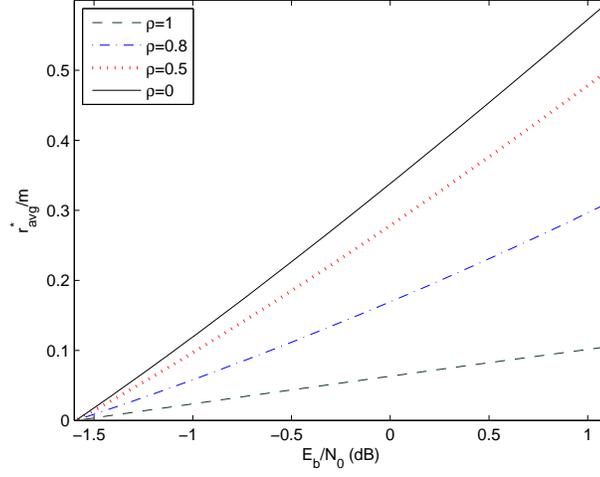


Fig. 3.14: Normalized maximum average arrival rate $\frac{1}{m}r_{\text{avg}}^*$ vs. energy per bit $\frac{E_b}{N_0}$ for different fading correlations. $\theta = 1$, $m = 10$.

Theorem 3.2.2. *Assume that the source arrival rate is random and follows the described discrete-time ON-OFF Markov model. Then, the minimum energy per bit and wideband slope expressions as a function of the QoS exponent θ are given, respectively, by*

$$\frac{E_b}{N_{0 \min}} = \frac{\log_e 2}{\mathbb{E}\{z\}} \quad (3.34)$$

and

$$\mathcal{S}_0 = \frac{2(\mathbb{E}\{z\})^2}{\eta \frac{\theta m}{\log_e 2} (\mathbb{E}\{z\})^2 + \frac{\theta}{m \log_e 2} \sum_{i,j=1}^m \text{cov}\{z_i, z_j\} + \mathbb{E}\{z^2\}} \quad (3.35)$$

where η is defined in (3.9).

Proof: See Appendix A.9.

Interestingly, $\frac{E_b}{N_{0 \min}}$ again turns out to be a very robust quantity. Regardless of the buffer constraints, channel correlations, and randomness of the arrivals, the minimum received energy per bit is $\frac{E_b}{N_{0 \min}} = \log_e 2 = -1.59$ dB when $\mathbb{E}\{z\} = 1$. On the other hand, the impact of random arrivals on the wideband slope is perspicuous in (3.35). When compared

to (3.33), we immediately notice that having random arrivals leads to the introduction of the term $\eta \frac{\theta m}{\log_e 2} (\mathbb{E}\{z\})^2$ in the denominator of (3.35). Notice that when $p_{22} = 1$ and $p_{11} = 0$, we have $P_{\text{ON}} = 1$, meaning that we have a constant arrival rate. In this case, $\eta = 0$ and indeed (3.35) specializes to (3.33). More generally, we have $\eta \geq 0$ for all $p_{11}, p_{22} \in [0, 1]$. Therefore, random arrivals potentially decreases the wideband slope and increases the energy requirements.

This is more clearly seen again in the special case in which $p_{11} = 1 - s$ and $p_{22} = s$. Now, we have $\eta = \frac{p_{11}}{p_{22}} = \frac{1-s}{s}$ and the wideband slope is

$$\mathcal{S}_0 = \frac{2(\mathbb{E}\{z\})^2}{\frac{1-s}{s} \frac{\theta m}{\log_e 2} (\mathbb{E}\{z\})^2 + \frac{\theta}{m \log_e 2} \sum_{i,j=1}^m \text{cov}\{z_i, z_j\} + \mathbb{E}\{z^2\}}. \quad (3.36)$$

As $P_{\text{ON}} = s$ decreases, the wideband slope decreases as well. Therefore, the source becoming more bursty leads to increased energy per bit. This is illustrated in Fig. 3.15 where maximum average arrival rate vs. energy per bit is plotted and the same channel fading and correlation model as in Fig. 3.14 is used. In this figure, we assume $\theta = 1$ and $\rho = 0.75$. As predicted, the minimum bit energies are all the same. However, we have diminishing slopes with decreasing P_{ON} . Note that for a fixed average arrival rate, as P_{ON} gets smaller, source becomes more bursty. Data arrives less frequently but with a higher rate. This in turn increases energy per bit as implied by smaller wideband slopes.

Discrete-Time Markov Sources with n States

In this model, we assume that there are $n - 1$ sources, each having its own ON and OFF states. In the ON state, a source sends data to the buffer at the rate of λ bits/block. Otherwise, it is in OFF state in which no data is generated. In this set-up, depending on how many sources are active (i.e., are in ON state), data arrivals to the buffer can be regarded as a discrete-time Markov process with n states. In the i^{th} state of this model, $(i - 1)$ sources are active. For simplicity, we assume that the probability of each source being active in a

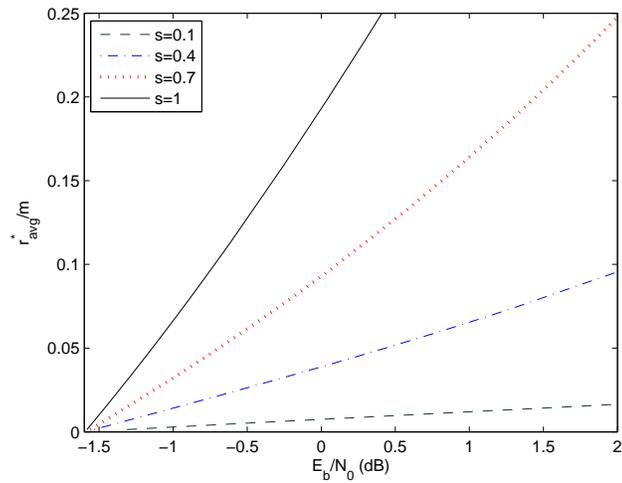


Fig. 3.15: Maximum average arrival rate $\frac{1}{m}r_{\text{avg}}^*$ vs. energy per bit $\frac{E_b}{N_0}$ for different values of $P_{\text{ON}} = s$ when $\rho = 0.75$, $\theta = 1$, and $m = 10$.

given block is s , independent of the previous states and of the other sources. Then, the state probabilities will be given by

$$\pi_i = \binom{n-1}{i-1} s^{i-1} (1-s)^{n-i} \text{ for } i = 1, 2, \dots, n. \quad (3.37)$$

Note that the system is essentially memoryless because each state is independent of the previous state. Hence, transition probability matrix becomes

$$\mathbf{J} = \begin{bmatrix} \pi_1 & \pi_2 & \cdots & \pi_n \\ \pi_1 & \pi_2 & \cdots & \pi_n \\ \vdots & \vdots & \vdots & \vdots \\ \pi_1 & \pi_2 & \cdots & \pi_n \end{bmatrix}. \quad (3.38)$$

Using (3.37) we can write the average rate expression as

$$r_{\text{avg}} = \sum_{i=1}^n \binom{n-1}{i-1} q^{i-1} (1-q)^{n-i} (i-1)\lambda = (n-1)q\lambda. \quad (3.39)$$

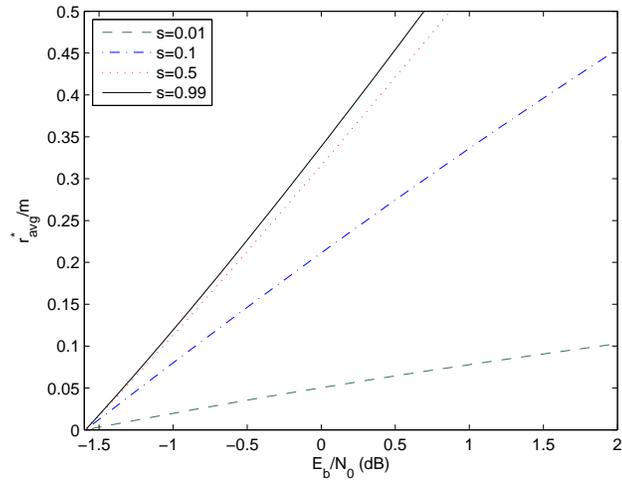


Fig. 3.16: Maximum average arrival rate $\frac{1}{m} r_{\text{avg}}^*$ vs. energy per bit $\frac{E_b}{N_0}$ for different values of s when channel blocks are uncorrelated and $\theta = 1$. Number of states for the arrival process is $n = 10$.

For this case, we do not have closed-form expressions. However, we can easily obtain the effective bandwidth and maximum average arrival rate numerically. In particular, by numerically solving (3.2) and using (3.39), we can determine the maximum average arrival rate r_{avg}^* as a function of SNR. In Fig. 3.16, we display the maximum average arrival rate as a function of energy per bit. Similarly as in the simple ON/OFF model, we observe that when source burstiness is decreased by increasing s , energy efficiency improves.

To have a better understanding of the effect of the QoS constraints, average arrival rate curves in Fig. 3.17 are obtained for different θ values. We first notice that QoS exponent θ does not have any effect on the minimum energy per bit because all curves merge at -1.59 dB which is again the minimum energy per bit for all θ values. However, energy efficiency degrades with stricter QoS conditions as increasing θ reduces the wideband slope.

Finally, for comparison purposes, we depict the throughput as a function of SNR in Fig. 3.18 for different source characteristics and θ values. The trends in the throughput vs. SNR curves for the considered n -state discrete Markov source are observed to be similar to those in Fig. 3.3 plotted for the ON-OFF discrete Markov source. For instance, again

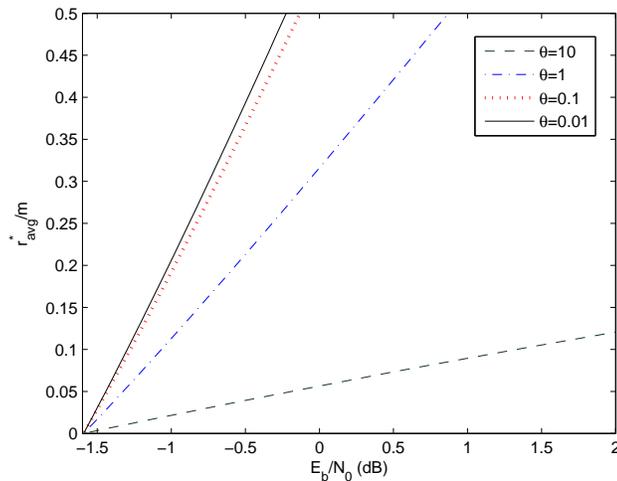


Fig. 3.17: Maximum average arrival rate $\frac{1}{m}r_{avg}^*$ vs. energy per bit $\frac{E_b}{N_0}$ for different θ values when channel blocks are uncorrelated and $s = 0.5$. Number of states for arrival process is $n = 10$.

increased source burstiness (i.e., lower values of s) and stricter queueing constraints (i.e., higher θ values) result in the degradation of the throughput for both two-state (ON/OFF) and n -state source models.

3.2.4 Energy Efficiency with Markov Fluid Sources

ON-OFF Markov Fluid Model

Now, we consider Markov fluid sources with two states, namely OFF state with no arrivals and ON state in which the arrival rate is λ . The generating matrix is defined in (2.21). Minimum energy per bit and wideband slope are derived in the following result.

Theorem 3.2.3. *Assume that the source arrival is modeled by a two-state (ON-OFF) continuous-time Markov chain. Then, the minimum energy per bit and wideband slope expressions as a function of the QoS exponent θ are given, respectively, by*

$$\frac{E_b}{N_{0min}} = \frac{\log_e 2}{\mathbb{E}\{z\}} \quad (3.40)$$

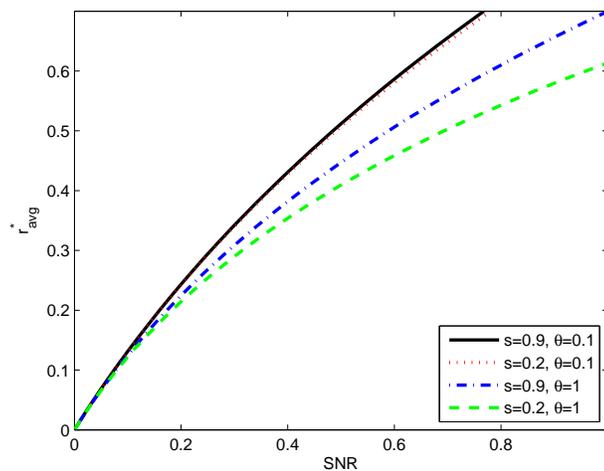


Fig. 3.18: Maximum average arrival rate r_{avg}^* vs. signal-to-noise ratio SNR when channel blocks are uncorrelated. Number of states of the arrival process is $n = 10$.

and

$$\mathcal{S}_0 = \frac{2(\mathbb{E}\{z\})^2}{\zeta \frac{\theta m}{\log_e 2} (\mathbb{E}\{z\})^2 + \frac{\theta}{m \log_e 2} \sum_{i,j=1}^m \text{cov}\{z_i, z_j\} + \mathbb{E}\{z^2\}} \quad (3.41)$$

where $\zeta = \frac{2\beta}{\alpha(\alpha+\beta)}$ as defined in (3.17).

Proof: See Appendix A.10.

Similarly as before, QoS constraints and source randomness do not affect the minimum energy per bit. On the other hand, it is seen in (3.41) that the impact of source arrival characteristics on the wideband slope is via the state transition rates α and β . For instance, larger the α value, the higher the wideband slope is. This is due to the fact that as α , which is the transition rate from OFF state to ON state, increases, the system is more likely to be in the ON state. Contrarily, wideband slope diminishes with increasing β . This is expected as well since larger β leads to higher OFF-state probabilities. The effect of α and β is illustrated in Fig. 3.19, where maximum average arrival rate vs. energy per bit is plotted. In this figure, we set $\theta = 1$ and $\rho = 0.75$. As predicted, the same minimum energy per bit is achieved for different values of α and β , while wideband slope increases with increasing

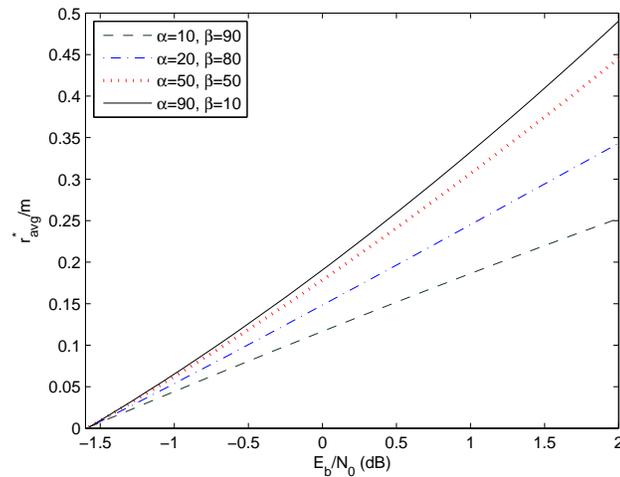


Fig. 3.19: Maximum average arrival rate $\frac{1}{m}r_{\text{avg}}^*$ vs. energy per bit $\frac{E_b}{N_0}$ when $\rho = 0.75$, $\theta = 1$, and $m = 10$.

α or decreasing β .

n-State Markov Fluid Birth-Death Process

In this subsection, we consider a birth-death process for the Markov fluid source. We assume that there are n states and the arrival rate in the i^{th} state is $(i-1)\lambda$ for $i = 1, \dots, n$.

The generating matrix for the birth-death process is in the form of

$$\mathbf{G} = \begin{bmatrix} -\alpha & \alpha & 0 & \cdots & \cdots & 0 \\ \beta & -(\alpha + \beta) & \alpha & 0 & \cdots & 0 \\ 0 & & \ddots & & & \vdots \\ \vdots & & & \ddots & & 0 \\ 0 & \cdots & 0 & \beta & -(\alpha + \beta) & \alpha \\ 0 & \cdots & \cdots & 0 & -\beta & \beta \end{bmatrix}. \quad (3.42)$$

Hence, the transition rate from state i to state $i+1$ is α whereas the transition rate from state i to state $i-1$ is β . The effective bandwidth of this source, which does not have a simple closed-form expression, can be found from (2.19). In order to conduct an energy

efficiency analysis, average arrival rate needs to be identified as well. Using (2.20) and (3.42), we can easily determine that the stationary distribution as

$$\pi_i = \frac{\left(\frac{\alpha}{\beta}\right)^{i-1} - \left(\frac{\alpha}{\beta}\right)^i}{1 - \left(\frac{\alpha}{\beta}\right)^n} \text{ for } i = 1, 2, \dots, n \quad (3.43)$$

when $\frac{\alpha}{\beta} \neq 1$, $\alpha \neq 0$ and $\beta \neq 0$. If any of these inequalities is not satisfied, state probabilities can be obtained by limiting functions.

Now, under the assumptions that $\alpha \neq \beta$ and the arrival rate in state i is $\lambda_i = (i - 1)\lambda(\text{SNR})$, the average arrival rate is given by

$$r_{\text{avg}} = \frac{\xi(1 - n\xi^{n-1} + (n-1)\xi^n)}{(1-\xi)(1-\xi^n)}\lambda(\text{SNR}) \quad (3.44)$$

where $\xi = \frac{\alpha}{\beta}$.

Remark 3. Note that (3.44) specializes to (2.24) if $n = 2$. When $\xi \rightarrow \infty$, the probability of the n^{th} state approaches 1, and the arrival rate will be $(n - 1)\lambda(\text{SNR})$ at steady state. On the other hand, for $\xi = 0$, the state of the source is stuck at the first state in which the arrival rate is zero.

Numerically, we can obtain the effective bandwidth of the n -state birth-death Markov fluid process using (2.19). Subsequently, solving (3.2) and incorporating (3.44), we can determine the maximum average arrival rate $r_{\text{avg}}^*(\text{SNR})$, which we further employ for characterizing the energy efficiency. The results of this numerical analysis are displayed in the following figures. In Fig. 3.20, we demonstrate the effect of α on the energy efficiency. In particular, when β is kept fixed, increasing α improves the energy efficiency as in the two-state case. We illustrate the effect of QoS constraints in Fig. 3.21. Similar conclusions as before readily apply. QoS exponent θ does not alter the minimum bit energy, which is -1.59 dB again, but the wideband slope is reduced with increasing θ .

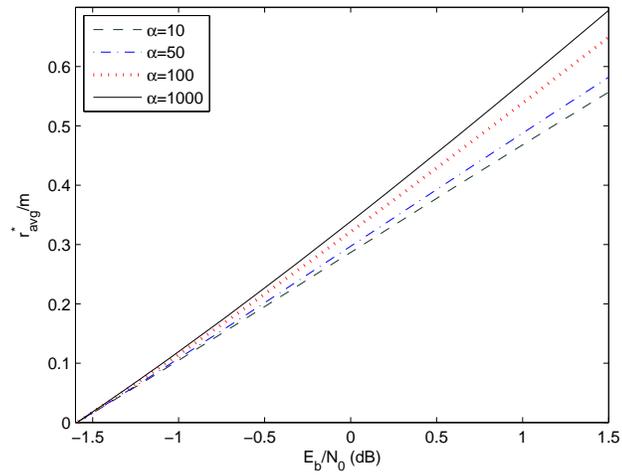


Fig. 3.20: Maximum average arrival rate $\frac{1}{m}r_{\text{avg}}^*$ vs. energy per bit $\frac{E_b}{N_0}$ when channel blocks are uncorrelated. $\theta = 1$, $\beta = 100$ and the number of states of the arrival process is $n = 10$.

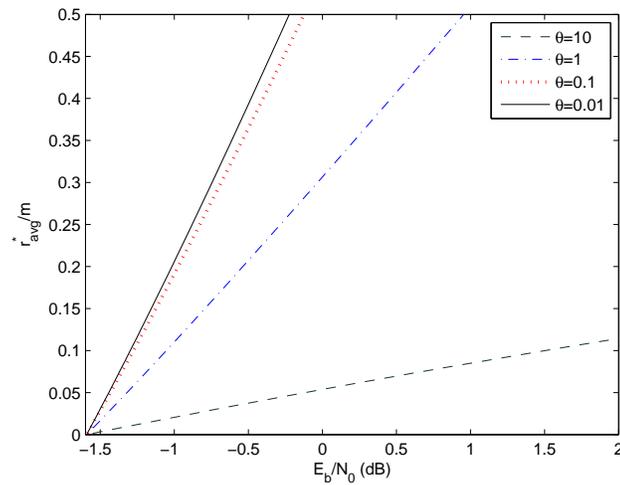


Fig. 3.21: Maximum average arrival rate r_{avg}^* vs. energy per bit $\frac{E_b}{N_0}$ when channel blocks are uncorrelated and $\alpha = \beta = 50$. Number of states of the arrival model is 10.

3.2.5 Energy Efficiency with Markov-Modulated Poisson Process

ON-OFF Markov-Modulated Poisson Process

Again, we initially address the two-state model in which there are no arrivals in the OFF state and the intensity of the Poisson arrival process is λ in the ON state. The generating matrix \mathbf{G} is the same as in (2.21).

Theorem 3.2.4. *Assume that the source arrival is modeled by a two-state (ON/OFF) Markov modulated Poisson process. Then, the minimum energy per bit and wideband slope expressions as a function of the QoS exponent θ are given, respectively, by*

$$\frac{E_b}{N_{0 \min}} = \frac{e^\theta - 1 \log_e 2}{\theta \mathbb{E}\{z\}} \quad (3.45)$$

and

$$S_0 = \frac{\frac{2\theta}{e^\theta - 1} (\mathbb{E}\{z\})^2}{\zeta \frac{\theta m}{\log_e 2} (\mathbb{E}\{z\})^2 + \frac{\theta}{m \log_e 2} \sum_{i,j=1}^m \text{cov}\{z_i, z_j\} + \mathbb{E}\{z^2\}} \quad (3.46)$$

where $\zeta = \frac{2\beta}{\alpha(\alpha+\beta)}$ as defined in (3.17).

Proof: See Appendix A.11.

Remark 4. *It is interesting to observe that, unlike the previous arrival models, minimum energy per bit in the case of MMPP source depends on the QoS exponent θ . More specifically, minimum energy per bit increases with $(e^\theta - 1)/\theta$ which is an increasing monotonic function of θ and always greater than one for $\theta > 0$. On the other hand, as $\theta \rightarrow 0$, $(e^\theta - 1)/\theta \rightarrow 1$. Therefore, $\frac{E_b}{N_{0 \min}} \geq \frac{\log_e 2}{\mathbb{E}\{z\}}$ with equality only if no QoS constraints are imposed (i.e., when $\theta = 0$). Furthermore, in addition to its significant impact on the minimum energy per bit, increasing θ leads to much quicker reduction in the wideband slope due to the presence of the term $\frac{\theta}{e^\theta - 1}$ in (3.46). Hence, overall, energy costs grow very fast as θ increases. This is again because of the additional randomness arising from Poisson*

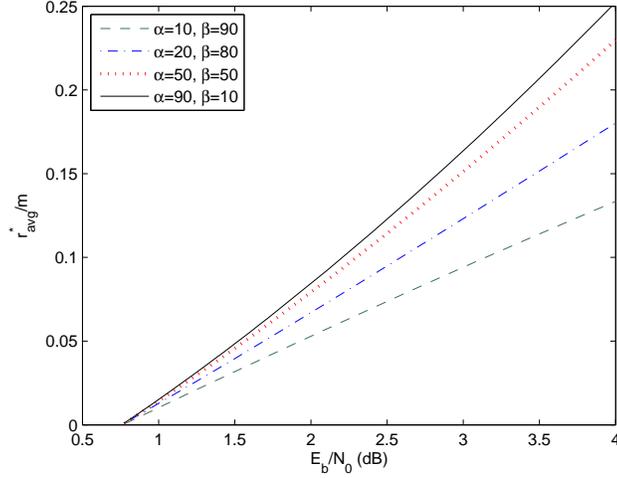


Fig. 3.22: Maximum average arrival rate $\frac{1}{m}r_{\text{avg}}^*$ vs. energy per bit $\frac{E_b}{N_0}$ when $\rho = 0.75$, $\theta = 1$, and $m = 10$.

arrivals in the ON state.

Remark 5. From (3.46), we note that the effect of the state transition rates α and β on the energy efficiency is the same as in the Markov fluid source model. Increasing α or decreasing β improves the energy efficiency of the system because the burstiness of the data arrivals is reduced and the buffer overflows can be avoided at lower energy costs.

We plot the maximum average arrival rate vs. energy per bit in Fig. 3.22. We set $\mathbb{E}\{z\} = 1$ and $\theta = 1$ for which the minimum energy per bit is 0.76 dB. The increase in bit energy with respect to -1.59 dB is due to $10 \log_{10}((e^\theta - 1)/\theta) = 10 \log_{10}(e - 1)$ for $\theta = 1$. From the figure, we can again infer that adjusting α or β to increase the ON state probability makes the system more energy efficient due to the increase in the wideband slope.

n-State Markov-Modulated Poisson Process

Finally, we consider an n -state MMPP process and assume that the intensity of the Poisson arrivals in the i^{th} state is $(i - 1)\lambda$. For the Markov transitions between states, we consider the birth-death process and adopt the transition rate matrix \mathbf{G} from (3.42). We solve

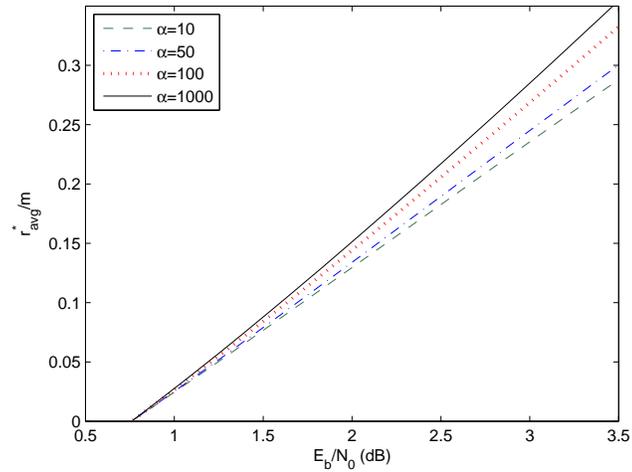


Fig. 3.23: Maximum average arrival rate $\frac{1}{m}r_{\text{avg}}^*$ vs. energy per bit $\frac{E_b}{N_0}$ when channel blocks are uncorrelated. $\theta = 1$, $\beta = 100$ and the number of Markov states for the arrival process is $n = 10$.

for the maximum intensity $\lambda^*(\text{SNR}, \theta)$ by incorporating (2.25) into (3.2). Then, using the expression in (3.44), we obtain $r_{\text{avg}}^*(\text{SNR}, \theta)$.

In Figs. 3.23 and 3.24, we depict the maximum average arrival rate as a function of the energy per bit with uncorrelated channel coefficients being assumed in each block. In Fig. 3.23, we set $n = 10$, $\theta = 1$ and $\beta = 100$, and demonstrate how α influences the energy efficiency of system. The observation has similarities with other Markovian sources regarding the source burstiness. Interestingly, the minimum energy per bit is again 0.76 dB as in the two-state case, leading to the conclusion that the number of states does not alter $\frac{E_b}{N_{0 \min}}$ in this case. The degradation in energy efficiency due to increased θ is shown in Fig. 3.24. As described in the two-state case, higher values of θ (i.e., stricter QoS constraints) result in higher $\frac{E_b}{N_{0 \min}}$ and smaller wideband slope. Therefore, even for relatively small increases in θ , we can have large gaps between curves, indicating significantly high energy costs.

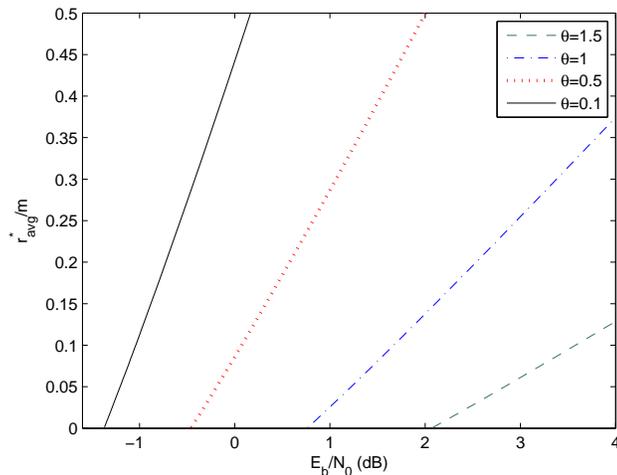


Fig. 3.24: Maximum average arrival rate $\frac{1}{m}r_{\text{avg}}^*$ vs. energy per bit $\frac{E_b}{N_0}$ when channel blocks are uncorrelated and $\alpha = \beta = 50$. Number of Markov states for the arrival process is $n = 10$.

3.3 Energy Efficiency in Multiple-Antenna Channels

3.3.1 Channel Model

We consider a flat-fading MIMO channel model between the transmitter and receiver. We assume that transmitter and receiver has n_T and n_R antennas, respectively. The channel input-output relation can be expressed as

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n}. \quad (3.47)$$

Note that, \mathbf{x} denotes the $n_T \times 1$ -dimensional channel input vector, and \mathbf{y} denotes the $n_R \times 1$ -dimensional channel output vector. Input average energy is $\mathbb{E}\{\|\mathbf{x}\|^2\} = \mathcal{E}$. \mathbf{n} denotes the zero-mean, circularly-symmetric, complex Gaussian noise with dimension $n_R \times 1$. The covariance matrix of noise is given as $\mathbb{E}\{\mathbf{n}\mathbf{n}^\dagger\} = N_0\mathbf{I}$, where \mathbf{I} is the identity matrix.

Hence, the signal-to-noise ratio is

$$\text{SNR} = \frac{\mathbb{E}\{\|\mathbf{x}\|^2\}}{\mathbb{E}\{\|\mathbf{n}\|^2\}} = \frac{\mathcal{E}}{n_R N_0}. \quad (3.48)$$

Furthermore, we define the normalized input covariance matrix as

$$\mathbf{K}_x = \frac{\mathbb{E}\{\mathbf{x}\mathbf{x}^\dagger\}}{\mathcal{E}}. \quad (3.49)$$

with trace $\text{tr}(\mathbf{K}_x) = 1$.

Finally, \mathbf{H} denotes the $n_R \times n_T$ -dimensional random channel matrix whose components are the fading coefficients of the channels between the corresponding transmitting and receiving antennas. We assume that the components of \mathbf{H} have arbitrary distributions with finite variances unless specified otherwise. Additionally, we assume block-fading scenario in which the realization of matrix \mathbf{H} remains fixed over a block and changes independently from one block to another.

3.3.2 Effective Capacity of Wireless Transmissions with MIMO Channels

For the MIMO fading Gaussian channel with channel perfectly known at the transmitter and receiver, the maximum instantaneous service rate with given input covariance matrix \mathbf{K}_x is the instantaneous channel capacity, which is expressed as

$$R = \log_2(\mathbf{I} + n_R \text{SNR} \mathbf{H} \mathbf{K}_x \mathbf{H}^\dagger). \quad (3.50)$$

With uniform power allocation across transmit antennas, the input covariance matrix is $\mathbf{K}_x = \frac{1}{n_T} \mathbf{I}$. Hence, we have the effective capacity expression as

$$C_E(\text{SNR}, \theta) = -\frac{1}{\theta} \log_e \mathbb{E} \left\{ \exp \left(-\theta \log_2 \det \left(\mathbf{I} + \frac{n_R}{n_T} \text{SNR} \mathbf{H} \mathbf{H}^\dagger \right) \right) \right\}. \quad (3.51)$$

If the input covariance matrix is optimized, then the effective capacity is given by

$$C_E(\text{SNR}, \theta) = -\frac{1}{\theta n_R} \log_e \mathbb{E} \left\{ \exp \left(-\theta \max_{\mathbf{K}_x \succeq \mathbf{0}} \log_2 \det(\mathbf{I} + n_R \text{SNR} \mathbf{H} \mathbf{K}_x \mathbf{H}^\dagger) \right) \right\}. \quad (3.52)$$

3.3.3 Energy Efficiency of MIMO Channels with Discrete Markov Sources

In this section, we consider discrete Markov sources and first characterize the energy efficiency metrics for uniform power allocation. Due to space limitations, we omit the proofs.

Theorem 3.3.1. *With uniform power allocation, i.e., when the input covariance matrix is $\mathbf{K}_x = \frac{1}{n_T} \mathbf{I}$, the minimum energy per bit and wideband slope are given by*

$$\frac{E_b}{N_{0 \min}} = \frac{n_T \log_e 2}{n_R \mathbb{E} \{ \text{tr}(\mathbf{H}^\dagger \mathbf{H}) \}} \quad (3.53)$$

$$\mathcal{S}_0 = \frac{2 \mathbb{E}^2 \{ \text{tr}(\mathbf{H}^\dagger \mathbf{H}) \}}{\frac{\theta \eta}{\log_e 2} \mathbb{E}^2 \{ \text{tr}(\mathbf{H}^\dagger \mathbf{H}) \} + \frac{\theta}{\log_e 2} \text{var}(\text{tr}(\mathbf{H}^\dagger \mathbf{H})) + \mathbb{E} \{ \text{tr}((\mathbf{H}^\dagger \mathbf{H})^2) \}} \quad (3.54)$$

where $\text{tr}(\cdot)$ denotes the trace of a matrix and

$$\eta = \frac{(1 - p_{22})(p_{11} + p_{22})}{(1 - p_{11})(2 - p_{11} - p_{22})}. \quad (3.55)$$

Remark 6. *When \mathbf{H} has independent zero mean unit variance complex Gaussian random entries, we have [8]*

$$\begin{aligned} \mathbb{E} \{ \text{tr}(\mathbf{H}^\dagger \mathbf{H}) \} &= n_R n_T \\ \mathbb{E} \{ \text{tr}^2(\mathbf{H}^\dagger \mathbf{H}) \} &= n_R n_T (n_R n_T + 1) \\ \mathbb{E} \{ \text{tr}((\mathbf{H}^\dagger \mathbf{H})^2) \} &= n_R n_T (n_R + n_T). \end{aligned} \quad (3.56)$$

Inserting these into (3.53) and (3.54), we obtain

$$\frac{E_b}{N_{0\min}} = \frac{\log_e 2}{n_R^2}, \quad (3.57)$$

$$\mathcal{S}_0 = \frac{2}{\frac{\theta\eta}{\log_e 2} + \frac{\theta}{\log_e 2} \frac{1}{n_R n_T} + \frac{n_R + n_T}{n_R n_T}}. \quad (3.58)$$

Above, the minimum energy per bit depends only on the number of receive antennas. This dependence is somewhat expected as having more receive antennas leads to higher received power and the performance improves. Interestingly, source randomness and queueing constraints do not have any impact on $\frac{E_b}{N_{0\min}}$, which is achieved as SNR vanishes. On the other hand, these play an important role at non-zero SNR values as seen in (3.58). More specifically, wideband slope depends on the source randomness/burstiness through η , which is a function of the source transition probabilities. Hence, η can be regarded as a measure of source burstiness in the low-SNR regime. For instance, if the source is always ON (i.e., $p_{11} = 0$ and $p_{22} = 1$), then data arrives at a constant rate and there is essentially no source burstiness. In such a case, we indeed have $\eta = 0$. Hence, the first term in the denominator of (3.58) vanishes. Otherwise, the presence of this nonnegative term lowers the wideband slope, which lets us conclude that random arrivals in general hurt the energy efficiency.

In (3.58), queueing constraints are reflected via the QoS exponent θ . In particular, we notice that higher values of θ , which imply stricter queueing constraints, result in smaller slopes, again deteriorating the energy efficiency. Another observation is that the decrease in \mathcal{S}_0 due to having random arrivals (i.e., having $\eta > 0$) is proportional to θ as seen from the presence of the term $\frac{\theta\eta}{\log_e 2}$. Hence, source burstiness is more detrimental under stricter queueing limitations. Finally, we note from the term $\frac{\theta}{\log_e 2} \frac{1}{n_R n_T}$ that having multiple antennas partially offsets the reduction in \mathcal{S}_0 due to stricter queueing constraints.

Next, we provide similar characterizations when the transmission power is optimally allocated across different antennas in the low-SNR regime.

Theorem 3.3.2. *With optimal power allocation, the minimum energy per bit and wideband*

slope are given by

$$\frac{E_b}{N_{0 \min}} = \frac{\log_e 2}{n_R \mathbb{E} \{ \lambda_{\max}(\mathbf{H}^\dagger \mathbf{H}) \}}, \quad (3.59)$$

$$\mathcal{S}_0 = \frac{2}{\frac{\theta \eta}{\log_e 2} + \frac{\theta}{\log_e 2} (\kappa(\sigma_{\max}(\mathbf{H})) - 1) + \frac{1}{l} \kappa(\sigma_{\max}(\mathbf{H}))}, \quad (3.60)$$

where l is the multiplicity of $\lambda_{\max}(\mathbb{E} \{ \mathbf{H}^\dagger \mathbf{H} \})$, and $\kappa(\sigma_{\max}(\mathbf{H}))$ is the kurtosis of the maximum singular value of the matrix \mathbf{H} and is defined as [53]

$$\kappa(\sigma_{\max}(\mathbf{H})) = \frac{\mathbb{E} \{ \lambda_{\max}^2(\mathbf{H}^\dagger \mathbf{H}) \}}{\mathbb{E}^2 \{ \lambda_{\max}(\mathbf{H}^\dagger \mathbf{H}) \}}. \quad (3.61)$$

Remark 7. Regarding the effect of source randomness and queueing constraints, we have identical observations as in the case of uniform power allocation. Here, the major difference is that $\frac{E_b}{N_{0 \min}}$ and \mathcal{S}_0 are achieved by transmitting in the maximal-eigenvalue eigenspace of $\mathbf{H}^\dagger \mathbf{H}$.

In the numerical results, we assume uniform power allocation and therefore set the input covariance matrix as $\mathbf{K}_x = 1/n_T \mathbf{I}$. We consider a fading model in which the components of \mathbf{H} are independent and identically distributed Gaussian random variables with zero mean and unit variance. In the figures, we plot maximum average arrival rate vs. energy per bit curves in the low-SNR regime to depict $\frac{E_b}{N_{0 \min}}$ and \mathcal{S}_0 and how they are affected by the number of antennas, QoS exponent θ , and source burstiness.

In Fig. 3.25, we plot r_{avg}^* vs. $\frac{E_b}{N_0}$ for different number of receive antennas n_R and different ON-state probabilities P_{ON} . We set $n_T = 5$. Verifying the analytical characterizations, we observe that $\frac{E_b}{N_{0 \min}}$ diminishes and the curves shift to the left as n_R increases. For the same number of antennas, as the source becomes more bursty (i.e., P_{ON} decreases from 0.5 to 0.2), wideband slope becomes smaller while $\frac{E_b}{N_{0 \min}}$ stays fixed. In Fig. 3.26, we set $n_R = 5$ and vary n_T and P_{ON} . As expected, $\frac{E_b}{N_{0 \min}}$ remains the same while the wideband

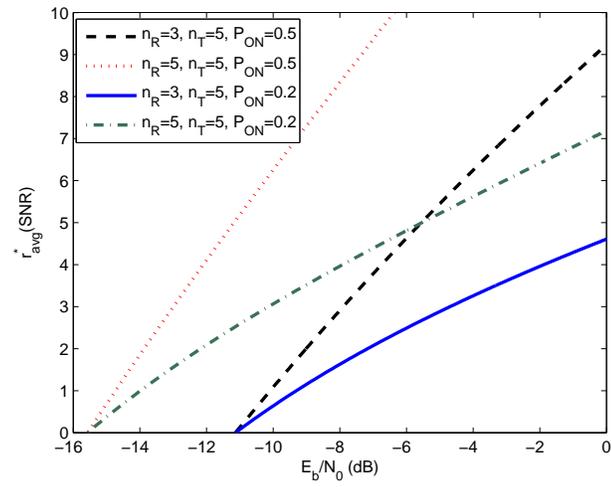


Fig. 3.25: Maximum average arrival rate r_{avg}^* vs. energy per bit $\frac{E_b}{N_0}$ when $\theta = 0.1$ and $n_T = 5$.

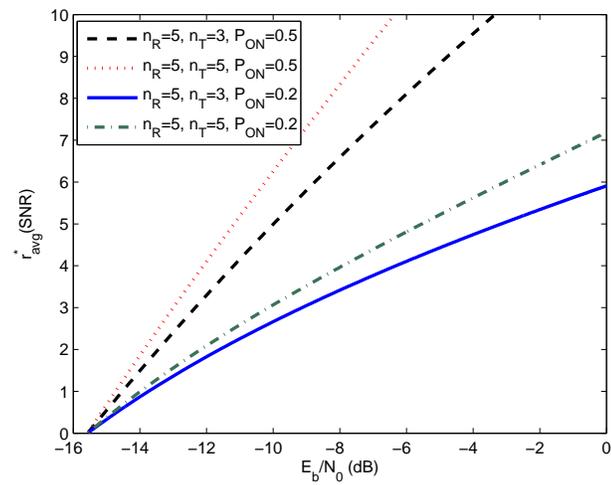


Fig. 3.26: Maximum average arrival rate r_{avg}^* vs. energy per bit $\frac{E_b}{N_0}$ when $\theta = 0.1$ and $n_R = 5$.

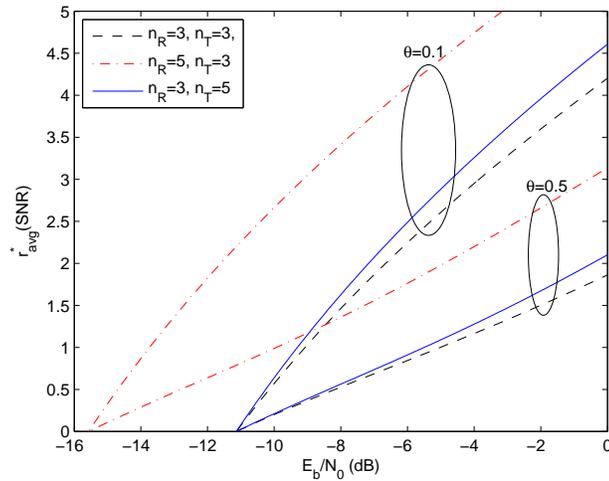


Fig. 3.27: Maximum average arrival rate r_{avg}^* vs. energy per bit $\frac{E_b}{N_0}$ when $P_{\text{ON}} = 0.2$.

slope diminishes with decreasing P_{ON} and n_T . Hence, beyond $\frac{E_b}{N_0}_{\text{min}}$, energy costs grow with more bursty sources, and some of this degradation can be overcome by increasing the number of transmit antennas. In Figs. 3.25 and 3.26, we have kept the QoS exponent fixed at $\theta = 0.1$. In Fig. 3.27, we set $P_{\text{ON}} = 0.2$ and analyze the impact of varying θ , n_T , and n_R on the energy efficiency. We notice that as θ increases from 0.1 to 0.5 and hence the queueing constraints become stricter, wideband slopes decrease significantly for the same set of antennas, so much so that all curves with $\theta = 0.5$ are below those with $\theta = 0.1$ for $\frac{E_b}{N_0} > -8$ dB regardless of the number of transmit and receive antennas.

3.3.4 Energy Efficiency of MIMO Channels with Markov Fluid Sources

In this section, we conduct a similar analysis for Markov fluid sources and obtain the following results.

Theorem 3.3.3. *When the input covariance matrix is $\mathbf{K}_x = \frac{1}{n_T} \mathbf{I}$, the minimum energy per*

bit and wideband slope are given by

$$\frac{E_b}{N_{0 \min}} = \frac{n_T \log_e 2}{n_R \mathbb{E} \{tr(\mathbf{H}^\dagger \mathbf{H})\}}. \quad (3.62)$$

$$\mathcal{S}_0 = \frac{2\mathbb{E}^2 \{tr(\mathbf{H}^\dagger \mathbf{H})\}}{\frac{\theta\zeta}{\log_e 2} \mathbb{E}^2 \{tr(\mathbf{H}^\dagger \mathbf{H})\} + \frac{\theta}{\log_e 2} \text{var}(tr(\mathbf{H}^\dagger \mathbf{H})) + \mathbb{E} \{tr((\mathbf{H}^\dagger \mathbf{H})^2)\}}. \quad (3.63)$$

Remark 8. When \mathbf{H} has independent zero mean unit variance complex Gaussian random entries, we have

$$\frac{E_b}{N_{0 \min}} = \frac{\log_e 2}{n_R^2}, \quad (3.64)$$

$$\mathcal{S}_0 = \frac{2}{\frac{\theta\zeta}{\log_e 2} + \frac{\theta}{\log_e 2} \frac{1}{n_R n_T} + \frac{n_R + n_T}{n_R n_T}}. \quad (3.65)$$

Theorem 3.3.4. With optimal power allocation, the minimum energy per bit and wideband slope are given by

$$\frac{E_b}{N_{0 \min}} = \frac{\log_e 2}{n_R \mathbb{E} \{\lambda_{\max}(\mathbf{H}^\dagger \mathbf{H})\}}, \quad (3.66)$$

$$\mathcal{S}_0 = \frac{2}{\frac{\theta\zeta}{\log_e 2} + \frac{\theta}{\log_e 2} (\kappa(\sigma_{\max}(\mathbf{H})) - 1) + \frac{1}{l} \kappa(\sigma_{\max}(\mathbf{H}))}, \quad (3.67)$$

where l is again the multiplicity of $\lambda_{\max}(\mathbb{E} \{\mathbf{H}^\dagger \mathbf{H}\})$, $\kappa(\sigma_{\max}(\mathbf{H}))$ is as defined in (3.61).

Remark 9. Regarding the impact of the parameters n_T , n_R , and θ , we have similar observations as for the discrete-Markov source. On the other hand, source characteristics appear in the formulations via the parameter $\zeta = \frac{2\beta}{\alpha(\alpha+\beta)}$, which essentially quantifies the source burstiness. Note that α is the transition rate from OFF to ON state whereas β is the transition rate from ON to OFF state. Note further that when $\alpha = 1$ and $\beta = 0$ and therefore $P_{ON} = 1$ and data arrival rate is constant, we have $\zeta = 0$.

Fig. 3.28 examines the effect of n_R and source characteristics on the energy efficiency.

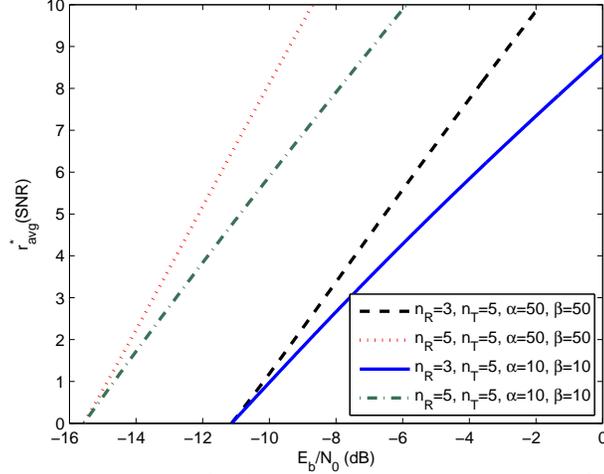


Fig. 3.28: Maximum average arrival rate r_{avg}^* vs. energy per bit $\frac{E_b}{N_0}$ when $\theta = 1$ and $n_T = 5$.

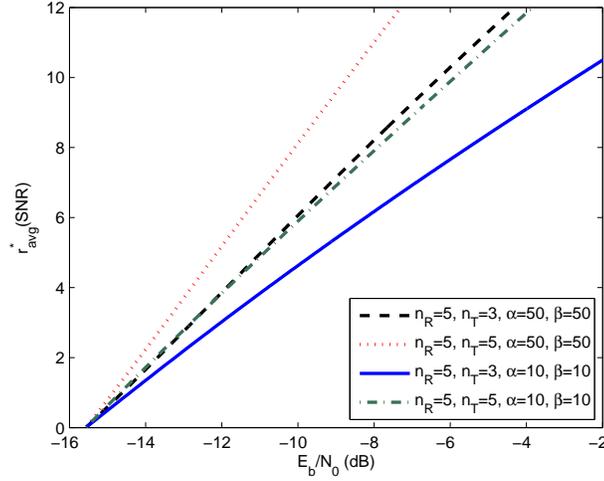


Fig. 3.29: Maximum average arrival rate r_{avg}^* vs. energy per bit $\frac{E_b}{N_0}$ when $\theta = 1$ and $n_R = 5$.

For all plots, we have $P_{\text{ON}} = \frac{\alpha}{\alpha+\beta} = 0.5$. We notice that lower transition rates α and β result in smaller wideband slopes, indicating higher energy costs. Note that low values of α and β imply longer ON and OFF durations on average. Hence, OFF state can be more persistent. As before, we observe that more receive antennas lead to smaller values of the minimum energy per bit.

In Fig. 3.29, we vary n_T , α , and β while keeping P_{ON} and n_R fixed. Minimum energy per bit is the same for all scenarios while wideband slope is reduced for smaller values of n_T , α , and β .

CHAPTER 4

ENERGY-EFFICIENT POWER CONTROL IN FADING CHANNELS WITH MARKOVIAN SOURCES AND QoS CONSTRAINTS

4.1 Channel Model

We consider a flat-fading channel between the transmitter and receiver. The channel input-output relation can be expressed as

$$y_i = h_i x_i + n_i \text{ for } i = 1, 2, \dots \quad (4.1)$$

where x_i and y_i are the channel input and output, respectively, and h_i denotes the channel fading coefficient. We assume that the transmitter, equipped with perfect channel side information (CSI), performs power control. Hence, the transmit power $P(\theta, z_i)$, where $z_i = |h_i|^2$ and θ is a QoS parameter described in the following section, varies with QoS

requirements and fading. Fading coefficients are assumed to be identically distributed, and the fading distribution can be arbitrary with finite variance. We consider a block-fading model and assume that the realizations of the fading coefficients stay fixed for a block of symbols and change independently for the next block. Finally, $\{n_i\}$ is a sequence of independent, zero-mean, circularly-symmetric, complex Gaussian noise components with variance $\mathbb{E}\{|n_i|^2\} = N_0$.

4.2 Energy-Efficient Power Control

In this chapter, we employ rate per unit energy (in bits/joule) as the performance metric of energy efficiency. In our setup, we define energy efficiency (EE) as

$$\text{EE} = \frac{r_{\text{avg}}^*(\theta)}{\left(\frac{1}{\epsilon}\mathbb{E}\{\mathbf{P}(\theta, z)\} + \mathbf{P}_c\right)/N_0B} = \frac{r_{\text{avg}}^*(\theta)}{\left(\frac{1}{\epsilon}\mathbb{E}\{\mu(\theta, z)\} + \mu_c\right)} \quad (4.2)$$

where \mathbf{P}_c is the circuit power and ϵ is the efficiency of the power amplifier, and $\mu_c = \mathbf{P}_c/N_0B$. Normalization with the noise power N_0B in the denominator above is performed in order to express EE in terms of the instantaneous SNR $\mu(\theta, z)$, and to perform optimization over $\mu(\theta, z)$ and have simplifications in the expressions. Furthermore, to be used in subsequent formulations, we define function $g(\theta)$ as

$$g(\theta) = \mathbb{E}\{[1 + \mu(\theta, z)z]^{-\varrho}\} \quad (4.3)$$

where again $\varrho = \theta TB \log_2 e$.

After formulating the energy efficiency, we can express the optimally energy-efficient

power control problem as¹

$$\max_{\mu(\theta, z)} \frac{r_{\text{avg}}^*(\theta)}{\left(\frac{1}{\epsilon} \mathbb{E}\{\mu(\theta, z)\} + \mu_c\right)}. \quad (4.4)$$

Next, we will address special cases of this optimization problem by considering specific arrival models and incorporating the corresponding average arrival rate expressions.

4.2.1 Discrete Markov Source

In this section, we consider ON-OFF discrete Markov arrival models and determine the optimal power adaptation strategy that maximizes the energy efficiency. After inserting the maximum average arrival rate expression in (3.6) into the optimization problem in (4.4) and simplifying the expressions by eliminating the constant terms, we can formulate the optimal power allocation problem as

$$\mu^*(\theta, z) = \arg \max_{\mu(\theta, z)} \frac{\log_e \left(\frac{1 - p_{11}g(\theta)}{(1 - p_{11} - p_{22})g^2(\theta) + p_{22}g(\theta)} \right)}{\frac{1}{\epsilon} \mathbb{E}\{\mu(\theta, z)\} + \mu_c} \quad (4.5)$$

where the function $g(\cdot)$ is defined in (4.3). Note that any function that can be expressed as the ratio of a convex function over a concave one is quasiconvex [31, Example 3.38] and the negative of a quasiconvex function is quasiconcave. Hence, the objective function in (4.5), being a concave function divided by an affine function of power allocation, is a quasiconcave function of the instantaneous SNR $\mu(\theta, z)$. By introducing an additional

¹Since the theory of effective bandwidth and effective capacity makes use of tools from large deviations and characterizes the performance in the large-queue-length regime, we consider a saturated buffer in our analysis, and the optimal power control policies are obtained under the assumption that there is always data to transmit from the buffer.

variable $\psi = \frac{1}{\mathbb{E}\{\frac{1}{\epsilon}\mu(\theta, z)\} + \mu_c}$, the problem can be transformed into

$$\min_{\mu(\theta, z) \geq 0} \quad -\psi \log_e \left(\frac{1 - p_{11}\mathbf{g}(\theta)}{(1 - p_{11} - p_{22})\mathbf{g}^2(\theta) + p_{22}\mathbf{g}(\theta)} \right) \quad (4.6)$$

$$\text{subject to} \quad \psi \left(\frac{1}{\epsilon} \mathbb{E} \{ \mu(\theta, z) \} + \mu_c \right) = 1. \quad (4.7)$$

The problem in (4.6) is a convex optimization problem. Therefore, we can use the convex optimization tools and determine the sufficient and necessary Karush-Kuhn-Tucker (KKT) conditions. By denoting the Lagrange multiplier by λ , we form the Lagrangian as

$$\begin{aligned} \mathcal{L}(\mu(\theta, z), \psi, \lambda) = & -\psi \log_e \left(\frac{1 - p_{11}\mathbf{g}(\theta)}{(1 - p_{11} - p_{22})\mathbf{g}^2(\theta) + p_{22}\mathbf{g}(\theta)} \right) \\ & + \lambda \left[\psi \left(\frac{1}{\epsilon} \mathbb{E} \{ \mu(\theta, z) \} + \mu_c \right) - 1 \right]. \end{aligned} \quad (4.8)$$

Now, the KKT conditions are given in (4.9)–(4.11)

$$\psi \left(\frac{1}{\epsilon} \mathbb{E} \{ \mu(\theta, z) \} + \mu_c \right) = 1, \quad (4.9)$$

$$-\psi \rho z [1 + \mu(\theta, z)z]^{-\rho-1} \left(\frac{(1 - p_{11})(1 - p_{22})}{(1 - p_{11}\mathbf{g}(\theta))((1 - p_{11} - p_{22})\mathbf{g}(\theta) + p_{22})} + \frac{1}{\mathbf{g}(\theta)} \right) + \frac{\lambda\psi}{\epsilon} = 0, \quad (4.10)$$

$$-\log_e \left(\frac{1 - p_{11}\mathbf{g}(\theta)}{(1 - p_{11} - p_{22})\mathbf{g}^2(\theta) + p_{22}\mathbf{g}(\theta)} \right) + \lambda \left(\frac{1}{\epsilon} \mathbb{E} \{ \mu(\theta, z) \} + \mu_c \right) = 0. \quad (4.11)$$

Note that (4.9) is due to the constraint in (4.7). (4.10) and (4.11) are obtained by taking the derivative of the Lagrangian in (4.8) with respect to $\mu(\theta, z)$ and ψ , respectively. After simplifying (4.10), we obtain

$$z[1 + \mu(\theta, z)z]^{-\rho-1} = \frac{\lambda/(\rho\epsilon)}{\frac{(1-p_{11})(1-p_{22})}{(1-p_{11}\mathbf{g}(\theta))((1-p_{11}-p_{22})\mathbf{g}(\theta)+p_{22})} + \frac{1}{\mathbf{g}(\theta)}}. \quad (4.12)$$

By solving (4.12) for $\mu(\theta, z)$, the optimal power allocation is found as

$$\mu^*(\theta, z) = \left[\frac{1}{\nu^{\frac{1}{1+e}} z^{\frac{e}{1+e}}} - \frac{1}{z} \right]^+ \quad (4.13)$$

where

$$\nu = \frac{\lambda/(\rho\epsilon)}{\frac{(1-p_{11})(1-p_{22})}{(1-p_{11}g(\theta))((1-p_{11}-p_{22})g(\theta)+p_{22})} + \frac{1}{g(\theta)}} \quad (4.14)$$

and $[c]^+ = \max(c, 0)$. We notice that $\mu^*(\theta, z) = 0$ when $z \leq \nu$. Hence, ν can be regarded as the fading gain threshold for transmission. When we consider the special case of constant-rate arrivals (i.e., when we have $P_{\text{ON}} = 1$), the above equation for ν specializes to the corresponding one in [19]. Note that the expression for ν in (4.14) depends on the Lagrange multiplier λ (and hence ν can also be considered as a scaled Lagrange multiplier). By combining (4.11) with (4.14), we obtain (4.15).

$$\begin{aligned} & \frac{1}{\frac{(1-p_{11})(1-p_{22})}{(1-p_{11}g(\theta))((1-p_{11}-p_{22})g(\theta)+p_{22})} + \frac{1}{g(\theta)}} \log_e \left(\frac{1 - p_{11}g(\theta)}{(1-p_{11}-p_{22})g^2(\theta) + p_{22}g(\theta)} \right) \\ & + \nu \rho \epsilon \left(\frac{1}{\epsilon} \mathbb{E}\{\mu(\theta, z)\} + \mu_c \right) = 0 \end{aligned} \quad (4.15)$$

Equation (4.15), which does not depend on λ , can be used to determine ν by incorporating the source statistics² and computing $\mathbb{E}\{\mu(\theta, z)\}$ and $g(\theta)$. For instance, in the case of Rayleigh fading, the fading power is exponentially distributed with density function $f_z(z) = e^{-z}$, and by using the expression for $\mu(\theta, z)$ in (4.13), these key expectations

²It is interesting to note that the optimal power control $\mu^*(\theta, z)$ depends on the source statistics (e.g., transition probabilities p_{11} and p_{22}) only through ν .

can be determined in closed-form as follows:

$$\begin{aligned}
\mathbb{E} \{ \mu(\theta, z) \} &= \int_{\nu}^{\infty} \left[\frac{1}{\nu^{\frac{1}{1+\varrho}} z^{\frac{\varrho}{1+\varrho}}} - \frac{1}{z} \right] e^{-z} dz \\
&= \left(\frac{1}{\nu} \right)^{\frac{1}{1+\varrho}} \int_{\nu}^{\infty} z^{-\frac{\varrho}{1+\varrho}} e^{-z} dz - \int_{\nu}^{\infty} \frac{e^{-z}}{z} dz \\
&= \left(\frac{1}{\nu} \right)^{\frac{1}{1+\varrho}} \Gamma \left(\frac{1}{1+\varrho}, \nu \right) + \text{Ei}(-\nu), \tag{4.16}
\end{aligned}$$

$$\begin{aligned}
\mathbf{g}(\theta) &= \mathbb{E} \{ [1 + \mu(\theta, z)z]^{-\varrho} \} \\
&= \int_{\nu}^{\infty} \left(\frac{z}{\nu} \right)^{-\frac{\varrho}{1+\varrho}} e^{-z} dz + \int_0^{\nu} e^{-z} dz \\
&= \nu^{\frac{\varrho}{1+\varrho}} \Gamma \left(\frac{1}{1+\varrho}, \nu \right) + 1 - e^{-\nu}. \tag{4.17}
\end{aligned}$$

Above, $\Gamma(s, w) = \int_w^{\infty} \tau^{s-1} e^{-\tau} d\tau$ is the upper incomplete gamma function and $\text{Ei}(w) = -\int_{-w}^{\infty} \frac{e^{-\tau}}{\tau} d\tau$ is the exponential integral. When $p_{11} = 1 - s, p_{22} = s$, we have a memoryless discrete source and power allocation problem becomes

$$\min_{\mu(\theta, z) \geq 0} -\psi \log_e \left(\frac{\frac{1}{\mathbf{g}(\theta)} - (1 - s)}{s} \right) \tag{4.18}$$

$$\text{subject to } \psi \left(\frac{1}{\epsilon} \mathbb{E} \{ \mu(\theta, z) \} + \mu_c \right) = 1. \tag{4.19}$$

Thus, Lagrangian function transforms into

$$\mathcal{L}(\mu(\theta, z), \psi, \lambda) = -\psi \log_e \left(\frac{\frac{1}{\mathbf{g}(\theta)} - (1 - s)}{s} \right) + \lambda \left[\psi \left(\frac{1}{\epsilon} \mathbb{E} \{ \mu(\theta, z) \} + \mu_c \right) - 1 \right]. \tag{4.20}$$

Now, the KKT conditions become

$$\psi \left(\frac{1}{\epsilon} \mathbb{E} \{ \mu(\theta, z) \} + \mu_c \right) = 1, \quad (4.21)$$

$$\psi \frac{-\rho z [1 + \mu(\theta, z)z]^{-\rho-1}}{\mathbf{g}(\theta) - (1-s)(\mathbf{g}(\theta))^2} + \frac{\lambda \psi}{\epsilon} = 0, \quad (4.22)$$

$$-\log_e \left(\frac{\frac{1}{\mathbf{g}(\theta)} - (1-s)}{s} \right) + \lambda \left(\frac{1}{\epsilon} \mathbb{E} \{ \mu(\theta, z) \} + \mu_c \right) = 0. \quad (4.23)$$

Power allocation policy formula is still as in (4.13) but now ν is determined from

$$- [\mathbf{g}(\theta) - (1-s)(\mathbf{g}(\theta))^2] \log_e \left(\frac{\frac{1}{\mathbf{g}(\theta)} - (1-s)}{s} \right) + \nu^* \rho \epsilon \left(\frac{1}{\epsilon} \mathbb{E} \{ \mu(\theta, z) \} + \mu_c \right) = 0. \quad (4.24)$$

Next, we provide numerical results for the general case of discrete Markov source with memory. For the numerical analysis, we set the values of the parameters as $\mathbf{P}_c = 1$, $N_0 = 1$, $\rho = 1$, $\epsilon = 1$. In Fig. 4.1, we plot energy efficiency (EE) vs. maximum average arrival rate r_{avg}^* with varying source parameters. Note that when $p_{22} = 1$ and $p_{11} = 0$ (and hence $P_{\text{ON}} = 1$), we have a source with constant arrival rate. Indeed, the best performance is achieved in this case and maximal EE value (or equivalently the peak of the EE curve) is the largest. We further notice in the figure that initially the maximal EE values diminish and are achieved at a lower value of r_{avg}^* when p_{22} and consequently P_{ON} decrease and therefore the source burstiness increases. However, interestingly when p_{22} is diminished from 0.5 to 0.2, maximal EE value slightly increases even though P_{ON} is smaller when $p_{22} = 0.2$. This is due to the fact that P_{ON} is not the only criterion to indicate the burstiness of the system. In fact, as we have shown in [24], a measure of burstiness at low SNRs is $\frac{(1-p_{22})(p_{11}+p_{22})}{(1-p_{11})(2-p_{11}-p_{22})}$ whose greater values imply a more bursty source. Indeed, this expression assumes a larger value when $p_{22} = 0.5$. On the other hand, as SNR increases and higher average arrival rates are supported, P_{ON} becomes a more relevant indicator of burstiness and the source with $p_{22} = 0.2$ starts leading to lower EE values, following a crossover between the two

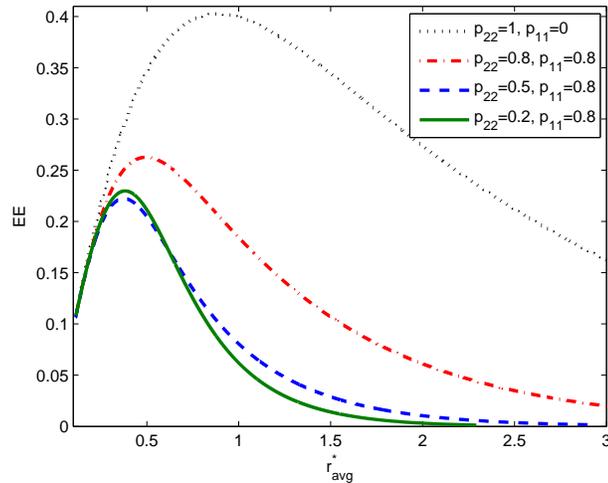


Fig. 4.1: Energy efficiency EE vs. maximum average arrival rate r_{avg}^* when $\varrho = 1$, $\mathbf{P}_c = 1$, $N_0 = 1$, $\epsilon = 1$.

curves.

In Fig. 4.2, we plot the optimal power control policy that maximizes the energy efficiency as a function of the instantaneous fading power values, z . We note in all cases that no power is allocated for transmission if z is below a threshold (i.e., ν). Power level initially increases as z increases above the threshold and then starts diminishing as z further grows. Hence, power control is essentially a combination of waterfilling policy (for low values of z) and channel inversion policy (for large values of z). We also observe that more power is consumed (and consequently average power consumption is larger) for a less bursty source at the maximal EE point.

In Fig. 4.3, we plot the EE vs. r_{avg}^* curves for different values of the circuit power \mathbf{P}_c . We readily notice that as \mathbf{P}_c diminishes, a higher level of EE is achieved at a lower value of r_{avg}^* . Indeed, if circuit power is not taken into account (i.e., if we set $\mathbf{P}_c = 0$), then maximum EE is achieved asymptotically as r_{avg}^* and hence SNR approach zero [24]. Hence, circuit power has significant impact on the performance.

In Fig. 4.4, we plot the maximum EE as a function of the QoS exponent θ for discrete Markov sources with different source statistics. We note in all cases that EE diminishes with increasing θ . Hence, more stringent buffer/delay constraints is detrimental to EE.

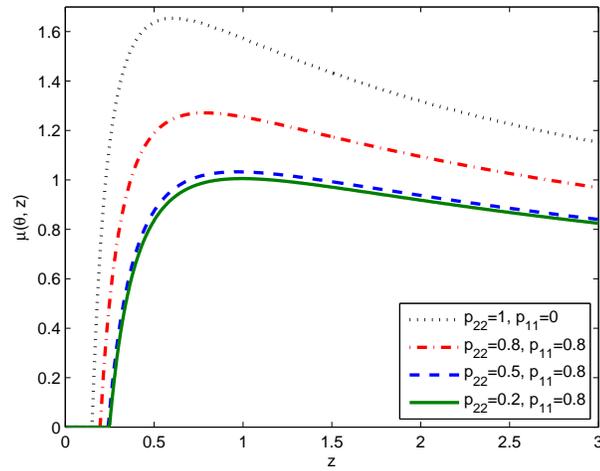


Fig. 4.2: Power control function $\mu(\theta, z)$ vs. z when $\varrho = 1$, $\mathbf{P}_c = 1$, $N_0 = 1$ and $\epsilon = 1$.

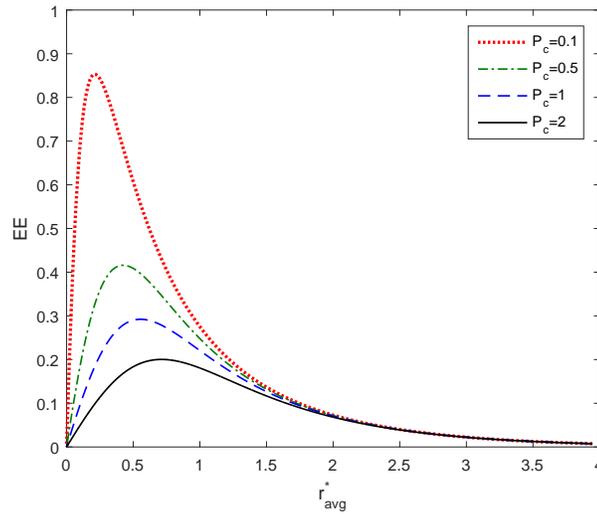


Fig. 4.3: Energy efficiency EE vs. maximum average arrival rate r_{avg}^* for different values of the circuit power P_c when $\varrho = 1$, $p_{11} = p_{22} = 0.5$, $N_0 = 1$, $\epsilon = 1$.

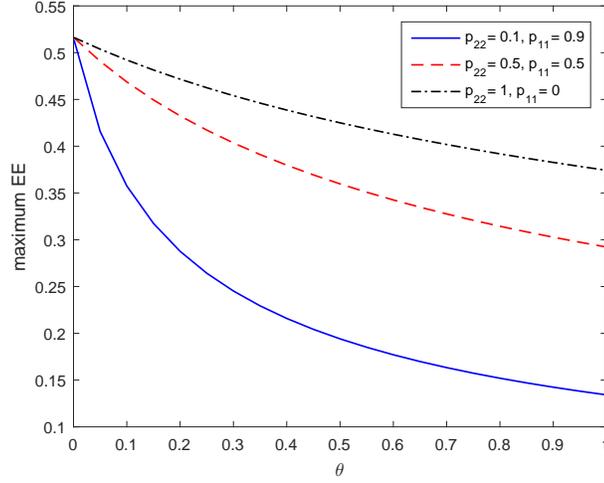


Fig. 4.4: Maximum EE vs. QoS exponent θ when $P_c = 1$, $N_0 = 1$, $\epsilon = 1$.

Also, similarly as before, the highest levels of EE are attained when the arrival rate is constant (i.e., when $p_{22} = 1$ and $p_{11} = 0$), and the EE diminishes as the sources become more bursty.

Finally, in Figs. 4.5 and 4.6, we display simulation results³. In particular, in Fig. 4.5, we have the buffer overflow probabilities $\Pr\{Q \geq q\}$ plotted as a function of the buffer threshold q with both optimal power control and no power control (i.e., with fixed transmission power). Note that we plot the buffer overflow probabilities in logarithmic scale. Note further from the approximation in (2.3) for large q that

$$\log \Pr\{Q \geq q\} \approx -\theta q + \log \varsigma. \quad (4.25)$$

Hence, the logarithm of the overflow probability is expected to decay linearly in θ for large q . Indeed, we observe this linear decay already even for rather small values of q . Moreover,

³We conduct the simulations as follows. We initially fix the value of the QoS exponent θ (e.g., $\theta = 2, 1$, or 0.5 in the figure) to provide a certain level of statistical QoS guarantee. Then, using the theoretical results from our analysis, we determine the EE-maximizing optimal power control and the value of r_{avg}^* at which EE is maximized. Subsequently, we generate random arrivals according to the discrete Markov process with average rate r_{avg}^* . We simulate the service process by generating random channel fading and using the optimal power control. Then, we have kept track of the buffer state among the arrivals and departures, and evaluated the frequency of exceeding a given threshold q to determine the values of overflow probabilities. Considering the same r_{avg}^* and the same average power and hence the same EE level, we have repeated the simulations with no power control.

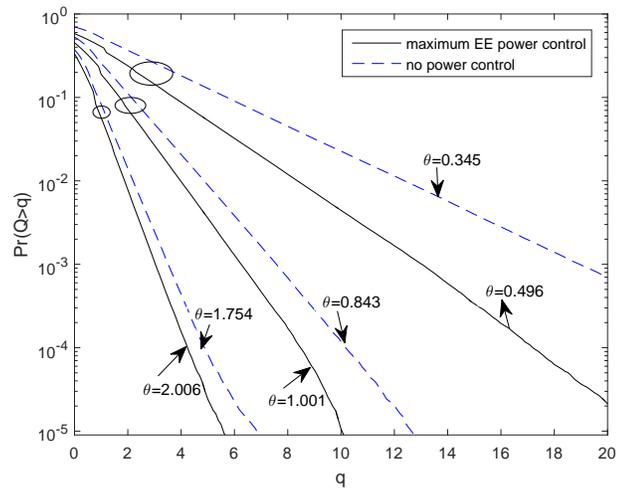


Fig. 4.5: Buffer overflow probability $\Pr\{Q \geq q\}$ vs. buffer threshold q for cases with optimal EE power control and fixed power when $\mathbf{P}_c = 1$, $N_0 = 1$, $\epsilon = 1$. Discrete Markov source with $p_{11} = p_{22} = 0.75$.

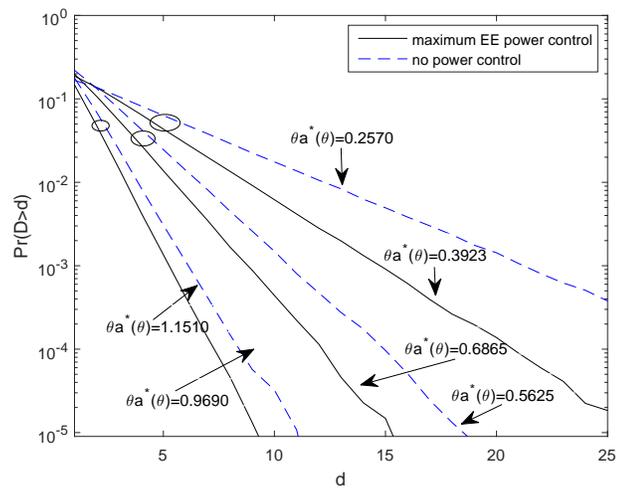


Fig. 4.6: Delay violation probability $\Pr\{D \geq d\}$ vs. delay threshold d for cases with optimal EE power control and fixed power when $\mathbf{P}_c = 1$, $N_0 = 1$, $\epsilon = 1$. Discrete Markov source with $p_{11} = p_{22} = 0.75$.

the simulations show excellent agreement with the theoretical analysis. Solid curves are for the case with optimal power control. We note that the simulated curves lead to simulated θ values of 2.006, 1.001, and 0.496 as indicated in the figure when we set $\theta = 2, 1,$ and $0.5,$ respectively, at the beginning of the simulations. Hence, the buffer overflow probabilities decay exponentially at the predicted rates even for very small values of q . In Fig. 4.5, the dashed curves next to the solid ones are the corresponding overflow probabilities at the same r_{avg}^* and EE levels but when no power control is employed. We immediately recognize that we have smaller values of θ in such cases (i.e., $\theta = 1.754, 0.843$ and 0.345 as opposed to having $\theta = 2.006, 1.001$ and $0.496,$ respectively, in the power control cases) meaning that for the same threshold q , the buffer overflow probabilities are higher when transmission power is fixed. Hence, the same EE can be attained but at the cost of having more frequent buffer overflows. Conversely, we can also say that for the same overflow probability, a higher EE is achieved when power control is adopted. These observations further demonstrate the benefits of power control in practical settings.

In Fig. 4.6, we plot the delay violation probability $\Pr\{D \geq d\}$ in logarithmic scale as a function of the delay threshold d again from the simulations. Note from the approximation in (2.5) that

$$\log \Pr\{D \geq d\} \approx -\theta a^*(\theta)d + \log \varsigma. \quad (4.26)$$

Hence, the logarithm of the delay violation probability is expected to decrease linearly in d with slope $-\theta a^*(\theta)$ where $a^*(\theta)$ is the effective bandwidth of the arrival process. We essentially have similar observations as in Fig. 4.5. Specifically, we again have excellent agreements with theory (e.g., the theoretical values in the power control cases are $\theta a^*(\theta) = 0.3927, 0.6944,$ and 1.2022 while the corresponding simulated values are $\theta a^*(\theta) = 0.3923, 0.6865,$ and $1.1510,$ respectively), and having no power control increases the frequency of delay violations at a given delay threshold.

$$\psi \left(\frac{1}{\epsilon} \mathbb{E} \{ \mu(\theta, z) \} + \mu_c \right) = 1, \quad (4.31)$$

$$- \psi \rho z [1 + \mu(\theta, z)z]^{-\rho-1} \left(\frac{\alpha\beta}{\mathbf{g}(\theta) (\alpha - \log_e \mathbf{g}(\theta))^2} + \frac{1}{\mathbf{g}(\theta)} \right) + \frac{\lambda\psi}{\epsilon} = 0, \quad (4.32)$$

$$\frac{\alpha + \beta - \log_e \mathbf{g}(\theta)}{\alpha - \log_e \mathbf{g}(\theta)} \log_e \mathbf{g}(\theta) + \lambda \left(\frac{1}{\epsilon} \mathbb{E} \{ \mu(\theta, z) \} + \mu_c \right) = 0. \quad (4.33)$$

4.2.2 Markov Fluid Source

Now, we consider the optimal power control with Markov fluid sources. By using the maximum average arrival rate expression in (3.14) in the objective function, eliminating the constant P_{ON} , and using the definition of $\mathbf{g}(\theta)$ in (4.3), we can recast the optimal power control problem in (4.4) as

$$\mu^*(\theta, z) = \arg \max_{\mu(\theta, z)} \frac{-\frac{\alpha + \beta - \log_e \mathbf{g}(\theta)}{\alpha - \log_e \mathbf{g}(\theta)} \log_e \mathbf{g}(\theta)}{\frac{1}{\epsilon} \mathbb{E} \{ \mu(\theta, z) \} + \mu_c}. \quad (4.27)$$

Again, by introducing the additional variable $\psi = \frac{1}{\mathbb{E} \{ \frac{1}{\epsilon} \mu(\theta, z) \} + \mu_c}$, the problem can be transformed into

$$\min_{\mu(\theta, z) \geq 0} \psi \frac{\alpha + \beta - \log_e \mathbf{g}(\theta)}{\alpha - \log_e \mathbf{g}(\theta)} \log_e \mathbf{g}(\theta) \quad (4.28)$$

$$\text{subject to} \quad \psi \left(\frac{1}{\epsilon} \mathbb{E} \{ \mu(\theta, z) \} + \mu_c \right) = 1. \quad (4.29)$$

By employing convex optimization tools, we can determine the sufficient and necessary KKT conditions. First, the Lagrangian function is given as

$$\begin{aligned} \mathcal{L}(\mu(\theta, z), \psi, \lambda) &= \psi \frac{\alpha + \beta - \log_e \mathbf{g}(\theta)}{\alpha - \log_e \mathbf{g}(\theta)} \log_e \mathbf{g}(\theta) \\ &\quad + \lambda \left[\psi \left(\frac{1}{\epsilon} \mathbb{E} \{ \mu(\theta, z) \} + \mu_c \right) - 1 \right]. \end{aligned} \quad (4.30)$$

The KKT conditions are given in (4.31)–(4.33) at the top of the next page. Similarly as for

the discrete Markov source, (4.31) is due to the constraint in (4.29). (4.32) and (4.33) are obtained by taking the derivative of the Lagrangian in (4.30) with respect to $\mu(\theta, z)$ and ψ , respectively. After simplifying (4.32), we obtain

$$z [1 + \mu(\theta, z)z]^{-\varrho-1} = \frac{\lambda/(\varrho\epsilon)}{\frac{\alpha\beta}{g(\theta)(\alpha-\log_e g(\theta))^2} + \frac{1}{g(\theta)}}. \quad (4.34)$$

Due to similarities between (4.12) and (4.34), the optimal power control function is obtained to be in the same form as for discrete Markov sources and is given by

$$\mu(\theta, z) = \left[\frac{1}{\nu^{\frac{1}{1+\varrho}} z^{\frac{\varrho}{1+\varrho}}} - \frac{1}{z} \right]^+ \quad (4.35)$$

but now with

$$\nu = \frac{\lambda/(\varrho\epsilon)}{\frac{\alpha\beta}{g(\theta)(\alpha-\log_e g(\theta))^2} + \frac{1}{g(\theta)}}. \quad (4.36)$$

Now, we can combine (4.33) with (4.36) to obtain

$$\begin{aligned} \frac{g(\theta)}{\frac{\alpha\beta}{(\alpha-\log_e g(\theta))^2} + 1} \frac{\alpha + \beta - \log_e g(\theta)}{\alpha - \log_e g(\theta)} \log_e g(\theta) \\ + \nu^* \varrho\epsilon \left(\frac{1}{\epsilon} \mathbb{E} \{ \mu(\theta, z) \} + \mu_c \right) = 0, \end{aligned} \quad (4.37)$$

which can further be used to numerically evaluate ν .

In Fig. 4.7, we plot the EE vs. maximum average arrival rate r_{avg}^* curve for different Markov fluid sources. As expected, the source with $\beta = 0$, being the constant arrival source, has the best performance in terms of energy efficiency. As α reduces, the source becomes more bursty and the performance degrades.

In Fig. 4.8, we again plot the EE vs. r_{avg}^* curves for a Markov fluid source with transition rates $\alpha = 2$ and $\beta = 8$, considering the optimal power control, suboptimal water-filling power control, and constant-power transmissions. As expected, optimal power control

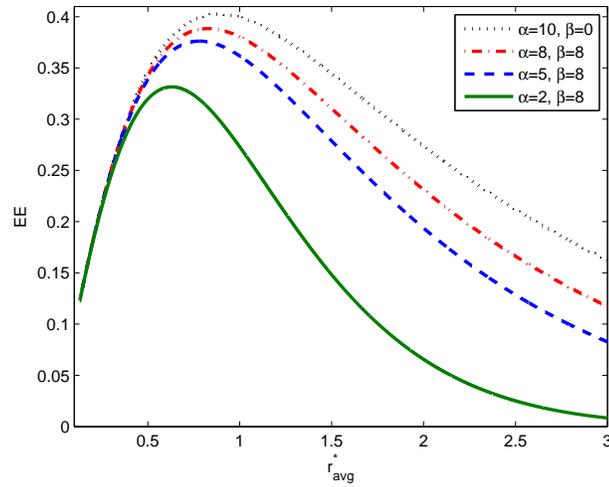


Fig. 4.7: Energy efficiency EE vs. maximum average arrival rate r_{avg}^* when $\varrho = 1$, $\mathbf{P}_c = 1$, $N_0 = 1$ $\epsilon = 1$.

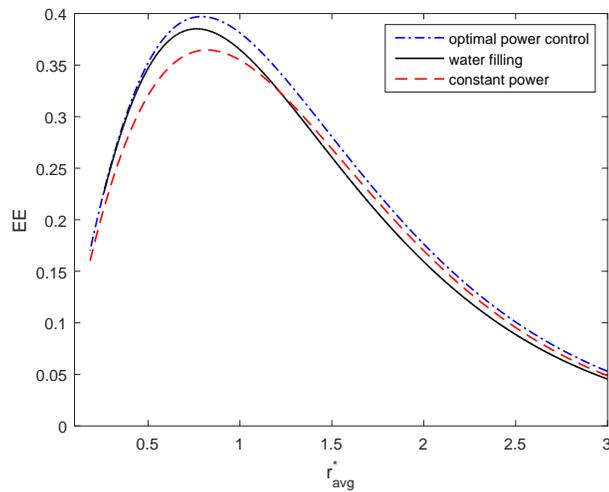


Fig. 4.8: Energy efficiency EE vs. maximum average arrival rate r_{avg}^* with different power control schemes when $\varrho = 0.5$, $\alpha = 2$, $\beta = 8$, $\mathbf{P}_c = N_0 = 1 = \epsilon = 1$.

leads to the maximum EE and outperforms the other two schemes uniformly over the entire range. Water-filling power control results in the second-highest EE level. At the same time, it is interesting to observe that transmission with constant power starts performing better than that with the water-filling policy as r_{avg}^* increases. This observation highlights the importance of identifying the optimal power control since water-filling takes into account neither source randomness nor QoS constraints.

4.2.3 Markov-Modulated Poisson Processes

The throughput expressions for discrete-time and continuous-time Markov-modulated Poisson sources have similarities to those for discrete-time Markov and Markov fluid sources, respectively. Particularly, (5.13) is obtained by scaling (3.6) with $\frac{\theta}{e^\theta - 1}$. The same observation holds regarding the comparison between (3.14) and (3.20). These scaling differences do not alter the optimal power control problem. Therefore, the optimal power control policies for the discrete-time and continuous-time MMPP sources are the same as for the cases of discrete and fluid Markov sources, respectively.

4.3 Optimal Power Control with EE Constraints

As noticed in the previous section, when the primary goal is the maximization of the energy efficiency, small throughput values can be attained especially if the source is bursty. On the other hand, in certain wireless systems, the goal is to maximize the throughput while being cognizant of the energy efficiency requirements. Motivated by such systems, we in this section assume that there is a minimum energy efficiency constraint on the system and we seek to find the optimal power allocation scheme to maximize the throughput. The optimal

power allocation problem is formulated as

$$\max_{\mu(\theta, z) \geq 0} r_{\text{avg}}^*(\theta) \quad (4.38)$$

$$\text{subject to} \quad \frac{r_{\text{avg}}^*(\theta)}{\frac{1}{\epsilon} \mathbb{E} \{ \mu(\theta, z) \} + \mu_c} \geq \zeta_{\text{min}}(\theta) \quad (4.39)$$

where EE_{min} represents the minimum required EE level. This optimization problem also enables us to characterize the tradeoff between the throughput and energy efficiency.

Note that the constraint can also be expressed as

$$-r_{\text{avg}}^*(\theta) + \text{EE}_{\text{min}}(\theta) \left[\frac{1}{\epsilon} \mathbb{E} \{ \mu(\theta, z) \} + \mu_c \right] \leq 0. \quad (4.40)$$

Again, we first demonstrate that the power allocation problem is convex and hence we can use convex optimization tools to solve the problem. As discussed at the end of Section 5.2, the objective function $r_{\text{avg}}^*(\theta)$ in (4.38) is a concave function of $\mu(\theta, z)$. It can be easily seen that the constraint in (4.40) is a convex function of $\mu(\theta, z)$ as it is the summation of a negative concave function and an affine function. Hence, the Lagrangian can be expressed as

$$\mathcal{L}(\mu(\theta, z), \lambda) = r_{\text{avg}}^*(\theta, z) - \lambda \left\{ -r_{\text{avg}}^*(\theta, z) + \text{EE}_{\text{min}}(\theta) \left[\frac{1}{\epsilon} \mathbb{E} \{ \mu(\theta, z) \} + \mu_c \right] \right\}. \quad (4.41)$$

4.3.1 Discrete Markov Source

The Lagrangian for discrete Markov source is simplified to

$$\mathcal{L}(\mu(\theta, z), \lambda) = (1 + \lambda) \frac{P_{\text{ON}}}{\theta} \log_e \left(\frac{1 - p_{11}g(\theta)}{(1 - p_{11} - p_{22})g^2(\theta) + p_{22}g(\theta)} \right) \lambda \text{EE}_{\text{min}}(\theta) \left[\frac{1}{\epsilon} \mathbb{E} \{ \mu(\theta, z) \} + \mu_c \right]. \quad (4.42)$$

After taking the first derivative, we obtain (4.43)

$$(1+\lambda)\frac{P_{\text{ON}}}{\theta}\varrho z [1+\mu(\theta, z)z]^{-\varrho-1}\left(\frac{(1-p_{11})(1-p_{22})}{(1-p_{11}g(\theta))((1-p_{11}-p_{22})g(\theta)+p_{22})} + \frac{1}{g(\theta)}\right) - \lambda EE_{\min}(\theta)\frac{1}{\epsilon} = 0. \quad (4.43)$$

(4.43) can further be expressed as

$$z [1 + \mu(\theta, z)z]^{-\varrho-1} = \nu, \quad (4.44)$$

where we define ν as

$$\nu = \frac{\lambda\theta EE_{\min}(\theta)}{\varrho\epsilon(1+\lambda)P_{\text{ON}}\left(\frac{(1-p_{11})(1-p_{22})}{(1-p_{11}g(\theta))((1-p_{11}-p_{22})g(\theta)+p_{22})} + \frac{1}{g(\theta)}\right)}. \quad (4.45)$$

Using (4.44), we can derive the power allocation formula as

$$\mu(\theta, z) = \left[\frac{1}{\nu^{\frac{1}{1+\varrho}} z^{\frac{\varrho}{1+\varrho}}} - \frac{1}{z} \right]^+. \quad (4.46)$$

We see that the power control formula for rate maximization under EE constraints is similar to that for maximizing EE. The key distinction lies in the formulation for ν in (4.45) which is different from (4.14).

In Fig. 4.9, we address the tradeoff between throughput and energy efficiency by solving the power control problem and determining the maximum throughput level under different energy efficiency constraints. More specifically, we plot the percentage gain in throughput by backing off from the maximal energy efficiency point, which is represented by the 100% EE gain⁴. The figure shows us that decreasing the energy efficiency leads to significant improvement in throughput. For instance, 20% reduction from the maximal

⁴Therefore, we can formulate EE gain percentage as $\frac{EE}{EE_{max}} \times 100\%$ where EE_{max} is the maximum energy efficiency. Similarly, throughput gain percentage is defined as $\frac{r_{\text{avg}}^*}{r_{\text{avg}, EE_{max}}^*} \times 100\%$ where $r_{\text{avg}, EE_{max}}^*$ is the average arrival rate at the maximum EE point.

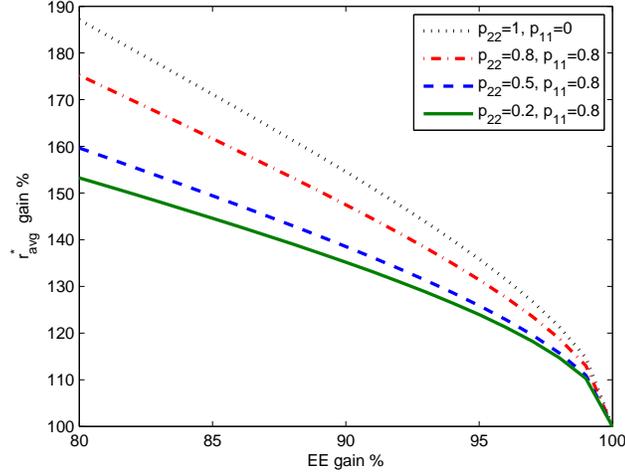


Fig. 4.9: Throughput r_{avg}^* gain % vs. EE gain %. $\theta = 1$. (Discrete Markov Source)

energy efficiency point results in 50 to 90% gain in throughput depending on the source characteristics. Even 1% decrease in energy efficiency generates about 10% gain on the throughput. We also note that the largest gain is realized in the case of constant arrival rate, and increasing burstiness reduces the throughput gain.

4.3.2 Markov Fluid Source

For Markov fluid source, Lagrangian function is given by

$$\mathcal{L}(\mu(\theta, z), \lambda) = (1 + \lambda) \frac{P_{\text{ON}}}{\theta} \frac{\alpha + \beta - \log_e \mathbf{g}(\theta)}{\alpha - \log_e \mathbf{g}(\theta)} \log_e \mathbf{g}(\theta) - \lambda \text{EE}_{\min}(\theta) \left[\frac{1}{\epsilon} \mathbb{E} \{ \mu(\theta, z) \} + \mu_c \right], \quad (4.47)$$

and the optimal power control policy has the same form as in (4.46) with ν defined as

$$\nu = \frac{\lambda \theta \text{EE}_{\min}(\theta)}{\varrho \epsilon (1 + \lambda) P_{\text{ON}} \left(\frac{\alpha \beta}{\mathbf{g}(\theta) (\alpha - \log_e \mathbf{g}(\theta))^2} + \frac{1}{\mathbf{g}(\theta)} \right)}. \quad (4.48)$$

In Fig. 4.10, we again demonstrate the tradeoff between energy efficiency and throughput when the source is modeled as a Markov fluid. We immediately observe that having a

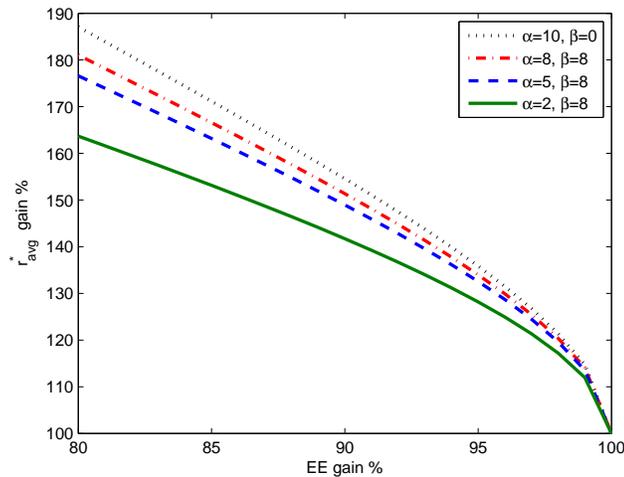


Fig. 4.10: Throughput r_{avg}^* gain % vs. EE gain %. $\theta = 1$. (Markov Fluid Source)

small reduction in the energy efficiency results in substantial gain in the throughput. On the other hand, the percentage of the gain decreases as the burstiness of the source increases (i.e. by decreasing α while β is fixed).

4.3.3 Markov Modulated Poisson Processes

The Lagrangian $\mathcal{L}(\mu(\theta, z), \lambda)$ and the scaled Lagrange multiplier ν for the discrete-time and continuous-time Markov-modulated Poisson sources can be immediately obtained by replacing θ with $(e^\theta - 1)$ in the corresponding expressions for discrete Markov and Markov fluid sources, respectively.

4.4 Optimal Power Control with Average Power Constraints

In this section, we consider a setting in which throughput maximization is the sole concern of the wireless system, and we study the optimal power control strategy that maximizes the

throughput under an average power constraint. The optimization problem is formulated as

$$\max_{\mathbf{P}(\theta, z) \geq 0} r_{\text{avg}}^*(\theta, z) \quad (4.49)$$

$$\text{subject to } \frac{1}{\epsilon} \mathbb{E} \{ \mathbf{P}(\theta, z) \} + \mathbf{P}_c \leq \bar{\mathbf{P}}. \quad (4.50)$$

Note that the optimization problems studied in previous sections have no explicit average power constraints. However, implicitly average power constraints are imposed through the energy efficiency requirements due to the fact that energy efficiency eventually starts diminishing with increasing average transmit power level. However, explicit average power constraints can be addressed without much difficulty as we demonstrate in this section.

The optimization problem in this section is again convex. Normalizing all the terms in the constraint in (4.50) with the noise power N_0B and denoting the average SNR = $\frac{\bar{\mathbf{P}}}{N_0B}$, the Lagrangian function can be written as

$$\mathcal{L}(\mu(\theta, z), \lambda) = r_{\text{avg}}^*(\theta, z) - \lambda \left\{ \frac{1}{\epsilon} \mathbb{E} \{ \mu(\theta, z) \} + \mu_c - \text{SNR} \right\}. \quad (4.51)$$

In the following analysis, we obtain the power allocation function for different source models using a similar approach as in previous sections. Specifically, we initially evaluate the first derivative of the Lagrangian function with respect to $\mu(\theta, z)$ and make it equal to 0. For all sources, the optimal power control is in the same form as in (4.46) with different ν expressions which we describe below for each source.

The first derivative of the Lagrangian for the discrete Markov source is obtained in (4.52) at the top of the next page.

$$\frac{P_{\text{ON}}}{\theta} \rho z [1 + \mu(\theta, z)z]^{-\rho-1} \left(\frac{(1 - p_{11})(1 - p_{22})}{(1 - p_{11}g(\theta))((1 - p_{11} - p_{22})g(\theta) + p_{22})} + \frac{1}{g(\theta)} \right) - \frac{\lambda}{\epsilon} = 0 \quad (4.52)$$

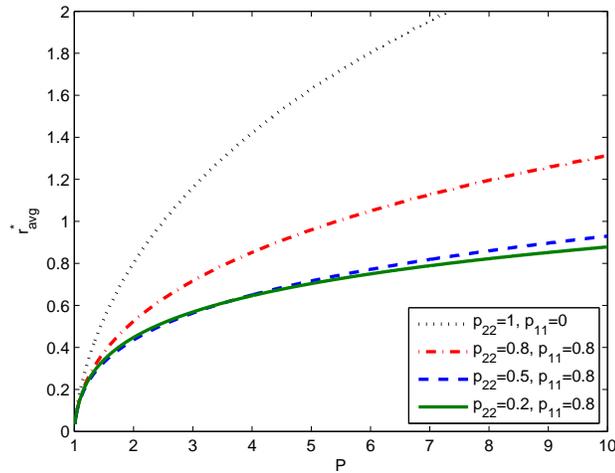


Fig. 4.11: Maximum average arrival rate r_{avg}^* vs. average power \bar{P} when $\theta = 1$. (Discrete Markov Source)

Using (4.52), ν is derived as

$$\nu = \frac{\lambda\theta}{\rho\epsilon P_{\text{ON}} \left(\frac{(1-p_{11})(1-p_{22})}{(1-p_{11}g(\theta))((1-p_{11}-p_{22})g(\theta)+p_{22})} + \frac{1}{g(\theta)} \right)}. \quad (4.53)$$

We plot the throughput vs. average power curve in Fig. 4.11 where we take into account different discrete Markov sources. As noted before, the best performance is realized for the case of constant arrival rates (i.e., when $p_{22} = 1$ and $p_{11} = 0$), and the throughput degrades with increased burstiness. Comparing the performances with source models with parameters $p_{22} = 0.5$, $p_{11} = 0.8$ and $p_{22} = 0.2$, $p_{11} = 0.8$, we observe that ON probability, P_{ON} , becomes a dominant factor on performance as average power increases. Source with smaller P_{ON} (i.e., the one with transition probabilities $p_{22} = 0.2$, $p_{11} = 0.8$) has lower performance. On the other hand, when the average power is relatively low, this source outperforms the one with parameters $p_{22} = 0.5$, $p_{11} = 0.8$ since the metric $\frac{(1-p_{22})(p_{11}+p_{22})}{(1-p_{11})(2-p_{11}-p_{22})}$ is a more critical burstiness factor at low SNR values (as also discussed in Section 4.2.1). Indeed, in this case, the transition probabilities $p_{22} = 0.5$, $p_{11} = 0.8$ result in a larger value for $\frac{(1-p_{22})(p_{11}+p_{22})}{(1-p_{11})(2-p_{11}-p_{22})}$ indicating a more bursty source in the low-SNR regime.

For the Markov fluid case, we obtain (4.54) given on the next page as the first derivative

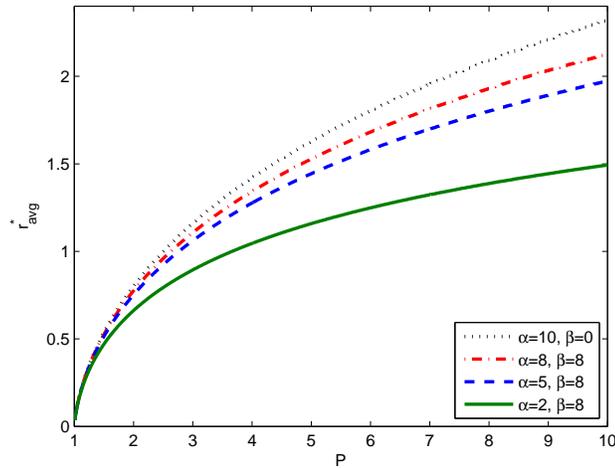


Fig. 4.12: Maximum average arrival rate r_{avg}^* vs. average power \bar{P} when $\theta = 1$. (Markov Fluid Source)

of the Lagrangian with respect to $\mu(\theta, z)$.

$$\frac{P_{\text{ON}}}{\theta} \rho z [1 + \mu(\theta, z)z]^{-e-1} \left(\frac{\alpha\beta}{g(\theta) (\alpha - \log_e g(\theta))^2} + \frac{1}{g(\theta)} \right) - \frac{\lambda}{\epsilon} = 0. \quad (4.54)$$

The parameter ν that we use in power allocation formula is given by

$$\nu = \frac{\lambda\theta}{\rho\epsilon P_{\text{ON}} \left(\frac{\alpha\beta}{g(\theta)(\alpha - \log_e g(\theta))^2} + \frac{1}{g(\theta)} \right)}. \quad (4.55)$$

For Markov fluid source we demonstrate the throughput as a function of average power in Fig. 4.12. Similarly as before, burstiness hurts the performance.

For discrete-time and continuous-time MMPP, the first derivatives of the Lagrangian functions with respect to $\mu(\theta, z)$ are given, respectively, by (4.52) and (4.54) with $\frac{P_{\text{ON}}}{\theta}$ replaced by $\frac{P_{\text{ON}}}{(e^\theta - 1)}$, and with the corresponding threshold parameters ν given, respectively, by (4.53) and (4.55) when $\lambda\theta$ is replaced by $\lambda(e^\theta - 1)$.

4.5 Optimal Power Control in Multichannel Systems

In this section, motivated by the fact that multicarrier channels employing orthogonal frequency division multiplexing (OFDM) can be regarded as multichannel systems, we extend our power control analysis to multichannel communication links. In a multichannel scenario (e.g., in multicarrier models), assuming that there are K subchannels each with bandwidth $\frac{B}{K}$, the instantaneous service rate becomes

$$R(\mathbf{z}) = \sum_{k=1}^K \frac{B}{K} \log_2(1 + \mu_k(\mathbf{z})z_k). \quad (4.56)$$

Above, we define $\mu_k(\mathbf{z}) = \frac{P_k(\theta, \mathbf{z})}{N_0 \frac{B}{K}}$, where $P_k(\theta, \mathbf{z})$ is the instantaneous transmission power in the k^{th} subchannel, and $\mathbf{z} = [z_1, \dots, z_K]$, where $z_k = |h_k|^2$ is the magnitude-square of the fading coefficient in the k^{th} subchannel. Under the block-fading assumption, the effective capacity with K subchannels can be expressed as

$$\begin{aligned} C_E(\theta) &= -\frac{1}{\theta} \log_e \mathbb{E} \{ e^{-\theta T R(\mathbf{z})} \} \\ &= -\frac{1}{\theta} \log_e \mathbb{E} \left\{ \prod_{k=1}^K e^{-\theta T B \frac{1}{K} \log_2(1 + \mu_k(\mathbf{z})z_k)} \right\} \\ &= -\frac{1}{\theta} \log_e \mathbb{E} \left\{ \prod_{k=1}^K (1 + \mu_k(\mathbf{z})z_k)^{-\frac{\theta}{K}} \right\}. \end{aligned} \quad (4.57)$$

In order to keep the analysis concise in this section, we only consider the problem of finding the optimal power allocation scheme that maximizes the throughput under a minimum energy efficiency constraint for discrete Markov and Markov fluid sources. The optimal power allocation problem can be expressed as the following convex optimization problem:

$$\max_{\mu(\theta, \mathbf{z}) \geq 0} r_{\text{avg}}^*(\theta, z) \quad (4.58)$$

$$\text{subject to} \quad \frac{r_{\text{avg}}^*(\theta, z)}{\frac{1}{\epsilon} \sum_{k=1}^K \mathbb{E} \{ \mu_k(\theta, \mathbf{z}) \} + \mu_c} \geq \text{EE}_{\min}(\theta) \quad (4.59)$$

where we define $\mu(\theta, \mathbf{z}) = [\mu_1(\theta, \mathbf{z}), \dots, \mu_K(\theta, \mathbf{z})]$, $\mu_k(\theta, \mathbf{z}) = \frac{\mathbf{P}_k(\theta, \mathbf{z})}{N_0 \frac{B}{K}}$ and $\mu_c = \frac{\mathbf{P}_c}{N_0 \frac{B}{K}} \cdot \zeta_{\min}$ is the minimum required energy efficiency level.

We can further rewrite the constraint as

$$-r_{\text{avg}}^*(\theta, z) + \text{EE}_{\min}(\theta) \left[\frac{1}{\epsilon} \sum_{k=1}^K \mathbb{E} \{ \mu_k(\theta, \mathbf{z}) \} + \mu_c \right] \leq 0. \quad (4.60)$$

Now, the Lagrangian becomes

$$\begin{aligned} \mathcal{L}(\mu(\theta, \mathbf{z}), \lambda) &= r_{\text{avg}}^*(\theta, z) \\ &\quad - \lambda \left\{ -r_{\text{avg}}^*(\theta, z) + \text{EE}_{\min}(\theta) \left[\frac{1}{\epsilon} \sum_{k=1}^K \mathbb{E} \{ \mu_k(\theta, \mathbf{z}) \} + \mu_c \right] \right\} \\ &\quad - \sum_{k=1}^K \lambda_k \mu_k(\theta, \mathbf{z}). \end{aligned} \quad (4.61)$$

To determine the optimal power control policy, we have to consider the solution of

$$\frac{\partial \mathcal{L}(\mu(\theta, \mathbf{z}), \lambda)}{\partial \mu(\theta, \mathbf{z})} = 0. \quad (4.62)$$

If we have $\mu_i(\theta, \mathbf{z}) > 0$ for $i \in \mathcal{N}_0 = \{1, \dots, K\}$, complementary slackness dictates that the corresponding Lagrangian multiplier λ_i is zero [31]. In the rest of the analysis we exploit this property.

4.5.1 Discrete Markov Source

First, let us define

$$\begin{aligned} \mathcal{L}(\mu(\theta, \mathbf{z}), \lambda) = & (1 + \lambda) \frac{P_{\text{ON}}}{\theta} \log_e \left(\frac{1 - p_{11} \mathbf{g}(\theta)}{(1 - p_{11} - p_{22}) \mathbf{g}^2(\theta) + p_{22} \mathbf{g}(\theta)} \right) \\ & - \lambda \mathbb{E} \mathbb{E}_{\min}(\theta) \left[\frac{1}{\epsilon} \sum_{k=1}^K \mathbb{E} \{ \mu_k(\mathbf{z}) \} + \mu_c \right] - \sum_{k=1}^K \lambda_k \mu_k(\mathbf{z}). \end{aligned} \quad (4.64)$$

$$\begin{aligned} (1 + \lambda) \frac{P_{\text{ON}}}{\theta} \frac{\varrho z_k}{K} [1 + \mu_k(\mathbf{z}) z_k]^{-\frac{\varrho}{K} - 1} \prod_{i \neq k} [1 + \mu_i(\mathbf{z}) z_i]^{-\frac{\varrho}{K}} \\ \times \left(\frac{(1 - p_{11})(1 - p_{22})}{(1 - p_{11} \mathbf{g}(\theta))((1 - p_{11} - p_{22}) \mathbf{g}(\theta) + p_{22})} + \frac{1}{\mathbf{g}(\theta)} \right) - \lambda \mathbb{E} \mathbb{E}_{\min}(\theta) \frac{1}{\epsilon} = 0 \end{aligned} \quad (4.65)$$

$$\mathbf{g}(\theta, \mathbf{z}) = \mathbb{E} \left\{ \prod_{k=1}^K [1 + \mu_k(\mathbf{z}) z_k]^{-\frac{\varrho}{K}} \right\} \quad (4.63)$$

where $\varrho = \theta T B \log_2 e$. The Lagrangian for the discrete Markov source is expressed in (4.64) on the next page.

Initially, we assume that we utilize all of the subchannels for transmission. Then, we can immediately state that $\lambda_i = 0$ for all i and derive the optimality equations as in (4.65) on the next page by calculating the derivatives with respect to μ_i for $i \in \mathcal{N}_0$.

We simplify (4.65) as

$$\nu = z_k [1 + \mu_k(\mathbf{z}) z_k]^{-\frac{\varrho}{K} - 1} \prod_{i \neq k} [1 + \mu_i(\mathbf{z}) z_i]^{-\frac{\varrho}{K}}, \quad k \in \mathcal{N}_0 \quad (4.66)$$

where ν is a scaled Lagrangian multiplier

$$\nu = \frac{\lambda K \mathbb{E} \mathbb{E}_{\min}(\theta) \log_e 2}{\epsilon (1 + \lambda) P_{\text{ON}} T B \left(\frac{(1 - p_{11})(1 - p_{22})}{(1 - p_{11} \mathbf{g}(\theta))((1 - p_{11} - p_{22}) \mathbf{g}(\theta) + p_{22})} + \frac{1}{\mathbf{g}(\theta)} \right)}. \quad (4.67)$$

By solving equations in (4.66), the optimal power allocation can be written as

$$\mu_k(\theta, \mathbf{z}) = \frac{1}{\nu^{\frac{1}{1+\varrho}} \prod_{i \in \mathcal{N}_0} z_i^{\frac{\varrho}{K(1+\varrho)}}} - \frac{1}{z_k}, \quad k \in \mathcal{N}_0. \quad (4.68)$$

Now we define \mathcal{N}_1 as

$$\mathcal{N}_1 = \left\{ k \in \mathcal{N}_0 \left| \frac{1}{\nu^{\frac{1}{1+\epsilon}} \prod_{i \in \mathcal{N}_0} z_i^{\frac{\epsilon}{K(1+\epsilon)}}} - \frac{1}{z_k} > 0 \right. \right\} \quad (4.69)$$

If $\mathcal{N}_1 = \mathcal{N}_0$ holds, then (4.68) is the optimal solution, otherwise we need to apply a recursive strategy which we describe as an algorithm in Table 1.

Algorithm 1 The optimal power control algorithm that maximizes throughput given a minimum EE constraint

- 1: Given \mathcal{N}_1 compute $N_1 = |\mathcal{N}_1|$;
 - 2: Initialize $k = 1$;
 - 3: **while** $\mathcal{N}_k \neq \mathcal{N}_{k-1}$ **do**
 - 4: $\mathcal{N}_{k+1} = \left\{ n \in \mathcal{N}_k \left| \frac{1}{\nu^{\frac{K}{K+N_k\epsilon}} \prod_{i \in \mathcal{N}_k} z_i^{\frac{\epsilon}{K+N_k\epsilon}}} - \frac{1}{z_n} > 0 \right. \right\}$;
 - 5: $N_{k+1} = |\mathcal{N}_{k+1}|$;
 - 6: $k = k + 1$;
 - 7: **end while**
 - 8: Define $\mathcal{N}^* = \mathcal{N}_k$ and $K^* = |\mathcal{N}_k|$;
 - 9: $\mu_n(\theta, \mathbf{z}) = \begin{cases} \frac{1}{\nu^{\frac{K}{K+K^*\epsilon}} \prod_{i \in \mathcal{N}^*} z_i^{\frac{\epsilon}{K+K^*\epsilon}}} - \frac{1}{z_k} & \text{for } n \in \mathcal{N}^*; \\ 0 & \text{otherwise.} \end{cases}$
-

Remark 10. *In the algorithm, basically, we first employ the formula in (4.68) for the power allocation. Then, if all power levels are above zero, we stop the algorithm. Otherwise, for subchannels with power levels less than zero, we do not allocate any power and we cease using these subchannels in the algorithm.*

Fig. 4.13 depicts the energy efficiency as a function of the maximum average arrival rate (or equivalently throughput) with varying source characteristics and the number of subchannels. The random arrivals are modeled as a discrete Markov process. When we have a higher ON probability or more subchannels, the system has better performance in terms of energy efficiency as the maximum EE point is the highest out of all scenarios when $P_{\text{ON}} = 1$ and $K = 4$. The source burstiness is in general an important factor in the presence of QoS constraints because increasing burstiness by decreasing the ON probability makes

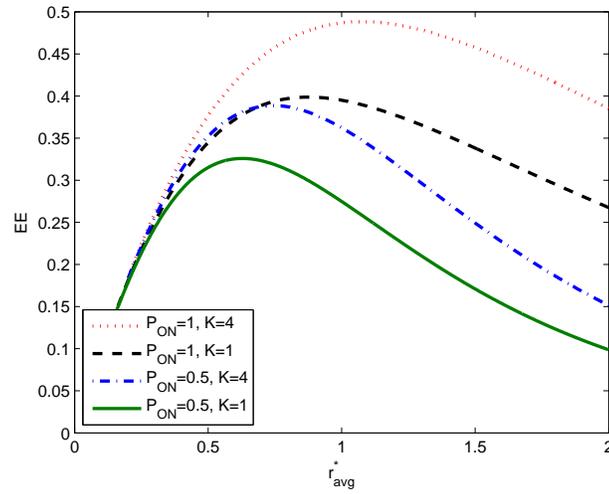


Fig. 4.13: Energy efficiency EE vs. maximum average arrival rate r_{avg}^* when $N_0 = 1$, $\varrho = 1$.

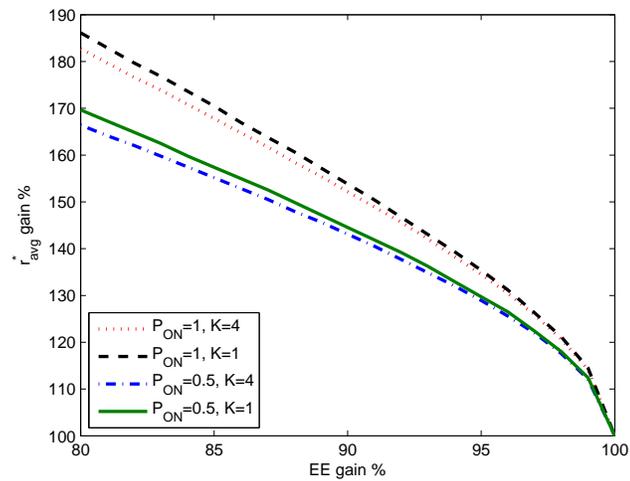


Fig. 4.14: Throughput r_{avg}^* gain % vs. EE gain % when $N_0 = 1$, $\varrho = 1$.

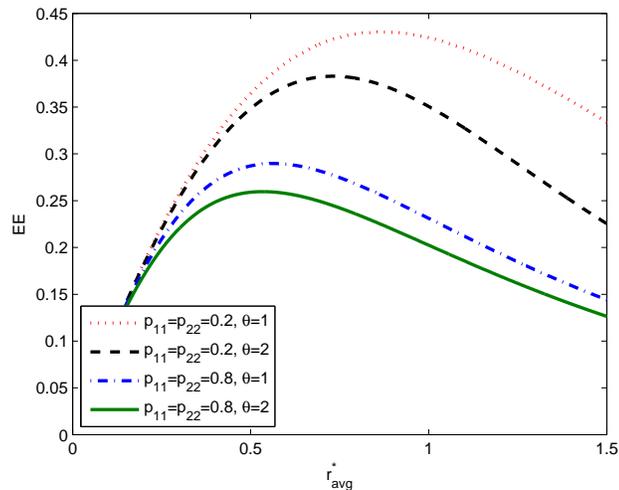


Fig. 4.15: Energy efficiency EE vs. maximum average arrival rate r_{avg}^* when $N_0 = 1$, $K = 4$.

the system more susceptible to buffer overflows. To avoid higher buffer overflow probabilities, the system supports smaller throughput with the same energy budget. As noted above, having more subchannels improves the energy efficiency. Essentially, when we use a single channel, the random variations in the wireless channel, which can be detrimental in the presence of buffer overflow constraints, have a more significant impact.

In Fig. 4.14, we analyze the tradeoff between throughput and energy efficiency. Similarly as in Fig. 4.9, we describe the maximum EE point (i.e., the peak of the bell-shaped EE curves in Fig. 4.13) as 100% on the x -axis and decrease the energy efficiency while computing the gain in the throughput. This figure shows how much throughput can be improved by sacrificing from the maximum energy efficiency. The observation from Fig. 4.14 is that throughput improvement for less bursty sources is higher. Also, with smaller number of subchannels, we observe a larger improvement.

In Fig. 4.15 where we again depict the energy efficiency as a function of the throughput, we investigate how the system performs under different levels of burstiness and QoS requirements. We notice that burstiness does not depend on P_{ON} only. Although systems with $p_{11} = p_{22} = 0.2$ and $p_{11} = p_{22} = 0.8$ have the same $P_{\text{ON}} = 0.5$, they perform

differently. Sources with higher transition probabilities from one state to a different state (i.e., higher p_{12} and p_{21}) exhibit reduced burstiness, and hence we have better performance with the source having $p_{11} = p_{22} = 0.2$. Additionally, more stringent QoS constraints (i.e. higher values of θ) clearly reduce the energy efficiency of the system no matter what the source characteristics are.

4.5.2 Markov Fluid Source

The Lagrangian function of the Markov fluid source is given by

$$\begin{aligned} \mathcal{L}(\mu(\theta, \mathbf{z}), \lambda) = (1 + \lambda) \frac{P_{\text{ON}} \alpha + \beta - \log_e \mathbf{g}(\theta)}{\theta \alpha - \log_e \mathbf{g}(\theta)} \log_e \mathbf{g}(\theta) - \lambda \mathbb{E} \mathbb{E}_{\min}(\theta) \left[\frac{1}{\epsilon} \sum_{k=1}^K \mathbb{E} \{ \mu_k(\theta, \mathbf{z}) \} + \mu_c \right] \\ - \sum_{k=1}^K \lambda_k \mu_k(\theta, \mathbf{z}). \end{aligned} \quad (4.70)$$

Again, we initially assume that we utilize all subchannels for transmission. Then, we can immediately state that $\lambda_i = 0$ for all i and derive the following optimality equation in (4.71) by calculating the derivatives with respect to μ_i for $i \in \mathcal{N}_0$:

$$\begin{aligned} (1 + \lambda) \frac{P_{\text{ON}}}{\theta} \frac{\rho z_k}{K} [1 + \mu_k(\mathbf{z}) z_k]^{-\frac{\rho}{K} - 1} \prod_{i \neq k} [1 + \mu_i(\mathbf{z}) z_i]^{-\frac{\rho}{K}} \left(\frac{\alpha \beta}{\mathbf{g}(\theta) (\alpha - \log_e \mathbf{g}(\theta))^2} + \frac{1}{\mathbf{g}(\theta)} \right) \\ - \lambda \mathbb{E} \mathbb{E}_{\min}(\theta) \frac{1}{\epsilon} = 0. \end{aligned} \quad (4.71)$$

We can simplify (4.71) as

$$\nu = z_k [1 + \mu_k(\mathbf{z}) z_k]^{-\frac{\rho}{K} - 1} \prod_{i \neq k} [1 + \mu_i(\mathbf{z}) z_i]^{-\frac{\rho}{K}}, \quad k \in \mathcal{N}_0 \quad (4.72)$$

where ν is defined as

$$\nu = \frac{\lambda K \mathbb{E} \mathbb{E}_{\min}(\theta) \log_e 2}{\epsilon (1 + \lambda) P_{\text{ON}} T B \left(\frac{\alpha \beta}{\mathbf{g}(\theta) (\alpha - \log_e \mathbf{g}(\theta))^2} + \frac{1}{\mathbf{g}(\theta)} \right)}. \quad (4.73)$$

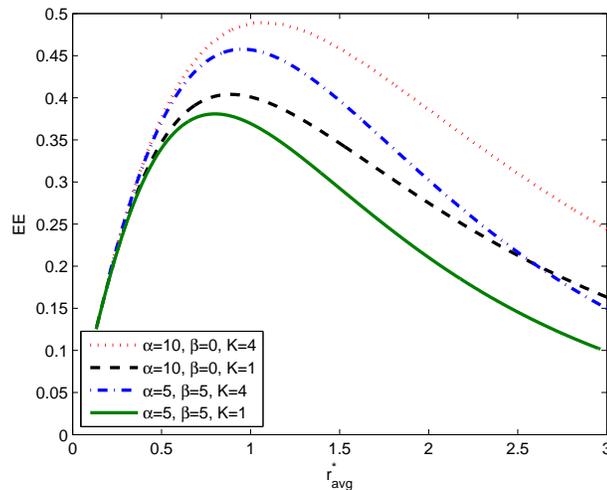


Fig. 4.16: Energy efficiency EE vs. maximum average arrival rate r_{avg}^* when $N_0 = 1$, $\rho = 1$.

Note that, the formulation of ν in (4.72) is exactly the same as in (4.66). Thus, the optimal power control for the case with the Markov fluid source follows from (4.68) and algorithm from Table 1.

In Fig. 4.16, we plot the energy efficiency vs. maximum average arrival rate curves for Markov fluid arrivals. We immediately observe that having more subchannels again improves the performance in terms of energy efficiency. Also, the maximum energy efficiency is achieved at a larger throughput level when P_{ON} or K (number of subchannels) increases. Additionally, we notice that at high SNR levels (equivalently for large r_{avg}^* values), increased burstiness can offset improvements due to the increased number of subchannels, as evidenced by the crossover between the dashed curve (for which $\alpha = 10, \beta = 0$ and hence the arrival rate is constant, and $K = 1$) and dot-dashed curve (for which $\alpha = 5, \beta = 5, K = 4$).

For Markov fluid sources, we analyze the energy efficiency and throughput tradeoff in Fig. 4.17. From Fig. 4.16, we observe the steep loss in energy efficiency for more bursty sources. This observation is further reflected in Fig. 4.17 as the throughput gain is lower for more bursty sources when we have the same percentage of sacrifice from the energy

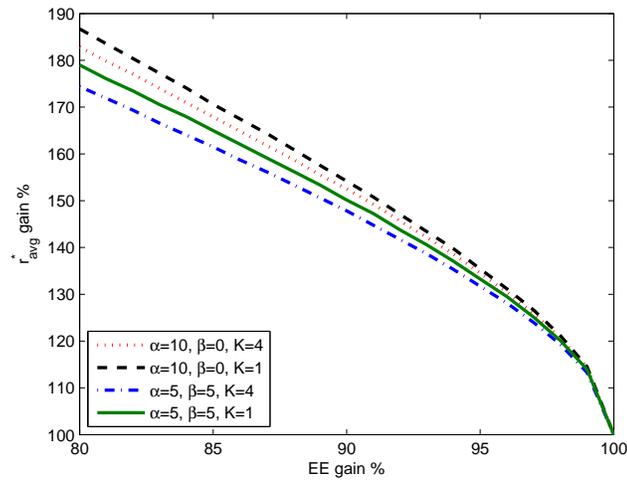


Fig. 4.17: Throughput r_{avg}^* gain % vs. EE gain % when $N_0 = 1$, $\varrho = 1$.

efficiency. Overall, we also note that instead of working at the optimal energy efficiency point, if we reduce the energy efficiency by about 20%, we can obtain gains, reaching up to almost twice the throughput levels.

Finally, in Fig. 4.18 we plot the energy efficiency curves with varying number of subchannels and QoS constraints for the Markov fluid source. Again, our previous observations are verified as increasing the number of subchannels K or decreasing the value of QoS exponent θ enhances the energy efficiency.

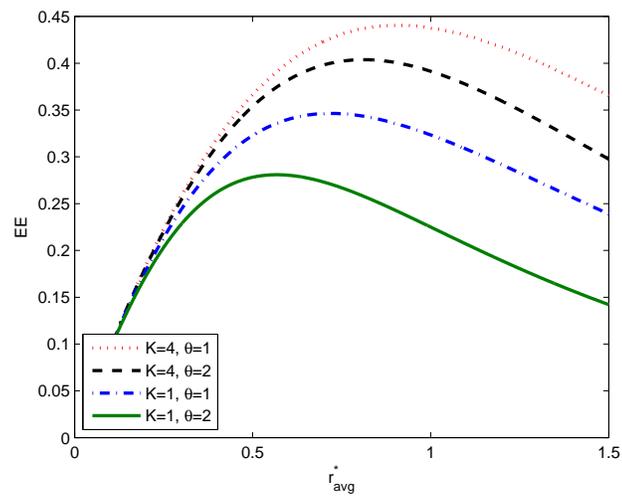


Fig. 4.18: Energy efficiency EE vs. maximum average arrival rate r_{avg}^* when $N_0 = 1$, $\alpha = \beta = 5$.

CHAPTER 5

SECURE TRANSMISSION OF DELAY-SENSITIVE DATA OVER WIRELESS FADING CHANNELS

5.1 Channel Model

As depicted in Figure 5.1, we consider a fading broadcast channel in which a transmitter sends common and confidential messages to two receivers. Messages are stored in buffers before being transmitted. Specifically, confidential messages intended for receiver 1 and receiver 2 are kept in buffers labeled 1 and 2, respectively, as shown in Fig. 5.1, and common messages are stored in buffer 0. Since delay-sensitive data traffic is considered,

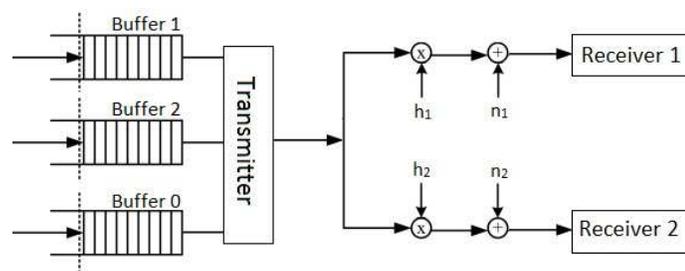


Fig. 5.1: Two-receiver broadcast channel model.

statistical queueing constraints are imposed in order to limit buffer overflows and delay violations. We assume flat-fading between the transmitter and receivers. The channel input-output relation can be expressed as

$$y_j = h_j x + n_j \text{ for } j = 1, 2 \quad (5.1)$$

where x is the channel input and y_j is the output at the j^{th} receiver for $j \in \{1, 2\}$. Input signal includes both confidential and common messages. Average transmitted signal energy is $\mathbb{E}\{|x|^2\} = \mathcal{E}$. Moreover, in (5.1), h_i denotes the fading coefficient in the channel between the transmitter and receiver j . Finally, n_j denotes the zero-mean, circularly-symmetric, complex Gaussian background noise at receiver j with variance $\mathbb{E}\{|n_j|^2\} = N_0$. Hence, the *input signal-to-noise ratio* (SNR) is

$$\text{SNR} = \frac{\mathbb{E}\{|x|^2\}}{\mathbb{E}\{|n_j|^2\}} = \frac{\mathcal{E}}{N_0} \quad j = 1, 2. \quad (5.2)$$

While fading coefficients can have arbitrary distributions with finite energies, we assume that block-fading is experienced. Hence, the realizations of the fading coefficients stay fixed for a block of symbols and change independently for the next block.

5.1.1 Instantaneous Secrecy Capacity of Confidential Messages and Capacity of Common Message Transmissions

In this section, we describe the secrecy capacity in detail in a general case in which the transmitter sends both common and confidential messages¹ to two receivers, and, with that, we identify the service rates of our queueing model. Confidential and common messages are sent simultaneously and it is assumed that common message is decoded at the receiver in the presence of the interference from the confidential message transmission. Confidential

¹Here, we consider standard information-theoretic arguments regarding the definition of messages and how they are encoded and transmitted over fading channels (see e.g., [67], [68]).

messages of two receivers are sent necessarily using time-division duplexing depending on the channel strengths. More specifically, confidential message is only sent to the receiver with the higher received SNR.

Secrecy capacity quantifies the maximum achievable rates of secure communication. For instance, it is well-known that the secrecy capacity of confidential message transmission with the signal-to-noise ratio denoted by SNR in the presence of an eavesdropper is given by

$$R(\text{SNR}) = [\log_2(1 + \text{SNR}z_m) - \log_2(1 + \text{SNR}z_e)]^+. \quad (5.3)$$

Note that the above formula of secrecy capacity is a generic one with z_m and z_e denoting the magnitude squares of the fading coefficients of channels of the intended user and eavesdropper, respectively. When the transmitter sends separate confidential messages to each user as we have assumed and described in Section 5.1, the unintended user can be regarded as an eavesdropper.

Having two confidential messages and one common message to send, transmitter allocates its power for the transmission of these messages. We assume that when confidential message intended for receiver i is being sent, δ_i portion of the power is used for confidential message transmission while $(1 - \delta_i)$ portion of the power is used for common message transmission. Additionally, we define the regions for time-division duplexing of confidential messages as

$$\begin{aligned} \Gamma_1 &= \{(z_1, z_2) \in \mathbb{R}^{2+} : z_1 \geq z_2\}, \\ \Gamma_2 &= \{(z_1, z_2) \in \mathbb{R}^{2+} : z_1 < z_2\}. \end{aligned}$$

For instance, when we have $(z_1, z_2) \in \Gamma_1$, only confidential message intended for receiver 1 is transmitted along with the common message². As previously stated, the common mes-

²We note that the event of $z_1 = z_2$ occurs with zero probability if the fading powers z_1 and z_2 have

sage is decoded in the presence of interference from confidential message transmissions. Both users can decode the common message when it is sent at a rate they both can decode, implying that the common message is sent at the minimum rate that both channels can support. Hence, the instantaneous transmission rate of the common message becomes

$$R_0(\text{SNR}) = \log_2 \left(1 + \frac{(1 - \delta_1)\text{SNR}z_2}{1 + \delta_1\text{SNR}z_2} \right) \mathbf{1}\{\Gamma_1\} + \log_2 \left(1 + \frac{(1 - \delta_2)\text{SNR}z_1}{1 + \delta_2\text{SNR}z_1} \right) \mathbf{1}\{\Gamma_2\}. \quad (5.4)$$

After subtracting the common message from the received signal, the receiver with the better channel can decode its confidential message without any interference from the common message. Therefore, we can express the instantaneous transmission rate of confidential messages intended for receivers 1 and 2, respectively, as

$$R_1(\text{SNR}) = \log_2 \left(\frac{1 + \delta_1\text{SNR}z_1}{1 + \delta_1\text{SNR}z_2} \right) \mathbf{1}\{\Gamma_1\} \quad (5.5)$$

$$R_2(\text{SNR}) = \log_2 \left(\frac{1 + \delta_2\text{SNR}z_2}{1 + \delta_2\text{SNR}z_1} \right) \mathbf{1}\{\Gamma_2\} \quad (5.6)$$

where $\mathbf{1}\{\cdot\}$ denotes the indicator function³.

continuous distributions, as frequently assumed in the statistical modeling of the wireless fading channel in the literature. However, in the case of discrete fading distributions, this event is in general a non-zero probability event. In such a case, the secrecy capacity is zero, and hence no confidential message transmission can be performed. All the power can be allocated to the transmission of the common message by setting $\delta_1 = 0$.

³The secrecy rate expressions in (5.5) and (5.6) are derived from the generic expression in (5.3). For instance, in (5.5), z_1 and z_2 correspond to z_m and z_e , respectively, and the signal-to-noise ratio is $\delta_1\text{SNR}$. Additionally, the indicator function essentially represents the operation $[\cdot]^+$, ensuring that the secrecy rate is zero if $z_1 < z_2$.

5.2 Throughput of Secure Transmissions with Random Data Arrivals Under QoS Constraints

In this section, we investigate the throughput of the transmission of confidential and common messages, considering different random source types introduced in Section 2.2. In order to highlight the impact of random arrivals, we also address the case of a source with a constant arrival rate. For each source type, we characterize the maximum *average* arrival rate as the maximum throughput. Thus, we determine the throughput by deriving the maximum average arrival rate in terms of SNR for both constant-rate arrivals and the four Markovian arrival models.

We note that our initial analysis considers perfect channel side information (CSI) at the transmitter. Hence, we assume that the transmitter knows the realizations of z_1 and z_2 . This is an accurate assumption, for instance, in a cellular scenario in which the base station knows the channel conditions and the users are not malicious but still the confidential messages are to be kept private from the unintended user. We address the case of no CSI subsequently in Section 5.4.

Constant-Rate Source

Throughput in the case of constant-rate arrival is given by the effective capacity. For each message, the effective capacity is given by

$$C_{Ei}(\text{SNR}, \theta_i) = -\frac{1}{\theta_i} \log_e \mathbb{E}\{e^{-\theta_i R_i(\text{SNR})}\} \text{ for } i = 0, 1, 2. \quad (5.7)$$

Note that for $i = 1$ and 2 , we have the maximum constant arrival rates of the confidential messages at the transmitter, which are intended for receivers 1 and 2, respectively. For $i = 0$, we have the maximum constant arrival rate of the common message at the transmitter. Note further that the QoS constraint θ_i of different messages can in general be different.

We also define the function $g_i(\text{SNR})$ as

$$g_i(\text{SNR}) = \mathbb{E} \left\{ e^{-\theta_i R_i(\text{SNR})} \right\} = e^{-\theta_i C_{Ei}(\text{SNR}, \theta_i)}. \quad (5.8)$$

Note that with this definition, we have

$$C_{Ei}(\text{SNR}, \theta_i) = -\frac{1}{\theta_i} \log_e g_i(\text{SNR}). \quad (5.9)$$

As it will be seen in subsequent subsections, maximum average arrival rates for random sources can also be concisely expressed using the function $g_i(\text{SNR})$.

Discrete Markov Source

In this case, we assume that (confidential and/or common) message arrivals to the buffers at the transmitter are according to a discrete-time Markov chain. In the case of ON-OFF discrete Markov source, introducing effective bandwidth expression in (2.16) into (3.2), and solving for r , we can obtain the maximum arrival rate $r^*(\text{SNR}, \theta)$ and then express the maximum average arrival rate as a function of the effective capacity C_E as

$$\begin{aligned} r_{\text{avg}i}^*(\text{SNR}, \theta_i) &= \frac{P_{\text{ON}}}{\theta_i} \log_e \left(\frac{e^{2\theta_i C_{Ei}(\text{SNR}, \theta_i)} - p_{11} e^{\theta_i C_{Ei}(\text{SNR}, \theta_i)}}{1 - p_{11} - p_{22} + p_{22} e^{\theta_i C_{Ei}(\text{SNR}, \theta_i)}} \right) \\ &= \frac{P_{\text{ON}}}{\theta_i} \log_e \left(\frac{1 - p_{11} g_i(\text{SNR})}{(1 - p_{11} - p_{22}) g_i^2(\text{SNR}) + p_{22} g_i(\text{SNR})} \right) \end{aligned} \quad (5.10)$$

for $i = 0, 1, 2$, where $g_i(\text{SNR})$ is defined in (A.59).

Note that the probability of the ON state is given by $P_{\text{ON}} = \frac{1-p_{11}}{2-p_{11}-p_{22}}$. If we use the assumption $p_{11} = 1 - s$ and $p_{22} = s$ (and hence $P_{\text{ON}} = s$), the expression for average arrival rate can be simplified further as

$$r_{\text{avg}i}^*(\text{SNR}, \theta_i) = \frac{s}{\theta_i} \log_e \left(\frac{e^{\theta_i C_{Ei}(\text{SNR}, \theta_i)} - (1 - s)}{s} \right). \quad (5.11)$$

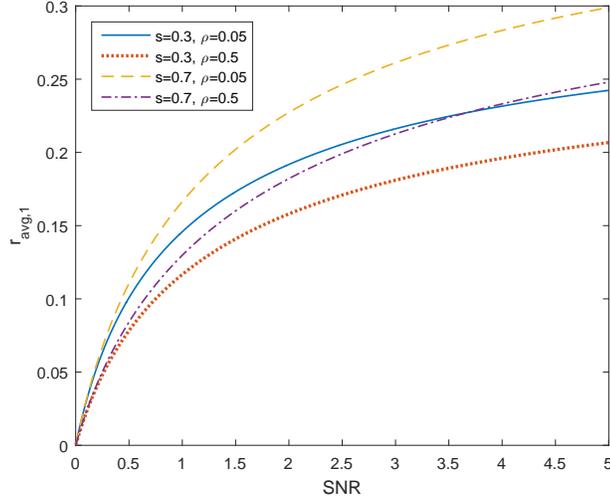


Fig. 5.2: Maximum average arrival rate of the confidential message of the first user $r_{\text{avg},1}^*$ vs. average signal-to-noise ratio SNR when $\theta_1 = 1$ and $\delta_1 = 0.5$.

It can be easily verified that $r_{\text{avg},i}^*$ is a monotonic function of s , i.e., as s (and hence ON-state probability $P_{\text{ON}} = s$) increases, the maximum average arrival rate increases. We see this effect in Fig. 5.2 where we plot the relationship between maximum average arrival rate of the confidential message of the first user vs. average SNR curves for different values of s and correlation coefficient ρ . We consider a Rayleigh fading environment and assume that the fading powers z_1 and z_2 are exponentially distributed with unit means, i.e., $\mathbb{E}\{z_1\} = \mathbb{E}\{z_2\} = 1$, and correlation coefficient $\rho = \frac{\text{cov}(z_1, z_2)}{\sqrt{\text{var}(z_1)\text{var}(z_2)}}$. Numerical evaluation verifies that as s increases, maximum average arrival rate increases for given SNR and ρ . Hence, as the source becomes less bursty, throughput improves. Also, the correlation between the channels of the legitimate user and eavesdropper has an impact on the throughput. Higher correlation values lead to diminished secrecy capacity, which results in smaller throughput values.

We have also performed buffer simulations to further verify our theoretical analysis. Initially, we set the values of the QoS exponent θ_i , SNR, source state transition probabilities p_{11} and p_{22} of the ON/OFF discrete Markov source, and determined the maximum average arrival rate the system can support using the theoretical characterizations in this section.

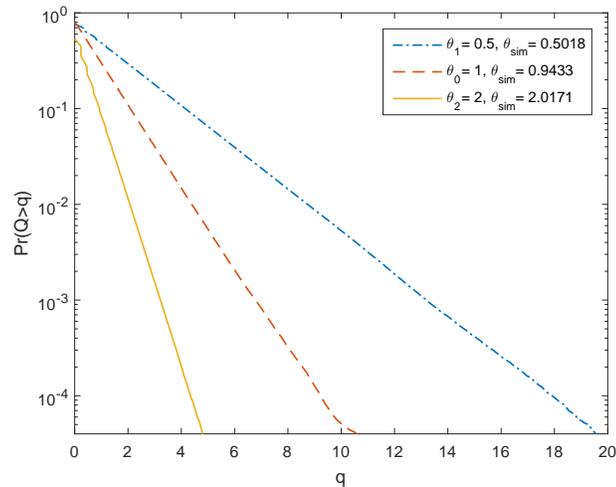


Fig. 5.3: Buffer overflow probability $\Pr\{Q > q\}$ vs. buffer threshold q for both confidential and common messages when $\theta_1 = 0.5, \theta_2 = 2, \theta_0 = 1, \delta_1 = \delta_2 = 0.7, p_{11} = p_{22} = 0.8$ and $\text{SNR} = 1$.

We also calculated the corresponding maximum data arrival rate r_i in the ON state. Then, we initiated the simulation by generating the random data arrivals according to the Markov source model, and generating the Gaussian fading coefficients for the service rates. In this process, we have kept track of the buffer length over 10^7 runs. We have compared the simulated buffer lengths with different thresholds to determine how frequently a threshold is exceeded and identify the overflow probabilities. In Fig. 5.3, we plot the buffer overflow probability (in logarithmic scale) vs. buffer threshold q . We obtain excellent results from these simulations. Specifically, we determined the simulated QoS exponent values θ_{sim} from the slopes of the buffer overflow probability curves in the figure⁴. The simulated θ_{sim} values were obtained as 2.0171, 0.9433, 0.5018 when the corresponding theoretical θ values were 2, 1, 0.5, respectively. Hence, if we originally set $\theta = 2$ and design the system accordingly, the buffer overflow probability decays with QoS exponent $\theta_{\text{sim}} = 2.0171$, matching the prediction very well.

⁴Note from (2.3) that the overflow probability is expected to behave in logarithmic scale as $\log \Pr\{Q \geq q\} \approx -\theta q + \log \zeta$. Hence, the slope of the logarithmic overflow probability vs. buffer threshold q curve is proportional to $-\theta$.

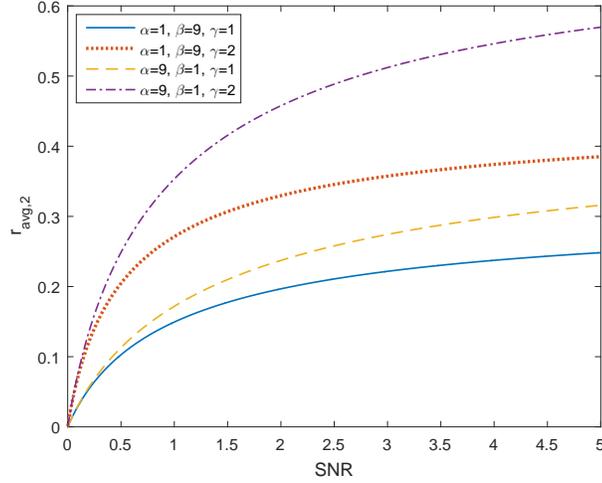


Fig. 5.4: Maximum average arrival rate of the confidential message of the second user $r_{\text{avg},2}^*$ vs. average signal-to-noise ratio SNR when $\theta_2 = 1$, $\rho = 0.05$ and $\delta_2 = 0.5$.

Markov Fluid Source

Similarly as in the case of discrete Markov source, for the ON-OFF Markov fluid source, incorporating (2.22) into (3.2), we determine the maximum average arrival rate as

$$\begin{aligned}
 r_{\text{avg},i}^*(\text{SNR}, \theta_i) &= P_{\text{ON}} \frac{\theta_i C_{E_i}(\text{SNR}, \theta_i) + \alpha + \beta}{\theta_i C_{E_i}(\text{SNR}, \theta_i) + \alpha} C_{E_i}(\text{SNR}, \theta_i) \\
 &= -\frac{P_{\text{ON}}}{\theta_i} \frac{\alpha + \beta - \log_e \mathbf{g}_i(\text{SNR})}{\alpha - \log_e \mathbf{g}_i(\text{SNR})} \log_e \mathbf{g}_i(\text{SNR}) \quad (5.12)
 \end{aligned}$$

for $i = 0, 1, 2$. Note that the probability of ON state is given as $P_{\text{ON}} = \frac{\alpha}{\alpha + \beta}$.

In Fig. 5.4, we plot the maximum average arrival rate of the confidential message of the second user as a function of average SNR while considering different channels and Markov fluid sources. Specifically, we assume different pairs of the source state transition rates α and β and different expected channel gains $\mathbb{E}\{z_2\} = \gamma$. As in Fig. 5.2, we still assume that z_1 and z_2 are exponentially distributed, and $\mathbb{E}\{z_1\} = 1$. It is observed that increasing α and decreasing β simultaneously increase the ON-state probability P_{ON} and reduce the burstiness of the source, and as a result, throughput increases. Furthermore, better channel conditions for the legitimate user lead to improved throughput due to increase in secrecy

capacity.

Discrete-Time Markov Modulated Poisson Source

In order to express the maximum average arrival rate in terms of C_E , we again insert the effective bandwidth expression in (2.30) into (3.2) and obtain

$$r_{\text{avg}i}^*(\text{SNR}, \theta_i) = \frac{P_{\text{ON}}}{(e^{\theta_i} - 1)} \log_e \left(\frac{1 - p_{11} \mathbf{g}_i(\text{SNR})}{(1 - p_{11} - p_{22}) \mathbf{g}_i^2(\text{SNR}) + p_{22} \mathbf{g}_i(\text{SNR})} \right). \quad (5.13)$$

Continuous-Time Markov Modulated Poisson Source

We find the following maximum average arrival rate r_{avg}^* by incorporating (2.26) into (3.2):

$$r_{\text{avg}i}^*(\text{SNR}, \theta_i) = -\frac{P_{\text{ON}}}{(e^{\theta_i} - 1)} \frac{\alpha + \beta - \log_e \mathbf{g}_i(\text{SNR})}{\alpha - \log_e \mathbf{g}_i(\text{SNR})} \log_e \mathbf{g}_i(\text{SNR}). \quad (5.14)$$

5.3 Energy Efficiency Of Secure Transmissions with Random Data Arrivals Under QoS Constraints

In this section, we investigate the energy efficiency of the transmission of confidential and common messages for various source types discussed previously. Using the throughput formulas we have obtained, we analyze the energy efficiency and derive closed-form expressions of the minimum energy per bit and wideband slope.

5.3.1 Minimum Energy per Bit

The minimum energy per bit in (3.29) characterizes the minimum energy needed to send one bit reliably over the wireless fading channel under statistical queueing constraints. Lower minimum energy per bit levels indicate higher energy efficiency. First, we formulate

the minimum energy per bit for the confidential messages as

$$\frac{E_b}{N_{0 \min, i}} = \lim_{\text{SNR} \rightarrow 0} \frac{\delta_i \Pr(\Gamma_i) \text{SNR}}{r_{\text{avg}i}^*(\text{SNR}, \theta_i)} = \frac{\delta_i \Pr(\Gamma_i)}{r_{\text{avg}i}^*(0)} \quad (5.15)$$

for $i = 1, 2$. Similarly for the common message, the minimum energy per bit becomes

$$\begin{aligned} \frac{E_b}{N_{0 \min, 0}} &= \lim_{\text{SNR} \rightarrow 0} \frac{[(1 - \delta_1) \Pr(\Gamma_1) + (1 - \delta_2) \Pr(\Gamma_2)] \text{SNR}}{r_{\text{avg}0}^*(\text{SNR}, \theta_0)} \\ &= \frac{(1 - \delta_1) \Pr(\Gamma_1) + (1 - \delta_2) \Pr(\Gamma_2)}{r_{\text{avg}0}^*(0)}. \end{aligned} \quad (5.16)$$

Below, we initially characterize the minimum energy per bit for the case of constant-rate arrivals, and subsequently show that the same minimum energy per bit levels are achieved when discrete-time Markov and Markov ON-OFF sources are considered.

Proposition 5.3.1. *When the data arrival rate is constant, the minimum energy per bit expressions for the confidential message transmissions to receivers 1 and 2 under QoS constraints are given, respectively, by*

$$\frac{E_b}{N_{0 \min, 1}} = \frac{\Pr(\Gamma_1) \log_e 2}{\mathbb{E}_{\Gamma_1}\{z_1 - z_2\}} \quad (5.17)$$

$$\frac{E_b}{N_{0 \min, 2}} = \frac{\Pr(\Gamma_2) \log_e 2}{\mathbb{E}_{\Gamma_2}\{z_2 - z_1\}}, \quad (5.18)$$

and the minimum energy per bit for the common message transmission under QoS constraints is given by

$$\frac{E_b}{N_{0 \min, 0}} = \frac{[(1 - \delta_1) \Pr(\Gamma_1) + (1 - \delta_2) \Pr(\Gamma_2)] \log_e 2}{(1 - \delta_1) \mathbb{E}_{\Gamma_1}\{z_2\} + (1 - \delta_2) \mathbb{E}_{\Gamma_2}\{z_1\}} \quad (5.19)$$

where $\Pr(\Gamma_1) = \Pr(z_1 < z_2)$, $\Pr(\Gamma_2) = \Pr(z_1 > z_2)$, and δ_i is the fraction of the power used for the transmission of the confidential message to receiver i . Moreover, \mathbb{E}_{Γ_1} denotes

the expectation in region Γ_1 while \mathbb{E}_{Γ_2} is similarly defined in the complement region Γ_2 .

Proof: See Appendix A.12.

When z_1 and z_2 are independent and exponentially distributed with $\mathbb{E}\{z_1\} = 1$ and $\mathbb{E}\{z_2\} = \gamma$, we have $\Pr(\Gamma_1) = \frac{1}{\gamma+1}$ and $\Pr(\Gamma_2) = \frac{\gamma}{\gamma+1}$, and we can get closed-form expressions for the minimum energy per bit formulations as follows:

$$\frac{E_b}{N_{0 \min,1}} = \log_e 2, \quad \frac{E_b}{N_{0 \min,2}} = \frac{\log_e 2}{\gamma} \quad (5.20)$$

$$\frac{E_b}{N_{0 \min,0}} = \frac{\gamma + 1}{\gamma} \log_e 2. \quad (5.21)$$

Interestingly, for both ON-OFF discrete-time Markov and Markov fluid sources, minimum energy per bit expressions are the same as those attained in the presence of constant-rate sources.

Proposition 5.3.2. *When data arrivals are modeled as ON-OFF discrete-time Markov or Markov fluid processes, the minimum energy per bit expressions for confidential and common message transmissions under QoS constraints remains the same as those for the constant arrival rate model and hence are given by (5.17), (5.18), and (5.19), respectively.*

Proof: See Appendix A.13.

Heretofore, we have seen that the minimum bit energy expressions do not depend on either the queueing constraints or the source randomness. More specifically, minimum bit energy of confidential/common message transmissions are the same regardless of the value of the QoS exponent θ and whether data arrives at a constant rate or according to an ON-OFF Markov process. However, this is not the case when we consider more bursty Markov-modulated Poisson arrivals, as shown in the result below.

Proposition 5.3.3. *When the source arrivals are modeled as ON-OFF discrete-time or continuous-time MMPPs, the minimum energy per bit expressions for confidential and com-*

mon message transmissions under QoS constraints are given, respectively, by

$$\frac{E_b}{N_{0\min,1}} = \frac{(e^{\theta_1} - 1) \Pr(\Gamma_1) \log_e 2}{\theta_1 \mathbb{E}_{\Gamma_1}\{z_1 - z_2\}} \quad (5.22)$$

$$\frac{E_b}{N_{0\min,2}} = \frac{(e^{\theta_2} - 1) \Pr(\Gamma_2) \log_e 2}{\theta_2 \mathbb{E}_{\Gamma_2}\{z_2 - z_1\}} \quad (5.23)$$

$$\frac{E_b}{N_{0\min,0}} = \frac{(e^{\theta_0} - 1)[(1 - \delta_1) \Pr(\Gamma_1) + (1 - \delta_2) \Pr(\Gamma_2)] \log_e 2}{\theta_0 [(1 - \delta_1) \mathbb{E}_{\Gamma_1}\{z_2\} + (1 - \delta_2) \mathbb{E}_{\Gamma_2}\{z_1\}]} \quad (5.24)$$

Proof: See Appendix A.14.

For MMPP sources, minimum energy per bit now depends on the QoS exponent through the term $\frac{e^\theta - 1}{\theta}$. Since $\frac{e^\theta - 1}{\theta} > 1$ for $\theta > 0$ and increases with increasing θ , a higher energy per bit is required for MMPP sources (compared to constant-rate and ON-OFF Markov sources) and energy cost grows as the QoS constraints become more stringent. Interestingly, energy per bit expressions still do not depend on the specific parameters of the random arrival model (such as transition probabilities/rates of the Markov chain and intensity of the Poisson arrivals).

As also noted before, Proposition 5.3.2 shows that the minimum energy per bit for discrete-time Markov and Markov fluid sources are the same as for the constant-rate source. The primary intuitive reasoning behind this result is that the minimum energy per bit is an asymptotic performance metric achieved as $\text{SNR} \rightarrow 0$, and the impact of source burstiness significantly diminishes at these asymptotically low SNR levels for discrete-time Markov and Markov fluid sources. Specifically, as SNR diminishes, the fixed arrival rate (in the ON-state of the Markov models) that can be supported by the wireless channel decreases as well, resulting in less and less impact on buffer overflows and delay violations.

On the other hand, if the arrival process is MMPP, the intensity of the Poisson process is reduced with decreasing SNR. However, the arrival process is still a Poisson process but

with a smaller intensity, meaning that there is still a probability, however small, for the instantaneous arrival rate in the ON state to be large since the arrival rate depends on the realization of a Poisson distributed random variable. Hence, MMPP source is more bursty in the low-SNR regime than discrete-time Markov and Markov fluid sources, and this is reflected in the larger minimum energy per bit values as shown in the results of Proposition 5.3.3.

5.3.2 Wideband Slope

Minimum energy per bit $\frac{E_b}{N_0 \min}$ is the ultimate performance limit of energy-efficient operation. At the same time, it is an asymptotic performance metric achieved in the limit as SNR vanishes. In this subsection, we complement the $\frac{E_b}{N_0 \min}$ analysis by characterizing the wideband slope of confidential and common message transmissions for different source models. Unlike the minimum energy per bit, wideband slope is distinct for each source and depends on the source statistics. In this subsection, we also provide numerical results to demonstrate the effectiveness of the linear approximation of the throughput in the low-SNR regime in terms of $\frac{E_b}{N_0 \min}$ and wideband slope \mathcal{S}_0 , and to identify the impact of secrecy requirements, source randomness, QoS constraints, and channel correlation on energy efficiency.

Constant-Rate Sources

Proposition 5.3.4. *For constant-rate arrivals, the wideband slope expressions for common and confidential message transmissions under QoS constraint are given by*

$$\mathcal{S}_{0,i} = \frac{2 \left(\mathbb{E} \left\{ \dot{f}_i(0) \right\} \right)^2}{\frac{\theta_i}{\log_e 2} \text{var} \left(\dot{f}_i(0) \right) - \mathbb{E} \left\{ \ddot{f}_i(0) \right\}} \quad (5.25)$$

for $i = 0, 1, 2$ where we have defined $f_i(\text{SNR}) = R_i(\text{SNR}) \log_e 2$ with $R_i(\text{SNR})$ being the instantaneous rate of confidential or common message given in (5.4)–(5.6), and the first

and second derivatives of $f_i(\text{SNR})$ at $\text{SNR} = 0$ are given by

$$\begin{aligned}
\dot{f}_1(0) &= \delta_1 (z_1 - z_2) \mathbf{1}\{z_1 \geq z_2\}, \\
\dot{f}_2(0) &= \delta_2 (z_2 - z_1) \mathbf{1}\{z_1 < z_2\}, \\
\dot{f}_0(0) &= (1 - \delta_1)z_2 \mathbf{1}\{z_1 \geq z_2\} + (1 - \delta_2)z_1 \mathbf{1}\{z_1 < z_2\}, \\
\ddot{f}_1(0) &= -\delta_1^2 [z_1^2 - z_2^2] \mathbf{1}\{z_1 \geq z_2\}, \\
\ddot{f}_2(0) &= -\delta_2^2 [z_2^2 - z_1^2] \mathbf{1}\{z_1 < z_2\}, \\
\ddot{f}_0(0) &= -(1 - \delta_1^2)z_2^2 \mathbf{1}\{z_1 \geq z_2\} - (1 - \delta_2^2)z_1^2 \mathbf{1}\{z_1 < z_2\}.
\end{aligned} \tag{5.26}$$

Proof: See Appendix A.15.

Above, $\mathcal{S}_{0,0}$ is the wideband slope for common message transmission while $\mathcal{S}_{0,1}$ and $\mathcal{S}_{0,2}$ denote the wideband slope of confidential message transmissions to receivers 1 and 2, respectively.

For independent and exponentially distributed z_1 and z_2 with $\mathbb{E}\{z_1\} = 1$ and $\mathbb{E}\{z_2\} = \gamma$, the wideband slope expressions simplify to

$$\mathcal{S}_{0,1} = \frac{2}{\frac{\theta_1}{\log_e 2} (1 + 2\gamma) + 4\gamma + 2} \tag{5.27}$$

$$\mathcal{S}_{0,2} = \frac{2}{\frac{\theta_2}{\log_e 2} \left(1 + \frac{2}{\gamma}\right) + \frac{4}{\gamma} + 2}. \tag{5.28}$$

If we further assume that $\delta_1 = \delta_2 = \delta$, then the wideband slope for common message becomes

$$\mathcal{S}_{0,0} = \frac{2}{\frac{\theta_0}{\log_e 2} + \frac{1-\delta^2}{(1-\delta)^2}}. \tag{5.29}$$

Discrete-Time Markov Sources

Next, we consider ON-OFF discrete-time Markov sources with transition probabilities denoted by p_{ij} for $i, j \in \{1, 2\}$.

Proposition 5.3.5. *The wideband slope expressions for confidential and common message transmissions under QoS constraint are given by*

$$\mathcal{S}_{0,i} = \frac{2 \left(\mathbb{E} \left\{ \dot{f}_i(0) \right\} \right)^2}{\eta \frac{\theta_i}{\log_e 2} \left(\mathbb{E} \left\{ \dot{f}_i(0) \right\} \right)^2 + \frac{\theta_i}{\log_e 2} \text{var} \left(\dot{f}_i(0) \right) - \mathbb{E} \left\{ \ddot{f}_i(0) \right\}} \quad (5.30)$$

for $i = 0, 1, 2$, where $\dot{f}_i(0)$ and $\ddot{f}_i(0)$ are given in (5.26). Additionally, η previously is defined in (3.9) as $\eta = \frac{(1-p_{22})(p_{11}+p_{22})}{(1-p_{11})(2-p_{11}-p_{22})}$.

Proof: See Appendix A.16

Again, for independent and exponentially distributed z_1 and z_2 with $\mathbb{E}\{z_1\} = 1$ and $\mathbb{E}\{z_2\} = \gamma$, the wideband slope expressions are given as

$$\mathcal{S}_{0,1} = \frac{2}{\frac{\theta_1 \eta}{\log_e 2} + \frac{\theta_1}{\log_e 2} (1 + 2\gamma) + 4\gamma + 2}, \quad (5.31)$$

$$\mathcal{S}_{0,2} = \frac{2}{\frac{\theta_2 \eta}{\log_e 2} + \frac{\theta_2}{\log_e 2} \left(1 + \frac{2}{\gamma} \right) + \frac{4}{\gamma} + 2}. \quad (5.32)$$

If we further assume that $\delta_1 = \delta_2 = \delta$, then the wideband slope for common message becomes

$$\mathcal{S}_{0,0} = \frac{2}{\frac{\theta_0 \eta}{\log_e 2} + \frac{\theta_0}{\log_e 2} + \frac{1-\delta^2}{(1-\delta)^2}}. \quad (5.33)$$

When compared with the corresponding wideband slope expressions in (5.27)–(5.29) for the constant-rate source, we notice that wideband slope formulas above in (5.31)–(5.33) for the discrete Markov source differ only due to the presence of the term $\frac{\theta \eta}{\log_e 2}$, which reflects essentially the source randomness with the parameter η . This additional term leads

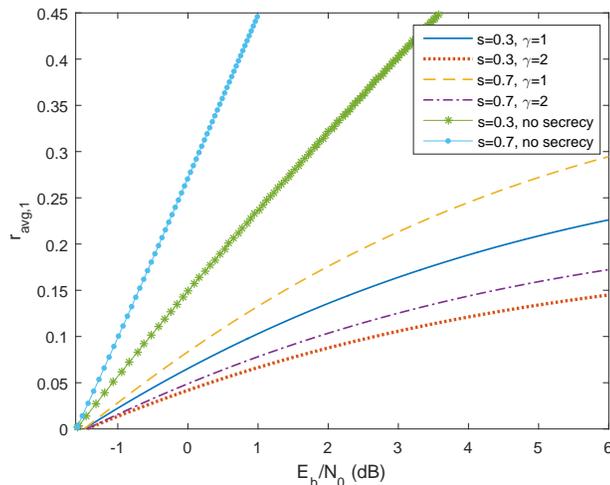


Fig. 5.5: Maximum average arrival rate of first user's confidential message $r_{\text{avg}1}^*$ vs. energy per bit $\frac{E_b}{N_0}$ in dB when $\theta_1 = 1$, $\rho = 0.05$ and $\delta_1 = 0.5$.

to smaller wideband slopes, indicating the detrimental impact of source randomness on energy efficiency. Note also that when $p_{11} = 0$ and $p_{22} = 1$, discrete Markov essentially becomes a constant-rate source and we have $\eta = 0$.

In Fig. 5.5, the maximum average arrival rate of the confidential message for the first user vs. energy per bit is plotted. We consider an ON-OFF discrete Markov source with $p_{11} = 1 - s$ and $p_{22} = s$ (and hence $P_{\text{ON}} = s$). We assume $\theta = 1$ and $\delta_1 = 0.5$. The channel power gains z_1 and z_2 are exponentially distributed with $\mathbb{E}\{z_1\} = 1$, $\mathbb{E}\{z_2\} = \gamma$ and correlation coefficient $\rho = 0.05$. As predicted, the minimum energy per bit does not depend on source burstiness or the second user channel statistics, i.e., γ . There is a slight increase in the minimum energy per bit values achieved in the cases of secrecy as compared to no secrecy. The main reason for this is the correlation in the channel conditions of the two users. Without any correlation, the minimum energy per bit becomes equal to -1.59 dB. As a result of similar minimum energy per bit values, wideband slope becomes a critical performance indicator in the low-SNR regime. We notice that wideband slope diminishes when secrecy requirements are imposed and also when source burstiness increases with diminishing ON-state probability $P_{\text{ON}} = s$. We also observe that, as the second user (or equivalently eavesdropper) channel conditions improve, i.e., as γ increases,

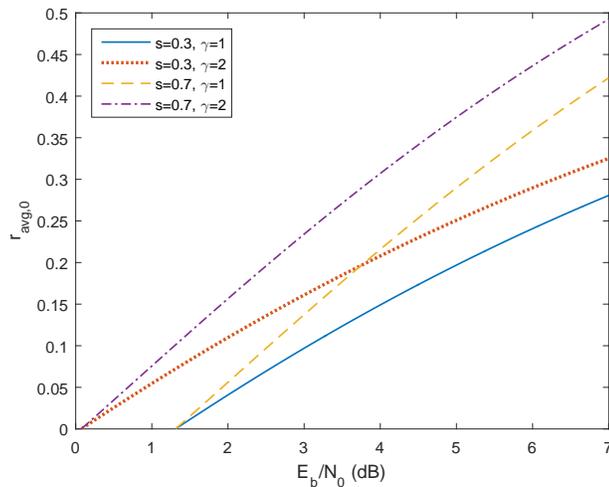


Fig. 5.6: Maximum average arrival rate of common message $r_{\text{avg}0}^*$ vs. energy per bit $\frac{E_b}{N_0}$ in dB when $\theta_0 = 1$, $\rho = 0.05$ and $\delta_1 = \delta_2 = 0.5$.

we have smaller wideband slopes.

We illustrate the maximum average arrival rate of the common message vs. energy per bit in Fig. 5.6, assuming the parameter setting $\theta = 1$, $\rho = 0.05$ and $\delta_1 = \delta_2 = 0.5$. We again verify that source characteristics do not play a role in the value of the minimum energy per bit. Better channel conditions for the second user improve the overall energy efficiency of the transmission of the common message by improving the minimum energy per bit. We also notice that wideband slope is the same when we alter the channel conditions. However, source burstiness has a negative impact on the wideband slope, thus, on the energy efficiency as well.

Markov Fluid Sources

In the following, we characterize the wideband slope in the case of ON-OFF Markov fluid arrivals with transition rates α and β .

Proposition 5.3.6. *The wideband slope expressions for confidential and common message*

transmissions under QoS constraint are given by

$$\mathcal{S}_{0,i} = \frac{2 \left(\mathbb{E} \left\{ \dot{f}_i(0) \right\} \right)^2}{\zeta \frac{\theta_i}{\log_e 2} \left(\mathbb{E} \left\{ \dot{f}_i(0) \right\} \right)^2 + \frac{\theta_i}{\log_e 2} \text{var} \left(\dot{f}_i(0) \right) + \mathbb{E} \left\{ \ddot{f}_i(0) \right\}} \quad (5.34)$$

for $i = 0, 1, 2$, where $\dot{f}_i(0)$ and $\ddot{f}_i(0)$ are defined in (5.26). Note further that ζ is defined in (3.17) as $\zeta = \frac{2\beta}{\alpha(\alpha+\beta)}$.

Proof: See Appendix A.17.

Similarly as for the previous arrival models, we can simplify the wideband expressions for the confidential message transmissions to the following when we have independent and exponentially distributed z_1 and z_2 with $\mathbb{E}\{z_1\} = 1$ and $\mathbb{E}\{z_2\} = \gamma$:

$$\mathcal{S}_{0,1} = \frac{2}{\frac{\theta_1 \zeta}{\log_e 2} + \frac{\theta_1}{\log_e 2} (1 + 2\gamma) + 4\gamma + 2}, \quad (5.35)$$

$$\mathcal{S}_{0,2} = \frac{2}{\frac{\theta_2 \zeta}{\log_e 2} + \frac{\theta_2}{\log_e 2} \left(1 + \frac{2}{\gamma} \right) + \frac{4}{\gamma} + 2}. \quad (5.36)$$

If we further assume that $\delta_1 = \delta_2 = \delta$, then the wideband slope for common message becomes

$$\mathcal{S}_{0,0} = \frac{2}{\frac{\theta_0 \zeta}{\log_e 2} + \frac{\theta_0}{\log_e 2} + \frac{1-\delta^2}{(1-\delta)^2}}. \quad (5.37)$$

The common theme in the above expressions and the ones corresponding to other source types (i.e., expressions in (5.27)–(5.29) and (5.31)–(5.33)) is that wideband slope expressions depend on three critical factors: QoS exponent θ , source burstiness parameter (ζ in the case of Markov fluid source and η in the case discrete Markov source, which both become zero when the arrival rate is constant), and channel statistics through $\mathbb{E}\{z_2\} = \gamma$. For instance, wideband slopes diminish as θ increases and more stringent buffer/delay constraints are imposed.

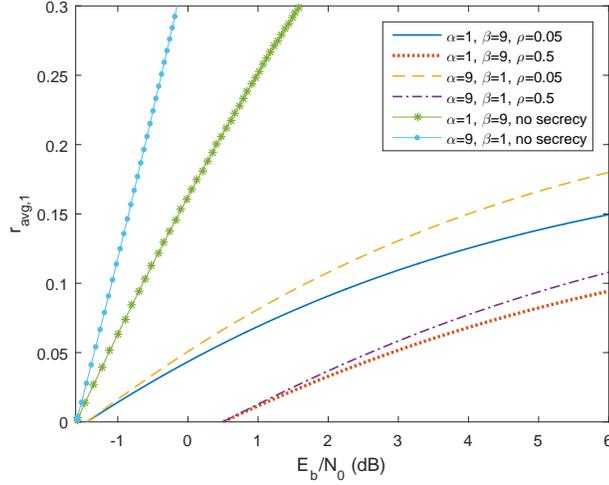


Fig. 5.7: Maximum average arrival rate of first user's confidential message $r_{\text{avg},1}^*$ vs. energy per bit $\frac{E_b}{N_0}$ when $\theta_1 = 1$, $\gamma = 1$ and $\delta_1 = 0.5$.

We depict, in Fig. 5.7, the maximum average arrival rate of the confidential message for the first user vs. energy per bit for Markov fluid sources with different values of α and β . We assume $\theta = 1$, $\gamma = 1$ and $\delta_1 = 0.5$. In the case of no secrecy, the minimum energy per bit is equal to -1.59 dB and it remains unchanged under different source characteristics. With secrecy, source burstiness again does not impact the minimum energy per bit. However, as channel correlation increases, the energy efficiency degrades due to higher minimum energy per bit. Additionally, the source characteristics have significant impact on the wideband slope e.g., wideband slope decreases as source becomes more bursty (i.e., as we change the state transition rates from $\alpha = 9$ and $\beta = 1$ to $\alpha = 1$ and $\beta = 9$).

MMPP Sources

Next, we address ON-OFF MMPP sources.

Proposition 5.3.7. *When the source is modeled as discrete-time MMPP the wideband slope expressions for confidential and common message transmissions under QoS constraint are*

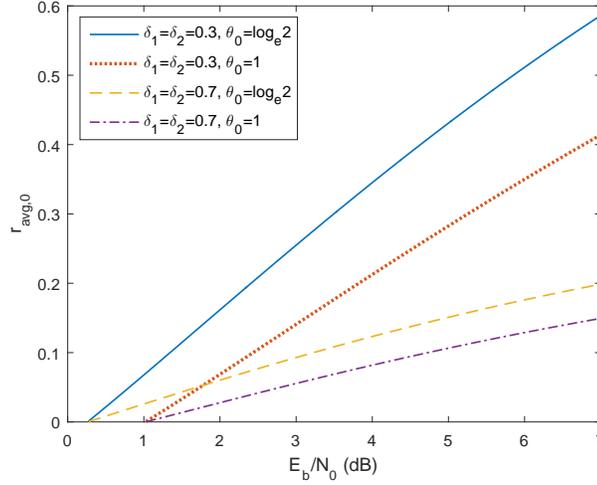


Fig. 5.8: Maximum average arrival rate of common message $r_{\text{avg}0}^*$ vs. energy per bit $\frac{E_b}{N_0}$ in dB when $\rho = 0.8$, $\gamma = 1$, $p_{11} = 0.1$ and $p_{22} = 0.9$.

given by

$$\mathcal{S}_{0,i} = \frac{\frac{2\theta_i}{e^{\theta_i}-1} \left(\mathbb{E} \left\{ \dot{f}_i(0) \right\} \right)^2}{\eta \frac{\theta_i}{\log_e 2} \left(\mathbb{E} \left\{ \dot{f}_i(0) \right\} \right)^2 + \frac{\theta_i}{\log_e 2} \text{var} \left(\dot{f}_i(0) \right) + \mathbb{E} \left\{ \ddot{f}_i(0) \right\}}. \quad (5.38)$$

for $i = 0, 1, 2$ where $\dot{f}_i(0)$ and $\ddot{f}_i(0)$ are defined in (5.26) and η is defined in (3.9).

When the source is modeled as continuous-time MMPP the wideband slope expressions for confidential and common message transmissions under QoS constraint are given by

$$\mathcal{S}_{0,i} = \frac{\frac{2\theta_i}{e^{\theta_i}-1} \left(\mathbb{E} \left\{ \dot{f}_i(0) \right\} \right)^2}{\zeta \frac{\theta_i}{\log_e 2} \left(\mathbb{E} \left\{ \dot{f}_i(0) \right\} \right)^2 + \frac{\theta_i}{\log_e 2} \text{var} \left(\dot{f}_i(0) \right) + \mathbb{E} \left\{ \ddot{f}_i(0) \right\}}. \quad (5.39)$$

for $i = 0, 1, 2$ where ζ is defined in (3.17).

We omit the proof as it is rather straightforward due to the relationship between the throughputs of the Markov and MMPP sources.

In Fig. 5.8, we illustrate the maximum average arrival rate of the common message vs. energy per bit when the source is ON-OFF discrete-time MMPP. We set $\rho = 0.8$, $\gamma = 1$, $p_{11} = 0.1$ and $p_{22} = 0.9$, and study the impact of different values of θ_0 and δ_i .

For the MMPP source, the minimum energy per bit depends on the QoS exponent θ_0 and it decreases as θ_0 diminishes, indicating less stringent queueing constraints. Power allocation has no impact on the minimum energy per bit. However, with more power allocated to the common message, the wideband slope becomes higher.

5.4 Throughput and Energy Efficiency with no Channel Knowledge at the Transmitter

In this section, we depart from the perfect transmitter CSI assumption of the previous sections and consider a scenario in which the transmitter has no CSI. Specifically, we assume that the transmitter does not know the realizations of the channel fading coefficients, which is relevant in cases in which the eavesdropper is passive and malicious. This also represents a worst-case scenario due to the fact that even the legitimate channel is not known. Treating the eavesdropper as malicious, we address a special case of the previously treated system model. In particular, we do not consider common message transmission and assume that the transmitter just intends to send confidential messages to receiver 1 while keeping them private from receiver 2 (which is regarded as the eavesdropper).

Not knowing the realizations of the channel fading coefficients h_1 and h_2 , the transmitter sends the data at the fixed rate of λ bits/s/Hz. As before, instantaneous secrecy capacity $R(\text{SNR}) = [\log_2(1 + \text{SNR}z_1) - \log_2(1 + \gamma\text{SNR}z_2)]^+$ quantifies the maximum achievable rates of secure communication where $z_i = |h_i|^2$. Hence, if $\lambda \leq R(\text{SNR})$, then reliable and secure communication is attained and therefore the transmitted message is decoded correctly while eavesdropper is being kept ignorant of the message. If, on the other hand, $\lambda > R(\text{SNR})$, secrecy outage occurs. Under these assumptions, the wireless link can be modeled as a two-state discrete-time Markov chain. Specifically, the channel is assumed to be in the ON state if $\lambda \leq R(\text{SNR})$, while the channel is in the OFF state when $\lambda > R(\text{SNR})$.

The steady-state probability for the ON state can be easily obtained as

$$\mathbf{P}\{\Gamma\} = \mathbf{P}\{R(\text{SNR}) > \lambda\} = \mathbf{P}\left\{z_1 > 2^\lambda \gamma z_2 + \frac{2^\lambda - 1}{\text{SNR}}\right\} \quad (5.40)$$

$$= \int_0^\infty \int_{2^\lambda \gamma z_2 + \frac{2^\lambda - 1}{\text{SNR}}}^\infty p(z_1, z_2) dz_1 dz_2 \quad (5.41)$$

where we define $\Gamma = \{(z_1, z_2) \in \mathbb{R}^+ : \lambda < R(\text{SNR})\}$.

5.4.1 Effective Capacity with no Channel Knowledge at the Transmitter

In [9, Chap. 7, Example 7.2.7], it is shown for Markov modulated processes that

$$\frac{\Lambda(\theta)}{\theta} = \frac{1}{\theta} \log_e \mathbb{E}\{\rho(\phi(\theta)\mathbf{M})\}. \quad (5.42)$$

Above, \mathbf{M} is the transition matrix of the underlying Markov process, and $\phi(\theta)$ is a diagonal matrix whose components are the moment generating functions of the processes in the Markov states. We assume that the fading coefficients $\{h_i\}$ change independently from one block to another. Under this assumption, the effective capacity can be obtained as

$$C_E(\text{SNR}, \theta) = \frac{\Lambda(-\theta)}{-\theta} = -\frac{1}{\theta} \log_e [1 - \mathbf{P}\{\Gamma\}(1 - e^{-\theta\lambda})] \quad (5.43)$$

where $\mathbf{P}\{\Gamma\}$ is the channel ON-state probability given in (5.41).

5.4.2 Energy Efficiency with Discrete Markov Sources

First, we consider ON-OFF discrete Markov sources. We also assume that channel fading powers z_1 and z_2 are independent exponentially distributed with means 1 and γ , respectively. In the following result, we characterize the considered energy efficiency metrics under these assumptions.

Proposition 5.4.1. *The minimum energy per bit and wideband slope achieved with fixed-rate secure transmissions in the presence of an eavesdropper with ON-OFF discrete Markov data arrivals and statistical QoS constraints are given by*

$$\frac{E_b}{N_{0\min}} = e(\gamma + 1) \log_e 2, \quad (5.44)$$

$$\mathcal{S}_0 = \frac{1}{\frac{\theta(\eta-1)}{2\log_e 2} + \frac{\theta e(\gamma+1)}{2\log_e 2} + e\gamma + \frac{e(\gamma+1)}{2}}, \quad (5.45)$$

respectively, with η defined in (3.9).

Proof: See Appendix A.18.

As in the perfect CSI case, the minimum energy per bit in (5.44) does not depend on the QoS exponent θ and source statistics while the wideband slope in (5.45) depends on both. Specifically, wideband slope decreases with stricter QoS limitations (i.e., with increasing θ) and increased source burstiness (i.e., with larger η).

It is also interesting to compare the minimum energy per bit expressions achieved with perfect CSI and no CSI. Recall from (5.20) that with perfect CSI, the minimum energy per bit for the confidential message transmission to receiver 1 assuming exponentially distributed fading powers with $\mathbb{E}\{z_1\} = 1$ and $\mathbb{E}\{z_2\} = \gamma$ is

$$\frac{E_b}{N_{0\min}} = \log_e 2. \quad (5.46)$$

Comparing this with (5.44), we immediately identify the additional energy cost per bit of not having channel knowledge at the transmitter as $[e(\gamma + 1) - 1] \log_e 2$. Hence, the characterization in Proposition 5.4.1 nicely quantifies the energy cost of not having transmitter CSI in secure wireless transmissions.

Following the same methodology as described in the discussion of Fig. 5.3, we have again performed simulations in the case of no transmitter CSI. In Fig. 5.9, we plot the buffer overflow probability vs. buffer threshold q . We again have very good agreement

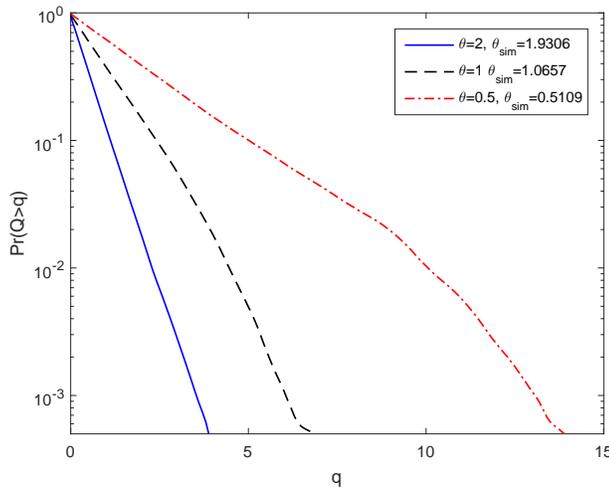


Fig. 5.9: Buffer overflow probability $\Pr\{Q > q\}$ vs. buffer threshold q for different values of θ . $p_{11} = p_{22} = 0.8$, $\text{SNR} = 0.05$

with theoretical predictions. In particular, the simulated θ_{sim} values were obtained as 1.9306, 1.0657, 0.5109 when the corresponding theoretical θ values were 2, 1, 0.5, respectively.

As also noted above, Proposition 5.4.1 shows that while the minimum energy per bit does not depend on the source statistics and QoS exponent θ , the wideband slope depends on both and decreases as burstiness parameter η increases. We see these clearly in Fig. 5.10, where we plot the maximum average arrival rate vs. energy per bit for discrete Markov sources with varying statistics. As predicted, the minimum energy per bit stays the same at 5.76 dB, which is more than 7 dB larger than the minimum energy per bit of -1.59 dB achieved in the case of perfect CSI. We also observe that source with smaller p_{11} and greater p_{22} (while keeping $p_{11} + p_{22} = 1$) has a smaller η value and correspondingly larger wideband slope. Hence, lower source burstiness benefits the energy efficiency.

5.4.3 Energy Efficiency with Markov Fluid Sources

In this section, we consider ON-OFF Markov fluid sources and similarly as in the previous section identify the energy efficiency metrics.

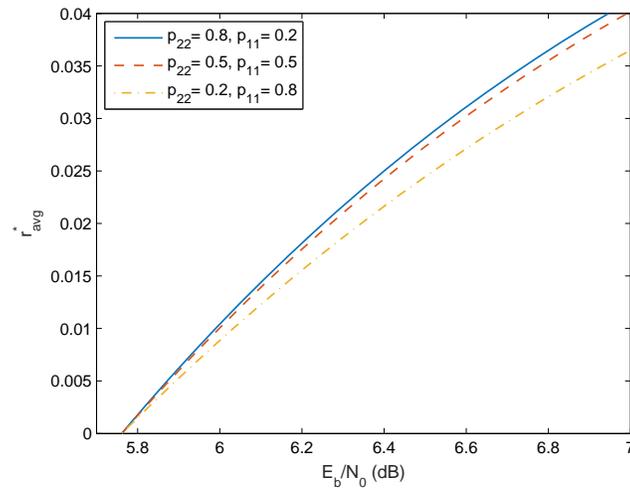


Fig. 5.10: Maximum average arrival rate r_{avg}^* vs. energy per bit $\frac{E_b}{N_0}$ with various source statistics when $\theta = 0.5$.

Proposition 5.4.2. *The minimum energy per bit and wideband slope achieved with fixed-rate secure transmissions in the presence of an eavesdropper with ON-OFF Markov fluid data arrivals and statistical QoS constraints are given by*

$$\frac{E_b}{N_{0min}} = e(\gamma + 1) \log_e 2, \quad (5.47)$$

$$S_0 = \frac{1}{\frac{\theta(\zeta-1)}{2 \log_e 2} + \frac{\theta e(\gamma+1)}{2 \log_e 2} + e\gamma + \frac{e(\gamma+1)}{2}}, \quad (5.48)$$

respectively, where ζ is defined in (3.17)

Proof: See Appendix A.19.

CHAPTER 6

ENERGY EFFICIENCY OF FIXED-RATE TRANSMISSIONS WITH MARKOV ARRIVALS UNDER QUEUEING CONSTRAINTS

6.1 Channel Model and Fixed-Rate Transmissions

We consider a flat-fading channel between the transmitter and receiver. The channel input-output relation can be expressed as

$$y(t) = h(t)x(t) + n(t) \quad (6.1)$$

where $x(t)$ and $y(t)$ are the complex-valued (i.e., low-pass equivalent) input and output signals, respectively, and $n(t)$ denotes the zero-mean, circularly-symmetric, complex Gaussian noise. The signal-to-noise ratio is defined as $\text{SNR} = \frac{P}{N_0 B}$, where P denotes the power of the input signal, $N_0/2$ is the power spectral density of the noise and B is the channel

bandwidth. Above in (6.1), $h(t)$ denotes the multiplicative fading component representing the attenuation and phase shift experienced in the channel. We consider a Rayleigh fading channel and assume that $h(t)$ is a zero-mean complex Gaussian process. Therefore, $z(t) = |h(t)|^2$ has an exponential distribution.

Not knowing the channel conditions, the transmitter sends the data at the fixed rate of R bits/s/Hz. If the wireless channel changes slowly and hence $h(t)$ stays almost a constant over a coding block, the instantaneous channel capacity of the fading Gaussian channel can be formulated in bits/sec/Hz as

$$C(t) = \log_2 \left(1 + \frac{P}{N_0 B} z(t) \right) = \log_2 (1 + \text{SNR} z(t)). \quad (6.2)$$

Then, we assume that if $R < C(t)$, reliable communication is attained and hence the transmitted message is decoded correctly. If, on the other hand, $R \geq C(t)$, outage occurs and retransmission is needed. Under these assumptions, we, following the approach in [27], model the wireless link as a two-state continuous-time Markov chain. The channel is assumed to be in the ON state if $R < C(t)$ or equivalently $z(t) > \zeta$, where $\zeta = \frac{2^R - 1}{\text{SNR}}$. The channel is in the OFF state when $z(t) \leq \zeta$. We denote transition rates from ON to OFF state as λ and from OFF to ON state as μ . Now, the transition rate matrix can be expressed as $\mathbf{Q} = \begin{bmatrix} -\lambda & \lambda \\ \mu & -\mu \end{bmatrix}$.

These transition rates need to be consistent with the properties of the channel. The stationary probabilities are easily obtained as $\frac{\lambda}{\lambda + \mu}$ for the ON state and as $\frac{\mu}{\lambda + \mu}$ for the OFF state. Without loss of generality, we assume that $z(t)$ has unit variance. Then, we can write

$$\Pr \{z(t) > \zeta\} = \int_{\zeta}^{\infty} e^{-z} dz = e^{-\zeta} = \frac{\lambda}{\lambda + \mu}, \quad (6.3)$$

$$\Pr \{z(t) \leq \zeta\} = \int_0^{\zeta} e^{-z} dz = 1 - e^{-\zeta} = \frac{\mu}{\lambda + \mu}. \quad (6.4)$$

Hence, we have $\lambda = \kappa e^{-\zeta}$ and $\mu = \kappa(1 - e^{-\zeta})$ where $\kappa = \lambda + \mu$ can be seen as the

exponential decay rate of the memory of the underlying Rayleigh channel as discussed in [27] and can be determined from the channel statistics.

6.2 Energy Efficiency of Fixed-Rate Transmission of ON-OFF Markov Sources

6.2.1 Markov Fluid Sources

For the two-state Markov fluid source, the average arrival rate is

$$r_{\text{avg}} = P_{\text{ON}} r = \frac{\alpha}{\alpha + \beta} r \quad (6.5)$$

where $P_{\text{ON}} = \alpha/(\alpha + \beta)$ is obtained from the equations in (2.20) and the generating matrix in (2.21). In the Markov fluid model, maximum arrival rate that can be supported by fixed-rate transmissions in the presence of buffer constraints can be obtained by solving (3.2) and the maximum *average* arrival rate $r_{\text{avg}}^*(\text{SNR}, \theta)$ can be determined from (6.5). In the following result, we characterize this maximum average arrival rate in the low-SNR regime and find the minimum energy per bit requirement and the wideband slope.

Theorem 6.2.1. *Assume that the source arrivals and fixed-rate transmissions over the Rayleigh-fading channel are both modeled as ON-OFF continuous-time Markov processes. The decay rate of the memory of the Rayleigh channel is denoted by κ . Then, the minimum energy per bit and wideband slope expressions as a function of the channel and source parameters and the QoS exponent θ are given, respectively, by*

$$\frac{E_b}{N_{0 \min}} = e \log_e 2 = 2.7512 \text{ dB}, \text{ and} \quad (6.6)$$

$$\mathcal{S}_0 = \frac{1}{\frac{\theta}{\log_e 2} \left[\frac{e-1}{\kappa} + \frac{\beta}{\alpha(\alpha+\beta)} \right] + \frac{e}{2}}. \quad (6.7)$$

Proof: We first consider the condition in (3.2) and express it for Markov fluid transmission and source models as

$$\frac{\theta r - (\alpha + \beta) + \sqrt{(\theta r - (\alpha + \beta))^2 + 4\alpha\theta r}}{2\theta} = \frac{\theta R + (\lambda + \mu) - \sqrt{(\theta R + (\lambda + \mu))^2 - 4\lambda\theta R}}{2\theta}. \quad (6.8)$$

Note that the equality in (6.8) enables us to determine the maximum arrival rate, r^* , in the Markov fluid source model and the corresponding optimal fixed transmission rate R^* for the given channel parameters (e.g., the transition rates λ and μ) and the imposed queueing constraints specified by the QoS exponent θ . Note further that as seen in (3.29) and (3.31), we have to determine the first and second derivatives of r^* with respect to SNR at SNR = 0 in order to identify the minimum energy per bit and wideband slope. In (6.8), we have dependence on SNR through λ and μ . It is important to also note that optimum arrival and transmission rates r^* and R^* in general depend on SNR as well.

Initially, we consider an arbitrary fixed-rate transmission strategy $R(\text{SNR})$ for any given SNR. After multiplying both sides of (6.8) with 2θ and taking the derivative with respect to SNR, we obtain the equation in (6.9), given at the top of the next page, where we have defined $g(\text{SNR}) = e^{-\frac{2R-1}{\text{SNR}}}$.

$$\begin{aligned} \theta \dot{r}(\text{SNR}) + \frac{(\theta r(\text{SNR}) - (\alpha + \beta))\theta \dot{r}(\text{SNR}) + 2\alpha\theta \dot{r}(\text{SNR})}{\sqrt{(\theta r(\text{SNR}) - (\alpha + \beta))^2 + 4\alpha\theta r(\text{SNR})}} = \\ \theta \dot{R}(\text{SNR}) + \frac{(\theta R(\text{SNR}) + \kappa)\theta \dot{R}(\text{SNR}) - 2\kappa g(\text{SNR})\theta \dot{R}(\text{SNR}) - 2\kappa \dot{g}(\text{SNR})\theta R(\text{SNR})}{\sqrt{(\theta R(\text{SNR}) + \kappa)^2 - 4\kappa g(\text{SNR})\theta R(\text{SNR})}} \end{aligned} \quad (6.9)$$

Next, we let SNR $\rightarrow 0$. Noting that the arrival rate $r(0) = 0$ and transmission rate $R(0) = 0$ at SNR = 0 and we have $g(0) = 2^{-\dot{R}(0)}$, the equality in (6.9) simplifies to

$$\dot{r}_{\text{avg}}(0) = \dot{r}(0) \frac{\alpha}{\alpha + \beta} = \dot{R}(0) 2^{-\dot{R}(0)}. \quad (6.10)$$

Assume that $R(\text{SNR})$ has the following second-order expansion at SNR = 0:

$$R(\text{SNR}) = a\text{SNR} + b\text{SNR}^2 + o(\text{SNR}) \quad (6.11)$$

for some constants a and b . Then, plugging the result in (6.10) into the formula of minimum energy per bit, we immediately obtain

$$\frac{E_b}{N_{0 \min}} = \frac{2^a}{a}, \quad (6.12)$$

which characterizes the minimum energy per bit for a given transmission rate with $\dot{R}(0) = a$. The smallest value of $\frac{E_b}{N_{0 \min}}$ can be obtained by optimizing over the choice of a . It can be easily seen that $a^* = 1/\log_e 2$ is the optimized value which we use in (6.12) in order to obtain the minimum energy per bit expression given in (6.6). As another equivalent approach, note that this optimal $a^* = 1/\log_e 2$ indeed maximizes $\dot{R}(0)2^{-\dot{R}(0)} = a2^{-a} = \frac{a}{2^a}$. Since maximizing $r(\text{SNR}, \theta, R) = \dot{r}(0)\text{SNR} + o(\text{SNR})$ in the low-SNR regime up to first order is equivalent to maximizing $\dot{r}(0)$, we readily conclude that $\dot{R}^*(0) = a^* = 1/\log_e 2$. Hence, from (6.10), we have $\dot{r}_{\text{avg}}^*(0) = a^* 2^{-a^*}$. Plugging this $\dot{r}_{\text{avg}}^*(0)$ into (3.29), we again obtain the desired result in (6.6).

In order to determine the wideband slope, we first take the second derivative of both sides of (6.8) with respect to SNR and evaluate them at $\text{SNR} = 0$. With further simplification we can easily derive the second derivative of the maximum average arrival rate with respect to SNR at $\text{SNR} = 0$ as in (6.13).

$$\begin{aligned} \ddot{r}_{\text{avg}}(0) = & \theta(\dot{R}(0))^2 \left[\frac{2}{\kappa} \left(2^{-2\dot{R}(0)} - 2^{-\dot{R}(0)} \right) - \frac{2\beta}{\alpha(\alpha + \beta)} 2^{-2\dot{R}(0)} \right] \\ & + 2^{-\dot{R}(0)} \left[2\ddot{R}(0) \left(1 - \dot{R}(0) \log_e 2 \right) - (\dot{R}(0))^3 (\log_e 2)^2 \right]. \end{aligned} \quad (6.13)$$

When we use the optimal $\dot{R}^*(0) = a^* = 1/\log_e 2$ value, we notice that $\ddot{r}_{\text{avg}}^*(0)$ does not depend on $b = \ddot{R}(0)/2$.

Finally, inserting (6.10) and (6.13) into (3.31) and using $a^* = 1/\log_e 2$, we obtain the wideband slope expression in (6.7). ■

Remark 11. *Note that the minimum energy per bit in (6.6) does not depend on the QoS ex-*

ponent θ and hence does not get affected by the presence of the buffer constraints. However, when compared with the ultimate limit of $\frac{E_b}{N_0}_{min} = \log_e 2 = -1.59$ dB achieved when the transmission rate is given by the Shannon capacity, we notice that fixed-rate transmissions incur a certain cost and the minimum energy per bit has significantly increased to 2.7512 dB.

Remark 12. The wideband slope expression in (6.7) depends on the QoS exponent θ , channel memory κ , and the Markov source characteristics through the transition rates α and β . In particular, we see that as θ increases (i.e., more strict buffer constraints are imposed), or κ decreases meaning that channel memory decays more slowly, we have smaller wideband slopes, resulting in smaller average arrival rates at the same energy per bit level or equivalently higher energy per bit to support the same arrival rate. Hence, stricter queuing constraints and/or more correlated channel adversely affect the energy efficiency in the low-SNR regime. Furthermore, increasing source burstiness, for instance by decreasing α and increasing β while keeping $\alpha + \beta$ constant, also lowers the wideband slope and degrades the energy efficiency. Note that smaller α with $\alpha + \beta$ constant means that the stationary distribution of the ON state, P_{ON} , is smaller. Hence, data arrivals occur in less frequent bursts.

Remark 13. It is interesting to note that in the absence of buffer constraints, i.e., when $\theta = 0$, wideband slope expression becomes $S_0 = 2/e = 0.7358$. Hence, we have no dependence on channel memory and source characteristics. We also notice that the wideband slope is smaller compared to $S_0 = 1$ achieved in Rayleigh fading channels when Shannon capacity is considered [57]. Hence, the cost of fixed-rate transmissions is reflected in the wideband slope as well.

In Figure 6.1, we plot the maximum average arrival rate r_{avg}^* as a function of the energy per bit $\frac{E_b}{N_0}$ when $\theta = 1, 10$. For given θ , different curves are obtained for different values of κ , α , and β while $\alpha + \beta$ is fixed. Note that the special case in which $\beta = 0$ corresponds to

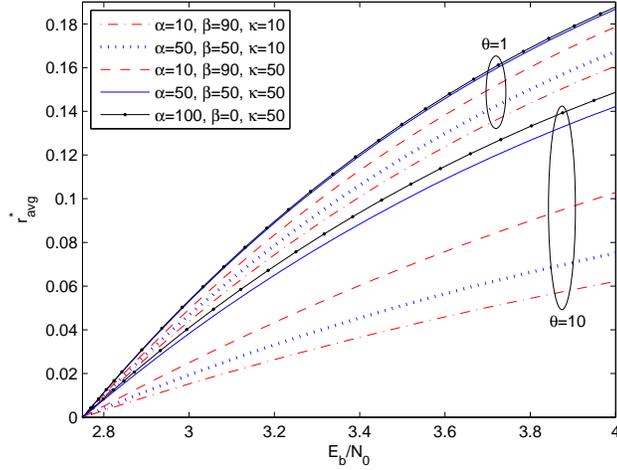


Fig. 6.1: Maximum average arrival rate r_{avg}^* vs. energy per bit $\frac{E_b}{N_0}$ when $\theta = 1, 10$.

constant arrival rate. Confirming our discussions above, we observe that, regardless of the buffer constraints and source characteristics, all curves approach the same minimum energy per bit level of 2.7512 dB. However, smaller κ and hence more slowly decaying channel memory, lower α and larger β and hence more bursty source, and larger θ and hence stricter buffer constraints, all lower the wideband slope and hence result in degradations in the energy efficiency.

6.2.2 Discrete-Time Markov Sources

Finally, we note that we provide above a general framework for energy-efficiency analysis in the low power regime with time-varying sources. While we primarily apply this framework to Markov fluid sources, other source models can be analyzed by following a similar approach. For instance, for a discrete-time Markov ON-OFF source for which

the state transition probability matrix is $\mathbf{J} = \begin{bmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{bmatrix}$, the effective bandwidth is given by $a(\theta, r) = \frac{1}{\theta} \log_e \left(\frac{p_{11} + p_{22} e^{r\theta} + \sqrt{(p_{11} + p_{22} e^{r\theta})^2 - 4(p_{11} + p_{22} - 1)e^{r\theta}}}{2} \right)$. Using the techniques of the

proof of Theorem 6.2.1, we readily have the following characterization for the discrete-time Markov source model with the same transmission and channel assumptions.

Theorem 6.2.2. *Assume now that the source arrival follows the discrete-time ON-OFF model described above. Then, the minimum energy per bit and wideband slope expressions as a function of the channel and source parameters and the QoS exponent θ are given, respectively, by*

$$\frac{E_b}{N_{0\min}} = e \log_e 2 = 2.7512 \text{ dB}, \text{ and}$$

$$\mathcal{S}_0 = \frac{1}{\frac{\theta}{\log_e 2} \left[\frac{e-1}{\kappa} + \frac{(\tilde{\varpi}-1)}{2} \right] + \frac{e}{2}}$$

where we have defined $\tilde{\varpi} = \varpi \left(\frac{1}{P_{ON}} \right)^2$ and

$$\varpi = \frac{p_{22}^2 + 2(1-p_{11})}{2(2-p_{11}-p_{22})} - \frac{[p_{22}(p_{11}+p_{22}) - 2(p_{11}+p_{22}-1)]^2}{2(2-p_{11}-p_{22})^3}.$$

Based on Theorem 6.2.2, similar conclusions as in the Markov fluid source model can immediately be drawn for the discrete-time Markov source as well.

CHAPTER 7

ANALYSIS OF MULTIUSER CHANNELS WITH MARKOV ARRIVALS UNDER QoS CONSTRAINTS

7.1 Throughput Regions of Multiple-Access Fading Channels with Markov Arrivals and QoS Constraints

7.1.1 Channel Model

We consider a multiple-access fading channel in which M users transmit to a common receiver. We assume that each user experiences Markov data arrivals. Randomly arriving data is initially buffered at each user before transmission over the multiple-access channel. For each random source traffic, certain statistical QoS constraints are imposed at each user in order to limit the buffer violation probability.

In the considered multiple-access channel, each link experiences flat-fading and the channel input-output relation can be expressed as

$$y = \sum_{i=1}^M h_i x_i + n \quad (7.1)$$

where x_i is the channel input of the i^{th} user and y is the output at the receiver. Average transmitted signal energy of the i^{th} user is $\mathbb{E}\{|x_i|^2\} = \mathcal{E}_i$. Moreover, n denotes the zero-mean, circularly-symmetric, complex Gaussian background noise at the receiver with variance $\mathbb{E}\{|n|^2\} = N_0$. Hence, the signal-to-noise ratio (SNR) of the i^{th} user is $\text{SNR}_i = \frac{\mathbb{E}\{|x_i|^2\}}{\mathbb{E}\{|n|^2\}} = \frac{\mathcal{E}_i}{N_0}$ for $i = 1, \dots, M$. Finally, in (7.1), h_i denotes the fading coefficient in the channel between the user i and the receiver. While fading coefficients can have arbitrary distributions with finite energies, we assume that block-fading is experienced. Hence, the realizations of the fading coefficients stay fixed for a block of symbols and change independently for the next block.

7.1.2 MAC Throughput Region

In this section, we initially describe our throughput metric as the maximum average arrival rate that can be supported in a setting in which arrivals are modeled as ON-OFF Markov processes, service rates are given by the instantaneous channel capacities, and buffer overflow probabilities are limited as described in Section 2.1. In particular, we formulate the maximum average arrival rates by using both effective bandwidth and effective capacity formulas. We subsequently consider three different strategies for communication in multiple-access fading channels, namely time-division with power control, superposition coding with fixed decoding order, and superposition coding with variable decoding order. Each scheme leads to different service rates at different users and results in different throughput regions.

TDMA

Time division is a simple strategy in which the users send their signals in non-overlapping intervals. Hence, interference is avoided in this case at the cost of reduced transmission

rates. User i transmits τ_i fraction of the time with energy \mathcal{E}_i/τ_i . Therefore, instantaneous service rates in bits/channel use are

$$R_i(\text{SNR}_i) = \tau_i \log_2 \left(1 + \frac{\text{SNR}_i z_i}{\tau_i} \right) \text{ for } i = 1, \dots, M \quad (7.2)$$

where again $z_i = |h_i|^2$. With these service rates, the effective capacity expressions of the users become

$$C_{Ei}(\text{SNR}_i) = -\frac{1}{\theta_i} \log_e \mathbb{E} \left\{ e^{-\frac{\theta_i \tau_i}{\log_e 2} \log_e \left(1 + \frac{\text{SNR}_i z_i}{\tau_i} \right)} \right\}. \quad (7.3)$$

Superposition Coding with Fixed Decoding Order (SC-FDO)

In this strategy, transmitters simultaneously send the data and the receiver decodes the received sum-signal in a fixed-order denoted by π_k (for $k = 1, \dots, M!$) during τ_k fraction of the time. Note that signals of M users can be decoded in $M!$ different orders. Note also that the time fractions $\{\tau_k\}$ satisfy $\tau_k \geq 0$ and $\sum_{k=1}^{M!} \tau_k = 1$. The throughput region is characterized by varying the values of $\{\tau_k\}$. In τ_k fraction of the time, instantaneous service rate of user i in bits/channel use is given by

$$R_{\pi_k(i)} = \log_2 \left(1 + \frac{\text{SNR}_i z_i}{1 + \sum_{\pi_k(j) > \pi_k(i)} \text{SNR}_j z_j} \right). \quad (7.4)$$

Note from the above rate expression that user j with $\pi_k(j) > \pi_k(i)$ is decoded later than user i when decoding order π_k is employed and hence user i sees user j 's signal as interference. Through successive interference cancelation, the signals of the previously-decoded users do not interfere. Accordingly, the effective capacity expression is given by

$$C_{Ei}(\text{SNR}_i) = -\frac{1}{\theta_i} \log_e \mathbb{E} \left\{ e^{-\theta_i \sum_{m=1}^{M!} \tau_m R_{\pi_m(i)}} \right\}. \quad (7.5)$$

Superposition Coding with Variable Decoding Order (SC-VDO)

In this method, users again transmit simultaneously. However, differently from the previous scheme in which the decoding is fixed in each fraction of time, we now consider varying the decoding order depending on the channel states or more specifically channel fading magnitude-squares $\mathbf{z} = [z_1, \dots, z_M] \in \mathbb{R}_+^M$. Assume that the space of fading powers \mathbb{R}_+^M is partitioned into $M!$ regions denoted by $\{\Gamma_k\}_{k=1}^{M!}$. If $\mathbf{z} \in \Gamma_k$, the decoding order π_k is used at the receiver. For a given partition, the effective capacity expression of the i^{th} user can now be expressed as

$$C_{Ei}(\text{SNR}_i) = -\frac{1}{\theta_i} \log_e \mathbb{E} \left\{ e^{-\theta_i \sum_{k=1}^{M!} R_{\pi_k(i)} \mathbf{1}\{\mathbf{z} \in \Gamma_k\}} \right\} \quad (7.6)$$

where $\mathbf{1}\{\cdot\}$ is the indicator function, and $R_{\pi_k(i)}$ is given in (7.4). Since determining the optimal partition of the fading state space is in general a difficult task, we consider suboptimal strategies in order to demonstrate the possible improvements of adopting variable decoding order. In particular, one strategy is

$$\frac{\lambda_{\pi(1)}}{\text{SNR}_{\pi(1)} z_{\pi(1)}} \leq \frac{\lambda_{\pi(2)}}{\text{SNR}_{\pi(2)} z_{\pi(2)}} \leq \dots \leq \frac{\lambda_{\pi(M)}}{\text{SNR}_{\pi(M)} z_{\pi(M)}} \quad (7.7)$$

whose performance was shown in [60] to be close to that of the optimal one, which maximizes the weighted sum-throughput in the special case of two users and constant arrival rates.

7.1.3 Numerical Results

While the analysis above is general, we in this section provide numerical results considering the case of two users, i.e., $M = 2$, and assuming a symmetric setting in which $\text{SNR}_1 = \text{SNR}_2 = \text{SNR}$ and $\theta_1 = \theta_2 = \theta$. The channel fading magnitude-squares z_1 and z_2 are exponentially distributed with arbitrary correlation ρ . Also, for discrete Markov sources,

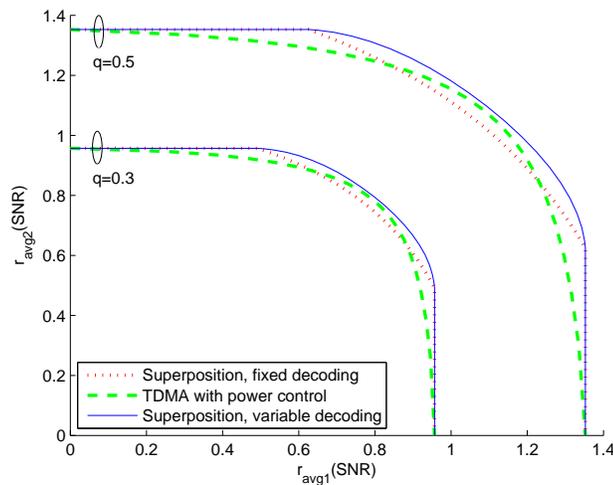


Fig. 7.1: Throughput regions for discrete Markov sources when $\theta = 1$, $\text{SNR} = 10$ and $\rho = 0$.

we set $p_{11} = 1 - q$ and $p_{22} = q$. In all figures, we plot the regions of average arrival rates (or equivalently the throughput regions) achieved by TDMA, SC-FDO and SC-VDO.

In Figure 7.1, we plot the throughput regions for $p_{22} = q = 0.5$ and $q = 0.3$ when the arrivals are modeled as a discrete Markov process and we have $\theta = 1$, $\text{SNR} = 10$, and $\rho = 0$. For these parameter values, we notice that the SC-VDO using the strategy in (7.7) provides the largest throughput region, demonstrating the benefits of variable-decoding order. Interestingly, sum-rate achieved by TDMA exceeds that of SC-FDO. In this figure, we also observe the impact of burstiness on rate regions. When $p_{22} = q$, which is the probability for ON state, has a lower value, data arrivals of given rate r occur less frequently and hence the source is more bursty. It is clearly seen that increased source burstiness reduces the throughput regions of all strategies in a similar fashion.

In Figs. 7.2 and 7.3, we plot the throughput regions again considering discrete Markov arrivals. In Fig. 7.2, we observe that increasing θ from 0.1 to 1 (or equivalently imposing more stringent QoS constraints) reduces the throughput regions. We also notice that while TDMA results in the smallest throughput region for less strict QoS constraints (i.e., when $\theta = 0.1$), TDMA sum rate exceeds that of SC-FDO when θ is increased to 1. Hence, buffer constraints have significant impact on the performance of different communication strate-

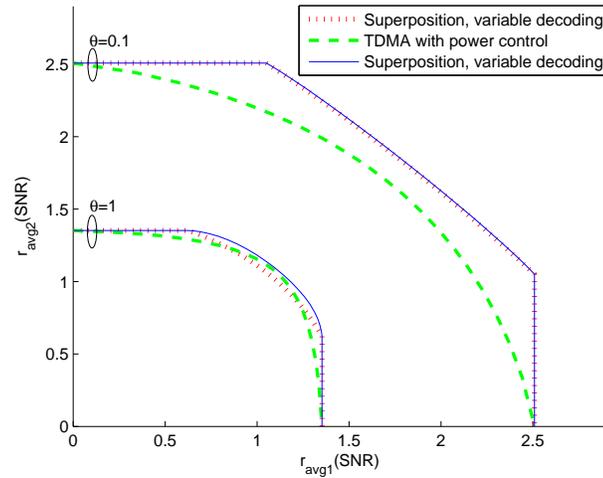


Fig. 7.2: Throughput regions for discrete Markov sources when $q = 0.5$, $\text{SNR} = 10$ and $\rho = 0$.

gies. In Fig. 7.3, we see that increasing SNR expectedly improves the throughput regions. Surprisingly, TDMA sum-rate becomes the largest at $\text{SNR} = 30$, which is in stark contrast to the results in the absence of buffer constraints in which TDMA is always suboptimal with respect to superposition transmissions.

In Fig. 7.4, we consider Markov fluid arrivals. Fading correlation is $\rho = 0.1$. We demonstrate the effect of different values of the transition rates α and β on the throughput region. Having α small and β large (with $\alpha + \beta$ fixed) results in a smaller probability for the ON state, representing a more bursty source. Again, as in Fig. 7.1, we note that increased burstiness hurts the throughput.

7.2 Power Control in Fading Broadcast Channels with Random Arrivals and QoS Constraints

7.2.1 Channel Model

We consider a fading broadcast channel model with one common transmitter and M receivers or users as depicted in Figure 7.5. The transmitter experiences M data flows gener-

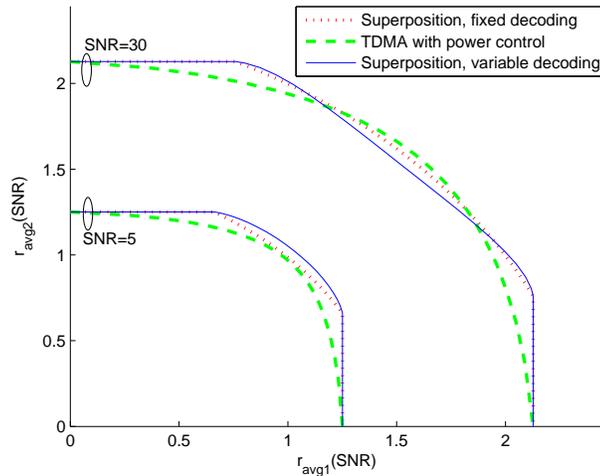


Fig. 7.3: Throughput regions for discrete Markov sources when $q = 0.5$, $\theta = 0.7$ and $\rho = 0$.

ated by Markovian sources. Each flow is intended for a different user and is buffered before transmission in a separate queue. In the fading broadcast channel, the channel input-output relation between the transmitter and the i^{th} user can be expressed as

$$y_i = h_i x + n_i \text{ for } i = 1, \dots, M \quad (7.8)$$

where x is the transmitted signal and y_i is the received signal at the i^{th} receiver. The transmitter operates under an average power constraint of \bar{P} . Hence, the average transmitted signal energy is $\mathbb{E}\{|x|^2\} = \frac{\bar{P}}{B} = \mathcal{E}$ where B denotes the system bandwidth and it is assumed that the symbol rate is B complex symbols/s. Moreover, n_i denotes the zero-mean, circularly-symmetric, complex Gaussian background noise at the i^{th} receiver with variance $\mathbb{E}\{|n_i|^2\} = N_{0i}$. Hence, the signal-to-noise ratio (SNR), defined with respect to the noise level of the first receiver, is

$$\text{SNR} = \frac{\mathbb{E}\{|x|^2\}}{\mathbb{E}\{|n|^2\}} = \frac{\bar{P}}{BN_{01}} = \frac{\mathcal{E}}{N_0} \quad (7.9)$$

where we $N_{01} = N_0$.

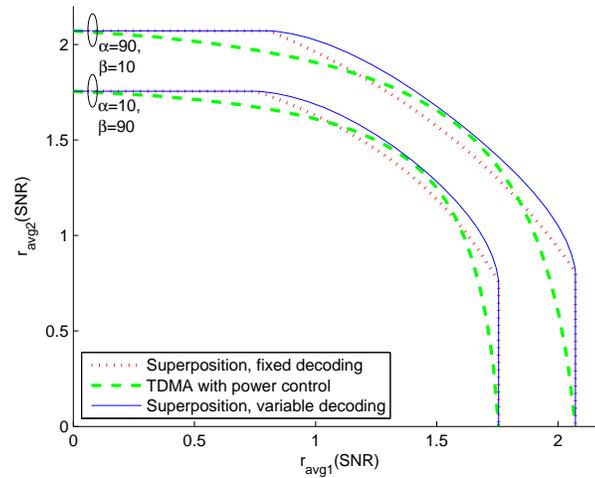


Fig. 7.4: Throughput regions for Markov fluid sources when $\theta = 1$, $\text{SNR} = 10$ and $\rho = 0.1$.

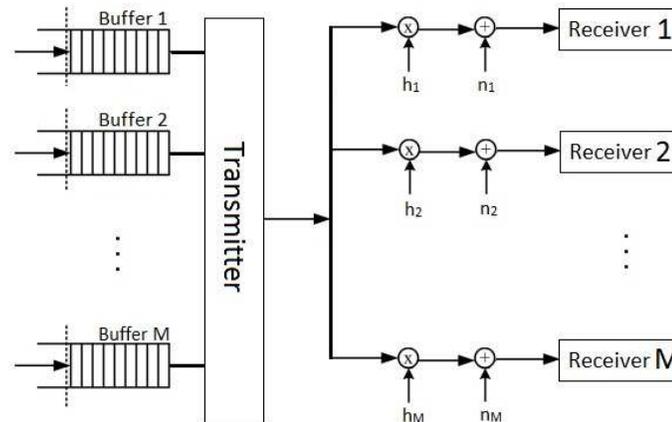


Fig. 7.5: Fading broadcast channel with random arrivals.

Furthermore, in (7.8), h_i denotes the fading coefficient in the channel between the transmitter and receiver i . While fading coefficients can have arbitrary distributions with finite energies, we assume that block-fading is experienced. Hence, the realizations of the fading coefficients stay fixed for a block of symbols and change independently for the next block.

Finally, we assume that power control is employed at the transmitter. Hence, the instantaneous power transmitted to each user is a function of the current fading state $\mathbf{z} = (z_1, \dots, z_M)$ where $z_i = |h_i|^2$ denotes the fading power. We can express the instantaneous transmit power to user i as $P_i(\mathbf{z})$ and denote the instantaneous transmitted SNR level to

user i as

$$\mu_i(\mathbf{z}) = \frac{P_i(\mathbf{z})}{BN_{0i}}. \quad (7.10)$$

Now, the average power constraint at the transmitter becomes

$$\mathbb{E} \left\{ \sum_{i=1}^M \gamma_i \mu_i \right\} = \text{SNR} \quad (7.11)$$

where $\gamma_i = \frac{N_{0i}}{N_{01}}$.

7.2.2 Throughput Regions of Fading Broadcast Channels with Power Control

In this section, we first identify the transmission rates in fading broadcast channels achieved with superposition coding and successive interference cancellation, and determine the throughput regions by formulating the maximum average arrival rates that can be supported in the broadcast channel. Subsequently, we address the optimal power control for the two-user case, describe the Lagrangian optimality conditions, and provide an optimization algorithm.

Transmission Rates with Superposition Coding and Effective Capacity

In order to find the throughput region, we first determine the effective capacities of the transmissions to users in the broadcast channel. The instantaneous service rate to user i achieved by superposition coding at the transmitter and successive interference cancellation at the receiver is given by [94] [93]

$$R_i = \log_2 \left(1 + \frac{\mu_i(\mathbf{z}) z_i}{1 + \sum_{k=1}^M \mu_k(\mathbf{z}) \frac{\gamma_k z_i}{\gamma_i} \mathbf{1}_{\{\gamma_i/z_i > \gamma_k z_k\}}} \right) \quad (7.12)$$

where $\mathbf{1}\{\cdot\}$ is the indicator function. Now, using (2.8) we can express the effective capacity of the transmission to the i^{th} user as

$$C_E(\theta_i) = -\frac{1}{\theta_i} \log_e \mathbb{E} \{e^{-\theta_i R_i}\} \quad i = 1, \dots, M. \quad (7.13)$$

Maximum Average Arrival Rates under QoS Constraints

In this section, we formulate the maximum average arrival rates of Markovian sources that can be supported by transmissions over the fading broadcast channel under statistical queueing constraints. Specifically, we consider two-state Markov arrival models in which the arrival rates are r and 0 in the ON and OFF states, respectively. Stationary distribution of ON state is denoted as P_{ON} . Therefore, the average arrival rate of user i is simply

$$r_{\text{avg}i} = P_{\text{ON}}^i r_i. \quad (7.14)$$

Next, we seek to determine the maximum average arrival rate $r_{\text{avg}i}^*$ that can be supported while satisfying the statistical QoS limitations given in the form in (2.1). As shown in [8, Theorem 2.1], if the effective bandwidth of the arrival process is equal to the effective capacity of the service process, i.e.,

$$a_i(\theta_i, r_i) = C_{Ei}(\theta_i), \quad (7.15)$$

then, (2.1) is satisfied, i.e., buffer violation probability decays exponentially fast with rate controlled by the QoS exponent θ_i . Hence, the solution of (7.15) provides the maximum arrival rate $r_i^*(\theta_i)$ of the data flow intended for user i , which can be supported in the broadcast channel for the given QoS exponent θ_i . Then, the maximum average arrival rate is

$$r_{\text{avg}i}^*(\theta_i) = r_i^*(\theta_i) P_{\text{ON}}^i. \quad (7.16)$$

We adopt this maximum arrival rate as our throughput metric since average arrival rate is equal to the average departure rate when the queue is in steady state [8]. We first show that the throughput is concave in each μ_i . In [18, Lemma 1], it is stated that effective capacity is a concave function of $\boldsymbol{\mu}$ where $\boldsymbol{\mu} = \{\mu_1, \mu_2, \dots, \mu_M\}$ represents the vector composed of the power allocation policies. Reference [11] shows that effective bandwidth of the source is strictly monotonically increasing and is also convex in source arrival rates. Therefore, the inverse function of the effective bandwidth $a^{*-1}(C_E(\theta))$ is a nondecreasing concave function of the effective capacity, which is concave in $\boldsymbol{\mu}$. Using the concavity properties of the composition of functions [31], we realize that the maximum average arrival rate $r_{\text{avg}}^*(\theta) = P_{\text{ON}} a^{*-1}(C_E(\theta))$ is concave.

With this concavity property, we present the optimal power control problem as the following convex optimization problem in which the weighted sum of average arrival rates is maximized over all power allocation policies $\boldsymbol{\mu}$ satisfying the average sum power constraint at the transmitter:

$$\max \sum_{i=1}^M \lambda_i r_{\text{avg}i}^*(\theta_i), \quad (7.17)$$

$$\text{subject to } \mathbb{E} \left\{ \sum_{i=1}^M \gamma_i \mu_i \right\} = \text{SNR} \quad (7.18)$$

where $\{\lambda_i\}$ are the weights satisfying $\sum_{i=1}^M \lambda_i = 1$.

Optimal Power Control in the Two-User Case

For the broadcast channel with two users, there are two different decoding orders. According to the rate expression in (7.12), the user with the better channel can decode the information of the other user and cancel the interference. Given the ratio of the noise powers as $\gamma = \frac{N_{02}}{N_{01}}$, we define two regions for the decoding orders. When the channel conditions are such that $\mathbf{z} \in \Gamma = \{\mathbf{z} : \gamma z_1 > z_2\}$, first user can decode and eliminate the message intended for the second user. When $\mathbf{z} \in \Gamma^c = \{\mathbf{z} : \gamma z_1 < z_2\}$, first user decodes

its message in the presence of the interference from the signal intended for the second user.

Hence the instantaneous service rates of both users are given by

$$R_1 = \begin{cases} \log_2(1 + \mu_1(\mathbf{z})z_1) & \mathbf{z} \in \Gamma \\ \log_2\left(1 + \frac{\mu_1(\mathbf{z})z_1}{1 + \gamma\mu_2(\mathbf{z})z_1}\right) & \mathbf{z} \in \Gamma^c \end{cases}, \quad (7.19)$$

$$R_2 = \begin{cases} \log_2\left(1 + \frac{\mu_2(\mathbf{z})z_2}{1 + \mu_1(\mathbf{z})z_2/\gamma}\right) & \mathbf{z} \in \Gamma \\ \log_2(1 + \mu_2(\mathbf{z})z_2) & \mathbf{z} \in \Gamma^c \end{cases}. \quad (7.20)$$

For the case of two users, the Lagrangian of the convex optimization problem in (7.17)-(7.18) is given by

$$\mathcal{L}(\mu_1, \mu_2, \kappa) = \lambda_1 r_{\text{avg}1}^*(\theta_1) + \lambda_2 r_{\text{avg}2}^*(\theta_2) - \kappa(\mathbb{E}\{\mu_1 + \gamma\mu_2\} - \text{SNR}) \quad (7.21)$$

where κ is the Lagrange multiplier.

Next, we determine the optimality conditions. By taking the derivative of the Lagrangian with respect to μ_1 and μ_2 in regions Γ and Γ^c , we obtain the optimality conditions given in (7.22)–(7.25), where $\varrho_i = \frac{\theta_i}{\log_e 2}$.

$$\begin{aligned} & \frac{\lambda_1}{\psi_1 \log_e 2} (1 + \mu_1 z_1)^{-\varrho_1 - 1} z_1 \\ & - \frac{\lambda_2}{\psi_2 \log_e 2} \left(1 + \frac{\mu_2(\mathbf{z})z_2}{1 + \mu_1(\mathbf{z})z_2/\gamma}\right)^{-\varrho_2 - 1} \frac{\mu_2 z_2^2 / \gamma}{(1 + \mu_1(\mathbf{z})z_2/\gamma)^2} - \kappa = 0 \quad \forall \mathbf{z} \in \mathcal{Z} \end{aligned} \quad (7.22)$$

$$\frac{\lambda_2}{\psi_2 \log_e 2} \left(1 + \frac{\mu_2(\mathbf{z})z_2}{1 + \mu_1(\mathbf{z})z_2/\gamma}\right)^{-\varrho_2 - 1} \frac{z_2}{1 + \mu_1(\mathbf{z})z_2/\gamma} - \gamma\kappa = 0 \quad \forall \mathbf{z} \in \mathcal{Z} \quad (7.23)$$

$$\frac{\lambda_1}{\psi_1 \log_e 2} \left(1 + \frac{\mu_1 z_1}{1 + \mu_2 z_1 \gamma}\right)^{-\varrho_1 - 1} \frac{z_1}{1 + \mu_2 z_1 \gamma} - \kappa = 0 \quad \forall \mathbf{z} \in \mathcal{Z}^c \quad (7.24)$$

$$\begin{aligned}
& - \frac{\lambda_1}{\psi_1 \log_e 2} \left(1 + \frac{\mu_1 z_1}{1 + \mu_2 z_1 \gamma} \right)^{-\varrho_1 - 1} \frac{\mu_1 z_1^2 \gamma}{(1 + \mu_2 z_1 \gamma)^2} \\
& + \frac{\lambda_2}{\psi_2 \log_e 2} (1 + \mu_2(\mathbf{z})z_2)^{-\varrho_2 - 1} z_2 - \gamma \kappa = 0 \quad \forall \mathbf{z} \in \mathcal{Z}^c
\end{aligned} \tag{7.25}$$

These are the conditions that need to be satisfied by the optimal power control policy. In these optimality conditions, ψ_1 and ψ_2 vary according to the source type and characteristics. Assuming $g_i(\theta_i) = \mathbb{E} \{ e^{-\theta_i R_i} \}$, expressions for ψ_i for different arrival models (including the Markov arrivals) are given by the following:

- Constant Arrival:

$$\psi_i = g_i(\theta_i). \tag{7.26}$$

- Discrete Memoryless Source: $p_{22}^i = q_i$ and $p_{11}^i = 1 - q_i$.

$$\psi_i = \frac{g_i(\theta_i) - (1 - q_i)[g_i(\theta_i)]^2}{q_i} \tag{7.27}$$

- Discrete Markov Source:

$$\psi_i = \frac{1/P_{\text{ON}i}}{\frac{1}{g_i(\theta_i) - p_{11}^i [g_i(\theta_i)]^2} + \frac{1 - p_{11}^i - p_{22}^i}{(1 - p_{11}^i - p_{22}^i) g_i(\theta_i) + p_{22}^i}}. \tag{7.28}$$

- Markov Fluid Source:

$$\psi_i = \frac{g_i(\theta_i) [\log_e(g_i(\theta_i)) - \alpha]^2 / P_{\text{ON}i}}{[\log_e(g_i(\theta_i)) - \alpha] [\log_e(g_i(\theta_i)) - \alpha - \beta] - \beta}. \tag{7.29}$$

Now, for some special cases, we can obtain the following relationships for the optimal power control policies utilizing the optimality conditions. When $\mathbf{z} \in \Gamma$, using (7.23), we

obtain the optimal power control policy for the second user as

$$\mu_2 = \left[\frac{1}{\nu_2^{\frac{1}{e_2+1}} \left(\frac{z_2}{1+\mu_1 z_2/\gamma} \right)^{\frac{e_2}{e_2+1}}} - \frac{1 + \mu_1 z_2/\gamma}{z_2} \right]^+ \quad (7.30)$$

where $\nu_2 = \frac{\gamma \kappa \psi_2 \log_e 2}{\lambda_2}$ and $[c]^+ = \max(c, 0)$. Hence, $\mu_2 = 0$ when $\mu_1 > \gamma \left(\frac{1}{\nu_2} - \frac{1}{z_2} \right)$. From (7.22), we can derive μ_1 when $\mu_2 = 0$ as

$$\mu_1 = \left[\frac{1}{\nu_1^{\frac{1}{e_1+1}} z_1^{\frac{e_1}{e_1+1}}} - \frac{1}{z_1} \right]^+ \quad (7.31)$$

where $\nu_1 = \frac{\kappa \psi_1 \log_e 2}{\lambda_1}$. If $\mu_2 \neq 0$, then the (μ_1, μ_2) pair can be found by solving (7.22) and (7.23). Therefore, μ_1 is the positive solution to the equation below

$$\frac{z_1}{\nu_1} (1 + \mu_1 z_1)^{-e_1-1} - \left(\frac{z_2}{\nu_2 (1 + \mu_1 z_2/\gamma)} \right)^{\frac{1}{e_2+1}} = 0. \quad (7.32)$$

Similar method can be applied to power control policies when $\mathbf{z} \in \Gamma^c$. Therefore, by simplifying the equations in (7.24) and (7.25) we obtain

$$\frac{z_2}{\nu_2} (1 + \mu_2 z_2)^{-e_2-1} - \left(\frac{z_1}{\nu_1 (1 + \mu_2 z_1/\gamma)} \right)^{\frac{1}{e_1+1}} = 0. \quad (7.33)$$

As we do not have any closed-form expressions for the optimal power control policies in general, we resort to numerical computations to find optimal strategies. We propose the algorithm in Table 1 to determine the optimal power control policies. This algorithm is similar to that in [95] (in which only constant-rate arrivals are addressed) with the major difference that we employ more general formulations for ψ_1 and ψ_2 in order to take into account the random arrivals and Markov properties. Essentially, in this algorithm there are two loops which we search for the ψ_1 , ψ_2 (outer loop) and κ (inner loop). First we initialize these values, and until we satisfy the average power constraint, we update κ .

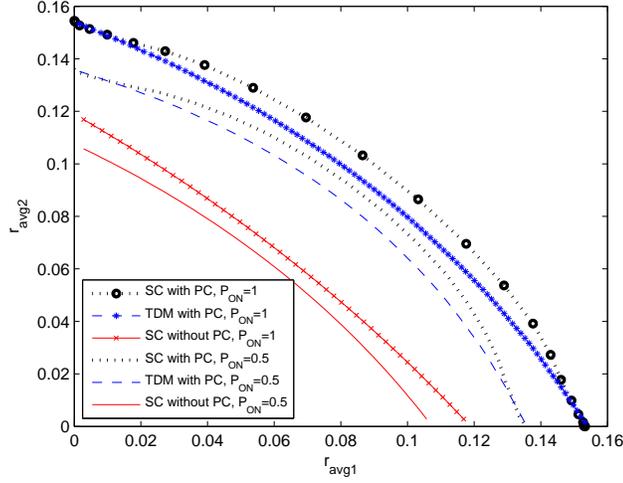


Fig. 7.6: Maximum average arrival rate regions $(r_{\text{avg}1}^*, r_{\text{avg}2}^*)$ of the two users when $\theta = 2$, SNR = 0.1, $\gamma = 1$ and source has discrete Markov property when $P_{\text{ON}} = 0.5$.

After satisfying average power constraint, we update the ψ_1 and ψ_2 values and return to inner loop again to find the fitting value of κ . We continue the algorithm until the change in ψ_1 and ψ_2 values is small and below a threshold.

In the application of the algorithm, we have the used subgradient method for updating ψ_1 , ψ_2 and κ in order to have convergence. Gaussian quadrature methods are employed in the computation of the integrals.

7.2.3 Numerical Results and Discussion

For the numerical analysis, we assume that z_1 and z_2 are independent exponential random variables with unit mean. Also, we assume $\gamma = 1$, SNR = 0.1 and $\theta = 2$ unless stated otherwise. Discrete Markov sources are considered in Figs. 7.6 and 7.7, and Markov fluid sources are addressed in Figs. 7.8 and 7.9.

Furthermore, we consider three different strategies, namely superposition coding with optimal power control (SC with PC), time-division multiplexing with power control (TDM with PC), and superposition coding without power control (SC without PC). SC with PC is the case we have concentrated and described in the chapter. In the relatively simple strategy

Algorithm 2 The optimal power control algorithm that maximizes the weighted sum of throughput expressions

1: Given λ_1, λ_2 , Initialize ψ_1, ψ_2 ;
2: Initialize κ
3: Determine $\nu_1 = \frac{\kappa\psi_1 \log_e 2}{\lambda_1}, \nu_2 = \frac{\gamma\kappa\psi_2 \log_e 2}{\lambda_2}$;
4: **if** $\gamma z_1 > z_2$ **then**
5: **if** $z_1 > \nu_1$ **then**
6: $\mu_1 = \left[\frac{1}{\frac{1}{\nu_1^{\frac{1}{e_1+1}}} \frac{e_1}{z_1^{\frac{1}{e_1+1}}} - \frac{1}{z_1}} \right]^+$;
7: **if** $\mu_1 > \gamma \left(\frac{1}{\nu_2} - \frac{1}{z_2} \right)$ or $z_2 < \nu_2$ **then**
8: $\mu_2 = 0$;
9: **else**
10: **if** (7.32) returns positive solution **then**
11: Compute μ_1, μ_2 from (7.22) and (7.23);
12: **end if**
13: **end if**
14: **else**
15: $\mu_1 = 0, \mu_2 = \left[\frac{1}{\frac{1}{\nu_2^{\frac{1}{e_2+1}}} \frac{e_2}{z_2^{\frac{1}{e_2+1}}} - \frac{1}{z_2}} \right]^+$;
16: **end if**
17: **else**
18: **if** $z_2 > \nu_2$ **then**
19: $\mu_2 = \left[\frac{1}{\frac{1}{\nu_2^{\frac{1}{e_2+1}}} \frac{e_2}{z_2^{\frac{1}{e_2+1}}} - \frac{1}{z_2}} \right]^+$;
20: **if** $\mu_2 > \frac{1}{\gamma} \left(\frac{1}{\nu_1} - \frac{1}{z_1} \right)$ or $z_1 < \nu_1$ **then**
21: $\mu_1 = 0$;
22: **else**
23: **if** (7.33) returns positive solution **then**
24: Compute μ_1, μ_2 from (7.24) and (7.25);
25: **end if**
26: **end if**
27: **else**
28: $\mu_2 = 0, \mu_1 = \left[\frac{1}{\frac{1}{\nu_1^{\frac{1}{e_1+1}}} \frac{e_1}{z_1^{\frac{1}{e_1+1}}} - \frac{1}{z_1}} \right]^+$;
29: **end if**
30: **end if**
31: **if** μ_1 and μ_2 do not satisfy the average power constraint **then**
32: Update κ and return to Step 3;
33: **end if**
34: Update ψ_1 and ψ_2 using μ_1 and μ_2 ;
35: **if** new values of ψ_1 and ψ_2 do not agree with the previous values **then**
36: Return to Step 3;
37: **end if**
38: Declare μ_1 and μ_2 as the optimal power allocation policies.

of TDM with PC, transmitter sends the data to one user at a time using time-sharing and employ the optimal power control policies for the single-user case given by

$$\mu_i = \left[\frac{1}{\frac{1}{\nu_i^{e_i+1}} z_i \frac{\theta_i}{e_i^{e_i+1}}} - \frac{1}{z_i} \right]^+ \quad \text{for } i = 1, 2. \quad (7.34)$$

In SC without PC, transmitter sends the data to both users simultaneously but no power adaptation is considered. Hence, transmission occurs at fixed power levels.

In Fig. 7.6 we plot the throughput region (or equivalently the region of maximum arrival rates) in the two-users case for two different ON-state probabilities. When $P_{\text{ON}} = 1$, source arrival rate is constant. On the other hand, discrete Markov arrivals are experienced when $P_{\text{ON}} = 0.5$. SC with PC provides the largest throughput region out of all three strategies. Each point on the boundary of this region is obtained by varying the weights $\{\lambda\}$ in the optimization problem in (7.17) and obtaining the optimal power control policy for each case. TDM with PC leads to the second largest throughput region. However, transmission with constant power gives a much smaller rate region for the same parameters, demonstrating the effectiveness of power control. In this figure, we also observe the impact of the source burstiness. We see that when the source becomes bursty with $P_{\text{ON}} = 0.5$, the rate region shrinks for all strategies compared to the case of constant arrivals, i.e., when $P_{\text{ON}} = 1$.

For the same six scenarios, we plot the sum rate as a function of SNR in Fig. 7.7. Similar conclusions as in the previous figure apply. However, as SNR increases, we notice that the gap in performance between SC with PC and TDM with PC diminishes. Hence, TDM becomes an effective strategy at low SNR levels. Moreover, for low SNR values, transmission with constant power affects the performance more significantly than having a bursty source.

In Fig. 7.8, we plot the throughput region when the data arrivals to the transmitter are from Markov fluid sources. Burstiness in Markov fluid sources can be described by

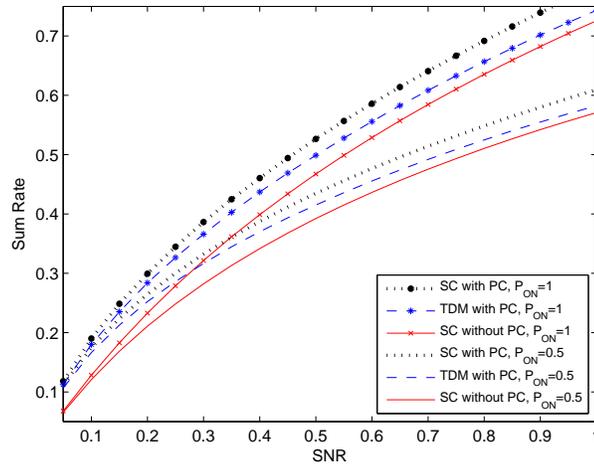


Fig. 7.7: Sum rate vs. SNR when $\theta = 2$ and $\gamma = 1$ with discrete Markov source.

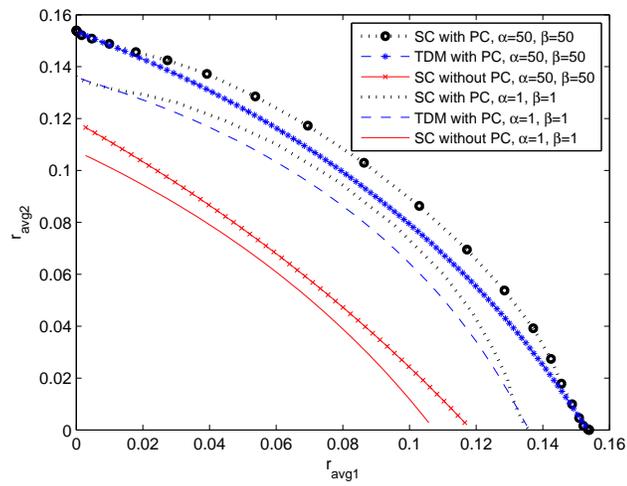


Fig. 7.8: Maximum average arrival rate regions $(r_{\text{avg1}}^*, r_{\text{avg2}}^*)$ of the two users when $\theta = 2$, SNR = 0.1, $\gamma = 1$ and source has Markov fluid property.

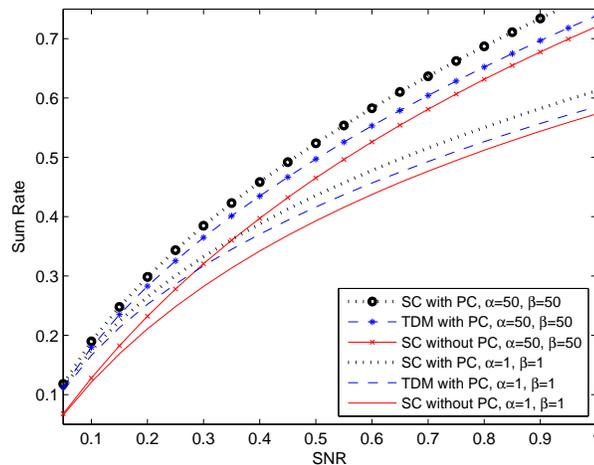


Fig. 7.9: Sum rate vs. SNR when $\theta = 2$ and $\gamma = 1$ with Markov fluid source.

the ON-state probability P_{ON} and the parameters of the generating matrix, α and β , which describe the rate of change between states. When these parameters have higher values, states fluctuate more frequently. In the figure, we use two different source models with the same $P_{\text{ON}} = 0.5$. When $\alpha_i + \beta_i = 100$, rate regions are larger. In this case, switching between ON and OFF states happens relatively fast. Hence, long-duration data flows to the buffers are often avoided. For sources with $\alpha_i + \beta_i = 2$, ON-state can persist with higher probability because probability of switching from one state to another is small. This can lead to large bursts of data, which is detrimental in the presence of buffer constraints. Hence, we see that the rate region shrinks. Similarly as in Fig. 7.6, we observe that SC with PC provides the largest rate region while SC without PC gives the worst rate region.

In Fig. 7.9, we use the same scenarios as in Fig. 7.8, but we plot the sum rate vs. SNR curves. Again, we can immediately draw similar conclusions. Clearly, decreasing the values of α and β hurts the system by reducing the throughput. The SC without PC strategy leads to the smallest sum-throughput out of all strategies. However, as SNR increases, the performance gap between TDMA with PC and SC without PC tends to diminish.

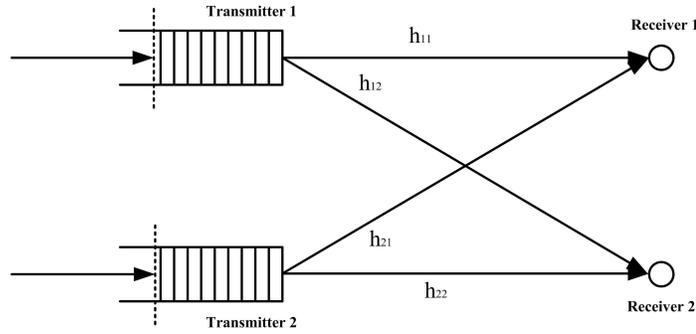


Fig. 7.10: The system model.

7.3 Energy Efficiency in Fading Interference Channels under QoS Constraints

7.3.1 Channel Model

As depicted in Figure 7.10, we consider a two-user flat-fading interference channel model in which each transmitter, operating under buffer constraints, sends information to its intended receiver while causing interference to the other through the cross-links. In this interference channel, the input-output relationships can be expressed as

$$\begin{aligned} y_1 &= h_{11}x_1 + h_{21}x_2 + n_1 \\ y_2 &= h_{12}x_1 + h_{22}x_2 + n_2 \end{aligned} \quad (7.35)$$

where x_i is the channel input from the i^{th} transmitter and y_j is the received signal at the j^{th} receiver, for $i, j \in \{1, 2\}$. Average transmitted signal energy is $\mathbb{E}\{|x_i|^2\} = \mathcal{E}_i$. Moreover, n_i denotes the zero-mean, circularly-symmetric, complex Gaussian noise with variance $\mathbb{E}\{|n_i|^2\} = N_0$. Hence, the i^{th} transmitter's signal-to-noise ratio (SNR) is

$$\text{SNR}_i = \frac{\mathbb{E}\{|x_i|^2\}}{\mathbb{E}\{|n_i|^2\}} = \frac{\mathcal{E}_i}{N_0} \quad i = 1, 2. \quad (7.36)$$

We denote the ratio of the SNRs of first and second users as $\gamma = \frac{\text{SNR}_2}{\text{SNR}_1}$.

Finally, in (7.35), h_{ij} denotes the random fading coefficient of the channel between the i^{th} transmitter and j^{th} receiver, and $z_{ij} = |h_{ij}|^2$ is the magnitude-square of the fading coefficient. While fading coefficients can have arbitrary distributions with finite energies, we assume that block-fading is experienced. Hence, the realizations of the fading coefficients stay fixed for a block of symbols and change independently for the next block.

7.3.2 Energy Efficiency in Fading Interference Channels

In this section, we consider three different strategies for communication over interference channels, namely time-division with power control, treating interference as noise, and simultaneous decoding. We investigate the energy efficiency of these schemes by determining the corresponding minimum energy per bit and wideband slope expressions.

Time Division with Power Control

Time division is a simple strategy in which the transmitters send their signals in non-overlapping intervals. Hence, interference is avoided in this case at the cost of reduced transmission rates. In the two-user model, transmitter 1 sends the data α fraction of the time with energy \mathcal{E}_1/α . Consequently, the remaining $(1 - \alpha)$ fraction of the time is dedicated to transmitter 2. The following result provides a characterization of the energy efficiency in the low-SNR regime.

Proposition 7.3.1. *For time division with power control, the minimum energy per bit and wideband slope expressions of both pairs of users as a function of the fading statistics and the QoS exponent θ are given, respectively, by*

$$\frac{E_b}{N_{0 \min,1}} = \frac{\log_e 2}{\mathbb{E}\{z_{11}\}} \text{ and } \frac{E_b}{N_{0 \min,2}} = \frac{\log_e 2}{\mathbb{E}\{z_{22}\}}, \quad (7.37)$$

$$\mathcal{S}_{0,1} = \frac{2(\mathbb{E}\{z_{11}\})^2}{\frac{\theta_1}{\log_e 2} \text{var}(z_{11}) + \frac{1}{\alpha} \mathbb{E}\{z_{11}^2\}} \text{ and} \quad (7.38)$$

$$\mathcal{S}_{0,2} = \frac{2(\mathbb{E}\{z_{22}\})^2}{\frac{\theta_2}{\log_e 2} \text{var}(z_{22}) + \frac{1}{(1-\alpha)} \mathbb{E}\{z_{22}^2\}}. \quad (7.39)$$

where α is the fraction of time allocated to transmitter 1, and $\text{var}(x)$ denotes the variance of the random variable x .

Proof: When time division is employed for transmission, the instantaneous service rates of the two transmitters in bits/channel use are

$$R_1 = \alpha \log_2 \left(1 + \frac{\text{SNR}_1 z_{11}}{\alpha} \right), \quad \text{and} \quad (7.40)$$

$$R_2 = (1 - \alpha) \log_2 \left(1 + \frac{\text{SNR}_1 z_{22}}{1 - \alpha} \right), \quad (7.41)$$

respectively. With these service rates, the effective capacity expressions, i.e., the arrival rates that can be supported by the two transmitters, become

$$C_{E1}(\text{SNR}_1) = -\frac{1}{\theta_1} \log_e \mathbb{E} \left\{ e^{-\frac{\theta_1 \alpha}{\log_e 2} \log_e \left(1 + \frac{\text{SNR}_1 z_{11}}{\alpha} \right)} \right\}, \quad (7.42)$$

$$C_{E2}(\text{SNR}_2) = -\frac{1}{\theta_2} \log_e \mathbb{E} \left\{ e^{-\frac{\theta_2 (1-\alpha)}{\log_e 2} \log_e \left(1 + \frac{\text{SNR}_2 z_{22}}{1-\alpha} \right)} \right\}. \quad (7.43)$$

Now, the first and second derivatives of the effective capacities at $\text{SNR}_1 = 0$ and $\text{SNR}_2 = 0$, respectively, are

$$\dot{C}_{E1}(0) = \frac{\mathbb{E}\{z_{11}\}}{\log_e 2} \text{ and } \dot{C}_{E2}(0) = \frac{\mathbb{E}\{z_{22}\}}{\log_e 2} \quad (7.44)$$

$$\ddot{C}_{E1}(0) = - \left[\frac{\theta_1}{(\log_e 2)^2} \text{var}(z_{11}) + \frac{1}{\alpha \log_e 2} \mathbb{E}\{z_{11}^2\} \right] \quad (7.45)$$

$$\ddot{C}_{E2}(0) = - \left[\frac{\theta_2}{(\log_e 2)^2} \text{var}(z_{22}) + \frac{1}{(1 - \alpha) \log_e 2} \mathbb{E} \{ z_{22}^2 \} \right]. \quad (7.46)$$

Inserting the expressions in (7.44), (7.45), and (7.46) into 3.29 and 3.31, we obtain the minimum bit energy and wideband slope expressions. \square

Remark 14. *It is seen that the minimum bit energies are functions of only the mean of the fading magnitude-squares of the direct links and are independent of the QoS exponent θ and time-sharing parameter α , which generally affect the wideband slopes. For instance, the wideband slopes diminish with increasingly more strict QoS constraints or equivalently larger values of the QoS exponent θ .*

Treating Interference as Noise

In this strategy, transmitters simultaneously send the data and the receivers regard the interference as noise.

Proposition 7.3.2. *If interference is treated as noise, the minimum energy per bit and wideband slope expressions as a function of the fading statistics and the QoS exponent θ are given, respectively, by*

$$\frac{E_b}{N_{0 \min,1}} = \frac{\log_e 2}{\mathbb{E} \{ z_{11} \}} \text{ and } \frac{E_b}{N_{0 \min,2}} = \frac{\log_e 2}{\mathbb{E} \{ z_{22} \}}, \quad (7.47)$$

$$\mathcal{S}_{0,1} = \frac{2(\mathbb{E} \{ z_{11} \})^2}{\frac{\theta_1}{\log_e 2} \text{var}(z_{11}) + \mathbb{E} \{ z_{11}^2 + 2\gamma z_{11} z_{21} \}} \text{ and} \quad (7.48)$$

$$\mathcal{S}_{0,2} = \frac{2(\mathbb{E} \{ z_{22} \})^2}{\frac{\theta_2}{\log_e 2} \text{var}(z_{22}) + \mathbb{E} \left\{ z_{22}^2 + \frac{2}{\gamma} z_{22} z_{12} \right\}} \quad (7.49)$$

where $\gamma = \frac{SNR_2}{SNR_1}$ is the ratio that is kept fixed as both SNR_1 and SNR_2 approach zero.

Proof: In this case, the maximum instantaneous service rates in bits/channel use are

$$R_1 = \log_2 \left(1 + \frac{\text{SNR}_1 z_{11}}{1 + \text{SNR}_2 z_{21}} \right), \quad \text{and} \quad (7.50)$$

$$R_2 = \log_2 \left(1 + \frac{\text{SNR}_2 z_{22}}{1 + \text{SNR}_1 z_{12}} \right). \quad (7.51)$$

Accordingly, the effective capacity expressions are given by

$$C_{E1}(\text{SNR}_1) = -\frac{1}{\theta_1} \log_e \mathbb{E} \left\{ e^{-\frac{\theta_1}{\log_e 2} \log_e \left(1 + \frac{\text{SNR}_1 z_{11}}{1 + \text{SNR}_2 z_{21}} \right)} \right\} \quad (7.52)$$

$$C_{E2}(\text{SNR}_2) = -\frac{1}{\theta_2} \log_e \mathbb{E} \left\{ e^{-\frac{\theta_2}{\log_e 2} \log_e \left(1 + \frac{\text{SNR}_2 z_{22}}{1 + \text{SNR}_1 z_{12}} \right)} \right\}. \quad (7.53)$$

First and second derivatives of the effective capacities at $\text{SNR}_1 = 0$ and $\text{SNR}_2 = 0$, respectively, can easily be determined as

$$\dot{C}_{E1}(0) = \frac{\mathbb{E}\{z_{11}\}}{\log_e 2} \quad \text{and} \quad \dot{C}_{E2}(0) = \frac{\mathbb{E}\{z_{22}\}}{\log_e 2} \quad (7.54)$$

$$\ddot{C}_{E1}(0) = - \left[\frac{\theta_1}{(\log_e 2)^2} \text{var}(z_{11}) + \frac{1}{\log_e 2} \mathbb{E} \{ z_{11}^2 + 2\gamma z_{11} z_{21} \} \right] \quad (7.55)$$

$$\ddot{C}_{E2}(0) = - \left[\frac{\theta_2}{(\log_e 2)^2} \text{var}(z_{22}) + \frac{1}{\log_e 2} \mathbb{E} \left\{ z_{22}^2 + \frac{2}{\gamma} z_{22} z_{12} \right\} \right]. \quad (7.56)$$

Inserting the above derivative expressions into 3.29 and 3.31, we obtain the desired result.

□

Remark 15. *We immediately notice that the minimum bit energy expressions are the same as in time division strategy. Hence, asymptotically as SNRs vanish, similar energy efficiency performances are achieved by both methods. On the other hand, wideband slopes*

are evidently different, leading to the conclusion that different levels of energy efficiency can be attained when SNRs are small but nonzero. Note also that we now expectedly have wideband slopes depending on the interference strength through z_{12} and z_{21} .

Simultaneous Decoding

Final method we consider is simultaneous decoding. In this scheme, transmitters again send the information simultaneously but as a key difference from the previous subsection, receivers attempt to decode both messages. Therefore, computational complexity of decoding is higher at the receivers.

Before presenting the result, we first define the region:

$$\Gamma = \{z_{11}, z_{21}, z_{12}, z_{22} \geq 0 : z_{11} + \gamma z_{21} < z_{12} + \gamma z_{22}\}. \quad (7.57)$$

Proposition 7.3.3. *For simultaneous decoding, the minimum energy per bit and wideband slope expressions as a function of the fading statistics and the QoS exponent θ are given, respectively, by*

$$\frac{E_b}{N_{0 \min, i}} = \frac{\log_e 2}{\mathbb{E}\{\dot{g}_i(0)\}} \text{ for } i = 1, 2 \text{ and} \quad (7.58)$$

$$\mathcal{S}_{0, i} = \frac{2(\mathbb{E}\{\dot{g}_i(0)\})^2}{\frac{\theta_i}{\log_e 2} \text{var}(\dot{g}_i(0)) + \mathbb{E}\{\ddot{g}_i(0)\}} \text{ for } i = 1, 2 \quad (7.59)$$

where

$$\dot{g}_1(0) = \begin{cases} \alpha_1 z_{11} + (1 - \alpha_1) [(z_{11} + \gamma z_{21} - \gamma z_{22})]^+ & \text{if } Z \in \Gamma \\ \alpha_2 z_{11} + (1 - \alpha_2) z_{12} & \text{if } Z \in \Gamma^c \end{cases} \quad (7.60)$$

$$\dot{g}_2(0) = \begin{cases} (1 - \alpha_1)z_{22} + \alpha_1 z_{21} & \text{if } Z \in \Gamma \\ (1 - \alpha_2)z_{22} + \alpha_2 \left[(z_{22} + \frac{z_{12}}{\gamma} - \frac{z_{11}}{\gamma}) \right]^+ & \text{if } Z \in \Gamma^c \end{cases} \quad (7.61)$$

$$\ddot{g}_1(0) = \begin{cases} \alpha_1 z_{11}^2 + (1 - \alpha_1) [(z_{11} + \gamma z_{21})^2 - (\gamma z_{22})^2]^+ & \text{if } Z \in \Gamma \\ \alpha_2 z_{11}^2 + (1 - \alpha_2) [(z_{12} + \gamma z_{22})^2 - (\gamma z_{22})^2] & \text{if } Z \in \Gamma^c \end{cases} \quad (7.62)$$

$$\ddot{g}_2(0) = \begin{cases} (1 - \alpha_1)(z_{22})^2 + \alpha_1 \left[(z_{21} + \frac{z_{11}}{\gamma})^2 - (\frac{z_{11}}{\gamma})^2 \right] & \text{if } Z \in \Gamma \\ (1 - \alpha_2)(z_{22})^2 + \alpha_2 \left[(\frac{z_{12}}{\gamma} + z_{22})^2 - (\frac{z_{11}}{\gamma})^2 \right]^+ & \text{if } Z \in \Gamma^c \end{cases} \quad (7.63)$$

Above, $\alpha_1 \in [0, 1]$ and $\alpha_2 \in [0, 1]$ are time-sharing parameters between different operating points on the achievable instantaneous rate region of simultaneous decoding, and Z represents the collection $\{z_{11}, z_{21}, z_{12}, z_{22}\}$.

Proof: To derive the effective capacity formulas, we first need to identify the instantaneous service rates achieved with simultaneous decoding. For given channel gains, the instantaneous rate region achieved with simultaneous decoding is [96]

$$\begin{aligned} R_1 &< \log_2(1 + \text{SNR}_1 z_{11}) \\ R_2 &< \log_2(1 + \text{SNR}_2 z_{22}) \\ R_1 + R_2 &< \min \{ \log_2(1 + \text{SNR}_1 z_{11} + \text{SNR}_2 z_{21}), \\ &\quad \log_2(1 + \text{SNR}_1 z_{12} + \text{SNR}_2 z_{22}) \}. \end{aligned} \quad (7.64)$$

Therefore, transmission rates have different regions depending on which term is the minimum in the sum rate constraint in (7.64) or equivalently whether $Z = \{z_{11}, z_{21}, z_{12}, z_{22}\}$ is in region Γ or not. Transmission rates on the boundary of this region can be characterized as in (7.65) and (7.66) on the next page. In these expressions, $\alpha_1 \in [0, 1]$ and $\alpha_2 \in [0, 1]$ are the time sharing parameters (for when $Z \in \Gamma$ and $Z \in \Gamma^c$, respectively) between different operating points on the boundary of the achievable rate region. Using these maximum

$$\text{If } Z \in \Gamma \rightarrow \left\{ \begin{array}{l} R_1 = \alpha_1 \log_2(1 + \text{SNR}_1 z_{11}) + (1 - \alpha_1) \left[\log_2 \left(\frac{1 + \text{SNR}_1 z_{11} + \text{SNR}_2 z_{21}}{1 + \text{SNR}_2 z_{22}} \right) \right]^+ \\ R_2 = \alpha_1 \log_2 \left(1 + \frac{\text{SNR}_2 z_{21}}{1 + \text{SNR}_1 z_{11}} \right) + (1 - \alpha_1) \log_2(1 + \text{SNR}_2 z_{22}) \end{array} \right\} \quad (7.65)$$

$$\text{If } Z \in \Gamma^c \rightarrow \left\{ \begin{array}{l} R_1 = \alpha_2 \log_2(1 + \text{SNR}_1 z_{11}) + (1 - \alpha_2) \log_2 \left(1 + \frac{\text{SNR}_1 z_{12}}{1 + \text{SNR}_2 z_{22}} \right) \\ R_2 = \alpha_2 \left[\log_2 \left(\frac{1 + \text{SNR}_2 z_{22} + \text{SNR}_1 z_{12}}{1 + \text{SNR}_1 z_{11}} \right) \right]^+ + (1 - \alpha_2) \log_2(1 + \text{SNR}_2 z_{22}) \end{array} \right\} \quad (7.66)$$

service rate expressions, effective capacity formulas can be written as

$$C_{Ei}(\text{SNR}_i) = -\frac{1}{\theta_i} \log_e \mathbb{E} \{ e^{-\theta_i R_i} \} \text{ for } i = 1, 2. \quad (7.67)$$

We define $g_i(\text{SNR}_i) = R_i \log_e 2$. First and second derivative of the effective capacity at $\text{SNR}_1 = 0$ and $\text{SNR}_2 = 0$, respectively, are given by

$$\dot{C}_{Ei}(0) = \mathbb{E} \left\{ \frac{\dot{g}_i(0)}{\log_e 2} \right\} \text{ for } i = 1, 2 \quad (7.68)$$

$$\ddot{C}_{Ei}(0) = -\mathbb{E} \left\{ \theta_i \left(\frac{\dot{g}_i(0)}{\log_e 2} \right)^2 + \frac{\ddot{g}_i(0)}{\log_e 2} \right\} \text{ for } i = 1, 2 \quad (7.69)$$

where the derivatives $\dot{g}_i(0)$ and $\ddot{g}_i(0)$ are defined in the Proposition. Similarly as before, inserting these derivative expressions into 3.29 and 3.31, we have the result. \square

Remark 16. *We note that the minimum bit energy expressions are different from those achieved with time division and treating interference as noise. The comparison of minimum bit energies of different strategies is provided through numerical results in the next section. In general, we observe that improved energy efficiency is experienced in strong interference channels when simultaneous decoding methods are employed at the receivers.*

7.3.3 Numerical Results and Discussion

In this section, we provide numerical results obtained through Monte Carlo simulations. We set the SNR ratio as $\gamma = 1$ and assume that the QoS exponents are equal, i.e., $\theta_1 = \theta_2$. All channel coefficients are assumed to be independent. Moreover, channel fading coefficients of the direct links are identically distributed. Fading coefficients of the cross links are also pairwise identically distributed.

Time Division with Power Control vs Treating Interference as Noise

We first compare the energy efficiencies of time division multiplexing and treating interference as noise. Since both achieve the same minimum energy per bit, we investigate the wideband slope regions in Figs. 7.11 and 7.12. In these figures, we assume that $\mathbb{E}\{z_{11}^2\} = \mathbb{E}\{z_{22}^2\} = 2$ and $\mathbb{E}\{z_{11}\} = \mathbb{E}\{z_{22}\} = 1$. In Fig. 7.11, we have $\theta_1 = \theta_2 = 1$. In this figure, curved line represents the slope region of time division multiplexing. Rectangular regions are the slope regions of treating interference as noise. For strong, medium, and weak interference scenarios, we set $\mathbb{E}\{z_{11}z_{21}\} = \mathbb{E}\{z_{22}z_{12}\}$ equal to 4, 2, and 1, respectively. In the figure, we notice that the slope region of treating interference as noise grows expectedly as the interference weakens, and this strategy can attain points outside the slope region of time division multiplexing when the interference is weak. Hence, both users can achieve relatively high wideband slopes and operate more energy efficiently.

The effect of buffer constraints is demonstrated in Fig. 7.12. In Figs. 7.12(a) and 7.12(b), strong interference is considered, and hence treating interference as noise performs worse. In Fig. 7.12(a), we have $\theta_1 = \theta_2 = 0.1$, which is increased to 10 in Fig. 7.12(b). Hence, in the latter case, we have more strict QoS limitations. In such a scenario, we notice in Fig. 7.12(b) that both slope regions become smaller and also approach each other. Hence, under strict QoS constraints, energy efficiency degrades, and time division and treating interference as noise start providing comparable performances. Similar observations are noted in Figs. 7.12(c) and 7.12(d), in which weak interference is considered.

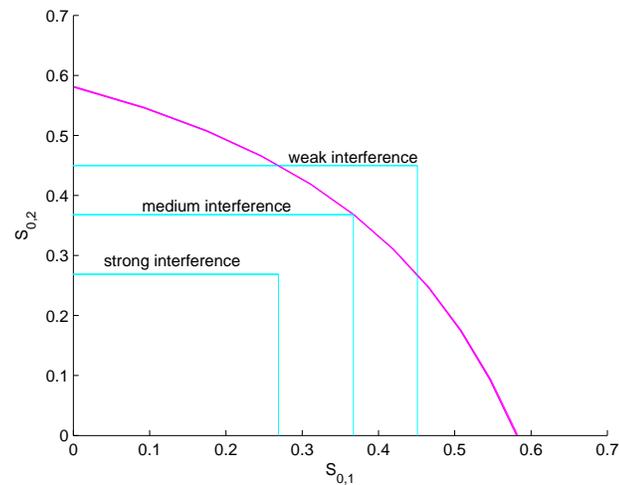


Fig. 7.11: $S_{0,1}$ vs. $S_{0,2}$ for different interference levels for time division with power control (curved line) and treating interference as noise (rectangular regions). $\theta_1 = \theta_2 = 1$.

Energy Efficiency of Simultaneous Decoding

Finally, we address the energy efficiency of simultaneous decoding. Our focus is the minimum energy per bit, which is possibly different from those attained by time division and treating interference as noise. In Figs. 7.13 and 7.14, we plot the minimum energy per bit achieved by transmitter 1 as a function of the time sharing parameters α_1 and α_2 . The flat planes represent the bit energy levels achieved by time division multiplexing and treating interference as noise. In Fig. 7.13, in which weak interference is considered, we observe that simultaneous decoding requires higher minimum bit energies and is therefore not favorable. On the other hand, we note in Fig. 7.14, where interference is strong, that simultaneous decoding can lead to significant gains in minimum energy per bit requirements.

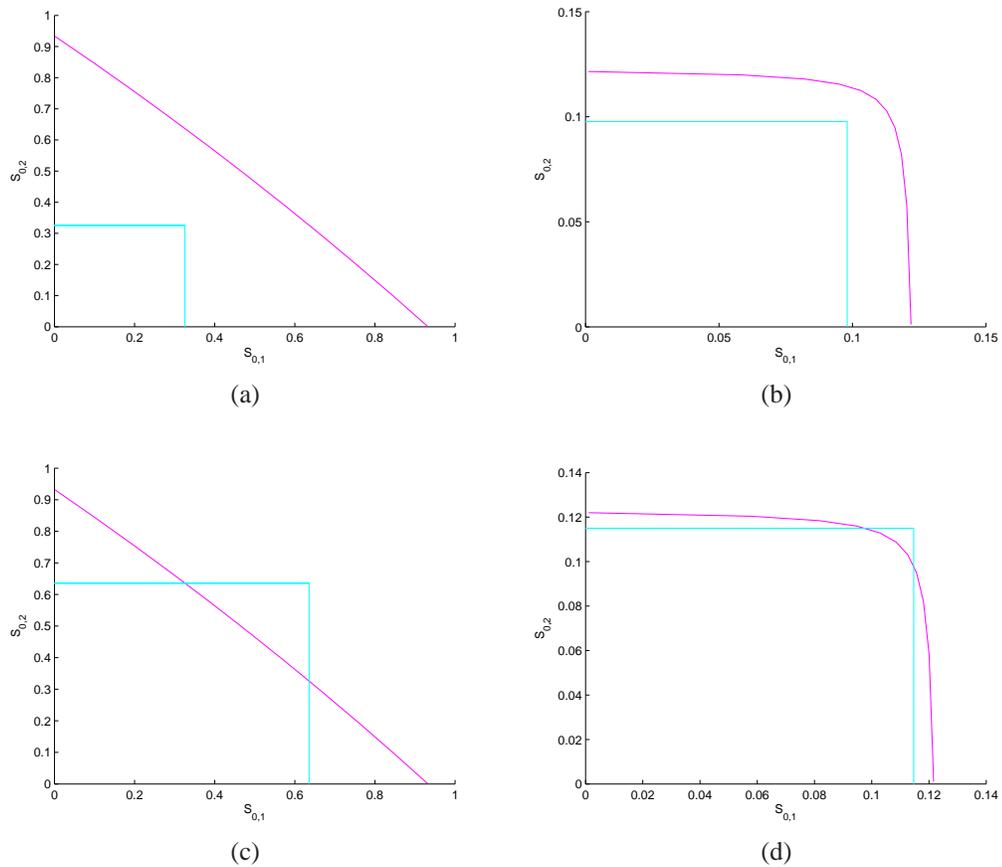


Fig. 7.12: $S_{0,1}$ vs. $S_{0,2}$ for different QoS exponents and different interference levels. for time division with power control (curved line) and treating interference as noise (rectangular regions).

7.4 Energy Efficiency in Fading Relay Channels under Secrecy and QoS Constraints

7.4.1 Channel Model

Fig. 7.15 depicts the two-hop communication link we consider in this section. In this model, destination **D** gets information from source **S** with the help of an intermediate relay node **R**, while the eavesdropper **E** listens both transmissions. Due to half-duplex communication, eavesdropper listens either the source or relay at any given time.

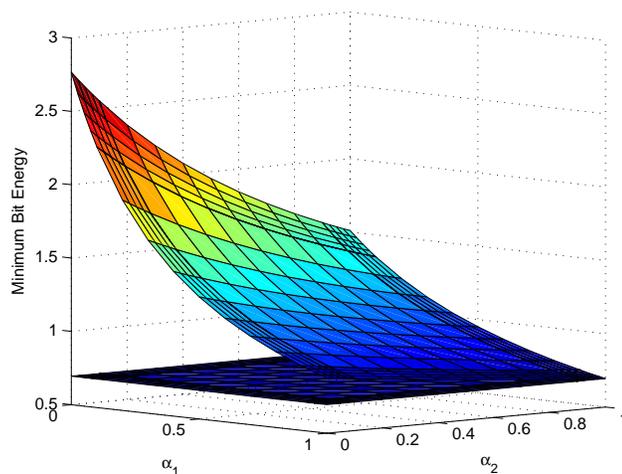


Fig. 7.13: Minimum energy per bit achieved by transmitter 1 vs. time-sharing parameters α_1 and α_2 .

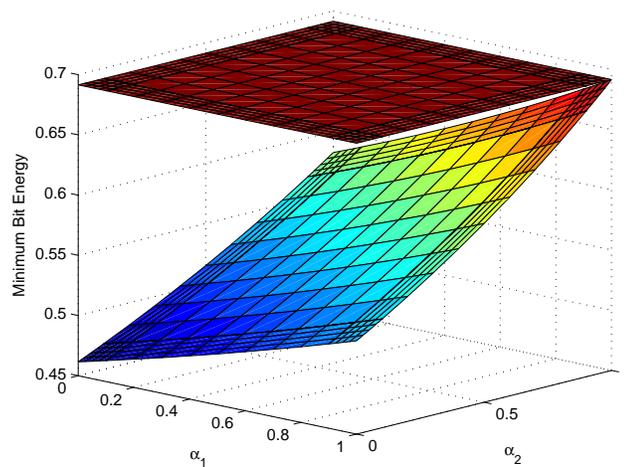


Fig. 7.14: Minimum energy per bit achieved by transmitter 1 vs. time-sharing parameters α_1 and α_2 .

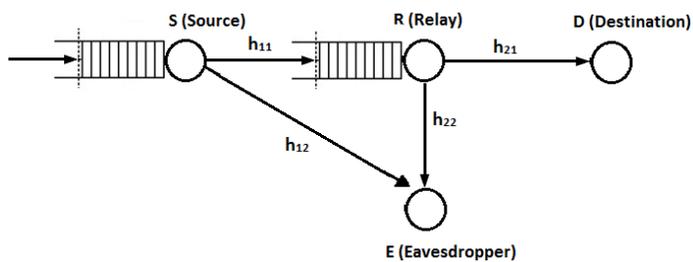


Fig. 7.15: Channel Model

When the source is transmitting, the channel input-output relationships are given as

$$y_r = h_{11}x_s + n_r, \quad (7.70)$$

$$y_{e1} = h_{12}x_s + n_{e1}. \quad (7.71)$$

Above, x_s denotes the input signal from the source. y_r is the received signal at the relay and y_{e1} is the received signal at the eavesdropper. h_{11} and h_{12} denote the fading coefficients in the channels between the source and relay and the source and eavesdropper, respectively. n_r and n_{e1} represent the zero-mean, circularly-symmetric, complex Gaussian noise samples with variances $\mathbb{E}\{|n_r|^2\} = N_0$ and $\mathbb{E}\{|n_{e1}|^2\} = N_{0e}$.

We consider DF relaying. Therefore, the relay decodes the message based on the received signal y_r and re-encodes and forwards it to the destination in the subsequent time interval. The input-output relationships are now given as

$$y_d = h_{21}x_r + n_d, \quad (7.72)$$

$$y_{e2} = h_{22}x_r + n_{e2}. \quad (7.73)$$

Above, x_r denotes the re-encoded input signal from the relay. y_d is the received signal at the destination and y_{e2} is the received signal at the eavesdropper. h_{11} and h_{12} denote the fading coefficients in the corresponding channels. Similarly as above, n_d and n_{e2} are again the zero-mean, circularly-symmetric, complex Gaussian noise components with variances $\mathbb{E}\{|n_d|^2\} = N_0$ and $\mathbb{E}\{|n_{e2}|^2\} = N_{0e}$. We assume that the average energies of the transmitted signals are $\mathbb{E}\{|x_s|^2\} = \mathbb{E}\{|x_r|^2\} = \mathcal{E}$. Hence, the signal-to-noise ratio between legitimate users is defined as

$$\text{SNR} = \frac{\mathbb{E}\{|x_s|^2\}}{\mathbb{E}\{|n_r|^2\}} = \frac{\mathbb{E}\{|x_r|^2\}}{\mathbb{E}\{|n_d|^2\}} = \frac{\mathcal{E}}{N_0}. \quad (7.74)$$

Due to possibly different noise power levels at the legitimate users and eavesdropper, the

signal-to-noise ratio at the eavesdropper is defined as $\text{SNR}_e = \frac{N_0}{N_{0e}}\text{SNR}$. For simplicity, we denote $\gamma = \frac{N_0}{N_{0e}}$.

7.4.2 Preliminaries

Effective Capacity

In the half-duplex system, $\tau_1 = \tau \in (0, 1)$ portion of the time source transmits, while $\tau_2 = 1 - \tau$ portion of the time relay transmits. Therefore, the effective capacities of the source and relay are given by

$$R_E(\text{SNR}, \theta_i) = -\frac{1}{\theta_i} \log_e \mathbb{E}\{e^{-\theta_i \tau_i \lambda_i}\} \text{ for } i = 1, 2. \quad (7.75)$$

If the channel input sequence is an independent and identically distributed (i.i.d.) sequence of Gaussian random variables with zero mean and variance \mathcal{E} , then the instantaneous service rate is

$$\lambda_i = \log_2(1 + \text{SNR}z_{i1}) \text{ for } i = 1, 2. \quad (7.76)$$

On the other hand, in the presence of an eavesdropper, the instantaneous secrecy rate is the service process (with which secrecy can be achieved) and is given by

$$\lambda_i = [\log_2(1 + \text{SNR}z_{i1}) - \log_2(1 + \gamma\text{SNR}z_{i2})]^+ \text{ for } i = 1, 2. \quad (7.77)$$

Remark 17. *We assume that $S - R$ and $R - D$ links are secured individually by transmitting at the rates given in (7.77). Note that as shown in [103], via independent randomization at the source and relay, securing each hop from the eavesdropper guarantees secrecy of the overall communication from the source to destination.*

Energy Efficiency Metrics

In this section, we employ energy per bit as the performance metric of energy efficiency.

In our setup, we define energy per bit as

$$\frac{E_b}{N_0} = \frac{2\mathcal{E}/N_0}{R_E(\theta_1, \theta_2, \text{SNR})} = \frac{2\text{SNR}}{R_E(\theta_1, \theta_2, \text{SNR})} \quad (7.78)$$

where $R_E(\theta_1, \theta_2, \text{SNR})$ is the throughput of a half-duplex relay channel. Note that we have $2\mathcal{E}$ in the numerator since we take into account the energy consumption by both the source transmitter and relay transmitter. In [102], the throughput of the half-duplex fading relay channel under statistical buffer constraints is characterized as follows:

$$\text{Case I } \theta_1 \geq \theta_2 : R_E(\theta_1, \theta_2, \text{SNR}) = -\frac{1}{\theta_1} \log_e \mathbb{E} \left\{ e^{-\theta_1 \tilde{\tau} \lambda_1} \right\}, \quad (7.79)$$

$$\text{Case II } \theta_1 < \theta_2 : R_E(\theta_1, \theta_2, \text{SNR}) = -\frac{1}{\theta_1} \log_e \mathbb{E} \left\{ e^{-\theta_1 \hat{\tau} \lambda_1} \right\}, \quad (7.80)$$

where $\tilde{\tau} = \min \{ \tau_0, \tau^* \}$ and $\hat{\tau} = \min \{ \tau_0, \tau' \}$. τ_0 is given by

$$\tau_0 = \frac{\mathbb{E} \{ \lambda_2(\text{SNR}) \}}{\mathbb{E} \{ \lambda_1(\text{SNR}) \} + \mathbb{E} \{ \lambda_2(\text{SNR}) \}}. \quad (7.81)$$

τ^* is the solution of the equality

$$-\frac{1}{\theta_1} \log_e \mathbb{E} \left\{ e^{-\theta_1 \tau^* \lambda_1} \right\} = -\frac{1}{\theta_2} \log_e \mathbb{E} \left\{ e^{-\theta_2 (1-\tau^*) \lambda_2} \right\}. \quad (7.82)$$

τ' is the solution of the equality

$$-\frac{1}{\theta_1} \log_e \mathbb{E} \left\{ e^{-\theta_1 \tau' \lambda_1} \right\} = -\frac{1}{\theta_1} \left(\log_e \mathbb{E} \left\{ e^{-\theta_2 (1-\tau') \lambda_2} \right\} + \log_e \mathbb{E} \left\{ e^{(\theta_2 - \theta_1) \tau' \lambda_1} \right\} \right). \quad (7.83)$$

In our analysis, we study the minimum energy per bit $\frac{E_b}{N_0}_{\min}$ under QoS constraints,

which can be obtained from [57]

$$\frac{E_b}{N_{0 \min}} = \lim_{\text{SNR} \rightarrow 0} \frac{2\text{SNR}}{R_E(\theta_1, \theta_2, \text{SNR})} = \frac{2}{\dot{R}_E(\theta_1, \theta_2, \text{SNR})}. \quad (7.84)$$

7.4.3 Energy Efficiency Of Two-Hop Wireless Communication

In this section, we analyze the energy efficiency of a two-hop wireless system using the minimum energy per bit. We first consider the case in which no secrecy constraints are imposed. Subsequently, we address the scenario in which information needs to be kept confidential from an eavesdropper.

No Secrecy Constraints

In this case, the system is a simple two-hop wireless channel and the instantaneous transmission rate for both source and relay is given by (7.76). Next, we determine the minimum energy per bit for this scenario.

Theorem 7.4.1. *Assume that SNR is the same for both S – R and R – D links. With block fading assumption, the minimum energy per bit is given by*

$$\frac{E_b}{N_{0 \min}} = \frac{2 (\mathbb{E} \{z_{11}\} + \mathbb{E} \{z_{21}\}) \log_e 2}{\mathbb{E} \{z_{11}\} \mathbb{E} \{z_{21}\}}. \quad (7.85)$$

Proof: Since we have two different throughput expressions according to θ_1 and θ_2 values, we determine the minimum energy per bit for both **Case I** and **Case II**.

When $\theta_1 \geq \theta_2$ the throughput expression is given in (7.79). We take the derivative of the throughput expression with respect to SNR, which is given by

$$\dot{R}_E(\theta_1, \theta_2, \text{SNR}) = \frac{\mathbb{E} \left\{ e^{-\theta_1 \tilde{\tau} \lambda_1(\text{SNR})} \left[\dot{\tilde{\tau}} \lambda_1(\text{SNR}) + \tilde{\tau} \dot{\lambda}_1(\text{SNR}) \right] \right\}}{\mathbb{E} \{ e^{-\theta_1 \tilde{\tau} \lambda_1(\text{SNR})} \}}. \quad (7.86)$$

Note that, $\dot{\tilde{\tau}}$ is the first derivative of the time sharing parameter in terms of SNR. When

SNR $\rightarrow 0$, (7.86) can be simplified to

$$\dot{R}_E(\theta_1, \theta_2, 0) = \tilde{\tau}(0)\mathbb{E}\left\{\dot{\lambda}_1(0)\right\}. \quad (7.87)$$

Since $\tilde{\tau} = \min\{\tau_0, \tau^*\}$, we need to determine $\tau_0(\text{SNR})$ and $\tau^*(\text{SNR})$ as SNR $\rightarrow 0$. We can easily see that

$$\tau_0(0) = \lim_{\text{SNR} \rightarrow 0} \frac{\mathbb{E}\{\lambda_2(\text{SNR})\}}{\mathbb{E}\{\lambda_1(\text{SNR})\} + \mathbb{E}\{\lambda_2(\text{SNR})\}} \quad (7.88)$$

$$= \lim_{\text{SNR} \rightarrow 0} \frac{\mathbb{E}\{\dot{\lambda}_2(\text{SNR})\}}{\mathbb{E}\{\dot{\lambda}_1(\text{SNR})\} + \mathbb{E}\{\dot{\lambda}_2(\text{SNR})\}} \quad (7.89)$$

$$= \frac{\mathbb{E}\{z_{21}\}}{\mathbb{E}\{z_{11}\} + \mathbb{E}\{z_{21}\}}, \quad (7.90)$$

where (7.89) follows by the application of the L'Hopital's rule. In order to find $\tau^*(0)$ we can use the first order approximations of the expressions on both sides of equality in (7.82). Thus, we obtain the equality below

$$\tau^*(0)\mathbb{E}\{z_{11}\} = (1 - \tau^*(0))\mathbb{E}\{z_{21}\}, \quad (7.91)$$

from which we find that $\tau^*(0) = \tau_0(0)$. We use this result in (7.87), and obtain the first derivative as

$$\dot{R}_E(\theta_1, \theta_2, 0) = \frac{\mathbb{E}\{z_{21}\}}{\mathbb{E}\{z_{11}\} + \mathbb{E}\{z_{21}\}}\mathbb{E}\{z_1\}\log_2 e. \quad (7.92)$$

Inserting (7.92) into (7.84), we obtain the minimum energy per bit expression in (7.85).

When $\theta_1 < \theta_2$ and SNR $\rightarrow 0$, we have a similar first derivative expression for the throughput as in (7.87):

$$\dot{R}_E(\theta_1, \theta_2, 0) = \hat{\tau}(0)\mathbb{E}\left\{\dot{\lambda}_1(0)\right\}. \quad (7.93)$$

We know that $\hat{\tau} = \min\{\tau', \tau_0\}$. By using the first order approximations of the both sides of the equality in (7.83), we can determine $\tau'(0)$. More specifically, in the limit as SNR $\rightarrow 0$,

this equality becomes

$$\tau'(0)\mathbb{E}\{z_{11}\} = \frac{\theta_2}{\theta_1}(1 - \tau'(0))\mathbb{E}\{z_{21}\} - \frac{\theta_2 - \theta_1}{\theta_1}\tau'(0)\mathbb{E}\{z_{11}\}, \quad (7.94)$$

from which we conclude that $\tau'(0) = \tau_0(0)$. Therefore, for both cases, first derivative of the throughput and hence minimum energy per bit are the same. ■

Secrecy Constraints

In this case, information needs to be kept confidential. The eavesdropper listens the communication over both $\mathbf{S} - \mathbf{R}$ and $\mathbf{R} - \mathbf{D}$ links. For this case, the minimum energy per bit expression is determined in the following result.

Theorem 7.4.2. *Assume that SNR is the same for both $\mathbf{S} - \mathbf{R}$ and $\mathbf{R} - \mathbf{D}$ links. Further assume that SNR of the eavesdropper is $SNR_e = \gamma SNR$. With block fading assumption, the minimum energy per bit is given by*

$$\frac{E_b}{N_{0min}} = \frac{2(\mathbb{E}\{[z_{11} - \gamma z_{12}]^+\} + \mathbb{E}\{[z_{21} - \gamma z_{22}]^+\}) \log_e 2}{\mathbb{E}\{[z_{11} - \gamma z_{12}]^+\} \mathbb{E}\{[z_{21} - \gamma z_{22}]^+\}}. \quad (7.95)$$

Proof: We first note that the expressions in (7.87) and (7.93) are applicable in the presence of secrecy constraints as well. Hence, although the instantaneous rates have changed, we can easily determine that $\tilde{\tau}$ and $\hat{\tau}$ are still equal to τ_0 at zero SNR. Therefore, considering (7.89) with the rate expression in (7.77), we obtain

$$\tilde{\tau}(0) = \hat{\tau}(0) = \frac{\mathbb{E}\{[z_{21} - \gamma z_{22}]^+\}}{\mathbb{E}\{[z_{11} - \gamma z_{12}]^+\} + \mathbb{E}\{[z_{21} - \gamma z_{22}]^+\}}. \quad (7.96)$$

Combining (7.96) with (7.87) and (7.93) and inserting them into (7.84), we obtain the minimum energy per bit expression given in Theorem 7.4.2.

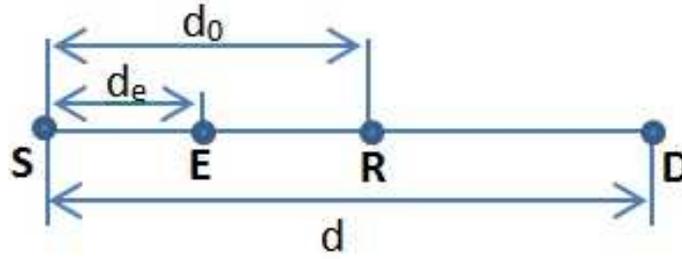


Fig. 7.16: Locations of source (S), eavesdropper (E), relay (R) and destination (D).

7.4.4 Numerical Results and Discussions

In this section, we provide numerical results in which we investigate, in addition to the minimum energy per bit, throughput vs. energy per bit and distance curves. In the numerical computations, we assume for ease of exposition that the nodes are linearly aligned as in Fig. 7.16. SNR level is the same in all channels (i.e., $\gamma = 1$). d denotes the distance between the source and destination, d_0 is the distance between the source and relay, and d_e is the distance between the source and eavesdropper. Without loss of generality, we set $d = 1$. While the analysis is applicable any fading distribution, we consider Rayleigh fading links. In this setting, by selecting the path-loss exponent as 4, the magnitude squares of the fading coefficients, which are exponentially distributed, have the following mean values:

- **S – R link:** $\mathbb{E}\{z_{11}\} = 1/d_0^4$,
- **S – E link:** $\mathbb{E}\{z_{12}\} = 1/d_e^4$,
- **R – D link:** $\mathbb{E}\{z_{21}\} = 1/(d - d_0)^4$,
- **R – E link:** $\mathbb{E}\{z_{22}\} = 1/(d_0 - d_e)^4$.

Fig. 7.17 plots the maximum arrival rates as a function of energy per bit when there is no eavesdropper present. We assume that the relay is located in the middle of the source and destination, meaning that $d_0 = 0.5d$. Our objective in this figure is to compare the impact of different values of QoS exponents (while assuming $\theta_1 = \theta_2$). As expected, the minimum energy per bit, which we have seen to be independent of θ in Theorems 7.4.1

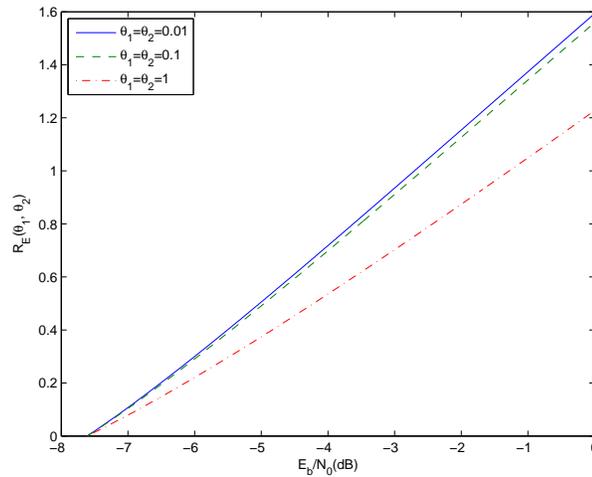


Fig. 7.17: Maximum arrival rate $R_E(\theta_1, \theta_2)$ vs. energy per bit E_b/N_0 without eavesdropper.

and 7.4.2, is not altered at different values of QoS exponents. For this case, the minimum energy per bit is computed as -7.61 dB and we can see all curves intersect at -7.61 dB when SNR is zero. We also observe that θ increases and buffer constraints become more stringent, the required E_b/N_0 increases and hence energy efficiency degrades for any given nonzero arrival rate R_E .

Fig. 7.18 depicts the maximum arrival rates that can be supported while keeping the data secret from an eavesdropper. We assume that eavesdropper is located between source and relay at a distance of $d_e = 0.3d$. The relay is again in the middle, i.e., $d_0 = 0.5d$. We also assume that channels between $\mathbf{R} - \mathbf{D}$ and $\mathbf{R} - \mathbf{E}$ to be independent. Therefore, the correlation coefficient between the fading coefficient of the relay to destination link and that of the relay to eavesdropper link is $\rho_2 = 0$. On the other hand, we assume that $\mathbf{S} - \mathbf{R}$ and $\mathbf{S} - \mathbf{E}$ links can possibly have correlated channel coefficients. Each curve in Fig. 7.18 is plotted for a different value of ρ_1 , which denotes the correlation coefficient of the fading coefficients of source to relay and source to eavesdropper links. We observe that the minimum energy per bit varies with the correlation coefficient. In particular, when the correlation coefficient increases, the energy requirement for the same throughput in-

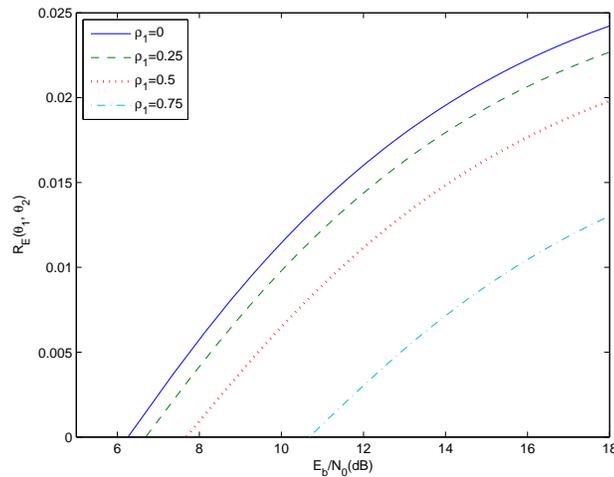


Fig. 7.18: Maximum arrival rate $R_E(\theta_1, \theta_2)$ vs. energy per bit E_b/N_0 in the presence of an eavesdropper when $\theta_1 = 0.1$, $\theta_2 = 0.01$.

creases. Thus, correlation hurts the energy efficiency. The reason behind this phenomenon is that secrecy rates diminish as correlation increases. Indeed, the secrecy rate is zero if the channels are fully correlated, in which case no amount of energy is sufficient for secure transmission of the arriving data. Moreover, comparison between Figs. 7.17 and 7.18 reveals the significant energy costs of secrecy. Note that QoS exponent does not have any effect on the minimum energy per bit. Thus, despite having different QoS constraints in these two figures, the penalty of secrecy on the energy efficiency is clear by just comparing the minimum energy per bit values.

Indeed, we can easily compute the minimum energy per bit formula in (7.95) when $\gamma = 1$ and there is no correlation. First, we need to compute the expectations in the formula.

Lemma 7.4.3. *Assume that w_1 and w_2 are uncorrelated exponential random variables with means μ_1 and μ_2 respectively. We can easily obtain*

$$\mathbb{E} \{ [w_1 - w_2]^+ \} = \frac{\mu_1^2}{\mu_1 + \mu_2}. \quad (7.97)$$

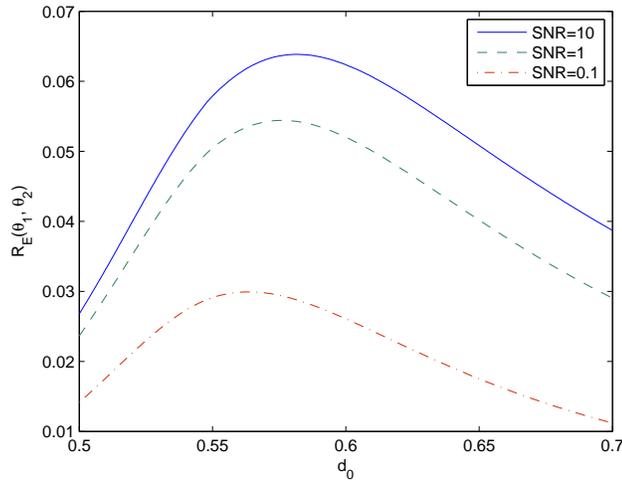


Fig. 7.19: Maximum arrival rate $R_E(\theta_1, \theta_2)$ vs. distance between source and relay d_0 when $\theta_1 = \theta_2 = 0.1$

When we use the formulation (7.97) in (7.95), the minimum energy per bit becomes

$$\frac{E_b}{N_{0 \min}} = 2 \left(d_0^4 + \frac{d_0^8}{d_e^4} + (d - d_0)^4 + \frac{(d - d_0)^8}{(d_0 - d_e)^4} \right) \log_e 2. \quad (7.98)$$

When we insert the distance values we use in Fig. 7.18, we obtain the minimum energy per bit as 6.26 dB, verifying the observation in this figure in the case of no correlation.

It is expected that the locations of the nodes are critical for the arrival rates that the system can support. To investigate the impact of locations, we plot the throughput vs. relay location curves in Figs. 7.19 and 7.20. We analyze the effect of different SNR values on the throughput and best location for the relay in Fig. 7.19. QoS exponents are chosen as $\theta_1 = \theta_2 = 0.1$ and there is no correlation between the channels. As expected, SNR affects the throughput positively. However, we have diminishing returns as SNR increases since eavesdropper shares the same SNR with the relay. Interestingly, the location that gives the highest throughput changes with the SNR as well. When SNR is reduced, the optimal relay location is closer to the source.

The effect of the QoS exponents is addressed in Fig. 7.20. We keep the constraint on the relay buffer (i.e., the relay QoS exponent) fixed while changing the constraint on the

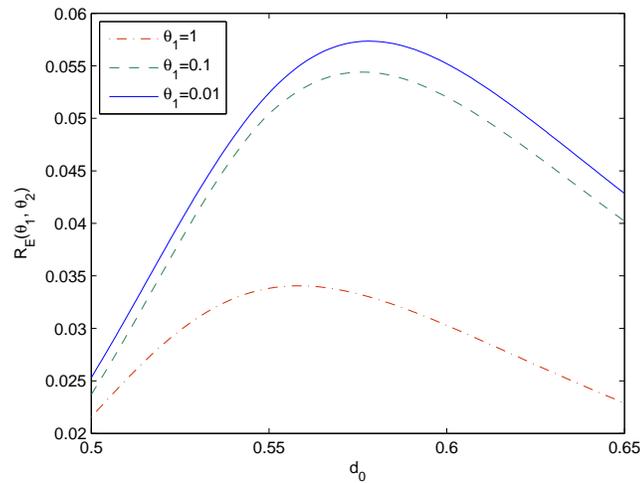


Fig. 7.20: Maximum arrival rate $R_E(\theta_1, \theta_2)$ vs. distance between source and relay d_0 when $\theta_2 = 0.1$ and $\text{SNR} = 1$.

source buffer. We also assume that $\text{SNR} = 1$ and no correlation exists between any pair of channels. It is immediate to see that stricter buffer constraints (i.e., higher values of θ_1) reduce the throughput. The important inference is that when the buffer constraints get stricter, the optimal location of the relay at which the throughput is maximized again gets closer to the source. Hence, low SNR and high θ levels have similar impact on the system. Notice that relay in this case becomes closer to the eavesdropper as well, which tends to adversely affect the security of the $\mathbf{R} - \mathbf{D}$ link. But, this is preferred as the $\mathbf{S} - \mathbf{R}$ link becomes the bottleneck.

Finally, in Fig. 7.21, we plot the minimum energy per bit as a function of d_0 for different values of d_e . Clearly, if the eavesdropper approaches the source or the destination, the minimum energy per bit increases. We further observe that the optimal location of the relay at which the lowest minimum energy per bit is attained changes in the same direction as that of the eavesdropper location. For instance, as the eavesdropper approaches the source, so does the relay. This is again due to the $\mathbf{S} - \mathbf{R}$ link becoming the bottleneck.

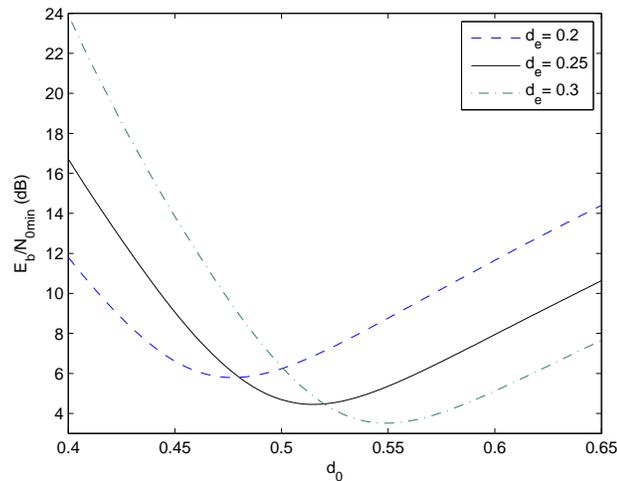


Fig. 7.21: Minimum energy per bit E_b/N_{0min} vs. distance between source and relay d_0 when $\theta_2 = 0.1$ and $\text{SNR} = 1$.

7.5 Energy Efficiency in Cognitive Radio Channels with Markov Arrivals

7.5.1 Channel Sensing

Secondary users are assumed to operate over blocks of m symbols. In each block, channel sensing is performed over the initial duration of n symbols to determine the primary users' activity (i.e., whether the channel is idle or busy). We assume that the primary users' activity remains unchanged over one block duration. However, activity between the blocks is modeled as a Markov chain with two states, denoted by B and I , corresponding to "busy" and "idle" channels, respectively. In state B , channel is occupied by the primary users whereas state I indicates no primary user activity in the channel. In the Markov chain, $P_{i,j}$ represents the transition probability from state i to state j where $i, j \in \{I, B\}$. Note that $\sum_j P_{i,j} = 1$ and the probabilities $P_{B,I}$ and $P_{I,B}$ are denoted by s and q , respectively.

As in [48], we formulate channel sensing as a binary hypothesis-testing problem:

$$\begin{aligned}\mathcal{H}_0 : y_i &= n_i & i &= 1, 2, \dots, n \\ \mathcal{H}_1 : y_i &= w_i + n_i & i &= 1, 2, \dots, n\end{aligned}\quad (7.99)$$

where hypotheses \mathcal{H}_0 and \mathcal{H}_1 describe the absence and presence of primary users, respectively. Above, n_i represents the circularly symmetric, zero-mean, complex background Gaussian noise samples with variance $\mathbb{E}\{|n_i|^2\} = N_0$ and w_i denotes the primary users' faded sum signal at the secondary receiver, which is independent and identically distributed according to circularly symmetric, zero-mean, complex Gaussian with variance $\mathbb{E}\{|w_i|^2\} = \sigma_w^2$. We assume that secondary users employ energy detector which compares the total energy gathered in the sensing duration with a threshold λ , i.e., we have

$$T(y) = \frac{1}{n} \sum_{i=1}^n |y_i|^2 \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\geq}} \lambda. \quad (7.100)$$

The test statistic $T(y)$ above follows χ^2 distribution with $2n$ degrees of freedom. Under this statistical assumption, the false alarm and detection probabilities can be expressed in terms of the regularized Gamma function $P(a, x)$ [62, eq. 6.5.1] as follows:

$$\begin{aligned}P_f &= \Pr\{T(y) > \lambda | \mathcal{H}_0\} = \Pr(\hat{\mathcal{H}}_1 | \mathcal{H}_0) = 1 - P\left(\frac{n\lambda}{N_0}, n\right), \\ P_d &= \Pr\{T(y) > \lambda | \mathcal{H}_1\} = \Pr(\hat{\mathcal{H}}_1 | \mathcal{H}_1) = 1 - P\left(\frac{n\lambda}{N_0 + \sigma_s^2}, n\right).\end{aligned}$$

Above, $P(a, x) = \frac{\gamma(a, x)}{\Gamma(a)}$, where $\gamma(a, x)$ is the lower incomplete Gamma function [62, eq. 6.5.2], and $\Gamma(a)$ is the Gamma function [62, eq. 6.1.1]. Additionally, $\hat{\mathcal{H}}_1$ and $\hat{\mathcal{H}}_0$ denote the busy and idle sensing decisions, respectively.

7.5.2 Cognitive Radio Channel Model

Following channel sensing, data transmission is initiated over a flat-fading channel in the remaining block duration of $(m - n)$ symbols. The transmission power levels are chosen depending on the sensing decision. More specifically, the average power is \bar{P}_1 in the case of channel being detected as busy, and it is \bar{P}_2 in the case of channel being detected as idle. In general, we have $\bar{P}_1 \leq \bar{P}_2$ in order to control the interference on the primary users.

We consider a block-fading channel model in which the fading coefficients remain constant over each block of m symbols. Under these assumptions, the complex channel input-output relation can be described as

$$\mathbf{y} = \begin{cases} h\mathbf{x} + \mathbf{n} & \text{under } \mathcal{H}_0 \\ h\mathbf{x} + \mathbf{n} + \mathbf{w} & \text{under } \mathcal{H}_1 \end{cases} \quad (7.101)$$

where h denotes the circularly-symmetric, complex fading coefficient with finite variance, i.e., $\mathbb{E}\{|h|^2\} < \infty$. Additionally, \mathbf{x} and \mathbf{y} are the complex channel input and output vectors with length $(m - n)$, respectively, and \mathbf{w} and \mathbf{n} again denote the primary users' received faded signal and background Gaussian noise, respectively.

As a result of different channel sensing decisions and the channel's true state, we have four possible scenarios, together with corresponding signal-to-noise ratio expressions and the instantaneous channel capacities listed in the following:

- *Scenario I* : Busy channel is sensed as busy (Correct-detection)

$$\text{SNR}_1 = \frac{\bar{P}_1}{N_0 + \sigma_w^2} \text{ and } C_1 = (m - n) \log(1 + \text{SNR}_1 z).$$

- *Scenario II* : Busy channel is sensed as idle (Miss-detection)

$$\text{SNR}_2 = \frac{\bar{P}_2}{N_0 + \sigma_w^2} \text{ and } C_2 = (m - n) \log(1 + \text{SNR}_2 z).$$

- *Scenario III* : Idle channel is sensed as busy (False-alarm)

$$\text{SNR}_3 = \frac{\bar{P}_1}{N_0} \text{ and } C_3 = (m - n) \log(1 + \text{SNR}_3 z).$$

- *Scenario IV* : Idle channel is sensed as idle (Correct-detection)

$$\text{SNR}_4 = \frac{\bar{P}_2}{N_0} \text{ and } C_4 = (m - n) \log(1 + \text{SNR}_4 z).$$

Above, we have defined $z = |h|^2$ and expressed the instantaneous channel capacities. Actual transmission rates will depend on the sensing decisions $\hat{\mathcal{H}}_0$ and $\hat{\mathcal{H}}_1$. More specifically, under $\hat{\mathcal{H}}_1$ (i.e., in scenarios 1 and 3), the transmission rate is

$$R_1(\text{SNR}_1, z) = (m - n) \log(1 + \text{SNR}_1 z). \quad (7.102)$$

On the other hand, under $\hat{\mathcal{H}}_0$ (i.e., in scenarios 2 and 4), the secondary users send data at the rate

$$R_2(\text{SNR}_4, z) = (m - n) \log(1 + \text{SNR}_4 z). \quad (7.103)$$

7.5.3 Effective Capacity of Cognitive Radio Channels

In this section, we formulate the effective capacity which characterizes the maximum constant arrival rates that can be supported in the presence of buffer constraints through cognitive radio transmissions. Before deriving the effective capacity expression for the cognitive radio channel, we initially construct a state transition model. In scenario 1, transmission rate $R_1(\text{SNR}_1, z)$ equals the channel capacity and in scenario 3 (in which we have false alarm), $R_1(\text{SNR}_1, z)$ is less than the channel capacity due to the fact that secondary users actually do not experience interference from the primary users, i.e., $\text{SNR}_1 < \text{SNR}_3$. Hence, in both scenarios, the channel is in the ON state where reliable transmission is achieved. In scenario 2 (in which we have miss detection), transmission rate $R_2(\text{SNR}_4, z)$ is greater than the channel capacity due to interference caused by the primary users, i.e., $\text{SNR}_4 > \text{SNR}_2$. Thus, the channel is in the OFF state and reliable communication is not attained due to errors. Therefore, the transmission rate is effectively zero and retransmissions are required. Finally, in scenario 4, $R_2(\text{SNR}_4, z)$ is equal to the channel capacity and therefore the chan-

nel is again in the ON state. As a result, we have four states in total in the state transition model for the cognitive radio channel as depicted in Fig. 7.22.

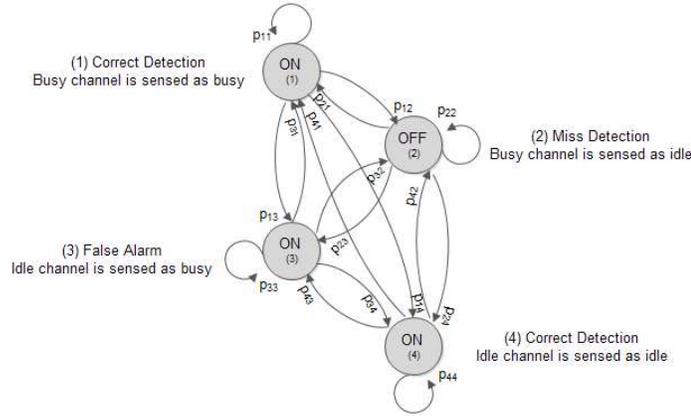


Fig. 7.22: The state-transition model for the cognitive radio channel with four states.

Next, we determine the transition probabilities from state i to state j , denoted by p_{ij} in the figure. Note that the channel is actually busy in the first two states and we see that the transition probabilities from these states to the first state are equal. The channel is actually idle in the last two states and their transition probabilities are the same. Hence, the transition probabilities can be grouped into two with respect to the channel's true state, i.e., being busy or idle. With this observation, the transition probabilities between each state can be derived as

$$\begin{aligned}
 p_{i1} &= (1 - s)P_d & p_{k1} &= qP_d, \\
 p_{i2} &= (1 - s)(1 - P_d) & p_{k2} &= q(1 - P_d), \\
 p_{i3} &= sP_f & p_{k3} &= (1 - q)P_f, \\
 p_{i4} &= s(1 - P_f) & p_{k4} &= (1 - q)(1 - P_f)
 \end{aligned} \tag{7.104}$$

for $i = 1, 2$ and $k = 3, 4$. Above, P_d and P_f denote the detection and false-alarm probabilities, respectively, in channel sensing, and s and q are the transition probabilities in the two-state Markov chain for primary user activity. The 4×4 state transition matrix is denoted by \mathbf{G} where $[\mathbf{G}]_{ij} = p_{ij}$. Note that \mathbf{G} has a rank of 2 due to having only two linearly

independent row vectors.

The effective capacity, which identifies the maximum constant arrival rate that a given service process can support in order to guarantee statistical QoS constraint given in (2.1), is formulated as [6] [10]

$$C_E(\theta) = - \lim_{t \rightarrow \infty} \frac{1}{\theta t} \log_e \mathbb{E}\{e^{-\theta S[t]}\} \triangleq - \frac{\Lambda(-\theta)}{\theta}, \quad (7.105)$$

where $S[t] \triangleq \sum_{j=1}^t R_i[j]$ is the time-accumulated service process and $R_i[j]$ is the discrete-time stationary and ergodic service process. Gartner-Ellis (GE) limit of the service process is defined by $\Lambda(\theta_i) = \lim_{t \rightarrow \infty} \frac{1}{t} \log_e \mathbb{E}\{e^{\theta_i S[t]}\}$. In [9, Chap. 7, Example 7.2.7], it is shown for Markov modulated processes that

$$\frac{\Lambda(\theta)}{\theta} = \frac{1}{\theta} \log_e \mathbb{E}\{\rho(\phi(\theta)\mathbf{G})\} \quad (7.106)$$

Above, \mathbf{G} is the transition matrix of the underlying Markov process, and $\phi(\theta) = \text{diag}(\phi_1(\theta), \dots, \phi_4(\theta))$ is a diagonal matrix whose components are the moment generating functions of the processes with 4 states. In our case, we have

$$\phi(\theta) = \text{diag}\{e^{\theta R_1(\text{SNR}_1)}, 1, e^{\theta R_1(\text{SNR}_1)}, e^{\theta R_2(\text{SNR}_4)}\}.$$

Since the rank of $\phi(\theta)\mathbf{G}$ is 2, spectral radius $\rho(\phi(\theta)\mathbf{G})$ is given by the maximum root of the characteristic polynomial of the matrix $\phi(\theta)\mathbf{G}$. Hence, we can derive the effective capacity expression of the cognitive radio channel as in (7.107) on the next page by combining (7.105), (7.106).

$$C_E(\theta) = - \frac{1}{\theta m} \log_e \mathbb{E} \left\{ \frac{1}{2} \left[\phi_1(\theta)p_{i1} + \phi_2(\theta)p_{i2} + \phi_3(\theta)p_{k3} + \phi_4(\theta)p_{k4} \right] \right. \\ \left. + \frac{1}{2} \sqrt{\left[\phi_1(\theta)p_{i1} + \phi_2(\theta)p_{i2} - \phi_3(\theta)p_{k3} - \phi_4(\theta)p_{k4} \right]^2 + 4(\phi_1(\theta)p_{k1} + \phi_2(\theta)p_{k2})(\phi_3(\theta)p_{i3} + \phi_4(\theta)p_{i4})} \right\} \quad (7.107)$$

7.5.4 Energy Efficiency Metrics

We have defined the average signal-to-noise ratio as

$$\text{SNR}_{\text{avg}} = \frac{\overline{P}_{\text{avg}}}{N_0} = \frac{\text{Pr}\{\hat{\mathcal{H}}_1\}\overline{P}_1 + \text{Pr}\{\hat{\mathcal{H}}_0\}\overline{P}_2}{N_0} \quad (7.108)$$

$$= \frac{\gamma(sP_d + qP_f) + s(1 - P_d) + q(1 - P_f)}{s + q} \frac{\overline{P}_2}{N_0} \quad (7.109)$$

$$= \varphi \frac{\overline{P}_2}{N_0} \quad (7.110)$$

where we have defined $\varphi = \frac{\gamma(qP_d + sP_f) + s(1 - P_f) + q(1 - P_d)}{s + q}$ and $\frac{\overline{P}_1}{\overline{P}_2} = \gamma$. Furthermore, we denote $\xi = \frac{\sigma_w^2}{N_0}$. With the help of these definitions, we can determine from (3.2) the maximum arrival rate $r^*(\text{SNR}_{\text{avg}}, \theta)$ that can be supported in the cognitive radio channel for given SNR_{avg} , QoS exponent θ . Then, the maximum average arrival rate is

$$r_{\text{avg}}^*(\text{SNR}_{\text{avg}}, \theta) = r^*(\text{SNR}_{\text{avg}}, \theta) P_{\text{ON}}. \quad (7.111)$$

In this section, we employ energy per bit as the performance metric of energy efficiency.

In our setup, we define energy per bit as

$$\frac{E_b}{N_0} = \frac{\text{SNR}_{\text{avg}}}{r_{\text{avg}}^*(\text{SNR}_{\text{avg}}, \theta)/m}. \quad (7.112)$$

In our analysis, following the approach in [57], we study the minimum energy per bit and the wideband slope, which is defined as the slope of the spectral efficiency curve at zero spectral efficiency. While the minimum bit energy is a performance measure in the limit as $\text{SNR}_{\text{avg}} \rightarrow 0$, wideband slope has emerged as a tool that enables us to analyze the energy efficiency at low but nonzero SNR_{avg} levels. The minimum energy per bit $\frac{E_b}{N_0}_{\text{min}}$ under QoS

constraints can be obtained from [57]

$$\frac{E_b}{N_{0 \min}} = \lim_{\text{SNR}_{\text{avg}} \rightarrow 0} \frac{\text{SNR}_{\text{avg}}}{r_{\text{avg}}^*(\text{SNR}_{\text{avg}}, \theta)/m} = \frac{1}{\dot{r}_{\text{avg}}^*(0, \theta)/m}. \quad (7.113)$$

At $\frac{E_b}{N_{0 \min}}$, the slope \mathcal{S}_0 of the spectral efficiency versus E_b/N_0 (in dB) curve is defined as [57]

$$\mathcal{S}_0 = \lim_{\frac{E_b}{N_0} \downarrow \frac{E_b}{N_{0 \min}}} \frac{r_{\text{avg}}^*(\text{SNR}_{\text{avg}}, \theta)/m}{10 \log_{10} \frac{E_b}{N_0} - 10 \log_{10} \frac{E_b}{N_{0 \min}}} 10 \log_{10} 2. \quad (7.114)$$

Considering the expression for normalized effective capacity, the wideband slope can be found from

$$\mathcal{S}_0 = -\frac{2(\dot{r}_{\text{avg}}^*(0)/m)^2}{\ddot{r}_{\text{avg}}^*(0)/m} \log_e 2 \quad (7.115)$$

where $\dot{r}_{\text{avg}}^*(0)$ and $\ddot{r}_{\text{avg}}^*(0)$ are the first and second derivatives, respectively, of r_{avg}^* with respect to SNR_{avg} at zero $\text{SNR}_{\text{avg}} \cdot \frac{E_b}{N_{0 \min}}$ and \mathcal{S}_0 provide a linear approximation of the spectral efficiency curve at low spectral efficiencies.

7.5.5 Energy Efficiency in Cognitive Radio Channels with Markov Arrivals

Having formulated the effective bandwidth of two-state Markov arrivals and effective capacity of the cognitive radio channel and having introduced the energy efficiency metrics, we now derive the minimum energy per bit and wideband slope in our setting.

Theorem 7.5.1. *Assume that the source arrival rate is random and follows the described ON-OFF model with the state transition matrix \mathbf{J} given in (2.15), and the cognitive radio channel is characterized by the state transition matrix \mathbf{G} . Then, the minimum energy per bit as a function of the sensing performance, channel conditions, and QoS exponent θ is*

given by

$$\frac{E_b}{N_{0 \min}} = \frac{\varphi \log_e 2}{\frac{m-n}{m} \frac{1}{s+q} \left((qP_d + sP_f) \frac{\gamma}{1+\xi} + s(1-P_f) \right) \mathbb{E}\{z\}} \quad (7.116)$$

where $\varphi = \frac{\gamma(qP_d + sP_f) + s(1-P_f) + q(1-P_d)}{s+q}$ as defined before. Moreover, if we set $s + q = 1$, the wideband slope is given by (7.117) on the next page.

$$\mathcal{S}_0 = \frac{2\left(\tilde{s} \frac{\gamma}{1+\xi} + s(1-P_f)\right)^2 (\mathbb{E}\{z\})^2}{(\tilde{\omega}-2) \frac{\theta m}{\log_e 2} \left(\tilde{s} \frac{\gamma}{1+\xi} + s(1-P_f)\right)^2 (\mathbb{E}\{z\})^2 + \frac{\theta m}{\log_e 2} \left(\tilde{s} \left(\frac{\gamma}{1+\xi}\right)^2 + s(1-P_f)\right) \mathbb{E}\{z^2\} + \frac{m}{m-n} \left(\tilde{s} \left(\frac{\gamma}{1+\xi}\right)^2 + s(1-P_f)\right) \mathbb{E}\{z^2\}} \quad (7.117)$$

In (7.117), we have defined $\tilde{s} = [(1-s)P_d + sP_f]$, $\tilde{\omega} = \varpi \left(\frac{1}{P_{ON}}\right)^2$, and

$$\varpi = \frac{\alpha_{22}^2 + 2(1 - \alpha_{11})}{2(2 - \alpha_{11} - \alpha_{22})} - \frac{[\alpha_{22}(\alpha_{11} + \alpha_{22}) - 2(\alpha_{11} + \alpha_{22} - 1)]^2}{2(2 - \alpha_{11} - \alpha_{22})^3}. \quad (7.118)$$

Proof: We set $s + q = 1$. We first simplify the relationship in (3.2) to

$$\begin{aligned} & \frac{\alpha_{11} + \alpha_{22}e^{r\theta} + \sqrt{(\alpha_{11} + \alpha_{22}e^{r\theta})^2 - 4(\alpha_{11} + \alpha_{22} - 1)e^{r\theta}}}{2} \\ &= \frac{1}{\mathbb{E}\left\{\tilde{s}e^{-\theta R_1(\text{SNR}_1)} + s(1-P_f)e^{-\theta R_2(\text{SNR}_4)} + (1-s)(1-P_d)\right\}}. \end{aligned} \quad (7.119)$$

We denote the right hand side of (7.119) as $\frac{1}{g(\text{SNR}_{\text{avg}})}$. Taking the derivative of both sides of (7.119) with respect to SNR_{avg} , we obtain

$$\begin{aligned} \dot{r}(\text{SNR}_{\text{avg}}) & \left[\frac{\theta \alpha_{22} e^{r\theta}}{2} + \frac{(\alpha_{11} + \alpha_{22} e^{r\theta}) \theta \alpha_{22} - 2(\alpha_{11} + \alpha_{22} - 1) \theta e^{r\theta}}{2 \sqrt{(\alpha_{11} + \alpha_{22} e^{r\theta})^2 - 4(\alpha_{11} + \alpha_{22} - 1) e^{r\theta}}} \right] \\ &= - \frac{\dot{g}(\text{SNR}_{\text{avg}})}{(g(\text{SNR}_{\text{avg}}))^2}. \end{aligned} \quad (7.120)$$

Next, we let $\text{SNR}_{\text{avg}} \rightarrow 0$. Then, we have $g(0) = 1$ and the first derivative of $g(\text{SNR}_{\text{avg}})$ at

$\text{SNR}_{\text{avg}} = 0$ becomes

$$\dot{g}(0) = -\frac{\theta(m-n)}{\varphi \log_e 2} \left(\tilde{s} \frac{\gamma}{1+\xi} + s(1-P_f) \right) \mathbb{E}\{z\}. \quad (7.121)$$

Since the arrival rate $r \rightarrow 0$ when $\text{SNR}_{\text{avg}} \rightarrow 0$, the equality in (7.120) becomes

$$\begin{aligned} \dot{r}(0)\theta & \left[\frac{\alpha_{22}}{2} + \frac{(\alpha_{11} + \alpha_{22})\alpha_{22} - 2(\alpha_{11} + \alpha_{22} - 1)}{2(2 - \alpha_{11} - \alpha_{22})} \right] \\ & = -\frac{\dot{g}(0)}{(g(0))^2} = \frac{\theta(m-n)}{\varphi \log_e 2} \left(\tilde{s} \frac{\gamma}{1+\xi} + s(1-P_f) \right) \mathbb{E}\{z\}, \end{aligned} \quad (7.122)$$

from which we can obtain an expression for $\dot{r}(0)$ and derive

$$\dot{r}_{\text{avg}}^*(0) = \dot{r}(0) \frac{1 - \alpha_{11}}{2 - \alpha_{11} - \alpha_{22}} \quad (7.123)$$

$$= \frac{(m-n)}{\varphi \log_e 2} \left(\tilde{s} \frac{\gamma}{1+\xi} + s(1-P_f) \right) \mathbb{E}\{z\}. \quad (7.124)$$

Plugging in the result in (7.124) into (3.29), we immediately obtain (7.116).

In order to determine the wideband slope, we first take the second derivative of both sides of (7.119) with respect to SNR_{avg} and evaluate them at $\text{SNR}_{\text{avg}} = 0$ as

$$\begin{aligned} \ddot{r}(0)\theta \frac{1 - \alpha_{11}}{2 - \alpha_{11} - \alpha_{22}} + (\dot{r}(0))^2 \theta^2 \eta & = \frac{2(\dot{g}(0))^2 - \ddot{g}(0)g(0)}{(g(0))^3} \\ & = 2(\dot{g}(0))^2 - \ddot{g}(0) \end{aligned} \quad (7.125)$$

where (7.125) follows from the fact that $g(0) = 1$, and η is defined in (3.9). Using the above equality and the expression for $\dot{r}(0)$ obtained from the equality in (7.122), we can easily derive the second derivative of the maximum average arrival rate with respect to SNR_{avg} at $\text{SNR}_{\text{avg}} = 0$ as

$$\ddot{r}_{\text{avg}}^*(0) = \frac{(2 - \tilde{\eta})(\dot{g}(0))^2 - \ddot{g}(0)}{\theta}. \quad (7.126)$$

where $\tilde{\eta} = \eta \left(\frac{1}{P_{\text{ON}}}\right)^2$ and $\ddot{g}(0)$ is given by

$$\ddot{g}(0) = \frac{1}{\varphi^2} \left[\left(\frac{\theta(m-n)}{\log_e 2} \right)^2 + \frac{\theta(m-n)}{\log_e 2} \right] \left(\tilde{s} \left(\frac{\gamma}{1+\xi} \right)^2 + s(1-P_f) \right) \mathbb{E}\{z^2\}.$$

Finally, inserting (7.124) and (7.126) into (3.31), the wideband slope expression in (3.35) is readily obtained. \square

Remark 18. *Minimum energy per bit does not depend on the QoS exponent θ or the burstiness of the source. Interestingly, burstiness of the primary user activity, signal power of the primary user, reliability of channel sensing have an impact on the minimum energy per bit. Increasing s or q when the other parameter is kept fixed results in increased minimum energy per bit. However, when $s + q = 1$, increasing s (i.e., increasing the probability of idle channel) improves the energy efficiency of the secondary user. As expected, higher interference from the primary user decreases the energy efficiency as it decreases the channel capacity for the secondary user when the primary user is active. Additionally, the minimum energy per bit formula depends on the sensing performance via detection and false alarm probabilities P_d and P_f whose impact is analyzed through numerical results below.*

Remark 19. *Although the QoS exponent θ does not have any effect on the minimum energy per bit, it introduces a penalty on the energy efficiency by reducing the wideband slope. The burstiness of the source also degrades the energy efficiency by increasing the value of η and again diminishing the wideband slope. Similarly as in [54], the effect of the source burstiness vanishes in the absence of QoS requirements.*

Next, we present numerical results in which we assume Rayleigh fading. Therefore, fading magnitude-square $z = |h|^2$ has an exponential distribution with $\mathbb{E}\{z\} = 1$. Furthermore, the noise power and the power of the primary users' received faded signal are assumed to be $N_0 = 0.01$ and $\sigma_w^2 = 0.01$, respectively, (i.e., $\xi = \sigma_w^2/N_0 = 1$).

In Fig. 7.23, we plot the maximum average arrival rate vs. energy per bit, $\frac{E_b}{N_0}$, and investigate the effect of source burstiness on the energy efficiency of cognitive radio trans-

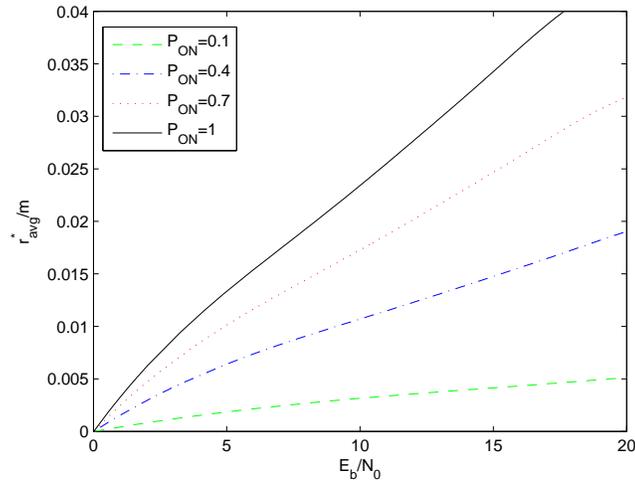


Fig. 7.23: Maximum average arrival rate r_{avg}^*/m vs. energy per bit $\frac{E_b}{N_0}$ with different source burstiness when $\theta = 1$.

missions under QoS constraints. The transition probabilities in the Markov chain of primary user activity are set to $s = 0.9$ and $q = 0.1$. Each block has a duration of $m = 100$ symbols and a duration of $n = 20$ symbols is used for channel sensing. The threshold of the energy detector is chosen as $\lambda = 0.014$. With these values of n and λ , the detection and false-alarm probabilities become $P_d = 0.9235$ and $P_f = 0.0478$. We observe in the figure that as the probability of the ON state of data arrivals diminishes and hence source burstiness increases, wideband slope is decreased, reducing the energy efficiency of the system.

In Fig. 7.24, we plot the minimum energy per bit, $\frac{E_b}{N_{0 \min}}$, as a function of the sensing duration n (while $\lambda = 0.014$) for different values of the transition probabilities s and q in the Markov chain of primary user activity. In the lower subfigure, we provide the corresponding detection and false alarm probabilities again as a function of n . We see that when the primary user activity decreases and as a result the probability of channel being idle increases, smallest $\frac{E_b}{N_{0 \min}}$ is attained at a lower value of n . Indeed, when the channel idle probability is $\Pr\{\mathcal{H}_0\} = \frac{s}{s+q} = 0.9$, the smallest value of $\frac{E_b}{N_{0 \min}}$ is achieved when $n = 2$. However, it is important to note that for this very short sensing duration,

the detection probability is small as well, which can lead to significant interference on the primary user.

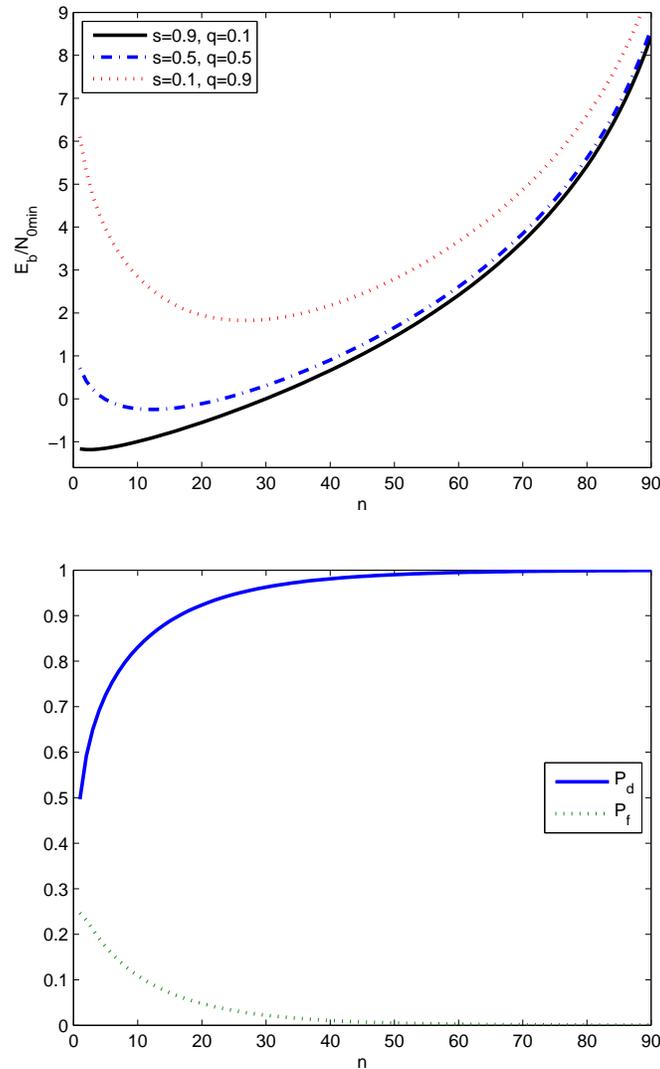


Fig. 7.24: Minimum energy per bit and P_d - P_f vs sensing duration n .

Finally, in Fig. 7.25, we plot the minimum energy per bit, $\frac{E_b}{N_{0\min}}$, and detection and false-alarm probabilities, P_d - P_f , as a function of the sensing threshold λ (while $n = 20$). We notice that for small λ , we have frequent false alarms and therefore sensing reliability is low, leading to high values of $\frac{E_b}{N_{0\min}}$. Increasing the threshold initially improves the sensing reliability and lowers the $\frac{E_b}{N_{0\min}}$. On the other hand, increasing the threshold beyond its optimum value at which $\frac{E_b}{N_{0\min}}$ is minimized, starts degrading sensing reliability by

reducing the detection probability and results in higher $\frac{E_b}{N_0 \min}$ values.

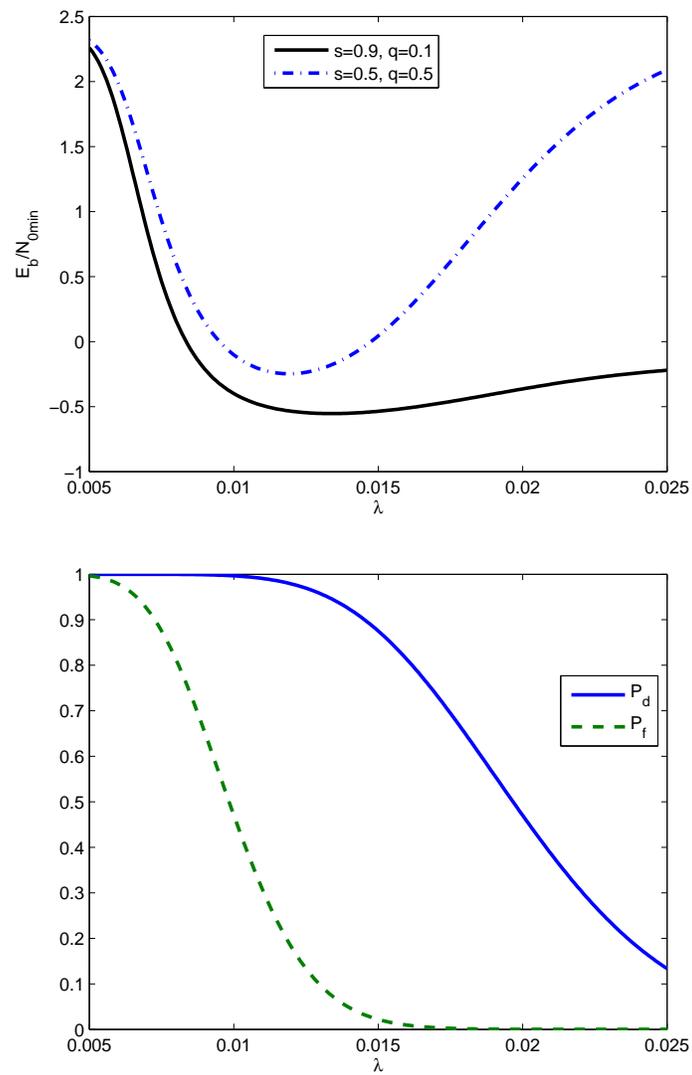


Fig. 7.25: Minimum energy per bit and P_d - P_f vs sensing threshold λ .

CHAPTER 8

FUTURE RESEARCH DIRECTIONS

8.1 QoS-Driven Energy-Efficient Power Control in Cognitive Radio Channels with Markov Arrivals

In Section 7.5.3, we analyzed the energy efficiency in cognitive radio channels with discrete Markov arrivals under QoS constraints. This work can be extended by addressing all of the source arrival models studied in this thesis. Another promising approach is to design optimal power control algorithms for cognitive user by taking the circuit power into account. This problem can focus on maximizing the throughput or energy efficiency of cognitive users in the presence of one or more of the following constraints:

- Maximum average and/or peak power constraint
- Maximum average and/or peak interference constraint
- Constraint on sensing duration
- Constraint on false alarm probability or detection probability
- Constraint on energy efficiency

8.2 Throughput and Optimal Resource Allocation in the Finite Blocklength Regime under QoS Constraints

In this thesis, one of the main assumptions has been that the coding blocklength is sufficiently long for the transmission rates to be accurately approximated by the Shannon capacity and data transmission is reliable. An interesting future research direction is to utilize recent characterizations in the literature and consider the finite blocklength coding regime, in which the data transmission is no longer arbitrarily reliable. Especially when the blocklength is short, the error probability (due to noise) becomes significant even if the rate is selected below the Shannon limit. By formulating the effective capacity in the FBL regime, we can express the throughput, analyze the impact of random arrivals on the performance, and design optimal resource allocation schemes.

APPENDIX A

APPENDIX

A.1 Proof of Theorem 3.1.1:

Using the effective bandwidth formulation in (2.16), we can express (3.2) in the following equivalent form:

$$\frac{1}{\theta} \log_e \left(\frac{p_{11} + p_{22} e^{\lambda\theta} + \sqrt{(p_{11} + p_{22} e^{\lambda\theta})^2 - 4(p_{11} + p_{22} - 1)e^{\lambda\theta}}}{2} \right) = C_E. \quad (\text{A.1})$$

Then, we rewrite the above equality as

$$p_{11} + p_{22} e^{\lambda\theta} + \sqrt{(p_{11} + p_{22} e^{\lambda\theta})^2 - 4(p_{11} + p_{22} - 1)e^{\lambda\theta}} = 2e^{\theta C_E}, \quad (\text{A.2})$$

from which, after moving the first two terms on the left-hand side to the right-hand side and taking the square of both sides, we obtain

$$(p_{11} + p_{22} e^{\lambda\theta})^2 - 4(p_{11} + p_{22} - 1)e^{\lambda\theta} = (2e^{\theta C_E} - p_{11} - p_{22} e^{\lambda\theta})^2. \quad (\text{A.3})$$

Now, by simply exchanging the second term on the left-hand side with the term on the right-hand side, we have

$$(p_{11} + p_{22}e^{\lambda\theta})^2 - (2e^{\theta C_E} - p_{11} - p_{22}e^{\lambda\theta})^2 = 4(p_{11} + p_{22} - 1)e^{\lambda\theta}, \quad (\text{A.4})$$

$$(2p_{11} + 2p_{22}e^{\lambda\theta} - 2e^{\theta C_E}) 2e^{\theta C_E} = 4(p_{11} + p_{22} - 1)e^{\lambda\theta}. \quad (\text{A.5})$$

After further rearrangements, we have

$$(p_{11} + p_{22} - 1 - p_{22}e^{\theta C_E})e^{\lambda\theta} = p_{11}e^{\theta C_E} - e^{2\theta C_E}. \quad (\text{A.6})$$

Solving the equation for λ , we get

$$\lambda^*(\text{SNR}, \theta) = \frac{1}{\theta} \log_e \left(\frac{e^{2\theta C_E(\text{SNR}, \theta)} - p_{11}e^{\theta C_E(\text{SNR}, \theta)}}{1 - p_{11} - p_{22} + p_{22}e^{\theta C_E(\text{SNR}, \theta)}} \right) \quad (\text{A.7})$$

which provides the maximum arrival rate in the ON state. We can now express the maximum arrival rate as $r^*(\text{SNR}, \theta) = P_{\text{ON}}\lambda^*(\text{SNR}, \theta)$ and obtain the expression in (3.6). \square

A.2 Proof of Theorem 3.1.2:

Let us define

$$\psi(\theta) = e^{-\theta C_E(\text{SNR}, \theta)} = \mathbb{E} \left\{ e^{-\frac{\theta}{\log_e 2} \sum_{i=1}^m \log_e(1 + \text{SNR}z_i)} \right\}. \quad (\text{A.8})$$

The following properties of ψ can be verified easily:

$$\psi(0) = 1, \quad (\text{A.9})$$

$$\dot{\psi}(\theta) = \mathbb{E} \left\{ - \sum_{i=1}^m \log_2(1 + \text{SNR}z_i) e^{-\theta \sum_{i=1}^m \log_2(1 + \text{SNR}z_i)} \right\}, \quad (\text{A.10})$$

$$\ddot{\psi}(\theta) = \mathbb{E} \left\{ \left(\sum_{i=1}^m \log_2(1 + \text{SNR}z_i) \right)^2 e^{-\theta \sum_{i=1}^m \log_2(1 + \text{SNR}z_i)} \right\}, \quad (\text{A.11})$$

and

$$\dot{\psi}(0) = - \mathbb{E} \left\{ \sum_{i=1}^m \log_2(1 + \text{SNR}z_i) \right\}, \quad (\text{A.12})$$

$$\ddot{\psi}(0) = \mathbb{E} \left\{ \left(\sum_{i=1}^m \log_2(1 + \text{SNR}z_i) \right)^2 \right\} \quad (\text{A.13})$$

where $\dot{\psi}$ and $\ddot{\psi}$ denote the first and second derivatives of ψ with respect to θ , respectively.

Additionally, we define $r_{\text{avg}}^*(\text{SNR}, \theta)$ as

$$r_{\text{avg}}^*(\text{SNR}, \theta) = \frac{f_1(\theta)}{\theta}. \quad (\text{A.14})$$

Therefore, by applying L'Hopital's rule and letting $\theta \rightarrow 0$, maximum average arrival rate and its slope can be easily found as

$$\lim_{\theta \rightarrow 0} r_{\text{avg}}^*(\text{SNR}, \theta) = \dot{f}_1(0), \quad (\text{A.15})$$

$$\left. \frac{\partial r_{\text{avg}}^*(\text{SNR}, \theta)}{\partial \theta} \right|_{\theta=0} = \frac{\ddot{f}_1(0)}{2}. \quad (\text{A.16})$$

Now, replacing $e^{-\theta C_E(\text{SNR}, \theta)}$ with $\psi(\theta)$ in the expression of $r_{\text{avg}}^*(\text{SNR}, \theta)$ in (3.6), we can express $f_1(\theta)$ as

$$f_1(\theta) = P_{\text{ON}} [\log_e(1 - p_{11}\psi(\theta)) - \log_e((1 - p_{11} - p_{22})\psi^2(\theta) + p_{22}\psi(\theta))]. \quad (\text{A.17})$$

$$\ddot{f}_1(0) = P_{\text{ON}} \frac{d}{d\theta} \left[\frac{-p_{11}\dot{\psi}(0)}{1-p_{11}\psi(0)} - \frac{(1-p_{11}-p_{22})\dot{\psi}(0)}{(1-p_{11}-p_{22})\psi(0)+p_{22}} - \frac{\dot{\psi}(0)}{\psi(0)} \right] \Big|_{\theta=0} \quad (\text{A.21})$$

$$= P_{\text{ON}} \left\{ \left[\frac{-p_{11}}{1-p_{11}} - \frac{1-p_{11}-p_{22}}{1-p_{11}} - 1 \right] \ddot{\psi}(0) + \left[-\frac{p_{11}^2}{(1-p_{11})^2} + \frac{(1-p_{11}-p_{22})^2}{(1-p_{11})^2} + 1 \right] (\dot{\psi}(0))^2 \right\} \quad (\text{A.22})$$

$$= -\ddot{\psi}(0) + (1-\eta) (\dot{\psi}(0))^2 \quad (\text{A.23})$$

$$= -\mathbb{E} \left\{ \left(\sum_{i=1}^m \log_2(1 + \text{SNR}z_i) \right)^2 \right\} + (1-\eta) \left[\mathbb{E} \left\{ \sum_{i=1}^m \log_2(1 + \text{SNR}z_i) \right\} \right]^2 \quad (\text{A.24})$$

Therefore, we derive $\dot{f}_1(\theta)$ when $\theta \rightarrow 0$ as

$$\dot{f}_1(0) = P_{\text{ON}} \left[\frac{-p_{11}\dot{\psi}(0)}{1-p_{11}\psi(0)} - \frac{(1-p_{11}-p_{22})\dot{\psi}(0)}{(1-p_{11}-p_{22})\psi(0)+p_{22}} - \frac{\dot{\psi}(0)}{\psi(0)} \right] \quad (\text{A.18})$$

$$= P_{\text{ON}} \left[\frac{-p_{11}}{1-p_{11}} - \frac{1-p_{11}-p_{22}}{1-p_{11}} - 1 \right] \dot{\psi}(0) \quad (\text{A.19})$$

$$= \mathbb{E} \left\{ \sum_{i=1}^m \log_2(1 + \text{SNR}z_i) \right\}. \quad (\text{A.20})$$

Note that (A.18) follows by taking the first derivative of the expression in (A.17) with respect to θ , and (A.19) is obtained using the property that $\psi(0) = 1$. Finally, (A.20) and hence the result in (3.7) immediately follow from (2.17), (A.12) and (A.15).

Next, we determine the slope of the throughput in (3.8) as the QoS exponent θ approaches zero. For this, we only need to derive the second derivative expression $\ddot{f}_1(0)$, which is done at the top of the next page. (A.21), (A.22) and (A.23) follow from straightforward algebraic steps. Inserting (A.12) and (A.13) into (A.23), we obtain (A.24). Finally, the result in (3.8) follows by combining (A.24) and (A.16). \square

A.3 Proof of Theorem 3.1.3:

In the analysis of the high-SNR slope of the effective capacity, it has been shown in [49] that

$$\begin{aligned} & -\frac{1}{\theta} \log_e \mathbb{E} \left\{ e^{-\theta \log_2(1+\text{SNR}z)} \right\} \\ &= \begin{cases} \frac{1}{\theta \log_2 e} \log_2 \text{SNR} + \mathcal{O}(1) & \text{if } \theta > \frac{1}{\log_2 e} \\ \log_2 \text{SNR} + \mathcal{O}(1) & \text{if } 0 < \theta < \frac{1}{\log_2 e} \end{cases} \end{aligned} \quad (\text{A.25})$$

where z is exponentially distributed with unit mean. If we assume that fading in each block is i.i.d., then the effective capacity expression in (2.10) becomes

$$C_E(\text{SNR}, \theta) = -\frac{1}{\theta} \log_e \mathbb{E} \left\{ e^{-\theta \sum_{i=1}^m \log(1+\text{SNR}z_i)} \right\} \quad (\text{A.26})$$

$$= -\frac{1}{\theta} \log_e \left(\prod_{i=1}^m \mathbb{E} \left\{ e^{-\theta \log(1+\text{SNR}z_i)} \right\} \right) \quad (\text{A.27})$$

$$= -\frac{m}{\theta} \log_e \mathbb{E} \left\{ e^{-\theta \log(1+\text{SNR}z)} \right\}. \quad (\text{A.28})$$

Furthermore, the maximum average arrival rate in (3.6) can be expressed as

$$\begin{aligned} & r_{\text{avg}}^*(\text{SNR}, \theta) \\ &= \frac{P_{\text{ON}}}{\theta} \log_e \left(\frac{e^{2\theta C_E(\text{SNR}, \theta)} (1 - p_{11} e^{-\theta C_E(\text{SNR}, \theta)})}{e^{\theta C_E(\text{SNR}, \theta)} ((1 - p_{11} - p_{22}) e^{-\theta C_E(\text{SNR}, \theta)} + p_{22})} \right) \end{aligned} \quad (\text{A.29})$$

$$= \frac{P_{\text{ON}}}{\theta} \log_e \left(\frac{e^{\theta C_E(\text{SNR}, \theta)} (1 - p_{11} e^{-\theta C_E(\text{SNR}, \theta)})}{(1 - p_{11} - p_{22}) e^{-\theta C_E(\text{SNR}, \theta)} + p_{22}} \right) \quad (\text{A.30})$$

$$\begin{aligned} &= \frac{P_{\text{ON}}}{\theta} \left(\log_e e^{\theta C_E(\text{SNR}, \theta)} + \log_e (1 - p_{11} e^{-\theta C_E(\text{SNR}, \theta)}) \right. \\ &\quad \left. - \log_e ((1 - p_{11} - p_{22}) e^{-\theta C_E(\text{SNR}, \theta)} + p_{22}) \right) \end{aligned} \quad (\text{A.31})$$

$$= \frac{P_{\text{ON}}}{\theta} \log_e e^{\theta C_E(\text{SNR}, \theta)} + \mathcal{O}(1) \quad (\text{A.32})$$

$$= P_{\text{ON}} C_E(\text{SNR}, \theta) + \mathcal{O}(1) \quad (\text{A.33})$$

where (A.30) and (A.31) follow from straightforward algebraic operations and (A.32) is due to the fact that $C_E(\text{SNR}, \theta)$ increases without bound as SNR increases and hence the term $e^{-\theta C_E(\text{SNR}, \theta)}$ vanishes asymptotically in the formulations.

Finally, combining (A.25), (A.28), and (A.33), we immediately obtain the desired result in (3.10) for the cases in which $\theta > 0$. When $\theta = 0$, the result follows from (3.7) in Theorem 3.1.2. \square

A.4 Proof of Theorem 3.1.4:

Using (2.22), we can rewrite (3.2) as

$$(\theta\lambda - (\alpha + \beta) - 2\theta C_E)^2 = (\theta\lambda - (\alpha + \beta))^2 + 4\alpha\theta\lambda \quad (\text{A.34})$$

which can further be simplified to

$$-2\theta C_E(2\theta\lambda - 2(\alpha + \beta) - 2\theta C_E) = 4\alpha\theta\lambda. \quad (\text{A.35})$$

Next, solving for λ , we obtain

$$\lambda^*(\text{SNR}, \theta) = \frac{\theta C_E(\text{SNR}, \theta) + \alpha + \beta}{\theta C_E(\text{SNR}, \theta) + \alpha} C_E(\text{SNR}, \theta). \quad (\text{A.36})$$

Finally, using the expression in (2.24), we derive the maximum average arrival rate given in (3.14). \square

$$\begin{aligned} \dot{f}_2(0) = \lim_{\theta \rightarrow 0} -P_{\text{ON}} \left\{ \left[\frac{-\frac{\dot{\psi}(\theta)}{\psi(\theta)}}{\alpha - \log_e \psi(\theta)} - \frac{(\alpha + \beta - \log_e \psi(\theta)) \left(-\frac{\dot{\psi}(\theta)}{\psi(\theta)}\right)}{(\alpha - \log_e \psi(\theta))^2} \right] \log_e \psi(\theta) \right. \\ \left. + \frac{\alpha + \beta - \log_e \psi(\theta)}{\alpha - \log_e \psi(\theta)} \frac{\dot{\psi}(\theta)}{\psi(\theta)} \right\} \end{aligned} \quad (\text{A.38})$$

$$= \lim_{\theta \rightarrow 0} -P_{\text{ON}} \left[1 + \frac{\alpha\beta}{(\alpha - \log_e \psi(\theta))^2} \right] \frac{\dot{\psi}(\theta)}{\psi(\theta)} \quad (\text{A.39})$$

$$= -P_{\text{ON}} \frac{\alpha + \beta}{\alpha} \dot{\psi}(0) \quad (\text{A.40})$$

$$= \mathbb{E} \left\{ \sum_{i=1}^m \log_2(1 + \text{SNR}z_i) \right\} \quad (\text{A.41})$$

A.5 Proof of Theorem 3.1.5:

Similar as in the Proof of Theorem 3.1.2 in Appendix A.2, we define $r_{\text{avg}}^*(\text{SNR}, \theta) = \frac{f_2(\theta)}{\theta}$ with

$$f_2(\theta) = -P_{\text{ON}} \frac{\alpha + \beta - \log_e \psi(\theta)}{\alpha - \log_e \psi(\theta)} \log_e \psi(\theta). \quad (\text{A.37})$$

Now, we have (A.15) and (A.16) hold with f_1 replaced with f_2 . The remainder of the proof requires only the determination of the first and second derivatives of $f_2(\theta)$ at $\theta = 0$. The first derivative $\dot{f}_2(0)$ is given at the top of the next page in (A.38)-(A.41). Note that (A.38) and (A.39) follow from straightforward algebraic steps, and (A.40) is obtained by noting the property that $\psi(0) = 1$. Finally, (A.41) and hence the result in (3.15) immediately follow from (2.23), (A.12) and (A.15).

Next, we obtain the slope expression in (3.16) in the limit as the QoS exponent θ approaches zero. For this, we characterize the second derivative expression $\ddot{f}_2(0)$ on the next page in (A.42)–(A.45). (A.42), (A.43) are readily obtained and (A.44) is determined by noting that $\psi(0) = 1$. We incorporate (A.12) and (A.13) into (A.44) to obtain (A.45). The result in (3.16) follows by combining (A.45) and (A.16) (with f_1 replaced with f_2). \square

$$\ddot{f}_2(0) = \lim_{\theta \rightarrow 0} -P_{\text{ON}} \frac{d}{d\theta} \left\{ \left[1 + \frac{\alpha\beta}{(\alpha - \log_e \psi(\theta))^2} \right] \frac{\dot{\psi}(\theta)}{\psi(\theta)} \right\} \quad (\text{A.42})$$

$$= \lim_{\theta \rightarrow 0} -P_{\text{ON}} \left\{ \frac{-2\alpha\beta \left(-\frac{\dot{\psi}(\theta)}{\psi(\theta)} \right)}{(\alpha - \log_e \psi(\theta))^3} \left(\frac{\dot{\psi}(\theta)}{\psi(\theta)} \right) + \left[1 + \frac{\alpha\beta}{(\alpha - \log_e \psi(\theta))^2} \right] \left[\frac{\ddot{\psi}(\theta)}{\psi(\theta)} - \left(\frac{\dot{\psi}(\theta)}{\psi(\theta)} \right)^2 \right] \right\} \quad (\text{A.43})$$

$$= -\ddot{\psi}(0) + \left(1 - \frac{2\beta}{\alpha(\alpha + \beta)} \right) \left(\dot{\psi}(0) \right)^2 \quad (\text{A.44})$$

$$= -\mathbb{E} \left\{ \left(\sum_{i=1}^m \log_2(1 + \text{SNR}z_i) \right)^2 \right\} + \left(1 - \frac{2\beta}{\alpha(\alpha + \beta)} \right) \left[\mathbb{E} \left\{ \sum_{i=1}^m \log_2(1 + \text{SNR}z_i) \right\} \right]^2 \quad (\text{A.45})$$

A.6 Proof of Theorem 3.1.7:

We find the maximum average arrival rate $r_{\text{avg}}^*(\text{SNR}, \theta)$ by incorporating (2.26) into (3.2) and expressing (3.2) as

$$\begin{aligned} & ((e^\theta - 1)\lambda - (\alpha + \beta) - 2\theta C_E)^2 \\ &= ((e^\theta - 1)\lambda - (\alpha + \beta))^2 + 4\alpha(e^\theta - 1)\lambda. \end{aligned} \quad (\text{A.46})$$

Similarly as in the proof of Theorem 3.1.4, we can simplify the above equality and solve for the maximum Poisson arrival intensity in the ON state to obtain

$$\lambda^*(\text{SNR}, \theta) = \frac{\theta [\theta C_E(\text{SNR}, \theta) + \alpha + \beta]}{(e^\theta - 1) [\theta C_E(\text{SNR}, \theta) + \alpha]} C_E(\text{SNR}, \theta). \quad (\text{A.47})$$

With this characterization, the maximum average arrival rate is readily obtained from (2.27). \square

A.7 Proof of Theorem 3.1.8:

Employing $f_2(\theta)$ defined in (A.37), we can express the maximum average arrival rate as

$$r_{\text{avg}}^*(\text{SNR}, \theta) = \frac{f_2(\theta)}{e^\theta - 1}. \quad (\text{A.48})$$

Then, the throughput in the limit as θ approaches zero is given by

$$\lim_{\theta \rightarrow 0} r_{\text{avg}}^*(\text{SNR}, \theta) = \lim_{\theta \rightarrow 0} \frac{\dot{f}_2(\theta)}{e^\theta} = \dot{f}_2(0). \quad (\text{A.49})$$

Inserting the result from (A.41) into (A.49), we obtain (3.21). Next, we determine the slope of the throughput when θ approaches zero:

$$\left. \frac{\partial r_{\text{avg}}^*(\text{SNR}, \theta)}{\partial \theta} \right|_{\theta=0} = \lim_{\theta \rightarrow 0} \frac{\dot{f}_2(\theta)}{e^\theta - 1} - \frac{e^\theta f_2(\theta)}{(e^\theta - 1)^2} \quad (\text{A.50})$$

$$= \lim_{\theta \rightarrow 0} \frac{(e^\theta - 1) \dot{f}_2(\theta) - e^\theta f_2(\theta)}{(e^\theta - 1)^2} \quad (\text{A.51})$$

$$= \lim_{\theta \rightarrow 0} \frac{(e^\theta - 1) \ddot{f}_2(\theta) - e^\theta f_2(\theta)}{2(e^\theta - 1)e^\theta} \quad (\text{A.52})$$

$$= \frac{\ddot{f}_2(0)}{2} - \frac{1}{2} \lim_{\theta \rightarrow 0} \frac{f_2(\theta)}{e^\theta - 1} \quad (\text{A.53})$$

$$= \frac{\ddot{f}_2(0)}{2} - \frac{\dot{f}_2(0)}{2}. \quad (\text{A.54})$$

(A.50) follows by taking the derivative of the expression in (A.48) with respect to θ . (A.51) is obtained by simplifying (A.50). We apply L'Hopital's rule on (A.51) to get (A.52) and further simplify it in (A.53). Finally, we obtain (A.54), which we used to derive (3.22) by inserting (A.41) and (A.45) into (A.54).

A.8 Proof of Theorem 3.2.1:

When the arrival rate is fixed, the following equality holds:

$$r_{\text{avg}}^*(\text{SNR}, \theta) = C_E(\text{SNR}, \theta). \quad (\text{A.55})$$

Therefore, in formulas (3.29), (3.31), we can use $\dot{C}_E(0)$ and $\ddot{C}_E(0)$ instead of $\dot{r}_{\text{avg}}^*(0)/m$ and $\ddot{r}_{\text{avg}}^*(0)/m$ respectively, where we have defined $C_E(\text{SNR}, \theta) = C_E(\text{SNR}, \theta)/m$ as the normalized effective capacity. Minimum energy per bit and wideband slope becomes

$$\frac{E_b}{N_{0\text{min}}} = \frac{1}{\dot{C}_E(0)}, \quad (\text{A.56})$$

and

$$\mathcal{S}_0 = -\frac{2(\dot{C}_E(0))^2}{\ddot{C}_E(0)} \log_e 2. \quad (\text{A.57})$$

Thus, we only need to obtain the first and second derivatives of $C_E(\text{SNR}, \theta)$ with respect to SNR at SNR = 0 to determine the minimum energy per bit and wideband slope. We first express the effective capacity given in (2.10) as

$$C_E(\text{SNR}) = -\frac{1}{\theta m} \log_e g(\text{SNR}) \quad (\text{A.58})$$

where we have defined

$$g(\text{SNR}) = \mathbb{E} \left\{ e^{-\frac{\theta}{\log_e 2} \sum_{i=1}^m \log_e(1+\text{SNR}z_i)} \right\}. \quad (\text{A.59})$$

Now, the first and second derivatives of $C_E(\text{SNR})$ with respect to SNR are easily seen to be given by

$$\dot{C}_E(\text{SNR}) = -\frac{1}{\theta m} \frac{\dot{g}(\text{SNR})}{g(\text{SNR})}, \text{ and} \quad (\text{A.60})$$

$$\ddot{C}_E(\text{SNR}) = -\frac{1}{\theta m} \frac{\ddot{g}(\text{SNR})g(\text{SNR}) - [\dot{g}(\text{SNR})]^2}{[g(\text{SNR})]^2}, \quad (\text{A.61})$$

where \dot{g} and \ddot{g} denote the first and second derivatives of the function g with respect to SNR and can be expressed as

$$\dot{g}(\text{SNR}) = -\frac{\theta}{\log_e 2} \mathbb{E} \left\{ \sum_{i=1}^m \frac{z_i}{1 + \text{SNR}z_i} e^{-\frac{\theta}{m \log_e 2} \sum_{i=1}^m \log_e(1 + \text{SNR}z_i)} \right\} \quad (\text{A.62})$$

and

$$\begin{aligned} \ddot{g}(\text{SNR}) &= \frac{\theta}{\log_e 2} \mathbb{E} \left\{ \left(\sum_{i=1}^m \frac{z_i^2}{(1 + \text{SNR}z_i)^2} + \frac{\theta}{\log_e 2} \sum_{i,j=1}^m \frac{z_i z_j}{(1 + \text{SNR}z_i)(1 + \text{SNR}z_j)} \right) \right. \\ &\quad \left. \times e^{-\frac{\theta}{\log_e 2} \sum_{i=1}^m \log_e(1 + \text{SNR}z_i)} \right\}. \end{aligned} \quad (\text{A.63})$$

Then, at $\text{SNR} = 0$, we have

$$\dot{C}_E(0) = \frac{\sum_{i=1}^m \mathbb{E}\{z_i\}}{m \log_e 2} = \frac{\mathbb{E}\{z\}}{\log_e 2} \quad (\text{A.64})$$

and

$$\begin{aligned} \ddot{C}_E(0) &= -\frac{\theta \sum_{i=1}^m \sum_{j=1}^m \text{cov}\{z_i, z_j\} + \log_e 2 \sum_{i=1}^m \mathbb{E}\{z_i^2\}}{m(\log_e 2)^2} \\ &= -\frac{\theta \sum_{i,j=1}^m \sum_{j=1}^m \text{cov}\{z_i, z_j\} + m \log_e 2 \mathbb{E}\{z^2\}}{m(\log_e 2)^2} \end{aligned} \quad (\text{A.65})$$

where we have used the facts that $\sum_{i=1}^m \mathbb{E}\{z_i\} = m\mathbb{E}\{z\}$ and $\sum_{i=1}^m \mathbb{E}\{z_i^2\} = m\mathbb{E}\{z^2\}$ due to our assumption that the fading coefficients and therefore $\{z_i\}$'s are identically distributed.

Plugging the expressions in (A.64) and (A.65) into those in (A.56) and (A.57), we readily obtain the minimum energy per bit and wideband slope expressions in (3.32) and (3.33). ■

A.9 Proof of Theorem 7.5.1:

To show the result, we need to obtain the first and second derivatives of $r_{\text{avg}}^*(\text{SNR})$. We first express the maximum average arrival rate in (3.6) as

$$r_{\text{avg}}^*(\text{SNR}, \theta) = \frac{P_{\text{ON}}}{\theta} \left[\log_e(1 - p_{11}g(\text{SNR})) - \log_e(g(\text{SNR})) - \log_e((1 - p_{11} - p_{22})g(\text{SNR}) + p_{22}) \right] \quad (\text{A.66})$$

where we have used the definition that $e^{\theta C_E(\text{SNR}, \theta)} = \frac{1}{g(\text{SNR})}$ with $g(\text{SNR})$ defined in (A.59).

Taking the first derivative with respect to SNR, we obtain

$$\dot{r}_{\text{avg}}^*(\text{SNR}, \theta) = \frac{P_{\text{ON}}}{\theta} \left[\frac{-p_{11}\dot{g}(\text{SNR})}{1 - p_{11}g(\text{SNR})} - \frac{\dot{g}(\text{SNR})}{g(\text{SNR})} - \frac{(1 - p_{11} - p_{22})\dot{g}(\text{SNR})}{(1 - p_{11} - p_{22})g(\text{SNR}) + p_{22}} \right]. \quad (\text{A.67})$$

Next, we let $\text{SNR} \rightarrow 0$. Since the arrival rate $\lambda \rightarrow 0$ when $\text{SNR} \rightarrow 0$, the equality in (A.67) becomes

$$\dot{r}_{\text{avg}}^*(0, \theta) = \frac{\dot{g}(0)}{\theta} P_{\text{ON}} \left[-\frac{p_{11}}{1 - p_{11}} - 1 - \frac{1 - p_{11} - p_{22}}{1 - p_{11}} \right] \quad (\text{A.68})$$

$$= -\frac{\dot{g}(0)}{\theta} = \frac{1}{\log_e 2} \sum_{i=1}^m \mathbb{E}\{z_i\} = \frac{m\mathbb{E}\{z\}}{\log_e 2} \quad (\text{A.69})$$

where $P_{\text{ON}} = \frac{1-p_{11}}{2-p_{11}-p_{22}}$. Plugging the result in (A.69) into (3.29), we immediately obtain (7.116).

In order to find the wideband slope, we first determine the second derivative of the maximum average arrival rate with respect to SNR and then evaluate it at $\text{SNR} = 0$ as follows:

$$\begin{aligned} \dot{r}_{\text{avg}}^*(0, \theta) &= \frac{\ddot{g}(0)}{\theta} P_{\text{ON}} \left[-\frac{p_{11}}{1-p_{11}} - 1 - \frac{(1-p_{11}-p_{22})}{1-p_{11}} \right] \\ &\quad + \frac{[\dot{g}(0)]^2}{\theta} P_{\text{ON}} \left[-\frac{p_{11}^2}{(1-p_{11})^2} + 1 + \frac{(1-p_{11}-p_{22})^2}{(1-p_{11})^2} \right] \\ &= -\frac{\ddot{g}(0)}{\theta} + (1-\eta) \frac{[\dot{g}(0)]^2}{\theta}. \end{aligned} \quad (\text{A.70})$$

(A.70) follows from the fact that $g(0) = 1$, and η is defined in (3.9). Finally, inserting (A.69) and (A.70) into (3.31), the wideband slope expression in (3.35) is readily obtained.

■

A.10 Proof of Theorem 3.3.3:

We differentiate the maximum average arrival rate expression in (3.14) with respect to SNR and obtain

$$\dot{r}_{\text{avg}}^*(\text{SNR}, \theta) = P_{\text{ON}} \left\{ \frac{2\theta \dot{C}_E(\text{SNR}) C_E(\text{SNR}) + (\alpha + \beta) \dot{C}_E(\text{SNR})}{\theta C_E(\text{SNR}) + \alpha} - \frac{[\theta C_E^2(\text{SNR}) + (\alpha + \beta) C_E(\text{SNR})] \theta \dot{C}_E(\text{SNR})}{(\theta C_E(\text{SNR}) + \alpha)^2} \right\}. \quad (\text{A.71})$$

As $\text{SNR} \rightarrow 0$, we can easily derive

$$\dot{r}_{\text{avg}}^*(0, \theta) = P_{\text{ON}} \frac{\alpha + \beta}{\alpha} \dot{C}_E(0) = \dot{C}_E(0) = \frac{m\mathbb{E}\{z\}}{\log_e 2} \quad (\text{A.72})$$

where we use the facts that $C_E(0) = 0$ and $P_{\text{ON}} = \frac{\alpha}{\alpha + \beta}$. Plugging the result in (A.72) into (3.29), we immediately obtain (6.6).

In order to determine the wideband slope, we additionally take the second derivative of the maximum average arrival rate with respect to SNR and evaluate it at $\text{SNR} = 0$ as

$$\dot{r}_{\text{avg}}^*(0, \theta) = \ddot{C}_E(0) - \frac{2\theta\beta}{\alpha(\alpha + \beta)} \left(\dot{C}_E(0) \right)^2. \quad (\text{A.73})$$

Now, inserting the results in (A.72) and (A.73) into (3.31) and using the formulations in (A.64) and (A.65), we obtain (3.41). ■

A.11 Proof of Theorem 3.2.4:

The proof is rather straightforward after realizing that $r_{\text{avg}}^*(\text{SNR}, \theta)$ of the MMPP source given in (3.20) is equal to the maximum average arrival rate of the Markov fluid source in (3.14) scaled with $\frac{\theta}{e^\theta - 1}$. Therefore, making use of the results in (A.72) and (A.73), we can immediately express the first and second derivatives of $r_{\text{avg}}^*(\text{SNR}, \theta)$ at $\text{SNR} = 0$ as

$$\dot{r}_{\text{avg}}^*(0, \theta) = \frac{\theta \dot{C}_E(0)}{(e^\theta - 1)} = \frac{\theta m \mathbb{E}\{z\}}{(e^\theta - 1) \log_e 2}, \quad (\text{A.74})$$

$$\ddot{r}_{\text{avg}}^*(0, \theta) = \frac{\theta}{(e^\theta - 1)} \left[\ddot{C}_E(0) - \frac{2\theta\beta}{\alpha(\alpha + \beta)} \left(\dot{C}_E(0) \right)^2 \right]. \quad (\text{A.75})$$

Then, the expressions in (3.45) and (3.46) are obtained by plugging (A.74) and (A.75) into (3.29) and (3.31). ■

A.12 Proof of Proposition 5.3.1

First, we define minimum energy per bit for the confidential messages as

$$\frac{E_b}{N_{0 \min, i}} = \frac{\delta_i \Pr(\Gamma_i)}{\dot{r}_{\text{avg}i}^*(0)} \quad (\text{A.76})$$

where $i = 1, 2$. Similarly for the common message, the minimum energy per bit becomes

$$\frac{E_b}{N_{0 \min, 0}} = \frac{(1 - \delta_1) \Pr(\Gamma_1) + (1 - \delta_2) \Pr(\Gamma_2)}{\dot{r}_{\text{avg}0}^*(0)}. \quad (\text{A.77})$$

As the arrival rate is constant, we can use effective capacity as the throughput formula. Therefore, we can exchange $\dot{r}_{\text{avg}}^*(0)$ with $\dot{C}_E(0)$ in the minimum energy per bit equation. For the proofs, we primarily focus on the $g(\text{SNR})$ function that is defined in (A.59).

Now, the first derivative of $C_E(\text{SNR})$ with respect to SNR is easily seen to be given by

$$\dot{C}_{Ei}(\text{SNR}) = -\frac{1}{\theta_i} \frac{\dot{g}_i(\text{SNR})}{g_i(\text{SNR})} \quad (\text{A.78})$$

where $\dot{g}_i(\text{SNR})$ denote the first derivative of the function $g_i(\text{SNR})$ with respect to SNR. It can be readily seen that $g_i(0) = 1$. If we use $f_i(\text{SNR})$ as the instantaneous service rate in nats (i.e. $R_i(\text{SNR}) = f_i(\text{SNR}) \log_e 2$), then we have the relation

$$\dot{g}_i(0) = -\frac{\theta_i}{\log_e 2} \mathbb{E} \left\{ \dot{f}_i(0) \right\} \quad (\text{A.79})$$

where the first derivative expressions $\dot{f}_i(0)$ for $i = 0, 1, 2$ are given by

$$\begin{aligned} \dot{f}_1(0) &= \delta_1 (z_1 - z_2) \mathbf{1}\{z_1 \geq z_2\}, \\ \dot{f}_2(0) &= \delta_2 (z_2 - z_1) \mathbf{1}\{z_1 < z_2\}, \\ \dot{f}_0(0) &= (1 - \delta_1) z_2 \mathbf{1}\{z_1 \geq z_2\} + (1 - \delta_2) z_1 \mathbf{1}\{z_1 < z_2\}. \end{aligned} \quad (\text{A.80})$$

By inserting $\dot{f}_i(0)$ formulations above to (A.79), and then $\dot{g}_i(0)$ to (A.78) consecutively, we obtain the minimum energy per bit expressions for confidential and common messages in (5.17) - (5.19) using (A.76) and (A.77).

A.13 Proof of Proposition 5.3.2

First, we prove the result for the discrete Markov source. We need to obtain the first derivative of $r_{\text{avg},i}^*(\text{SNR})$. Let us rewrite the maximum average arrival rate in (3.6) as

$$r_{\text{avg},i}^*(\text{SNR}, \theta_i) = \frac{P_{\text{ON}}}{\theta_i} \left[\log_e(1 - p_{11}g_i(\text{SNR})) - \log_e(g_i(\text{SNR})) - \log_e((1 - p_{11} - p_{22})g_i(\text{SNR}) + p_{22}) \right] \quad (\text{A.81})$$

where $g_i(\text{SNR})$ is defined in (A.59). Taking the first derivative with respect to SNR, we obtain

$$\dot{r}_{\text{avg},i}^*(\text{SNR}, \theta_i) = \frac{P_{\text{ON}}}{\theta_i} \left[\frac{-p_{11}\dot{g}_i(\text{SNR})}{1 - p_{11}g_i(\text{SNR})} - \frac{\dot{g}_i(\text{SNR})}{g_i(\text{SNR})} - \frac{(1 - p_{11} - p_{22})\dot{g}_i(\text{SNR})}{(1 - p_{11} - p_{22})g_i(\text{SNR}) + p_{22}} \right]. \quad (\text{A.82})$$

When we let $\text{SNR} \rightarrow 0$, the first derivative expression becomes

$$\dot{r}_{\text{avg},i}^*(0) = \frac{\dot{g}_i(0)}{\theta_i} P_{\text{ON}} \left[-\frac{p_{11}}{1 - p_{11}} - 1 - \frac{1 - p_{11} - p_{22}}{1 - p_{11}} \right] = \frac{\dot{f}_i(0)}{\log_e 2} \quad (\text{A.83})$$

where $P_{\text{ON}} = \frac{1 - p_{11}}{2 - p_{11} - p_{22}}$. Note that $g(0) = 1$. Plugging the result in (A.80) and (A.93) into (A.76) and (A.77), we immediately obtain (5.17) - (5.19).

Now, we show the proof for the Markov fluid source. We evaluate the derivative of $r_{\text{avg},i}^*(\text{SNR})$ in (3.14) with respect to SNR and obtain (A.84) given at the top of the next page.

When we let $\text{SNR} \rightarrow 0$, the first derivative expression simplifies to

$$\dot{r}_{\text{avg},i}^*(\text{SNR}, \theta_i) = -\frac{P_{\text{ON}}}{\theta_i} \left\{ \log_e \mathbf{g}_i(\text{SNR}) \frac{d}{d\text{SNR}} \left[\frac{\alpha + \beta - \log_e \mathbf{g}_i(\text{SNR})}{\alpha - \log_e \mathbf{g}_i(\text{SNR})} \right] + \frac{\alpha + \beta - \log_e \mathbf{g}_i(\text{SNR})}{\alpha - \log_e \mathbf{g}_i(\text{SNR})} \frac{\dot{\mathbf{g}}_i(\text{SNR})}{\mathbf{g}_i(\text{SNR})} \right\} \quad (\text{A.84})$$

$$\dot{r}_{\text{avg},i}^*(0) = -\frac{P_{\text{ON}}}{\theta_i} \frac{\alpha + \beta}{\alpha} \dot{\mathbf{g}}_i(0) = \frac{\dot{f}_i(0)}{\log_e 2} \quad (\text{A.85})$$

where $P_{\text{ON}} = \frac{\alpha}{\alpha + \beta}$. Note that $\mathbf{g}(0) = 1$. Plugging the result in (A.80) and (A.85) into (A.76) and (A.77), we immediately obtain (5.17) - (5.19).

A.14 Proof of Proposition 5.3.3

The proof is straightforward as we note that the maximum average arrival rate $r_{\text{avg},i}^*(\text{SNR})$ of discrete-time MMPP source in (5.13) is the scaled version of that of the discrete Markov source in (3.6). The scaling factor is $\frac{\theta_i}{e_i^{\theta_i} - 1}$. The same assertion can be made for the relationship between the maximum average arrival rates of continuous-time MMPP in (A.47) and Markov fluid source in (3.14). Therefore, the minimum energy per bit expressions for discrete-time and continuous-time MMPP sources can be obtained by scaling the formulations in (5.17)-(5.19) with $\frac{e_i^{\theta_i} - 1}{\theta_i}$.

A.15 Proof of Proposition 5.3.4

Let us recall that the wideband slope is given by

$$\mathcal{S}_0 = -\frac{2(\dot{r}_{\text{avg}}^*(0))^2}{\ddot{r}_{\text{avg}}^*(0)} \log_e 2. \quad (\text{A.86})$$

When the arrival rate is constant, we can exchange $r_{\text{avg},i}^*(\text{SNR})$ with $C_{Ei}(\text{SNR})$. For the wideband slope, in addition to the first derivative of the throughput, we also need to obtain the second derivative of the throughput. Second derivatives of the effective capacity at

SNR = 0 can be computed as

$$\ddot{C}_{Ei}(\text{SNR}) = -\frac{1}{\theta_i} \left[\frac{\ddot{g}_i(\text{SNR})}{g_i(\text{SNR})} - \left(\frac{\dot{g}_i(\text{SNR})}{g_i(\text{SNR})} \right)^2 \right]. \quad (\text{A.87})$$

To simplify this equation, we derive the second derivative of $g_i(\text{SNR})$ at SNR = 0 as

$$\ddot{g}_i(0) = -\frac{\theta_i}{\log_e 2} \mathbb{E} \left\{ \ddot{f}_i(0) \right\} + \left(\frac{\theta_i}{\log_e 2} \mathbb{E} \left\{ \dot{f}_i(0) \right\} \right)^2, \quad (\text{A.88})$$

where the second derivative expressions $\ddot{f}_i(0)$ for $i = 0, 1, 2$ are given by

$$\begin{aligned} \ddot{f}_1(0) &= -\delta_1^2 [z_1^2 - z_2^2] \mathbf{1}\{z_1 \geq z_2\}, \\ \ddot{f}_2(0) &= -\delta_2^2 [z_2^2 - z_1^2] \mathbf{1}\{z_1 < z_2\}, \\ \ddot{f}_0(0) &= -(1 - \delta_1^2) z_2^2 \mathbf{1}\{z_1 \geq z_2\} - (1 - \delta_2^2) z_1^2 \mathbf{1}\{z_1 < z_2\}. \end{aligned} \quad (\text{A.89})$$

We insert $\dot{f}_i(0)$ in (A.80) and $\ddot{f}_i(0)$ in (A.89) onto $\dot{g}_i(0)$ in (A.79) and $\ddot{g}_i(0)$ in (A.88) in order to obtain first and second derivative expressions of the effective capacity. By incorporating the (A.78) and (A.87) on (A.86) we obtain wideband slope expression in (5.39).

A.16 Proof of Proposition 5.3.5

In order to find the wideband slope, we need to determine the second derivative of the maximum average arrival rate with respect to SNR. As SNR $\rightarrow 0$ the second derivative

expression is given by

$$\begin{aligned}
\dot{r}_{\text{avg},i}^*(0, \theta_i) &= \frac{\ddot{g}_i(0)}{\theta_i} P_{\text{ON}} \left[-\frac{p_{11}}{1-p_{11}} - 1 - \frac{(1-p_{11}-p_{22})}{1-p_{11}} \right] \\
&\quad + \frac{[\dot{g}_i(0)]^2}{\theta_i} P_{\text{ON}} \left[-\frac{p_{11}^2}{(1-p_{11})^2} + 1 + \frac{(1-p_{11}-p_{22})^2}{(1-p_{11})^2} \right] \\
&= -\frac{\ddot{g}_i(0)}{\theta_i} + (1-\eta) \frac{[\dot{g}_i(0)]^2}{\theta_i} \tag{A.90}
\end{aligned}$$

where η is defined in (3.9). The fact that $g_i(0) = 1$ is taken into account in (A.90). Finally, inserting (A.69) and (A.90) into (A.86), the wideband slope expression in (5.30) is readily obtained.

A.17 Proof of Proposition 5.3.6

In order to find the wideband slope, we need to determine the second derivative of the maximum average arrival rate with respect to SNR. When $\text{SNR} \rightarrow 0$, the second derivative expression is given by

$$\begin{aligned}
\dot{r}_{\text{avg},i}^*(0, \theta_i) &= -\frac{P_{\text{ON}}}{\theta_i} \left\{ \frac{2\beta}{\alpha^2} \dot{g}_i(0) + \frac{\alpha + \beta}{\alpha} (\ddot{g}_i(0) - \dot{g}_i(0)) \right\} \\
&= -\frac{\ddot{g}_i(0)}{\theta_i} + (1-\zeta) \frac{[\dot{g}_i(0)]^2}{\theta_i} \tag{A.91}
\end{aligned}$$

where ζ is defined in (3.17) and we again use the fact that $g_i(0) = 1$. Finally, inserting (A.85) and (A.91) into (A.86), we obtain the wideband slope expression in (5.34).

A.18 Proof of Proposition 5.4.1

First, we define $g(\text{SNR}) = 1 - \mathbf{P}\{\Gamma_1\} (1 - e^{-\theta\lambda})$. For the ON-OFF discrete Markov source the maximum average arrival rate can be rewritten as

$$r_{\text{avg}}^*(\text{SNR}) = \frac{P_{\text{ON}}}{\theta} \log_e \left(\frac{1 - p_{11}g(\text{SNR})}{(1 - p_{11} - p_{22})g^2(\text{SNR}) + p_{22}g(\text{SNR})} \right). \quad (\text{A.92})$$

In order to find the minimum energy per bit and wideband slope, we need to determine the first and second derivatives of the maximum average arrival rate with respect to SNR. Initially, we take the first derivative of maximum average arrival rate and let $\text{SNR} \rightarrow 0$ as follows:

$$\dot{r}_{\text{avg}}^*(0) = -\frac{\dot{g}(0)}{\theta}. \quad (\text{A.93})$$

For this, we also need to characterize the first derivative of $g(\text{SNR})$. We start with the Taylor series expansion of the fixed rate λ in the low-SNR regime:

$$\lambda = \frac{a}{\log_e 2} \text{SNR} + \frac{b}{\log_e 2} \text{SNR}^2 + o(\text{SNR}^2). \quad (\text{A.94})$$

Now, the first derivative of $g(\text{SNR})$ is given by

$$\dot{g}(\text{SNR}) = -\frac{\partial}{\partial \text{SNR}} \mathbf{P}\{\Gamma_1\} (1 - e^{-\theta\lambda}) + \mathbf{P}\{\Gamma_1\} \frac{\partial e^{-\theta\lambda}}{\partial \text{SNR}}. \quad (\text{A.95})$$

As $\text{SNR} \rightarrow 0$, we have $\lambda \rightarrow 0$. Therefore at $\text{SNR} = 0$, we have

$$\dot{g}(0) = \lim_{\text{SNR} \rightarrow 0} \mathbf{P}\{\Gamma_1\} (-\theta) e^{-\theta\lambda} \frac{\partial \lambda}{\partial \text{SNR}}. \quad (\text{A.96})$$

To proceed we need to obtain the probability expression $\mathbf{P}\{\Gamma_1\}$. For independent and

exponentially distributed z_1 and z_2 with unit mean, we can obtain

$$\mathbf{P}\{\Gamma\} = \int_0^\infty e^{-z_2} \int_{2^{\lambda\gamma z_2 + \frac{2^\lambda - 1}{\text{SNR}}}}^\infty e^{-z_1} dz_1 dz_2 \quad (\text{A.97})$$

$$= e^{-\frac{2^\lambda - 1}{\text{SNR}}} \frac{1}{2^{\lambda\gamma + 1}}. \quad (\text{A.98})$$

Now, we can simplify the expression in (A.96) as

$$\dot{g}(0) = -\frac{e^{-a}}{\gamma + 1} \theta \frac{a}{\log_e 2}, \quad (\text{A.99})$$

and inserting this expression into (3.29), we obtain the minimum energy per bit as

$$\frac{E_b}{N_{0 \min}} = -\frac{\theta}{\dot{g}(0)} = \frac{(\gamma + 1) \log_e 2}{ae^{-a}}. \quad (\text{A.100})$$

Finally, we want to determine the smallest possible minimum energy per bit expression. It can be easily seen that the smallest value for the minimum energy per bit is obtained when $a = 1$, leading to the minimum energy per bit expression in (5.44).

In order to find the wideband slope, we first determine the second derivative of the maximum average arrival rate with respect to SNR and then evaluate it at $\text{SNR} = 0$ as follows:

$$\ddot{r}_{\text{avg}}^*(0) = -\frac{\ddot{g}(0)}{\theta} + (1 - \eta) \frac{[\dot{g}(0)]^2}{\theta}. \quad (\text{A.101})$$

Note that, η is defined in (3.9). The first derivative of $g(\text{SNR})$ at $\text{SNR} = 0$ is given by (A.96), and the second derivative is

$$\ddot{g}(0) = \lim_{\text{SNR} \rightarrow 0} 2 \frac{\partial \mathbf{P}\{\Gamma_1\}}{\partial \text{SNR}} \frac{\partial e^{-\theta\lambda}}{\partial \text{SNR}} + \mathbf{P}\{\Gamma_1\} \frac{\partial e^{-\theta\lambda}}{\partial \text{SNR}^2} \quad (\text{A.102})$$

$$= \frac{e^{-a}}{\gamma + 1} \frac{\theta}{\log_e 2} \left[a^3 + \frac{2a^2\gamma}{\gamma + 1} + \frac{\theta a^2}{\log_e 2} + 2b(a - 1) \right]. \quad (\text{A.103})$$

The wideband slope expression can be determined inserting the first and second derivative expressions in (A.93) and (A.101) into (3.31):

$$\mathcal{S}_0 = \frac{2 \frac{(\dot{g}(0))^2}{\theta^2}}{\frac{\ddot{g}(0)}{\theta} + \frac{\eta-1}{\theta} (\dot{g}(0))^2} \log_e 2 \quad (\text{A.104})$$

$$= \frac{1}{\frac{\theta(\eta-1)}{2 \log_e 2} + \frac{\theta}{\log_e 2 e^{-a}} + \frac{a+1}{e^{-a}} + \frac{2b(a-1)}{a^2 e^{-a}}}. \quad (\text{A.105})$$

Since the wideband slope is defined as the slope at the minimum energy per bit, we set $a = 1$. Note that with this choice, parameter b vanishes as $2b(a - 1) \rightarrow 0$ in (A.105). Thus, we obtain the formulation in (5.45).

A.19 Proof of Proposition 5.4.2

The maximum average arrival rate of Markov fluid source can be rewritten as

$$r_{\text{avg}}^*(\text{SNR}) = -\frac{P_{\text{ON}}}{\theta} \left[1 + \frac{\beta}{\alpha - \log_e(\mathbf{g}(\text{SNR}))} \right] \log_e(\mathbf{g}(\text{SNR})). \quad (\text{A.106})$$

By taking the first derivative of the expression in (A.106) and letting $\text{SNR} \rightarrow 0$, we obtain the following:

$$\dot{r}_{\text{avg}}^*(0) = -\frac{P_{\text{ON}}}{\theta} \left[1 + \frac{\beta}{\alpha} \right] \dot{g}(0) = -\frac{\dot{g}(0)}{\theta}. \quad (\text{A.107})$$

By combining (A.107) with (A.99) as $a \rightarrow 1$, and inserting into (3.29), we obtain the minimum energy per bit given in (5.47).

Now, we take the second derivative of the maximum average arrival rate with respect to SNR and then evaluate it as $\text{SNR} \rightarrow 0$

$$\ddot{r}_{\text{avg}}^*(0) = -\frac{P_{\text{ON}}}{\theta} \left[\left(\frac{2\beta}{\alpha^2} - 1 - \frac{\beta}{\alpha} \right) (\dot{g}(0))^2 + \left(1 + \frac{\beta}{\alpha} \right) \ddot{g}(0) \right] \quad (\text{A.108})$$

$$= -\frac{\ddot{g}(0)}{\theta} + (1 - \zeta) \frac{[\dot{g}(0)]^2}{\theta}. \quad (\text{A.109})$$

Note that, ζ is defined in (3.17). We derive the wideband slope expression by using (A.107), (A.109) and (3.31). Again, since the wideband slope is defined at the minimum energy per bit, we set $a = 1$ and obtain the formulation in (5.48).

REFERENCES

- [1] M. Peterson, "ICT at 10% of global electricity consumption?" <http://www.vertatique.com/ict-10-global-energy-consumption>, September 2013.
- [2] D. Feng, C. Jiang, G. Lim, L. J. Cimini, Jr., G. Feng, and G. Y. Li, "A survey of energy-efficient wireless communications," *IEEE Commun. Surveys & Tutorials*, vol. 15, no. 1, pp. 167–178, 2013.
- [3] G. Li, Z. Xu, C. Xiong, C. Yang, S. Zhang, Y. Chen and S. Xu, "Energy efficient wireless communications: tutorial, survey, and open issues," *IEEE Wireless Communications*, vol. 18, no. 6, pp. 28-35, Dec. 2011.
- [4] S. Tanwir and H. Perros, "A survey of VBR video traffic models," *IEEE Commun. Surveys & Tutorials*, vol. 15, no. 4, pp. 1778–1802, 2013.
- [5] Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update, 2013-2018. Available online at <http://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/mobile-white-paper-c11-520862.html>, June 2017
- [6] C.-S. Chang, "Stability, queue length, and delay of deterministic and stochastic queuing networks," *IEEE Trans. Auto. Control*, vol. 39, no. 5, pp. 913-931, May 1994.
- [7] C.-S. Chang, "Effective bandwidth in high-speed digital networks," *IEEE J. Sel. Areas Commun.*, vol.13, pp. 1091-1100, Aug. 1995.

- [8] C.-S. Chang and T. Zajic, "Effective bandwidths of departure processes from queues with time varying capacities," Proc. of IEEE INFOCOM, pp. 1001-1009, 1995.
- [9] C.-S. Chang, *Performance Guarantees in Communication Networks*. Springer-Verlag, 2000.
- [10] D. Wu and R. Negi "Effective capacity: a wireless link model for support of quality of service," *IEEE Trans. Wireless Commun.*, vol.2,no. 4, pp.630-643. July 2003.
- [11] A. I. Elwalid and D. Mitra, "Effective bandwidth of general Markovian traffic sources and admission control of high speed networks," *IEEE/ACM Trans. on Networking.*, vol. 1, no. 3, pp. 329-343, Jun. 1993.
- [12] G. Kesidis, J. Walrand, and C.-S. Chang, "Effective bandwidths for multiclass Markov fluids and other ATM sources," *IEEE/ACM Trans. on Networking.*, vol. 1, no.4, pp. 424-428, Aug. 1993.
- [13] M. Ozmen and M. C. Gursoy, "Energy-efficient power control in fading channels with Markovian sources and QoS constraints," to appear in the *IEEE Trans. Commun.*, 2016.
- [14] M. Ozmen and M. C. Gursoy, "Energy-efficient power control policies in fading channels with Markov arrivals and QoS constraints" Proc. of IEEE Global Conference on Signal and Information Processing (GlobalSIP), 2013.
- [15] M. Ozmen and M. C. Gursoy, "Energy efficiency of fixed-rate transmissions with markov arrivals under queueing constraints,"*IEEE Commun. Letters*, vol. 18, no. 4, pp. 608-611, Apr. 2014.
- [16] M. Ozmen and M. C. Gursoy, "Throughput regions of multiple-access fading channels with markov arrivals and QoS constraints,"*IEEE Wireless Commun. Letters*, vol. 2, no. 5, pp. 499-502, Oct. 2013.

- [17] Q. Du and X. Zhang, "Statistical QoS provisioning for wireless unicast/multicast of multi-layer video streams," *IEEE J. Sel. Areas Commun.*, vol. 28, no. 3, pp. 420–433, Apr. 2010.
- [18] M. C. Gursoy, D. Qiao, and S. Velipasalar, "Analysis of energy efficiency in fading channels under QoS constraints," *IEEE Trans. Wireless Commun.*, vol. 8, no. 8, pp. 4252–4263, Aug. 2009.
- [19] L. Musavian and T. Le-Ngoc, "Energy-efficient power allocation for delay-constrained systems," Proc. of the IEEE GLOBECOM, 2012.
- [20] G. Ru, H. Li, L. Liu, Z. Hu, and Y. Gan, "Energy efficiency of hybrid cellular with heterogeneous QoS provisions," *IEEE Commun. Letters*, vol. 18, no. 6, pp. 1003–1006, June 2014.
- [21] L. Liu, Y. Yang, J. F. Chamberland, and J. Zhang, "Energy-efficient power allocation for delay-sensitive multimedia traffic over wireless systems," *IEEE Trans. Vehicular Techn.*, vol. 63, no. 5, pp. 2038–2047, June 2014.
- [22] A. Helmy, L. Musavian, and T. Le-Ngoc, "Energy-efficient power adaptation over a frequency-selective fading channel with delay and power constraints," *IEEE Trans. Wireless Commun.* vol. 12, no. 9, pp. 4529–4541, Sept. 2013.
- [23] C. Xiong, G. Y. Li, Y. Liu, Y. Chen and S. Xu, "Energy-Efficient Design for Downlink OFDMA with Delay-Sensitive Traffic," *IEEE Trans. Wireless Commun.* vol. 12, no. 6, pp. 3085 - 3095, June 2013.
- [24] M. Ozmen and M. C. Gursoy, "Wireless throughput and energy efficiency with random arrivals and statistical queueing constraints," *IEEE Trans. Inform. Theory*, vol. 62, no. 3, pp. 1375–1395, March 2016.

- [25] L. Musavian, and Q. Ni, "Effective capacity maximization with statistical delay and effective energy efficiency requirements," *IEEE Trans. Wireless Commun.* vol. 14, no. 7, pp. 3824 - 3835, July 2015.
- [26] C. She, C. Yang and L. Liu, "Energy-efficient resource allocation for MIMO-OFDM systems serving random sources with statistical QoS requirement", *IEEE Trans. Commun.*, vol. 63, no. 11, pp. 4125-4141, Nov. 2015.
- [27] L. Liu, P. Parag, J. Tang, W. Y. Chen and J. F. Chamberland, "Resource allocation and quality of service evaluation for wireless communication systems using fluid models," *IEEE Trans. Inform. Theory*, vol. 53, no. 5, pp. 1767-1777. May 2007.
- [28] J. Tang and X. Zhang, "Quality-of-service driven power and rate adaptation for multichannel communications over wireless links," *IEEE Trans. Wireless Commun.*, vol. 6, no. 12, pp. 4349-4360, Dec. 2007.
- [29] W. Cheng, X. Zhang and H. Zhang, "Heterogeneous statistical QoS provisioning over 5G wireless full-duplex networks," *Proc. of IEEE INFOCOM*, pp. 55-63, 2015.
- [30] X. Zhang, J. Tang, "Power-delay tradeoff over wireless networks," *IEEE Trans. Communications*, vol. 61, no. 9, pp. 3673–3684, Sep. 2013.
- [31] S. Boyd, L. Vandenberghe, *Convex Optimization*, Cambridge University Press, 2004.
- [32] D. Qiao, M. C. Gursoy, and S. Velipasalar, "Effective capacity of two-hop wireless communication systems," *IEEE Trans. Inform. Theory*, vol. 59, no. 2, pp. 873-885, Feb. 2013.
- [33] Z. Hasan, H. Boostanimehr, and V. K. Bhargava, "Green cellular networks: A survey, some research issues and challenges," *IEEE Commun. Surveys & Tutorials*, vol. 13, no. 4, pp. 524–540, 2011.

- [34] R. L. Cruz, "A calculus for network delay, part I: Network elements in isolation," *IEEE Trans. Inf. Theory*, vol. 37, pp. 114-131, Jan. 1991.
- [35] R. L. Cruz, "A calculus for network delay, part II: Network analysis," *IEEE Trans. Inf. Theory*, vol. 37 pp. 132-141, Jan. 1991.
- [36] R. L. Cruz, "Quality of service guarantees in virtual circuit switched networks," *IEEE J. Select. Areas Commun.*, vol. 13 pp. 1048-1056, Aug. 1995.
- [37] Y. Jiang and Y. Liu, *Stochastic Network Calculus*. Springer-Verlag, 2008.
- [38] F. Kelly, "Notes on effective bandwidths," *Stochastic Networks: Theory and Applications*, Royal Statistical Society Lecture Notes Series, 4., pp: 141-168, Oxford University Press, 1996.
- [39] C. Courcoubetis and R. Weber, "Effective bandwidth for stationary sources," *Prob. Eng. Inf. Sci.*, 9, pp. 285-296, 1995.
- [40] A. L. Stolyar, "Large deviations of queues sharing a randomly time-varying server," *Queueing Systems*, vol. 59, pp. 1-35, 2008.
- [41] V.J. Venkataramanan and X. Lin, "On wireless scheduling algorithms for minimizing the queue overflow probability," *IEEE/ACM Trans. Networking*, vol. 18, no. 3, pp. 788-801, June 2010.
- [42] B. Sadiq and G. de Veciana, "Large deviations sum-queue optimality of a radial sum-rate monotone opportunistic scheduler," *IEEE Trans. Inform. Theory*, vol. 56, no. 7, pp. 3395-3412, July 2010.
- [43] D. Wu and R. Negi, "Downlink scheduling in a cellular network for quality-of-service assurance," *IEEE Trans. Veh. Technol.*, vol.53, no.5, pp. 1547-1557, Sep. 2004.

- [44] J. Tang and X. Zhang, "Quality-of-Service Driven Power and Rate Adaptation over Wireless Links," *IEEE Trans. Wireless Commun.*, vol. 6, no. 8, pp. 3058-3068, Aug. 2007.
- [45] J. Tang and X. Zhang, "Cross-layer-model based adaptive resource allocation for statistical QoS guarantees in mobile wireless networks," *IEEE Trans. Wireless Commun.*, vol. 7, pp. 2318-2328, June 2008.
- [46] S. Shakkottai "Effective capacity and QoS for wireless scheduling," *IEEE Trans. Automatic Control*, vol.53, no. 3, pp. 749-761, April 2008.
- [47] L. Musavian and S. Aissa, "Effective capacity of delay-constrained cognitive radio in Nakagami fading channels," *IEEE Trans. Wireless Commun.* vol. 9, no. 3, pp. 1054-1062, March 2010.
- [48] S. Akin and M. C. Gursoy, "Effective capacity analysis of cognitive radio channels for quality of service provisioning," *IEEE Trans. Wireless Commun.*, vol. 9, no. 11, pp. 3354-3364, Nov. 2010.
- [49] M. C. Gursoy, "MIMO wireless communications under statistical queueing constraints," *IEEE Trans. Inform. Theory*, vol. 57, no. 9, pp. 5897-5917, Sept. 2011.
- [50] D. Qiao, M. C. Gursoy, and S. Velipasalar, "Transmission strategies in multiple access fading channels with statistical QoS constraints," *IEEE Trans. Inform. Theory*, vol. 58, no. 3, pp. 1578-1593, March 2012.
- [51] B. Soret, M. C. Aguayo-Torres, and J. T. Entrambasaguas, "Capacity with explicit delay guarantees for generic sources over correlated Rayleigh channel," *IEEE Trans. Wireless Commun.*, vol. 9, no. 6, pp. 1901–1911, June 2010.

- [52] Q. Du and X. Zhang, "Statistical QoS provisioning for wireless unicast/multicast of multi-layer video streams," *IEEE J. Sel. Areas Commun.*, vol. 28, no. 3, pp. 420–433, Apr. 2010.
- [53] M.C. Gursoy, "MIMO wireless communications under statistical queueing constraints," *IEEE Trans. Inform. Theory*, vol. 57, no. 9, pp.5897-5917 Sept. 2011.
- [54] M. Ozmen, M. C. Gursoy, "Impact of channel and source variations on the energy efficiency under QoS constraints," *IEEE ISIT*, pp. 806-810, 2012.
- [55] A. Lozano, A. M. Tulino, and S. Verdù, "High-SNR power offset in multiantenna communication," *IEEE Trans. Inform. Theory*, vol. 51, no. 12, pp. 4134-4151, Dec. 2005.
- [56] J.-W. So, "Performance analysis of VoIP services in the IEEE 802.16e OFDMA system with inband signaling," *IEEE Trans. Veh. Technol.*, vol. 57, no. 3, pp. 1876–1886, May 2008.
- [57] S. Verdú, "Spectral efficiency in the wideband regime," *IEEE Trans. Inform. Theory*, vol. 48, no. 6 pp. 1319-1343. Jun. 2002.
- [58] A. El Gamal and Y.-H. Kim, *Network Information Theory*. Cambridge University Press, 2012.
- [59] D. Tse and S. Hanly, "Multi-access fading channels—part I: polymatroid structure, optimal resource allocation and throughput capacities," *IEEE Trans. Inf. Theory*, vol. 44, no. 7, pp. 2796–2815, Nov. 1998.
- [60] D. Qiao, M. C. Gursoy, and S. Velipasalar, "Transmission strategies in multiple-access fading channels with statistical QoS constraints," *IEEE Trans. Inf. Theory*, vol. 58, no. 3, pp. 1578–1593. Mar. 2012.

- [61] E. M. Yeh, "Fundamental performance limits in cross-layer wireless optimization: throughput, delay, and energy," *Foundations and Trends in Commun. and Inf. Theory*, vol. 9, no. 1, pp. 1–112, 2012.
- [62] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*. 9th ed., New York: Dover, 1970.
- [63] M. Ozmen and M. C. Gursoy, "Throughput and energy efficiency under queueing and secrecy constraints," Proc. of the Asilomar Conference on Signals, Systems and Computers, Nov. 2012
- [64] M. Ozmen, G. Ozcan and M. C. Gursoy, "Energy efficiency in cognitive radio channels with Markov arrivals" Proc. of IEEE First International Black Sea Conference on Communications and Networking (BlackSeaCom), 2013.
- [65] A. D. Wyner, "The wire-tap channel," *Bell Syst. Tech. J.*, vol. 54, pp. 1355–1367, Oct. 1975
- [66] I. Csiszár and J. Körner, "Broadcast channels with confidential messages," *IEEE Trans. Inform. Theory*, vol. 3, pp. 339–348, May 1978.
- [67] P. K. Gopala, L. Lai, and H. El Gamal, "On the secrecy capacity of fading channels," *IEEE Trans. Inf. Theory*, vol. 54, no. 10, pp. 4687–4698, Oct. 2008.
- [68] Y. Liang, L. Lai, H. V. Poor, S. Shamai (Shitz), "A broadcast approach for fading wiretap channels," *IEEE Trans. Inform. Theory*, vol. 60, no. 2, pp. 842–858, Feb. 2014.
- [69] H. Jeon, N. Kim, J. Choi, H. Lee, J. Ha, "Bounds on secrecy capacity over correlated ergodic fading channels at high SNR," *IEEE Trans. Inform. Theory*, vol. 57, no. 4, pp. 1975–1983, Apr. 2011.

- [70] P. Baracca, N. Laurenti, S. Tomasin, "Physical layer authentication over MIMO fading wiretap channels," *IEEE Trans. Wireless Commun.*, vol. 11, no. 7, pp. 2564-2573, July. 2012.
- [71] J. Li, A. P. Petropulu, "Ergodic secrecy rate for multiple-antenna wiretap channels with Rician fading," *IEEE Trans. Inform. For. and Sec.*, vol. 6, no.3, pp. 861-867, Sept. 2011.
- [72] M. C. Gursoy "Secure communication in the low-SNR regime," *IEEE Trans. Commun.*, vol. 60, no. 4, pp. 1114-1123, Apr. 2012.
- [73] C. Comaniciu, H. V. Poor, "On energy-secrecy trade-offs for Gaussian wiretap channels," *IEEE Trans. Inform. For. and Sec.*, vol. 8, no.2, pp. 314-323, Feb. 2013.
- [74] H. Zhang, Y. Huang, S. Li and L. Yang, "Energy-efficient precoder design for MIMO wiretap channels," *IEEE Commun. Letters*, vol. 18, no.9, pp. 1559-1562, Sept. 2014.
- [75] D. W. K. Ng, E. S. Lo and R. Schober, "Energy-efficient resource allocation for secure OFDMA systems," *IEEE Trans. Vehic. Tech.*, vol. 61, no.6, pp. 2572 - 2585, July 2012.
- [76] A. Kalantari, S. Maleki, G. Zheng, S. Chatzinotas and B. Ottersen, "Joint power control in wiretap interference channels," *IEEE Trans. Wireless Commun.*, vol. 14, no.7, pp. 3810 - 3823, July 2015.
- [77] X. Chen, L. Lei, "Energy-efficient optimization for physical layer security in multi-antenna downlink networks with QoS guarantee," *IEEE Commun. Letters*, vol. 17, no.4, pp. 637-640, Apr. 2013.
- [78] X. Zhu, B. Yang, C. Chen, L. Xue, X. Guan and F. Wu, "Cross-layer scheduling for OFDMA-based cognitive radio systems with delay and security constraints," *IEEE Trans. Veh. Tech.*, vol. 64, no. 12, pp. 5919-5934, Dec. 2015.

- [79] X. Zhu, B. Yang and X. Guan, "Cross-layer scheduling with secrecy demands in delay-aware OFDMA network," *IEEE Wireless Communications and Networking Conference (WCNC)*, Shanghai, 2013, pp. 1339-1344.
- [80] A. El Shafie and N. Al-Dhahir, "On secure communications over a wiretap channel with fixed-rate transmission: Protocol design and queueing analysis," *IEEE Wireless Commun. Letters*, vol. 4, no. 4, pp. 453-456, Aug. 2015.
- [81] A. El Shafie and N. Al-Dhahir, "Secure communications in the presence of a buffer-aided wireless-powered relay with self-energy recycling," *IEEE Wireless Communications Letters*, vol. 5, no. 1, pp. 32-35, Feb. 2016.
- [82] A. El Shafie, T. Q. Duong and N. Al-Dhahir, "QoS-based design for security enhancement in TDMA-based wireless networks," *IEEE Globecom Workshops*, Washington D.C., 2016
- [83] Z. Mao, C. E. Koksal and N. B. Shroff, "Achieving full secrecy rate with low packet delays: An optimal control approach," *IEEE Journal on Selected Areas in Commun.*, vol. 31, no. 9, pp. 1944-1956, Sept. 2013.
- [84] K. Khalil, O. O. Koyluoglu, H. E. Gamal and M. Youssef, "Opportunistic secrecy with a strict delay constraint," *IEEE Trans. on Commun.*, vol. 61, no. 11, pp. 4700-4709, Nov. 2013.
- [85] Y. Liang, H. V. Poor, and S. Shamai (Shitz), "Information theoretic security," *Found. Trends Commun. Inform. Theory*, vol. 5, no. 4-5, pp. 355-580, 2009.
- [86] R. Liu and W. Trappe, *Securing Wireless Communications at the Physical Layer*, Springer-Verlag, 2010.
- [87] M. Block and J. Barros, *Physical-Layer Security: From Information Theory to Security Engineering*, Cambridge University Press, 2011.

- [88] R. Bassily, *et al.*, "Cooperative security at the physical layer: A summary of recent advances," *IEEE Signal Process. Mag.*, vol. 30, no. 5, pp. 16-28, Sept. 2013.
- [89] Y.-W. P. Hong, P.-C. Lan, and C.-C. J. Kuo, "Enhancing physical-layer secrecy in multiantenna wireless systems: An overview of signal processing approaches," *IEEE Signal Process. Mag.*, vol. 30, no. 5, pp. 29-40, Sept. 2013.
- [90] A. Yener and S. Ulukus, "Wireless physical-layer security: Lessons learned from information theory," *Proc. of the IEEE*, vol. 103, no. 10, pp. 1814-1825, Oct. 2015.
- [91] Y. Zou, J. Zhu, X. Wang, and L. Hanzo, "A survey on wireless security: Technical challenges, recent advances, and future trends," *Proc. of the IEEE*, vol. 104, no. 9, pp. 1727–1765, Sep. 2016.
- [92] D. Qiao, M. C. Gursoy, and S. Velipasalar "Secure wireless communication and optimal power control under statistical queueing constraints," *IEEE Trans. Inform. Forensics and Security*, vol. 6, no. 3, pp. 628–639, Sep. 2011.
- [93] A. Goldsmith, *Wireless Communications*, 1st ed. Cambridge, U.K.: Cambridge Univ. Press, 2005.
- [94] L. Li and A. Goldsmith, "Capacity and optimal resource allocation for fading broadcast channels - part I: ergodic capacity," *IEEE Trans. Inform. Theory*, vol. 47, no. 3, pp. 1083-1102, Mar. 2001.
- [95] D. Qiao, M. C. Gursoy and S. Velipasalar, "Achievable throughput regions of fading broadcast and interference channels under QoS constraints," *IEEE Trans. Communications*, vol.61, no. 9, pp.3730-3740, Sept. 2013.
- [96] A. El Gamal and Y. Kim, "Lecture Notes on Network Information Theory," arXiv:1001.3404v4 [cs.IT]. Jun.2010.

- [97] L. Sankar, X. Shang, E. Erkip, and H.V. Poor, "Ergodic fading interference channels: Sum-capacity and separability," *IEEE Trans. Inform. Theory*, vol. 57, no. 5, pp. 2605 - 2626, May 2011.
- [98] R. Bassily, E. Ekrem, X. He, E. Tekin, J. Xie, M. Bloch, S. Ulukus and A. Yener, "Cooperative security at the physical layer: A summary of recent advances," *IEEE Signal Process. Mag.*, vol. 30, no. 5, pp. 16-28, Sep. 2013.
- [99] L. Dong, Z. Han, A. Petropulu, and H. V. Poor, "Improving wireless physical layer security via cooperating relays," *IEEE Trans. Signal Process.*, vol. 58, no. 3, pp. 1875-1888, Mar. 2010.
- [100] R. Bassily and S. Ulukus, "Secure communication in multiple relay networks through decode-and-forward strategies," *J. Commun. Netw.*, vol. 14, no. 4, pp. 352-363, Aug. 2012.
- [101] E. Biglieri, A. Goldsmith, L. J. Greenstein, N. Mandayam, and H. V. Poor, *Principles of Cognitive Radio*, Cambridge University Press, 2012.
- [102] D. Qiao, M. C. Gursoy and S. Velipasalar, "Effective capacity of two-hop wireless communication systems," *IEEE Trans. Inform. Theory*, vol.59, no. 2, pp.873-885. Feb. 2013.
- [103] O. O. Koyluoglu , C. E. Koksall and H. El Gamal "On secrecy capacity scaling in wireless networks," *IEEE Trans. Inform. Theory*, vol. 58, no. 5, pp.3000-3015 May 2012.

VITA

NAME OF AUTHOR: Mustafa Ozmen

PLACE OF BIRTH: Antalya, Turkey

GRADUATE AND UNDERGRADUATE SCHOOLS ATTENDED:

SYRACUSE UNIVERSITY, Syracuse, New York; August 2011 – January 2017

BILKENT UNIVERSITY, Ankara, Turkey; 2007 – 2011

DEGREES AWARDED:

B.Sc. Electrical and Electronics Engineering (2011), BILKENT UNIVERSITY