# A Binary Channel with Additive Bursty Noise Based on a Finite Queue

by

**Libo Zhong**

A project submitted to the
Department of Mathematics and Statistics
in conformity with the requirements for
the degree of
Master of Science

Queen's University
Kingston, Ontario, Canada
November, 2000

## Acknowledgments

I wish to express my sincere appreciation to my supervisor, Dr. Fady Alajaji, and co-supervisor Dr. Glen Takahara, for their valuable suggestions and expert guidance throughout this study.

My special thanks are due to Hongyan Kuai and other members in the Communication Group for suggesting to me to attend Queen's and for their help in my computer programming tasks.

I am very grateful to my parents for reluctantly allowing me to travel this long distance and for their constant love, support and patience.

Finally, I would like to thank Dr. Fady Alajaji, Dr. Glen Takahara, the Department of Mathematics and Statistics and the School of Graduate Studies for their financial assistance over the past year.

# Abstract

In this work, we introduce a binary communication channel with memory. The noise process of the channel is based on a finite queue. More specifically, we consider the channel in two cases: a uniform queue-based mode where we experiment on the cells of the queue with equal probability, and a non-uniform queue-based mode where we experiment on the cells of the queue with different probabilities.

The statistical properties of the uniform queue-based channel are first investigated. The resulting channel noise is a stationary and ergodic Markov process of order $K$. The state of the channel is a one-step Markov process with $2^K$ states. Expressions for the channel capacity, the stationary distribution, the gap frequency and the burst frequency of the Markovian noise process are presented in terms of $K$. In the case of the non-uniform queue-based channel, we have no closed-form expression for the noise stationary distribution; hence, only numerical results are given.

Next, the capacity, burst frequency and average burst duration of the uniform and the non-uniform queue-based channel with memory are compared with those of the finite-memory contagion channel and the Gilbert-Elliott channel. It is shown (both numerically and analytically) that, surprisingly, the uniform queue-based channel and the finite-memory contagion channel have an identical block transition probability when they have the same memory, bit error rate and correlation coefficient; hence, they have identical capacities and burst frequencies. The non-uniform queue-based channel has larger capacities than the uniform case when the queue probability $q_1 < \frac{1}{K}$, and it has smaller capacities than the uniform case when $q_1 > \frac{1}{K}$. When $q_1 \to 1$, the non-uniform case converges to the uniform case with memory $K = 1$. The non-uniform queue-based channel has lower burst frequencies than the uniform channel for low correlation coefficients, and it has higher burst frequencies for high correlation coefficients.

# Contents

# List of Tables

# List of Figures

# Chapter 1

# Introduction

In most real-world communications channels, noise distortion may produce errors in a bursty fashion; i.e., errors occur in clusters or bunches separated by fairly long error-free segments of data. This phenomenon is commonly know as "memory"[7]. In the quest to develop models that adequately represent real channel behavior and that are mathematically tractable, we introduce a binary channel with additive bursty noise based on a finite queue. It offers an interesting alternative to the Gilbert model and others.

In this chapter we present a literature review of the articles upon which our research is based. We then specify the main contributions of this project. Finally, we outline the general flow of the project.

## 1.1 Literature Review

Gilbert [6] initiated the study of finite-state Markov models for channels with memory by proposing a two-state model. Elliott [3] suggested a modification to Gilbert's model by introducing a parameter, which denotes the probability of correct reception when the channel is in the good state. In [8], Mushkin and Bar-David introduced a method for calculating the capacity of the Gilbert-Elliott channel; they defined two Markov random processes, $\{q_l^*\}$ and $\{q_l\}$, and gave their initial distribution and transition probabilities. Also for the Gilbert-Elliott channel, Pimentel presented an equation in matrix form to calculate the probability of an error sequence in his Ph.D. thesis [10].

The Polya urn model and the Polya process are models for contagion where every accident effectively increases the probability of future accidents [4]. A detailed study of the Polya-contagion channel and the finite-memory contagion channel is given in [2] by Alajaji and Fuja.

## 1.2 Contribution

The contributions of this project are as follows:

- we calculate the statistical properties of the binary process $\{Z_i\}$ generated by the finite queue experiment;
- we obtain the capacity of the channel;
- we obtain the noise gap and burst process statistics for the uniform queue-based channel with memory.

As well, C++ programs perform the following tasks:

- the computation of the capacity for the uniform and non-uniform queue-based channels, the finite-memory contagion channel and the Gilbert-Elliott channel;
- the computation of the burst frequency for the uniform and non-uniform queue-based channel, the finite-memory contagion channel and the Gilbert-Elliott channel.

A matlab program performs the following task:

- the determination of the stationary distribution of the non-uniform queue-based channel with memory.

## 1.3 Project Overview

This project is organized in the following manner.

Chapter 2 presents background material needed to understand the arguments presented later. We start by introducing the finite-memory contagion channel [2] and the Gilbert-Elliott channel [3, 6], since we will later compare them to our proposed queue-based channel. Since we study the gap and the burst process of the channels, we also introduce the definitions of the gap process in Section 2.3.

In Chapter 3, we introduce a binary communication channel with additive bursty noise based on a finite queue. More specifically, we consider the channel in two cases: a uniform queue-based channel with memory where we experiment on the cells of the queue with equal probability, and a non-uniform queue-based channel with memory where we experiment on the cells of the queue with different probabilities. We begin by investigating the statistical properties of the uniform queue-based channel. The resulting channel noise is a stationary and ergodic Markov process of order $K$. The state of the channel is a one-step Markov process with $2^K$ states. We study the stationary distribution of the Markovian noise process and give our results in terms of $K$. We then derive the capacity and the gap and the burst statistics of the channel in terms of $K$. In the case of the non-uniform queue-based channel, we have no closed-form expression for the noise stationary distribution; hence, we only numerically calculate the capacity.

In Chapter 4, we compare the capacity, noise burst frequency and average burst duration of the uniform and the non-uniform queue-based channels with memory with those of the finite-memory contagion channel and the Gilbert-Elliott channel. We show (both numerically and analytically) that, surprisingly, the uniform queue-based channel and the finite-memory contagion channel have an identical block transition probability when they have the same memory, bit error rate and correlation coefficient; hence, they have identical capacities and burst frequencies. The non-uniform case is even more interesting. It has one more parameter $q_1$ which denotes the probability of selecting the first cell of the queue. When $q_1 < \frac{1}{K}$, the capacity values for the non-uniform case are larger than those for the uniform case. When $q_1 > \frac{1}{K}$, the capacity values for the non-uniform case are lower than those for the uniform case. The burst frequencies for the non-uniform case are lower than the uniform case for low correlation coefficients and larger for high correlation coefficients.

Finally, Chapter 5 summarizes the project.

# Chapter 2

# Preliminaries: Previous Binary Channels with Memory

In this chapter, we introduce two previous models for binary channels with memory; we will be comparing them with our proposed channel model in Chapter 4.

## 2.1 Finite-Memory Contagion Channel

The following comes from [2].

Consider a discrete, binary, additive noise communication channel; i.e., a channel for which the $i$th output $Y_i \in \{0, 1\}$ is the modulo-two sum of the $i$th input $X_i \in \{0, 1\}$ and the $i$th noise symbol $Z_i \in \{0, 1\}$. More succinctly, $Y_i = X_i \oplus Z_i$, for $i = 1, 2, 3, \cdots$ (see Fig. 2.1).

We assume that the input and noise sequences are independent of each other. The noise sequence $\{Z_i\}_{i=1}^{\infty}$ is drawn according to the following urn scheme: an urn originally contains $T$ balls, of which $R$ are red and $S$ are black ($T = R + S$). At the $j$th draw, $j = 1, 2, \cdots$, we select a ball from the urn and replace it with $1 + \Delta$ balls of the same

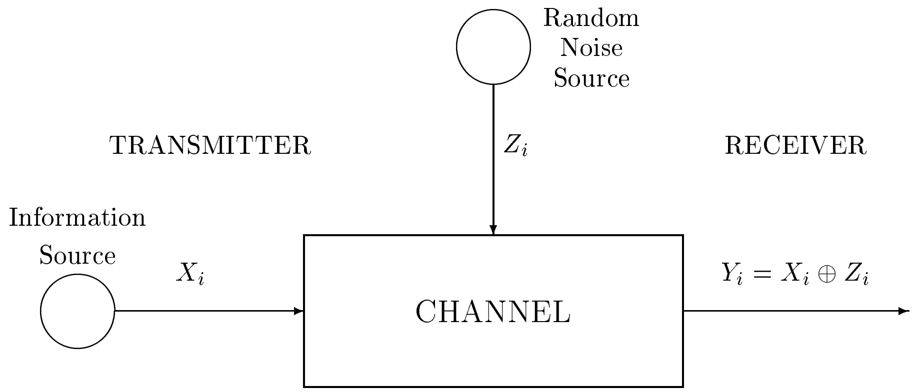



Figure 2.1: Digital communication channel.

color ($\Delta > 0$); then $M$ draws later - after the $(j+M)$th draw - we retrieve from the urn $\Delta$ balls of the color picked at time $j$. Let $\rho = R/T$ , $\sigma = 1 - \rho = S/T$ and $\delta = \Delta/T$. Then the noise process $\{Z_i\}$ corresponds to the outcomes of the draws from the urn, where

$$Z_i = \begin{cases} 1, & \text{if the } i^{th} \text{ drawn ball is red,} \\ 0, & \text{if the } i^{th} \text{ drawn ball is black.} \end{cases}$$

### 2.1.1 The distribution of the noise

The noise process $\{Z_i\}_{i=M+1}^{\infty}$ is a Markov process of order $M$. For an input block $\underline{X} = [X_1, X_2, \cdots, X_n]$ and an output block $\underline{Y} = [Y_1, Y_2, \cdots, Y_n]$, the block transition probability of the resulting binary channel is as follows.

- For block length $n \le M$, the block transition probability of this channel is

  $$P(\underline{Y} = \underline{y} | \underline{X} = \underline{x}) = \frac{\rho(\rho+\delta)\cdots[\rho+(d-1)\delta]\sigma(\sigma+\delta)\cdots[\sigma+(n-d-1)\delta]}{(1+\delta)(1+2\delta)\cdots[1+(n-1)\delta]}, \quad (2.1)$$

  where $d = d(\underline{y}, \underline{x})$ =weight($\underline{z} = \underline{y} \oplus \underline{x}$), and weight($\underline{a}$) denotes the Hamming weight of the vector $\underline{a}$ (i.e., the number of "ones" in $\underline{a}$).

- For $n \geq M+1$, the block transition probability of this channel is

$$\begin{aligned} P(\underline{Y}=\underline{y}|\underline{X}=\underline{x}) &= P(\underline{Z}=\underline{e}) \\ &= \prod_{i=1}^{n} P(Z_i = e_i | Z_{i-1} = e_{i-1}, \cdots, Z_{i-M} = e_{i-M}) \\ &= L \prod_{i=M+1}^{n} \left[\frac{\rho + \lambda_{i-1}\delta}{1+M\delta}\right]^{e_i} \left[\frac{\sigma + (M-\lambda_{i-1})\delta}{1+M\delta}\right]^{1-e_i}, \quad (2.2) \end{aligned}$$

  where

$$L = \frac{\prod_{s=0}^{k-1}(\rho + s\delta)\prod_{j=0}^{M-1-k}(\sigma + j\delta)}{\prod_{l=1}^{M-1}(1+l\delta)},$$

  $\prod_{i=0}^{a}(\cdot) \stackrel{\triangle}{=} 1$, if $a < 0$, $e_i = x_i \oplus y_i$, $k = e_1 + \cdots + e_M$ and $\lambda_{i-1} = e_{i-1} + \cdots + e_{i-M}$.

We thus see that the noise process is stationary.

We now consider the properties of the noise process $\{Z_i\}$. Define $\{W_n\}$ to be the process obtained by $M$-step blocking $\{Z_n\}$; i.e., $\{W_n\} = (Z_n, Z_{n+1}, Z_{n+2}, \cdots, Z_{n+M-1})$. Then $\{W_n\}$ is a one-step Markov process with $2^M$ states. We denote each state by its decimal representation; i.e., state 0 corresponds to state $(0\cdots00)$, state 1 corresponds to state $(0\cdots01)$, $\cdots$, and state $(2^M-1)$ corresponds to state $(1\cdots11)$.

The process $\{W_n\}$ has the following properties:

- $\{W_n\}$ is a stationary Markov process with stationary distribution $\boldsymbol{\pi} = [\pi_0, \pi_1, \cdots, \pi_{2^M-1}]$, where $\pi_i$ is computed as follows. Let $\omega(i)$ denote the number of 1's in the binary representation of the decimal integer $i$. Then

$$\pi_i = \frac{\prod_{j=0}^{\omega(i)-1}(\rho + j\delta)\prod_{k=0}^{M-1-\omega(i)}(\sigma + k\delta)}{\prod_{l=1}^{M-1}(1+l\delta)}, \quad (2.3)$$

  where $\prod_{i=0}^{a}(\cdot) \stackrel{\triangle}{=} 1$, if $a < 0$.

- If we let $p_{ij}$ be the one-step transition probabilities, then

$$p_{ij} = \begin{cases} \frac{\sigma+[M-\omega(i)]\delta}{1+M\delta}, & \text{if } j = 2i \text{ (modulo } 2^M), \\ \frac{\rho+\omega(i)\delta}{1+M\delta}, & \text{if } j = (2i+1) \text{ (modulo } 2^M), \\ 0, & \text{otherwise.} \end{cases} \tag{2.4}$$

The Markov process $\{W_n\}$ is irreducible and aperiodic; therefore it is ergodic.

### 2.1.2 Capacity of the finite-memory contagion channel

Using the results in the previous section, we arrive at the following proposition.

**Proposition 1** *The capacity $C_M$ of the M-memory contagion channel is nondecreasing in M and given by*

$$C_M = 1 - \sum_{k=0}^{M} \binom{M}{k} L_k h_b\left(\frac{\rho + k\delta}{1 + M\delta}\right), \tag{2.5}$$

*where*

$$L_k = \frac{\prod_{j=0}^{k-1}(\rho + j\delta) \prod_{l=0}^{M-1-k}(\sigma + l\delta)}{\prod_{m=1}^{M-1}(1 + m\delta)},$$

$\prod_{i=0}^{a}(\cdot) \stackrel{\triangle}{=} 1$, *if $a < 0$ and $h_b(\cdot)$ is the binary entropy function:*

$$h_b(q) = -q \log_2 q - (1-q)\log_2(1-q).$$

## 2.2 Gilbert-Elliott Channel

The following is from [8] and [10].

The Gilbert-Elliott channel is a time varying binary symmetric channel, the crossover probabilities of which are determined by the current state of a discrete-time stationary binary Markov process (see Fig. 2.2). The states are designated $G$ for good and $B$ for bad. Due to the underlying Markov nature of the channel, it has memory that depends on the transition probabilities between the states.

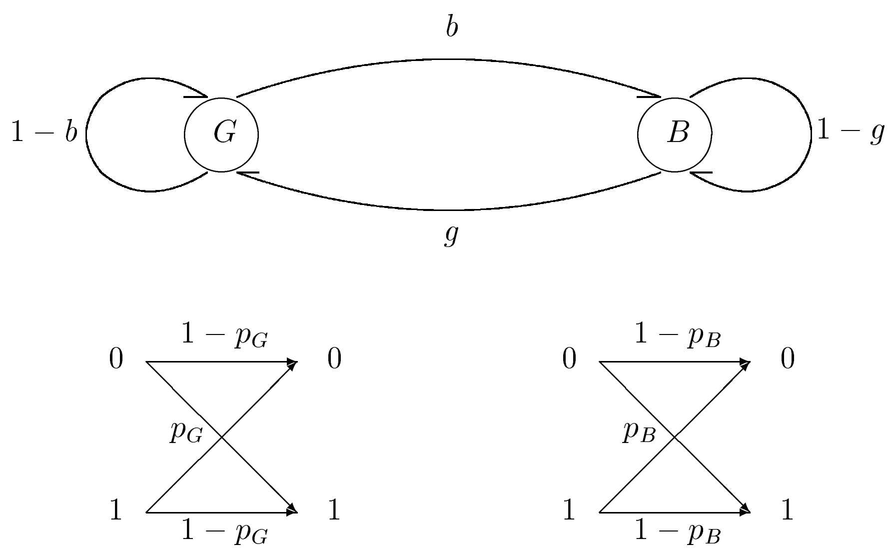



Figure 2.2: The Gilbert-Elliott channel model.

### 2.2.1 Basic properties of the conditional probabilities of channel errors

Let $X_l \in \{0,1\}$, $Z_l \in \{0,1\}$ and $Y_l \stackrel{\triangle}{=} X_l \oplus Z_l$ denote, respectively, the input, the noise, and the output of the channel at the $l$th use, $l = 1, 2, \cdots$. The error process $\{Z_l\}_{l=1}^{\infty}$ is independent of the channel input, that is

$$\Pr[\underline{Y}_l | \underline{X}_l] = \Pr[\underline{Z}_l]. \tag{2.6}$$

The error process has memory in the sense that it depends on an underlying state process $\{S_l\}_{l=0}^{\infty}$, $S_l \in \{G, B\}$, where $G$ stands for good and $B$ for bad. When conditioned on the state process, the error process is memoryless, that is

$$\Pr[\underline{Z}_l | \underline{S}_l] = \prod_{i=1}^{l} \Pr[Z_i | S_i]. \tag{2.7}$$

The conditional probabilities $\Pr[Z_l | S_l]$ do not depend on the time index $l$. The state process is a stationary first-order Markov process; that is,

$$\Pr[S_l|\underline{S}_{l-1}] = \Pr[S_l|S_{l-1}], \tag{2.8}$$

and the right side of (2.8) does not depend on the index $l$. The parameters of the channel are the crossover probabilities in the $G$ and $B$ states

$$p_G \triangleq \Pr[Z_l = 1|S_l = G], \tag{2.9}$$

$$p_B \triangleq \Pr[Z_l = 1|S_l = B], \tag{2.10}$$

and the transition probabilities

$$g \triangleq \Pr[S_l = G|S_{l-1} = B], \tag{2.11}$$

$$b \triangleq \Pr[S_l = B|S_{l-1} = G]. \tag{2.12}$$

The initial distribution of the state process is given by its stationary distribution

$$\Pr[S_0 = G] = g/(g+b) \tag{2.13}$$

and

$$\Pr[S_0 = B] = b/(g+b); \tag{2.14}$$

this ensures that $\{S_l\}$ is stationary. We define the good-to-bad ratio $\rho$ of the channel by

$$\rho \triangleq \Pr[S_l = G]/\Pr[S_l = B] = g/b, \tag{2.15}$$

and define the parameter $\mu$ by

$$\mu \triangleq 1 - g - b. \tag{2.16}$$

**Definition 1** *For sample paths such that* $\Pr[\underline{Z}_{l-1}, S_0] \neq 0$*, let* $q_l^*(\underline{Z}_{l-1}, S_0)$ *denote the probability for a channel error at the* $l$*th use, conditioned on the initial state, and the previous channel errors, that is,*

$$q_l^*(\underline{Z}_{l-1}, S_0) \stackrel{\triangle}{=} \Pr[Z_l = 1 | \underline{Z}_{l-1}, S_0], \tag{2.17}$$

*and let* $q_l(\underline{Z}_{l-1})$ *denote the same probability conditioned only on the previous channel errors, that is,*

$$q_l(\underline{Z}_{l-1}) \stackrel{\triangle}{=} E[q_l^*(\underline{Z}_{l-1}, S_0) | \underline{Z}_{l-1}] = \Pr[Z_l = 1 | \underline{Z}_{l-1}], \tag{2.18}$$

*where* $E[\cdot]$ *denotes expectation. The stationarity of the channel allows shifting the time reference and thus the following notation is valid*

$$q_m^*(\underline{Z}_{k+m-1}^{k+1}, S_k) = \Pr[Z_{k+m} = 1 | \underline{Z}_{k+m-1}^{k+1}, S_k] \tag{2.19}$$

*and*

$$q_m(\underline{Z}_{k+m-1}^{k+1}) = \Pr[Z_{k+m} = 1 | \underline{Z}_{k+m-1}^{k+1}], \; m = 0, 1, 2, \cdots, \tag{2.20}$$

*where the subscripts and superscripts to vectors are to be interpreted as follows:* $\underline{Z}_n^m \stackrel{\triangle}{=} (Z_m, Z_{m+1}, \cdots, Z_n)$ *for* $m \leq n$ *and* $\underline{Z}_n \stackrel{\triangle}{=} \underline{Z}_n^1$*.*

*The special cases* $m = 0, 1$ *mean that previous errors are not available, and therefore* $q_0^*(\underline{Z}_{k-1}^{k+1}, S_k)$ *and* $q_1(\underline{Z}_{k-1}^{k+1})$ *are denoted, for simplicity, by*

$$q_0^*(S_k) \stackrel{\triangle}{=} \Pr[Z_k = 1 | S_k] \tag{2.21}$$

*and by*

$$q_1 \stackrel{\triangle}{=} \Pr[Z_{k+1} = 1], \tag{2.22}$$

*which, due to stationarity, do not depend on* $k$*.*

The following proposition gives recursions for these functions.

**Proposition 2** *The following recursions hold:*

$$q_{l+1}^*(\underline{Z}_l, S_0) = \nu(Z_l, q_l^*(\underline{Z}_{l-1}, S_0)) \tag{2.23}$$

*and*

$$q_{l+1}(\underline{Z}_l) = \nu(Z_l, q_l(\underline{Z}_{l-1})), \tag{2.24}$$

*where the function* $\nu(\cdot,\cdot)$ *is defined by*

$$\nu(0,q) \triangleq \begin{cases} p_G + b(p_B - p_G) + \mu(q - p_G)[(1-p_B)/(1-q)], & p_B \neq 1 \\ (1-b)p_G + b, & p_B = 1,\ q \neq 1 \end{cases} \tag{2.25}$$

*and by*

$$\nu(1,q) \triangleq \begin{cases} p_G + b(p_B - p_G) + \mu(q - p_G)(p_B)/q), & p_G \neq 0 \\ (1-g)p_B, & p_G = 0,\ q \neq 0 \end{cases} \tag{2.26}$$

*for* $p_G \leq q \leq p_B$. *The initial values for these recursions*

$$q_0^*(S_0) = \begin{cases} p_G, & S_0 = G \\ p_B, & S_0 = B \end{cases} \tag{2.27}$$

*and*

$$q_1 = (\rho p_G + p_B)(\rho + 1)^{-1} \tag{2.28}$$

*follow directly from Definition 1.*

Considering $\{q_l^*\}_{l=0}^{\infty}$ and $\{q_l\}_{l=1}^{\infty}$ as random sequences, the following holds.

**Corollary 1** *The sequence* $\{q_l^*\}_{l=0}^{\infty}$ *is a Markov process with initial distribution*

$$\Pr[q_0^* = \alpha] = \begin{cases} \rho/(\rho+1), & \alpha = p_G \\ 1/(\rho+1), & \alpha = p_B \end{cases} \tag{2.29}$$

*and transition probabilities*

$$\Pr[q^*_{l+1} = \alpha | q^*_l = \beta] = \begin{cases} 1-\beta, & \alpha = \nu(0,\beta) \\ \beta, & \alpha = \nu(1,\beta) \end{cases} \tag{2.30}$$

*where $\nu(\cdot,\cdot)$ is as in (2.25) and (2.26).*

**Corollary 2** *The sequence $\{q_l\}_{l=1}^{\infty}$ is also a Markov process with the same transition probabilities (2.30) and with initial distribution*

$$\Pr[q_1 = (\rho p_G + p_B)(\rho+1)^{-1}] = 1. \tag{2.31}$$

### 2.2.2 Evaluation of capacity

Having established the basic properties of $q_l$ and $q^*_l$, we turn to the problem of evaluating the capacity. The issues involved in the definition of the capacity of a finite-state time-varying channel are discussed in [5]. The capacity $C$ of the Gilbert-Elliott channel can be formulated in terms of input and output sequences as follows:

$$C = \lim_{l\to\infty} \frac{1}{l} \max_{\Pr(\underline{X}_l)} I(\underline{X}_l; \underline{Y}_l) \tag{2.32}$$

where $I(\cdot;\cdot)$ denotes mutual information and where the maximum is taken over all possible probability functions $\Pr(\underline{X}_l)$ of the input sequence $\underline{X}_l$. It can be shown [8] that

$$C = 1 - \lim_{l\to\infty} \frac{1}{l} H(\underline{Z}_l) = 1 - \lim_{l\to\infty} E[h_b(q_l)],$$

where $h_b(\cdot)$ is the binary entropy. Furthermore, the following two propositions are proved in [8].

**Proposition 3** *Let $f(\cdot)$ be continuous over $[p_G, p_B]$. Then the following limits exist and are equal:*

$$\lim_{l\to\infty} E[f(q^*_l)] = \lim_{l\to\infty} E[f(q_l)], \tag{2.33}$$

**Proposition 4** *Let $f(\cdot)$ be convex over $[p_G, p_B]$. Then the sequence $\{E[f(q_l)]\}_{l=1}^{\infty}$ is monotonically increasing, the sequence $\{E[f(q_l^*)]\}_{l=1}^{\infty}$ is monotonically decreasing,*

$$E[f(q_l)] \leq E[f(q_{l+1})] \leq E[f(q_{l+1}^*)] \leq E[f(q_l^*)], \tag{2.34}$$

*and both converge to the same limit*

$$\lim_{l\to\infty} E[f(q_l^*)] = \lim_{l\to\infty} E[f(q_l)]. \tag{2.35}$$

Since the binary entropy function $h_b(q) = -qlog_2q - (1-q)log_2(1-q)$ is concave (and continuous on $[p_G, p_B]$), we thus have

$$E[h_b(q_l)] \geq E[h_b(q_{l+1})] \geq E[h_b(q_{l+1}^*)] \geq E[h_b(q_l^*)], \tag{2.36}$$

and

$$\lim_{l\to\infty} E[h_b(q_l^*)] = \lim_{l\to\infty} E[h_b(q_l)].$$

Thus the capacity of the Gilbert-Elliott channel, in bits per channel use, is given by

$$C = 1 - \lim_{l\to\infty} E[h_b(q_l^*)] = 1 - \lim_{l\to\infty} E[h_b(q_l)]. \tag{2.37}$$

Note also that it follows from (2.36) and (2.37) that

$$\underline{C}_l \leq C \leq \overline{C}_l, \quad \forall l \geq 1, \tag{2.38}$$

and

$$C = \lim_{l\to\infty} \underline{C}_l = \lim_{l\to\infty} \overline{C}_l,$$

where

$$\underline{C}_l \triangleq 1 - E[h_b(q_l)],$$

and

$$\overline{C}_l \triangleq 1 - E[h_b(q_l^*)].$$

Thus for $l \geq 1$, $\underline{C}_l$ and $\overline{C}_l$ can be computed to provide a lower and upper estimate, respectively, to the capacity $C$; as $l$ increases, the estimates become more accurate. We will use (2.38) with $l = 9$ to estimate $C$ in Chapter 4.

### 2.2.3 Calculation of the probability of an error sequence

A useful approach for calculating the probability of an error sequence for the Gilbert-Elliott Channel is discussed in [10]. By the law of total probability, the probability of an error sequence of length $n$, $\underline{e}_n = e_1 e_2 \cdots e_n$, may be expressed as:

$$\Pr(\underline{e}_n) = \boldsymbol{\pi}^T \Big(\prod_{k=1}^{n} \mathbf{P}(e_k)\Big)\mathbf{1}. \tag{2.39}$$

where $\mathbf{P}(e_k)$ is a $2 \times 2$ matrix whose $(i,j)^{th}$ entry is the probability that the output symbol is $e_k$ when the chain makes a transition from state $s_{k-1} = i$ to $s_k = j$.

The following expression is valid for the vector $\boldsymbol{\pi} = [\pi_0 \;\; \pi_1]^T$:

$$\boldsymbol{\pi}^T \mathbf{P} = \boldsymbol{\pi}^T; \tag{2.40}$$

where

$$\mathbf{P} = \sum_{e=0}^{1} \mathbf{P}(e), \tag{2.41}$$

$$\mathbf{P}\mathbf{1} = \mathbf{1}, \tag{2.42}$$

and

$$\boldsymbol{\pi}^T \mathbf{1} = 1. \tag{2.43}$$

From (2.40) and (2.42) we conclude that the vectors $\boldsymbol{\pi}$ and $\mathbf{1}$ are the left and right eigenvectors of the matrix $\mathbf{P}$, respectively, corresponding to the eigenvalue 1.

For the Gilbert-Elliot channel, the matrices $\mathbf{P}$, $\mathbf{P}(0)$, $\mathbf{P}(1)$ and the vectors $\mathbf{1}$ and $\boldsymbol{\pi}$ are given by:

$$\mathbf{P} = \begin{bmatrix} 1-b & b \\ g & 1-g \end{bmatrix}, \tag{2.44}$$

$$\mathbf{P}(0) = \begin{bmatrix} (1-b)(1-p_G) & b(1-p_B) \\ g(1-p_G) & (1-g)(1-p_B) \end{bmatrix}, \tag{2.45}$$

$$\mathbf{P}(1) = \begin{bmatrix} (1-b)p_G & bp_B \\ gp_G & (1-g)p_B \end{bmatrix}, \tag{2.46}$$

$$\mathbf{1} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \tag{2.47}$$

and

$$\boldsymbol{\pi} = \begin{bmatrix} \pi_0 \\ \pi_1 \end{bmatrix} = \begin{bmatrix} \frac{g}{b+g} \\ \frac{b}{b+g} \end{bmatrix}. \tag{2.48}$$

## 2.3 A Critical Statistic for Channels with Memory

The following is from [1].

Consider the channel model described by the diagram in Fig. 2.1. As before $Y_i$ and $X_i$ are, respectively, the output and input symbol and $Z_i$ is the noise symbol. Throughout it will be assumed that the input sequence and noise sequence are statistically independent of each other.

### 2.3.1 The error-gap process

The state space of the error process is composed of two states: error (or 1) and error free (or 0). A positive integer exponent is used to indicate the number of consecutive symbols of the same type. In this notation the sequence 100001001 is written as $10^4 10^2 1$.

In probability expressions we use the notation,

$$\Pr[Z_1 = 0, \cdots, Z_{n-1} = 0, Z_n = 1] = P(0^{n-1}1). \tag{2.49}$$

A conditional probability is also written in terms of sequences:

$$\Pr[\text{sequence } B \mid \text{sequence } A].$$

If no other indication is given in the conditional part, it is understood that sequence $B$ directly follows sequence $A$.

**Definition 2** *Sequences of zeros between two errors (1 states) are called error gaps (also error-free runs). The length of a gap is defined as one plus the total number of zeros in the sequence between two 1's.*

Let $G_n$ denote the length of the gap starting at $Z_n$, conditioned on $Z_n = 1$. Assuming stationarity, the following equations hold

$$\begin{aligned} \Pr[G_n = m] &= \Pr[Z_{n+1} = 0, \cdots, Z_{n+m-1} = 0, Z_{n+m} = 1 | Z_n = 1] \\ &= P(0^{m-1}1|1) \end{aligned} \tag{2.50}$$

and

$$E(G_n) = \frac{1}{P(1)} \tag{2.51}$$

for all positive integers $m$, where $P(1)$ is the probability of error and is equal to $E(Z_n)$.

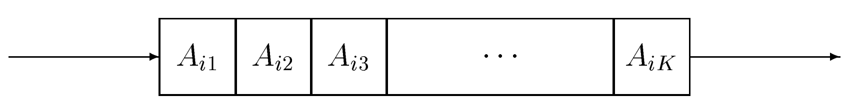
# Chapter 3

# A Queue-Based Channel with Memory

In this chapter, we present a binary channel with additive bursty noise based on a finite queue. Consider the binary channel given by $Y_i = X_i \oplus Z_i$, where the input stream is independent from the noise stream (see Fig. 2.1).

Consider the following two parcels.

- **Parcel 1** is a queue of length $K$, that contains *initially* $K$ balls.




Let the random variables $A_{ij}$ ($i$ is a time index refering to the $i^{th}$ experiment), $j = 1, 2, \cdots, K$, indicate the color of the ball in the corresponding cell of the queue at time $i$:

$$A_{ij} = \begin{cases} 1, & \text{if the } j^{th} \text{ cell contains a red ball at time } i, \\ 0, & \text{if the } j^{th} \text{ cell contains a black ball at time } i. \end{cases}$$

- **Parcel 2** is an urn that contains a very large number of balls where the proportion of black balls is $1-p$ and the proportion of red balls is $p$, where $p \in (0,1)$; usually $p \ll 1/2$.

Let the probability of selecting parcel 1 (the queue) be $\varepsilon$ and the probability of selecting parcel 2 (the urn) be $1-\varepsilon$; where $\varepsilon \in (0,1)$.

The noise process $\{Z_i\}$ is generated by one of the following mechanisms.

**Mechanism 1 Uniform queue-based channel with memory:** *By flipping a biased coin (with P(Head)=$\varepsilon$), we select one of the 2 parcels (select the queue if Head and the urn if Tail). Then a pointer randomly points at a ball from the selected parcel, and identifies its color.*

- If the selected ball is red, we introduce a red ball in cell 1 of the queue, pushing the last ball in cell $K$ out.
- If the selected ball is black, we introduce a black ball in cell 1 of the queue, pushing the last ball in cell $K$ out.

The noise process $\{Z_i\}$ is then modeled as follows:

$$Z_i = \begin{cases} 1, & \text{if the } i^{th} \text{ experiment points at a red ball;} \\ 0, & \text{if the } i^{th} \text{ experiment points at a black ball.} \end{cases}$$

**Mechanism 2 Non-uniform queue-based channel with memory:** *By flipping a biased coin (with P(Head)=$\varepsilon$), we select one of the 2 parcels (select the queue if Head and* *the urn if Tail). If parcel* 1 *(the queue) is selected, then a pointer points at the ball in cell* 1 *with probability* $q_1$ *and points at the ball in cell* $l$ *with probability* $q_l = (1-q_1)/(K-1)$, *for* $l = 2, 3, \cdots, K$, *and identifies its color. If parcel* 2 *(the urn) is selected, a pointer randomly points at a ball, and identifies its color.*

- If the selected ball is red, we introduce a red ball in cell 1 of the queue, pushing the last ball in cell $K$ out.
- If the selected ball is black, we introduce a black ball in cell 1 of the queue, pushing the last ball in cell $K$ out.

The noise process $\{Z_i\}$ is then modeled as follows:

$$Z_i = \begin{cases} 1, & \text{if the } i^{th} \text{ experiment points at a red ball;} \\ 0, & \text{if the } i^{th} \text{ experiment points at a black ball.} \end{cases}$$

**Definition 3** *For a given mechanism, define the state of the channel to be* $\underline{S}_i \triangleq (A_{i1}, A_{i2}, \cdots, A_{iK})$, *the binary* $K-$*tuple in the queue after the* $i^{th}$ *experiment is completed, where* $A_{ij}, j = 1, 2, \cdots, K$, *is an indicator random variable for cell* $j$ *in the queue after the* $i^{th}$ *experiment is performed. Note that, in terms of the noise process, the channel state at time* $i$ *can be written as* $\underline{S}_i = (Z_i, Z_{i-1}, \cdots, Z_{i-K+1})$, *for* $i \geq K$.

## 3.1 Properties of the Uniform Queue-Based Noise

We now investigate the properties of the binary noise process $\{Z_n\}_{n=1}^{\infty}$. To gain more insight , we first study the noise process for different values of $K$, and then give our results in terms of $K$ for the general case. We first observe that $\{Z_n\}_{n=1}^{\infty}$ is a homogeneous (on time-invariant [4]) Markov process of order $K$, since for $n \geq K+1$,

$$\begin{aligned} & \Pr[Z_n = 1 | Z_{n-1} = a_{n-1}, \cdots, Z_1 = a_1] = \varepsilon \frac{a_{n-1} + \cdots + a_{n-K}}{K} + (1-\varepsilon)p \\ = \ & \Pr[Z_n = 1 | Z_{n-1} = a_{n-1}, \cdots, Z_{n-K} = a_{n-K}], \end{aligned} \tag{3.1}$$

where $a_j \in \{0,1\}$, $j = 1, \cdots, n$.

Thoughout this work, we consider the case where the initial distribution of the Markov noise $\{Z_n\}$ is drawn according to its stationary distribution; hence the noise process $\{Z_n\}$ is stationary.

### 3.1.1 Case for $K = 1$

For $K = 1$, the noise process becomes a simple Markov process. After analyzing this process we get the following results.

- $\{Z_n\}$ is a binary homogeneous Markov process with $\Pr(Z_n = j | Z_{n-1} = i) \stackrel{\triangle}{=} p_{ij}$. The transition matrix of the process is:

$$\mathbf{Q} = \begin{bmatrix} p_{00} & p_{01} \\ p_{10} & p_{11} \end{bmatrix} = \begin{bmatrix} \varepsilon + (1-\varepsilon)(1-p) & (1-\varepsilon)p \\ (1-\varepsilon)(1-p) & \varepsilon + (1-\varepsilon)p \end{bmatrix}. \tag{3.2}$$

- We can observe from the transition matrix $\mathbf{Q}$ that it is possible to go with positive probability from any state to any other state in a finite number of steps. Therefore the Markov process is irreducible (and hence ergodic [9]).

- Since the process $\{Z_n\}$ is irreducible, its stationary distribution $\boldsymbol{\pi} \stackrel{\triangle}{=} (\pi(Z_\infty = 0); \pi(Z_\infty = 1)) \stackrel{\triangle}{=} (\pi_0; \pi_1)$ is unique, and it can be computed by solving $\boldsymbol{\pi} = \boldsymbol{\pi}\mathbf{Q}$. This yields $\boldsymbol{\pi} = (1-p; p)$.

#### Noise entropy rate

We know that if a stochastic process $W_n$ is stationary, then its conditional entropy $H(W_n | W_{n-1}, \cdots, W_1)$ is decreasing in $n$; and therefore $\lim_{n\to\infty} H(W_n | W_{n-1}, \cdots, W_1)$ exists. Therefore, we get for our Markov process $\{Z_n\}$:

$$\lim_{n\to\infty} H(Z_n | Z_{n-1}, \cdots, Z_1) = H(Z_2 | Z_1)$$

where

$$\begin{aligned} H(Z_2|Z_1) &= -\sum_{i,j=0}^{1} \pi_i p_{ij} \log_2 p_{ij} \\ &= (1-p)h_b[(1-\varepsilon)p] + ph_b[(1-\varepsilon)(1-p)], \end{aligned} \tag{3.3}$$

where $h_b(\cdot)$ is the binary entropy function.

Now by the Cesaro-mean average, we get that as $n \to \infty$,

$$\frac{1}{n}H(Z_1, \cdots, Z_n) = \frac{1}{n}\sum_{i=1}^{n} H(Z_n|Z_{n-1}, \cdots, Z_1) \to H(Z_2|Z_1).$$

Thus we can obtain the entropy rate $h(Z_\infty)$ of the Markov noise process:

$$h(Z_\infty) \triangleq \lim_{n\to\infty} \frac{1}{n} H(Z_1, \cdots, Z_n) = H(Z_2|Z_1). \tag{3.4}$$

### 3.1.2 Case for $K = 2$

For $K = 2$, the noise process becomes a $2^{nd}$ order Markov process. As we defined above, $\underline{S}_n \triangleq (A_{n1}, A_{n2}, \cdots, A_{nK})$; i.e., $\underline{S}_n = (Z_n, Z_{n-1})$, for $n \geq 2$.

We have that for $n \geq 2$,

$$\begin{aligned} & \Pr[\underline{S}_n = \underline{s}_n | \underline{S}_{n-1} = \underline{s}_{n-1}, \cdots, \underline{S}_1 = \underline{s}_1] \\ = \; & \Pr[Z_n = a_n, Z_{n-1} = a_{n-1} | Z_{n-1} = a_{n-1}, \cdots, Z_1 = a_1] \\ = \; & \Pr[Z_n = a_n | Z_{n-1} = a_{n-1}, Z_{n-2} = a_{n-2}] \\ = \; & \Pr[\underline{S}_n = \underline{s}_n | \underline{S}_{n-1} = \underline{s}_{n-1}] \end{aligned} \tag{3.5}$$

where $a_n \in \{0, 1\}$ and $\underline{s}_n \in \{(0,0), (0,1), (1,0), (1,1)\}$.

Thus $\{\underline{S}_n\}$ is a Markov process with 4 states. We denote each state by its decimal representation; i.e., state 0 corresponds to state $(0,0)$ or $(00)$, state 1 corresponds to state $(01)$, state 2 corresponds to state $(10)$ and state 3 corresponds to state $(11)$. We now get the following results about $\{\underline{S}_n\}$.

- Since $\{\underline{Z}_n\}$ is stationary, it is also 2-step block stationary, which implies that $\{\underline{S}_n\}$ is stationary.

- $\{\underline{S}_n\}$ is also a homogeneous Markov process with stationary (or initial) distribution

$$\boldsymbol{\pi} \stackrel{\triangle}{=} (\pi(\underline{S}_1 = 00); \pi(\underline{S}_1 = 01); \pi(\underline{S}_1 = 10); \pi(\underline{S}_1 = 11)) \stackrel{\triangle}{=} (\pi_0; \pi_1; \pi_2; \pi_3). \tag{3.6}$$

- If we denote by $p_{ij}$ the transition probability that $\underline{S}_n$ goes from state $i$ to state $j$, $i, j = 0, 1, 2, 3$, we can write the transition matrix of the process $\{\underline{S}_n\}$:

$$\mathbf{Q} = \begin{bmatrix} p_{00} & p_{01} & p_{02} & p_{03} \\ p_{10} & p_{11} & p_{12} & p_{13} \\ p_{20} & p_{21} & p_{22} & p_{23} \\ p_{30} & p_{31} & p_{32} & p_{33} \end{bmatrix}$$

$$= \begin{bmatrix} \varepsilon + (1-\varepsilon)(1-p) & 0 & (1-\varepsilon)p & 0 \\ \frac{\varepsilon}{2} + (1-\varepsilon)(1-p) & 0 & \frac{\varepsilon}{2} + (1-\varepsilon)p & 0 \\ 0 & \frac{\varepsilon}{2} + (1-\varepsilon)(1-p) & 0 & \frac{\varepsilon}{2} + (1-\varepsilon)p \\ 0 & (1-\varepsilon)(1-p) & 0 & \varepsilon + (1-\varepsilon)p \end{bmatrix}. \tag{3.7}$$

  From the transition matrix $\mathbf{Q}$, we can clearly see that any state can reach any other state with positive probability in a finite number of steps. Therefore the process $\underline{S}_n$ is irreducible (and hence ergodic).

- We can compute the stationary distribution $\boldsymbol{\pi} \stackrel{\triangle}{=} (\pi_0; \pi_1; \pi_2; \pi_3)$ of the process by solving $\boldsymbol{\pi} = \boldsymbol{\pi}\mathbf{Q}$. This yields

$$\begin{aligned} \boldsymbol{\pi} \;=\; & \left( \frac{[\frac{\varepsilon}{2} + (1-\varepsilon)(1-p)][(1-\varepsilon)(1-p)]}{[1-\frac{\varepsilon}{2}](1-\varepsilon)}; \frac{[(1-\varepsilon)(1-p)][(1-\varepsilon)p]}{[1-\frac{\varepsilon}{2}](1-\varepsilon)}; \right. \\ & \left. \frac{[(1-\varepsilon)(1-p)][(1-\varepsilon)p]}{[1-\frac{\varepsilon}{2}](1-\varepsilon)}; \frac{[(1-\varepsilon)p][\frac{\varepsilon}{2} + (1-\varepsilon)p]}{[1-\frac{\varepsilon}{2}](1-\varepsilon)} \right). \end{aligned} \tag{3.8}$$

#### Noise entropy rate

As we did in the previous section, we can obtain the entropy rate of the $2^{nd}$ order Markov noise process $\{\underline{Z}_n\}$ as follows.

$$\begin{aligned}
h(Z_\infty) &\stackrel{\triangle}{=} \lim_{n\to\infty} \frac{1}{n} H(Z_1, \cdots, Z_n) \\
&= H(Z_3|Z_2, Z_1) \\
&= H(\underline{S}_2|\underline{S}_1) \\
&= -\sum_{i,j=0}^{3} \pi_i p_{ij} \log_2 p_{ij} \\
&= \frac{[\frac{\varepsilon}{2} + (1-\varepsilon)(1-p)][(1-\varepsilon)(1-p)]}{[1-\frac{\varepsilon}{2}](1-\varepsilon)} h_b[(1-\varepsilon)p] \\
&+ \frac{2[(1-\varepsilon)(1-p)][(1-\varepsilon)p]}{[1-\frac{\varepsilon}{2}](1-\varepsilon)} h_b[\frac{\varepsilon}{2} + (1-\varepsilon)p] \\
&+ \frac{[(1-\varepsilon)p][\frac{\varepsilon}{2} + (1-\varepsilon)p]}{[1-\frac{\varepsilon}{2}](1-\varepsilon)} h_b[(1-\varepsilon)(1-p)],
\end{aligned} \tag{3.9}$$

where $h_b(\cdot)$ is the binary entropy function.

### 3.1.3 Case for $K = 3$

For $K = 3$, the noise process becomes a $3^{rd}$ order Markov process and $\underline{S}_n = (Z_n, Z_{n-1}, Z_{n-2})$ for $n \geq 3$.

We can easily show, as we did for the $K = 2$ case, that $\underline{S}_n$ is a Markov process with 8 states; i.e., $\underline{s}_n \in \{(000), (001), (010), (011), (100), (101), (110), (111)\}$. Again we denote each state by its decimal representation. We get the following results about $\{\underline{S}_n\}$:

- $\{\underline{S}_n\}$ is a homogeneous Markov process with stationary (or initial) distribution

$$\boldsymbol{\pi} \stackrel{\triangle}{=} (\pi_0; \pi_1; \pi_2; \pi_3; \pi_4; \pi_5; \pi_6; \pi_7). \tag{3.10}$$

- If we denote by $p_{ij}$ the transition probability that $\underline{S}_n$ goes from state $i$ to state $j$,

$i, j = 0, 1, 2, 3, 4, 5, 6, 7$, we can write the transition matrix of the process $\{\underline{S}_n\}$:

$$
\mathbf{Q} = \begin{bmatrix}
p_{00} & p_{01} & p_{02} & p_{03} & p_{04} & p_{05} & p_{06} & p_{07} \\
p_{10} & p_{11} & p_{12} & p_{13} & p_{14} & p_{15} & p_{16} & p_{17} \\
p_{20} & p_{21} & p_{22} & p_{23} & p_{24} & p_{25} & p_{26} & p_{27} \\
p_{30} & p_{31} & p_{32} & p_{33} & p_{34} & p_{35} & p_{36} & p_{37} \\
p_{40} & p_{41} & p_{42} & p_{43} & p_{44} & p_{45} & p_{46} & p_{47} \\
p_{50} & p_{51} & p_{52} & p_{53} & p_{54} & p_{55} & p_{56} & p_{57} \\
p_{60} & p_{61} & p_{62} & p_{63} & p_{64} & p_{65} & p_{66} & p_{67} \\
p_{70} & p_{71} & p_{72} & p_{73} & p_{74} & p_{75} & p_{76} & p_{77}
\end{bmatrix}
$$

$$
= \begin{bmatrix}
\varepsilon + (1-\varepsilon)(1-p) & 0 & 0 & 0 & (1-\varepsilon)p & 0 & 0 & 0 \\
\frac{2\varepsilon}{3} + (1-\varepsilon)(1-p) & 0 & 0 & 0 & \frac{\varepsilon}{3} + (1-\varepsilon)p & 0 & 0 & 0 \\
0 & \frac{2\varepsilon}{3} + (1-\varepsilon)(1-p) & 0 & 0 & 0 & \frac{\varepsilon}{3} + (1-\varepsilon)p & 0 & 0 \\
0 & \frac{\varepsilon}{3} + (1-\varepsilon)(1-p) & 0 & 0 & 0 & \frac{2\varepsilon}{3} + (1-\varepsilon)p & 0 & 0 \\
0 & 0 & \frac{2\varepsilon}{3} + (1-\varepsilon)(1-p) & 0 & 0 & 0 & \frac{\varepsilon}{3} + (1-\varepsilon)p & 0 \\
0 & 0 & \frac{\varepsilon}{3} + (1-\varepsilon)(1-p) & 0 & 0 & 0 & \frac{2\varepsilon}{3} + (1-\varepsilon)p & 0 \\
0 & 0 & 0 & \frac{\varepsilon}{3} + (1-\varepsilon)(1-p) & 0 & 0 & 0 & \frac{2\varepsilon}{3} + (1-\varepsilon)p \\
0 & 0 & 0 & (1-\varepsilon)(1-p) & 0 & 0 & 0 & \varepsilon + (1-\varepsilon)p
\end{bmatrix}.
$$

From the transition matrix $\mathbf{Q}$, we can clearly see that any state can reach any other state with positive probability in a finite number of steps. Therefore the process $\underline{S}_n$ is irreducible (and hence ergodic).

- We can compute the stationary distribution $\boldsymbol{\pi} \stackrel{\triangle}{=} (\pi_0; \pi_1; \pi_2; \pi_3; \pi_4; \pi_5; \pi_6; \pi_7)$ of the process by solving $\boldsymbol{\pi} = \boldsymbol{\pi}\mathbf{Q}$. This yields

$$
\begin{aligned}
\boldsymbol{\pi} = \Bigg( & \frac{[\frac{2\varepsilon}{3} + (1-\varepsilon)(1-p)][\frac{\varepsilon}{3} + (1-\varepsilon)(1-p)][(1-\varepsilon)(1-p)]}{[1-\frac{2\varepsilon}{3}][1-\frac{\varepsilon}{3}](1-\varepsilon)}; \\
& \frac{[\frac{\varepsilon}{3} + (1-\varepsilon)(1-p)][(1-\varepsilon)(1-p)][(1-\varepsilon)p]}{[1-\frac{2\varepsilon}{3}][1-\frac{\varepsilon}{3}](1-\varepsilon)}; \\
& \frac{[\frac{\varepsilon}{3} + (1-\varepsilon)(1-p)][(1-\varepsilon)(1-p)][(1-\varepsilon)p]}{[1-\frac{2\varepsilon}{3}][1-\frac{\varepsilon}{3}](1-\varepsilon)};
\end{aligned}
$$

$$\frac{[\frac{\varepsilon}{3}+(1-\varepsilon)(1-p)][(1-\varepsilon)(1-p)][(1-\varepsilon)p]}{[1-\frac{2\varepsilon}{3}][1-\frac{\varepsilon}{3}](1-\varepsilon)};$$
$$\frac{[(1-\varepsilon)(1-p)][(1-\varepsilon)p][\frac{\varepsilon}{3}+(1-\varepsilon)p]}{[1-\frac{2\varepsilon}{3}][1-\frac{\varepsilon}{3}](1-\varepsilon)};$$
$$\frac{[(1-\varepsilon)(1-p)][(1-\varepsilon)p][\frac{\varepsilon}{3}+(1-\varepsilon)p]}{[1-\frac{2\varepsilon}{3}][1-\frac{\varepsilon}{3}](1-\varepsilon)};$$
$$\frac{[(1-\varepsilon)(1-p)][(1-\varepsilon)p][\frac{\varepsilon}{3}+(1-\varepsilon)p]}{[1-\frac{2\varepsilon}{3}][1-\frac{\varepsilon}{3}](1-\varepsilon)};$$
$$\left.\frac{[(1-\varepsilon)p][\frac{\varepsilon}{3}+(1-\varepsilon)p][\frac{2\varepsilon}{3}+(1-\varepsilon)p]}{[1-\frac{2\varepsilon}{3}][1-\frac{\varepsilon}{3}](1-\varepsilon)}\right). \tag{3.11}$$

### Noise entropy rate

Similarly to the previous section, we can obtain the entropy rate of the $3^{rd}$ order Markov noise process $\{\underline{Z}_n\}$ as follows.

$$\begin{aligned}
h(Z_\infty) &\triangleq \lim_{n\to\infty}\frac{1}{n}H(Z_1,\cdots,Z_n)\\
&= H(Z_4|Z_3,Z_2,Z_1)\\
&= H(\underline{S}_2|\underline{S}_1)\\
&= -\sum_{i,j=0}^{7}\pi_i p_{ij}\log_2 p_{ij}\\
&= \frac{[\frac{2\varepsilon}{3}+(1-\varepsilon)(1-p)][\frac{\varepsilon}{3}+(1-\varepsilon)(1-p)][(1-\varepsilon)(1-p)]}{[1-\frac{2\varepsilon}{3}][1-\frac{\varepsilon}{3}](1-\varepsilon)}h_b[(1-\varepsilon)p]\\
&\quad+\frac{3[\frac{\varepsilon}{3}+(1-\varepsilon)(1-p)][(1-\varepsilon)(1-p)][(1-\varepsilon)p]}{[1-\frac{2\varepsilon}{3}][1-\frac{\varepsilon}{3}](1-\varepsilon)}h_b[\frac{\varepsilon}{3}+(1-\varepsilon)p]\\
&\quad+\frac{3[(1-\varepsilon)(1-p)][(1-\varepsilon)p][\frac{\varepsilon}{3}+(1-\varepsilon)p]}{[1-\frac{2\varepsilon}{3}][1-\frac{\varepsilon}{3}](1-\varepsilon)}h_b[\frac{2\varepsilon}{3}+(1-\varepsilon)p]\\
&\quad+\frac{[(1-\varepsilon)p][\frac{\varepsilon}{3}+(1-\varepsilon)p][\frac{2\varepsilon}{3}+(1-\varepsilon)p]}{[1-\frac{2\varepsilon}{3}][1-\frac{\varepsilon}{3}](1-\varepsilon)}h_b[(1-\varepsilon)(1-p)], \qquad (3.12)
\end{aligned}$$

where $h_b(\cdot)$ is the binary entropy function.

### 3.1.4 General case

We now analyze the noise process for the general case as a function of $K$. The noise process, as we showed earlier, is a $K^{th}$ order stationary Markov process and $\underline{S}_n = (Z_n, Z_{n-1}, \cdots, Z_{n-K+1})$ for $n \geq K$.

We can easily show, as we did for the previous cases, that $\{\underline{S}_n\}$ is a Markov process with $2^K$ states; i.e., $\underline{s}_n \in \{(0\cdots00), (0\cdots01), \cdots, (1\cdots11)\}$ where each string is of length $K$. We denote each state by its decimal representation. We get the following results about $\{\underline{S}_n\}$.

- $\{\underline{S}_n\}$ is a homogeneous Markov process with stationary (or initial) distribution

$$\boldsymbol{\pi} \stackrel{\triangle}{=} (\pi_0; \pi_1; \cdots; \pi_i; \cdots; \pi_{2^K-1}).$$

- If we denote by $p_{ij}$ the transition probability that $\underline{S}_n$ goes from state $i$ to state $j$, $i, j = 0, 1, 2, \cdots, 2^K - 1$, we can write the transition matrix of the process $\{\underline{S}_n\}$:

$$\mathbf{Q} = [p_{ij}]$$

with

$$p_{ij} = \begin{cases} \varepsilon \frac{K-\omega(i)}{K} + (1-\varepsilon)(1-p), & \text{if } j = \lfloor \frac{i}{2} \rfloor; \\ \varepsilon \frac{\omega(i)}{K} + (1-\varepsilon)p, & \text{if } j = \lfloor \frac{i+2^K}{2} \rfloor; \\ 0, & \text{otherwise;} \end{cases} \tag{3.13}$$

where $\omega(i)$ is the number of "ones" in the $K$-bit binary (or base 2) representation of the row number $i$.

The form of the $2^K * 2^K$ transition matrix $\mathbf{Q}$ will look as follows:

$$
\mathbf{Q} = \begin{bmatrix}
\varepsilon + (1-\varepsilon)(1-p) & 0 & \cdots & \cdots & \cdots & (1-\varepsilon)p & 0 & \cdots & \cdots & 0 \\
\substack{\varepsilon\frac{K-1}{K}+ \\ (1-\varepsilon)(1-p)} & 0 & \cdots & \cdots & \cdots & \substack{\varepsilon\frac{1}{K}+ \\ (1-\varepsilon)p} & 0 & \cdots & \cdots & 0 \\
0 & \substack{\varepsilon\frac{K-1}{K}+ \\ (1-\varepsilon)(1-p)} & 0 & \cdots & \cdots & \cdots & \substack{\varepsilon\frac{1}{K}+ \\ (1-\varepsilon)p} & 0 & \cdots & 0 \\
0 & \substack{\varepsilon\frac{K-2}{K}+ \\ (1-\varepsilon)(1-p)} & 0 & \cdots & \cdots & \cdots & \substack{\varepsilon\frac{2}{K}+ \\ (1-\varepsilon)p} & 0 & \cdots & 0 \\
\vdots & & & & & & & & \vdots & \vdots \\
\vdots & & & & & & & & \vdots & \vdots \\
\vdots & & & & & & & & \vdots & \vdots \\
\vdots & & & & & & & & \vdots & \vdots \\
0 & 0 & \cdots & \cdots & \substack{\varepsilon\frac{1}{K}+ \\ (1-\varepsilon)(1-p)} & 0 & \cdots & \cdots & \cdots & \substack{\varepsilon\frac{K-1}{K}+ \\ (1-\varepsilon)p} \\
0 & 0 & \cdots & \cdots & (1-\varepsilon)(1-p) & 0 & \cdots & \cdots & \cdots & \varepsilon + (1-\varepsilon)p
\end{bmatrix}.
$$

The last row in the above matrix, corresponds to the $(2^K-1)^{th}$ row. We note that the matrix has a sliding block form: all entries are "zeros" except for the positions $(i,j)$ such that $j = \lfloor \frac{i}{2} \rfloor$ and $j = \lfloor \frac{i+2^K}{2} \rfloor$, where $i$ is the row number and $j$ is the column number, $i, j = 0, 1, 2, \cdots, 2^K-1$. Also, we notice that there are $\binom{K}{l}$ rows for which the binary representation of the number of these rows contain $l$ "ones".

From the transition matrix $\mathbf{Q}$, we can clearly see that any state can reach any other state with positive probability in a finite number of steps. Therefore the process $\underline{S}_n$ is irreducible (and hence ergodic).

- We can compute the stationary distribution $\boldsymbol{\pi} \stackrel{\triangle}{=} (\pi_0; \pi_1; \cdots; \pi_i; \cdots; \pi_{2^K-1})$ of the process by solving $\boldsymbol{\pi} = \boldsymbol{\pi}\mathbf{Q}$. This yields

$$
\pi_i = \frac{\prod_{j=0}^{K-1-\omega(i)} [\varepsilon\frac{j}{K} + (1-\varepsilon)(1-p)] \prod_{l=0}^{\omega(i)-1} [\varepsilon\frac{l}{K} + (1-\varepsilon)p]}{\prod_{m=1}^{K} (1-\varepsilon\frac{m}{K})}, \tag{3.14}
$$

for $i = 0, 1, 2, \cdots, 2^K-1$, where $\omega(i)$ is the number of "ones" in the binary representation of the decimal integer $i$ and $\prod_{i=0}^{a}(\cdot) \stackrel{\triangle}{=} 1$, if $a < 0$.

**Noise entropy rate**

Similarly to the previous section, we can obtain the entropy rate of the $K^{th}$ order Markov noise process $\{\underline{Z}_n\}$ as follows.

$$
\begin{aligned}
h(Z_\infty) &\triangleq \lim_{n\to\infty}\frac{1}{n}H(Z_1,\cdots,Z_n) \\
&= H(Z_{K+1}|Z_K,Z_{K-1},\cdots,Z_1) \\
&= H(\underline{S}_2|\underline{S}_1) \\
&= -\sum_{i,j=0}^{2^K-1}\pi_i p_{ij}\log_2 p_{ij} \\
&= \sum_{i=0}^{2^K-1}\frac{\prod_{s=0}^{K-1-\omega(i)}[\varepsilon\frac{s}{K}+(1-\varepsilon)(1-p)]\prod_{l=0}^{\omega(i)-1}[\varepsilon\frac{l}{K}+(1-\varepsilon)p]}{\prod_{m=1}^{K}(1-\varepsilon\frac{m}{K})} \\
&\quad\cdot h_b\left(\varepsilon\frac{\omega(i)}{K}+(1-\varepsilon)p\right) \\
&= \sum_{\omega(i)=0}^{K}\binom{K}{\omega(i)}L_{\omega(i)}h_b\left(\varepsilon\frac{\omega(i)}{K}+(1-\varepsilon)p\right) \qquad (3.15)
\end{aligned}
$$

where

$$
L_{\omega(i)} = \frac{\prod_{s=0}^{K-1-\omega(i)}[\varepsilon\frac{s}{K}+(1-\varepsilon)(1-p)]\prod_{l=0}^{\omega(i)-1}[\varepsilon\frac{l}{K}+(1-\varepsilon)p]}{\prod_{m=1}^{K}(1-\varepsilon\frac{m}{K})},
$$

$\prod_{i=0}^{a}(\cdot)\triangleq 1$, if $a<0$, and $h_b(\cdot)$ is the binary entropy function.

## 3.2 Block Transition Probability for the Uniform Queue-Based Channel

For a given input block $\underline{X}=[X_1,\cdots,X_n]$ and a given output block $\underline{Y}=[Y_1,\cdots,Y_n]$, where $n$ is the block length, the block transition probability of the resulting binary channel is as follows.

- For block length $n \leq K$, the block transition probability of this channel is

$$\Pr(\underline{Y} = \underline{y} | \underline{X} = \underline{x}) = \Pr(\underline{Z} = \underline{a}) = \sum_{j \in \mathcal{S}_n} \pi_j,$$

  where

  - $a_i = x_i \oplus y_i$,
  - $\mathcal{S}_n$ is the set of decimal representations of the set of states

$$S_n^b \stackrel{\triangle}{=} \{(\underline{a}, i_{n+1}, \cdots, i_K) | i_t \in \{0,1\}, t = n+1, \cdots, K\},$$

  - $\pi_j$, $j \in \mathcal{S}_n$, are the stationary probabilities of the resulting states.

  After the $n^{th}$ experiment, there are $2^{K-n}$ states; i.e., $(a_n, \cdots, a_1, 0 \cdots 0), \cdots, (a_n, \cdots, a_1, 1 \cdots 1)$, in $S_n^b$. $\mathcal{S}_n$ is the set of decimal representations of these states. The block probability is the sum of the stationary probabilities of these states. If two states have the same binary representation except in the last digit, then the weights of the two states are different by 1 (suppose one is $\omega(j)$, which is the number of "ones" in state $j$ (the weight of state $j$), the other is $\omega(j) + 1$). Thus the formulas of the stationary distribution of the two states are different by only one term; i.e., one includes $[\varepsilon \frac{K-1-\omega(j)}{K} + (1-\varepsilon)(1-p)]$ and the other includes $[\varepsilon \frac{\omega(j)}{K} + (1-\varepsilon)p]$. We can calculate the sum of these two terms to get $(1 - \varepsilon \frac{1}{K})$, which can be cancelled by $(1 - \varepsilon \frac{1}{K})$ in the denominator. Then the sum of the two stationary probabilities is $\frac{\prod_{s=0}^{K-2-\omega(j)} [\varepsilon \frac{s}{K} + (1-\varepsilon)(1-p)] \prod_{t=0}^{\omega(j)-1} [\varepsilon \frac{t}{K} + (1-\varepsilon)p]}{\prod_{l=2}^{K} (1 - \varepsilon \frac{l}{K})}$. Taking pairwise summations of $2^{K-n}$ terms, the total number of terms descreases to $2^{K-n-1}$. By repeating these operations of taking pairwise summation, taking out the common factor, and cancelling another common factor in the numerator and denominator, we get a final single term which is the following.

$$\sum_{j \in \mathcal{S}_n} \pi_j = \frac{\prod_{s=0}^{n-d-1} [\varepsilon \frac{s}{K} + (1-\varepsilon)(1-p)] \prod_{t=0}^{d-1} [\varepsilon \frac{t}{K} + (1-\varepsilon)p]}{\prod_{l=K-n+1}^{K} (1 - \varepsilon \frac{l}{K})}, \tag{3.16}$$

  where $\prod_{i=0}^{a}(\cdot) \stackrel{\triangle}{=} 1$, if $a < 0$ and $d$ is the number of "ones" in the block $\underline{a}$.

- For block length $n \geq K+1$, the block transition probability of this channel is

$$\begin{aligned}
& \Pr(\underline{Y} = \underline{y} | \underline{X} = \underline{x}) \\
= \; & \Pr(\underline{Z} = \underline{a}) \\
= \; & \prod_{i=1}^{n} \Pr(Z_i = a_i | Z_{i-1} = a_{i-1}, \cdots, Z_1 = a_1) \\
= \; & L \prod_{i=K+1}^{n} \left[ \varepsilon \frac{\lambda_{i-1}}{K} + (1-\varepsilon)p \right]^{a_i} \left[ \varepsilon \frac{K - \lambda_{i-1}}{K} + (1-\varepsilon)(1-p) \right]^{1-a_i},
\end{aligned} \tag{3.17}$$

where

$$L = \frac{\prod_{j=0}^{K-1-\lambda_K} [\varepsilon \frac{j}{K} + (1-\varepsilon)(1-p)] \prod_{l=0}^{\lambda_K - 1} [\varepsilon \frac{l}{K} + (1-\varepsilon)p]}{\prod_{t=1}^{K} (1 - \varepsilon \frac{t}{K})},$$

$\prod_{i=0}^{a} (\cdot) \stackrel{\triangle}{=} 1$, if $a < 0$ and $\lambda_{i-1} = a_{i-1} + \cdots + a_{i-K}$.

## 3.3 Capacity of the Uniform Queue-Based Channel

Using the results in the previous sections, we arrive at the following proposition.

**Proposition 5** *The uniform queue-based channel with memory is a channel with stationary ergodic Markov additive noise of memory $K$. The capacity $C_K$ of the channel is positive and non-decreasing in $K$ and is given by*

$$C_K = 1 - \sum_{\omega(i)=0}^{K} \binom{K}{\omega(i)} L_{\omega(i)} h_b \left( \varepsilon \frac{\omega(i)}{K} + (1-\varepsilon)p \right) \tag{3.18}$$

*where*

$$L_{\omega(i)} = \frac{\prod_{j=0}^{K-1-\omega(i)} [\varepsilon \frac{j}{K} + (1-\varepsilon)(1-p)] \prod_{l=0}^{\omega(i)-1} [\varepsilon \frac{l}{K} + (1-\varepsilon)p]}{\prod_{m=1}^{K} (1 - \varepsilon \frac{m}{K})},$$

$\prod_{i=0}^{a} (\cdot) \stackrel{\triangle}{=} 1$ *if $a < 0$, and $h_b(\cdot)$ is the binary entropy function.*

**Proof** The binary channel has an additive noise process $Z_n$, where $Z_n$ is a stationary ergodic $K^{th}$ order Markov process. Its capacity is then:

$$\begin{aligned}
C_K &= \lim_{n\to\infty} \sup_{\underline{X}} \frac{1}{n} I(\underline{X};\underline{Y}) \\
&= 1 - \lim_{n\to\infty} \frac{1}{n} H(\underline{Z}) \\
&= 1 - h(Z_\infty) \\
&= 1 - H(Z_{K+1}|Z_K, Z_{K-1}, \cdots, Z_1) \\
&= 1 - \sum_{\omega(i)=0}^{K} \binom{K}{\omega(i)} L_{\omega(i)} h_b\left(\varepsilon \frac{\omega(i)}{K} + (1-\varepsilon)p\right) \qquad (3.19)
\end{aligned}$$

from (3.15).

$C_K > 0$ since the noise entropy rate is bounded above by 1. Finally, we know that conditioning decreases entropy; thus $H(Z_{K+1}|Z_K, Z_{K-1}, \cdots, Z_1)$ decreases as $K$ increases. Therefore the capacity $C_K$ increases as the memory $K$ increases. $\square$

## 3.4 The Error-Gap Process for the Uniform Queue-Based Channel

The error process $\{Z_n\}_{n=1}^{\infty}$ of the uniform queue-based channel is stationary. According to Definition 2 in Chapter 2, we obtain the gap frequency $\Pr[G_n = l]$ and the average duration of the gap $E(G_n)$ for the uniform queue-based channel:

- For $1 \le l \le K-1$, where $K > 1$,

$$\begin{aligned}
\Pr[G_n = l] &= \Pr[Z_{n+1} = 0, \cdots, Z_{n+l-1} = 0, Z_{n+l} = 1 | Z_n = 1] \\
&= \Pr[Z_2 = 0, \cdots, Z_l = 0, Z_{l+1} = 1 | Z_1 = 1] \\
&= \frac{\Pr[Z_1 = 1, Z_2 = 0, \cdots, Z_l = 0, Z_{l+1} = 1]}{\Pr[Z_1 = 1]}.
\end{aligned}$$

From (3.16) with $n = l+1$ and $d = 2$, we get

$$\Pr[G_n = l] = \frac{1}{p} \cdot \frac{\prod_{s=0}^{l-2}[\varepsilon\frac{s}{K} + (1-\varepsilon)(1-p)]\prod_{t=0}^{1}[\varepsilon\frac{t}{K} + (1-\varepsilon)p]}{\prod_{u=K-l}^{K}(1-\varepsilon\frac{u}{K})}, \tag{3.20}$$

where $\Pr[Z_1 = 1] = p$ from (3.16) with $n = d = 1$.

- Similarly, for $l = K$,

$$\begin{aligned}
&\Pr[G_n = l] \\
= \;& \Pr[Z_{n+1} = 0, \cdots, Z_{n+l-1} = 0, Z_{n+l} = 1 | Z_n = 1] \\
= \;& \Pr[Z_2 = 0, \cdots, Z_K = 0, Z_{K+1} = 1 | Z_1 = 1] \\
= \;& \frac{\Pr[Z_1 = 1, Z_2 = 0, \cdots, Z_K = 0, Z_{K+1} = 1]}{\Pr[Z_1 = 1]} \\
= \;& \frac{\Pr[Z_1 = 1, Z_2 = 0, \cdots, Z_K = 0]\Pr[Z_{K+1} = 1 | Z_K = 0, \cdots, Z_2 = 0, Z_1 = 1]}{p}
\end{aligned}$$

From (3.16) with $n = K$ and $d = 1$, and (3.13), we get

$$\Pr[G_n = K] = \frac{1}{p} \cdot \frac{\left\{\prod_{s=0}^{K-2}[\varepsilon\frac{s}{K} + (1-\varepsilon)(1-p)]\right\}[(1-\varepsilon)p]}{\prod_{u=1}^{K}(1-\varepsilon\frac{u}{K})} \cdot [\varepsilon\frac{1}{K} + (1-\varepsilon)p]. \tag{3.21}$$

- For $l \geq K+1$,

$$\begin{aligned}
&\Pr[G_n = l] \\
= \;& \Pr[Z_{n+1} = 0, \cdots, Z_{n+l-1} = 0, Z_{n+l} = 1 | Z_n = 1] \\
= \;& \Pr[Z_2 = 0, \cdots, Z_l = 0, Z_{l+1} = 1 | Z_1 = 1] \\
= \;& \frac{\Pr[Z_1 = 1, Z_2 = 0, \cdots, Z_l = 0, Z_{l+1} = 1]}{\Pr[Z_1 = 1]}. \\
= \;& \frac{\Pr[Z_1 = 1, Z_2 = 0, \cdots, Z_K = 0, Z_{K+1} = 0, \cdots, Z_l = 0, Z_{l+1} = 1]}{p}. \\
= \;& \frac{1}{p} \cdot \Pr[Z_1 = 1, Z_2 = 0, \cdots, Z_K = 0]\Pr[Z_{K+1} = 0 | Z_K = 0, \cdots, Z_2 = 0, Z_1 = 1] \\
&\cdot \left\{\prod_{i=K+2}^{l} \Pr[Z_i = 0 | Z_{i-1} = 0, \cdots, Z_{i-K} = 0]\right\} \\
&\cdot \Pr[Z_{l+1} = 1 | Z_l = 0, \cdots, Z_{l-K+1} = 0].
\end{aligned}$$

From (3.16) with $n = K$ and $d = 1$, and (3.13), we get

$$\begin{aligned}\Pr[G_n = l] &= \frac{1}{p} \cdot \frac{\left\{\prod_{s=0}^{K-2}[\varepsilon \frac{s}{K} + (1-\varepsilon)(1-p)]\right\}[(1-\varepsilon)p]}{\prod_{u=1}^{K}(1-\varepsilon\frac{u}{K})} \cdot [\varepsilon\frac{K-1}{K} + (1-\varepsilon)(1-p)] \\ &\quad \cdot [\varepsilon + (1-\varepsilon)(1-p)]^{l-K-1} \cdot [(1-\varepsilon)p]. \end{aligned} \tag{3.22}$$

Finally,

$$E(G_n) = \frac{1}{P(1)} = \frac{1}{\Pr[Z_n = 1]} = \frac{1}{p}. \tag{3.23}$$

## 3.5 The Error-Burst Process for the Uniform Queue-Based Channel

We give the definition of the error-burst process, which is analogous to the definition of the error-gap process.

**Definition 4** *Sequences of ones between two zeros (0 states) are called error bursts. The length of a burst is defined as one plus the total number of ones in the sequence between two 0's.*

Let $B_n$ denote the length of the burst starting at $Z_n$, conditioned on $Z_n = 0$. Assuming stationarity, the following equations hold:

$$\begin{aligned}\Pr[B_n = m] &= \Pr[Z_{n+1} = 1, \cdots, Z_{n+m-1} = 1, Z_{n+m} = 0 | Z_n = 0] \\ &= P(1^{m-1}0|0), \end{aligned} \tag{3.24}$$

and

$$E(B_n) = \frac{1}{P(0)} \tag{3.25}$$

for all positive integers $m$, where $P(0)$ is the probability of no error and is equal to $1 - E(Z_n)$.

Eq. (3.25) is shown as follows:

$$\begin{aligned} E(B_n) &= \sum_{l=1}^{\infty} lP(1^{l-1}0|0) \\ &= \sum_{l=0}^{\infty}(l+1)P(1^l0|0) \\ &= \sum_{l=0}^{\infty}(l+1)[P(1^l|0) - P(1^{l+1}|0)] \\ &= \sum_{l=0}^{\infty} P(1^l|0) \\ &= \sum_{l=0}^{\infty} \frac{P(01^l)}{P(0)} \\ &= \frac{\sum_{l=0}^{\infty} P(01^l)}{P(0)}, \end{aligned}$$

where

$$\begin{aligned} \sum_{l=0}^{\infty} P(01^l) &= P(0) + P(01) + P(011) + P(0111) + \cdots + \cdots \\ &= 1 - P(1) + P(01) + P(011) + P(0111) + \cdots + \cdots \\ &= 1 - [P(01) + P(11)] + P(01) + P(011) + P(0111) + \cdots + \cdots \\ &= 1 - P(11) + P(011) + P(0111) + \cdots + \cdots \\ &= 1 - [P(011) + P(111)] + P(011) + P(0111) + \cdots + \cdots \\ &= 1 - P(111) + P(0111) + \cdots + \cdots \\ &\vdots \\ &\to 1 - P(1^{\infty}), \end{aligned}$$

where

$$\begin{aligned} P(1^{\infty}) &= \Pr[Z_1 = 1, \cdots, Z_K = 1] \prod_{i=K+1}^{\infty} \Pr[Z_i = 1 | Z_{i-1} = 1, \cdots, Z_{i-K} = 1] \\ &= \pi_{2^K-1}[\varepsilon + (1-\varepsilon)p]^{\infty} = 0 \end{aligned}$$

since $\varepsilon + (1-\varepsilon)p < 1$.

So

$$E(B_n) = \frac{1 - P(1^\infty)}{P(0)} = \frac{1}{P(0)}.$$

Using the same approach as for the gap process, we get the burst frequency $\Pr[B_n = l]$ and the average duration of the burst $E(B_n)$ for the uniform queue-based channel:

- For $1 \le l \le K-1$, where $K > 1$,

$$\begin{aligned}
& \Pr[B_n = l] \\
= \; & \Pr[Z_{n+1} = 1, \cdots, Z_{n+l-1} = 1, Z_{n+l} = 0 | Z_n = 0] \\
= \; & \Pr[Z_2 = 1, \cdots, Z_l = 1, Z_{l+1} = 0 | Z_1 = 0] \\
= \; & \frac{\Pr[Z_1 = 0, Z_2 = 1, \cdots, Z_l = 1, Z_{l+1} = 0]}{\Pr[Z_1 = 0]}.
\end{aligned}$$

From (3.16) with $n = l+1$ and $d = l-1$, we get

$$\Pr[B_n = l] = \frac{1}{1-p} \cdot \frac{\prod_{s=0}^{1}[\varepsilon \frac{s}{K} + (1-\varepsilon)(1-p)] \prod_{t=0}^{l-2}[\varepsilon \frac{t}{K} + (1-\varepsilon)p]}{\prod_{u=K-l}^{K}(1 - \varepsilon \frac{u}{K})}. \qquad (3.26)$$

- For $l = K$,

$$\begin{aligned}
& \Pr[B_n = l] \\
= \; & \Pr[Z_{n+1} = 1, \cdots, Z_{n+K-1} = 1, Z_{n+K} = 0 | Z_n = 0] \\
= \; & \Pr[Z_2 = 1, \cdots, Z_K = 1, Z_{K+1} = 0 | Z_1 = 0] \\
= \; & \frac{\Pr[Z_1 = 0, Z_2 = 1, \cdots, Z_K = 1, Z_{K+1} = 0]}{\Pr[Z_1 = 0]} \\
= \; & \frac{\Pr[Z_1 = 0, Z_2 = 1, \cdots, Z_K = 1] \Pr[Z_{K+1} = 0 | Z_K = 1, \cdots, Z_2 = 1, Z_1 = 0]}{1-p}.
\end{aligned}$$

From (3.16) with $n = K$ and $d = K-1$, and (3.13), we get

$$\begin{aligned}
\Pr[B_n = K] \; = \; & \frac{1}{1-p} \cdot \frac{[(1-\varepsilon)(1-p)] \prod_{t=0}^{K-2}[\varepsilon \frac{t}{K} + [(1-\varepsilon)p]}{\prod_{u=1}^{K}(1 - \varepsilon \frac{u}{K})} \\
& \cdot [\varepsilon \frac{1}{K} + (1-\varepsilon)(1-p)]. \qquad (3.27)
\end{aligned}$$

- For $l \geq K+1$,

$$\begin{aligned}
& \Pr[B_n = l] \\
= \ & \Pr[Z_{n+1} = 1, \cdots, Z_{n+l-1} = 1, Z_{n+l} = 0 | Z_n = 0] \\
= \ & \Pr[Z_2 = 1, \cdots, Z_l = 1, Z_{l+1} = 0 | Z_1 = 0] \\
= \ & \frac{\Pr[Z_1 = 0, Z_2 = 1, \cdots, Z_l = 1, Z_{l+1} = 0]}{\Pr[Z_1 = 0]}. \\
= \ & \frac{\Pr[Z_1 = 0, Z_2 = 1, \cdots, Z_K = 1, Z_{K+1} = 1, \cdots, Z_l = 1, Z_{l+1} = 0]}{1-p}. \\
= \ & \frac{1}{1-p} \cdot \Pr[Z_1 = 0, Z_2 = 1, \cdots, Z_K = 1] \\
& \cdot \Pr[Z_{K+1} = 1 | Z_K = 1, \cdots, Z_2 = 1, Z_1 = 0] \\
& \cdot \left\{ \prod_{i=K+2}^{l} \Pr[Z_i = 1 | Z_{i-1} = 1, \cdots, Z_{i-K} = 1] \right\} \\
& \cdot \Pr[Z_{l+1} = 0 | Z_l = 1, \cdots, Z_{l-K+1} = 1].
\end{aligned}$$

From (3.16) with $n = K$ and $d = K-1$, and (3.13), we get

$$\begin{aligned}
& \Pr[B_n = l] \\
= \ & \frac{1}{1-p} \cdot \frac{[(1-\varepsilon)(1-p)] \prod_{t=0}^{K-2} [\varepsilon \frac{t}{K} + [(1-\varepsilon)p]}{\prod_{u=1}^{K} (1 - \varepsilon \frac{u}{K})} \\
& \cdot [\varepsilon \frac{K-1}{K} + (1-\varepsilon)p] \cdot [\varepsilon + (1-\varepsilon)p]^{l-K-1} \cdot [(1-\varepsilon)(1-p)]. \qquad (3.28)
\end{aligned}$$

Finally,

$$E(B_n) = \frac{1}{P(0)} = \frac{1}{\Pr[Z_n = 0]} = \frac{1}{1-p}. \qquad (3.29)$$

## 3.6 The Non-Uniform Queue-Based Channel

The non-uniform queue-based channel has a similar transition matrix as the uniform queue-based channel. For example, for the case $K = 3$, the transition matrix of the non-uniform queue-based channel is given by the following.

$$
\mathbf{Q}=\begin{bmatrix}
\varepsilon+(1-\varepsilon)(1-p) & 0 & 0 & 0 & (1-\varepsilon)p & 0 & 0 & 0 \\
\varepsilon(q_1+q_2)+(1-\varepsilon)(1-p) & 0 & 0 & 0 & \varepsilon\cdot q_3+(1-\varepsilon)p & 0 & 0 & 0 \\
0 & \varepsilon(q_1+q_3)+(1-\varepsilon)(1-p) & 0 & 0 & 0 & \varepsilon\cdot q_2+(1-\varepsilon)p & 0 & 0 \\
0 & \varepsilon\cdot q_1+(1-\varepsilon)(1-p) & 0 & 0 & 0 & \varepsilon(q_2+q_3)+(1-\varepsilon)p & 0 & 0 \\
0 & 0 & \varepsilon(q_2+q_3)+(1-\varepsilon)(1-p) & 0 & 0 & 0 & \varepsilon\cdot q_1+(1-\varepsilon)p & 0 \\
0 & 0 & \varepsilon\cdot q_2+(1-\varepsilon)(1-p) & 0 & 0 & 0 & \varepsilon(q_1+q_3)+(1-\varepsilon)p & 0 \\
0 & 0 & 0 & \varepsilon\cdot q_3+(1-\varepsilon)(1-p) & 0 & 0 & 0 & \varepsilon(q_1+q_2)+(1-\varepsilon)p \\
0 & 0 & 0 & (1-\varepsilon)(1-p) & 0 & 0 & 0 & \varepsilon+(1-\varepsilon)p
\end{bmatrix}.
$$

Given $\varepsilon$, $p$, $q_1$, we can calculate the stationary distribution $\boldsymbol{\pi}=[\pi_0,\pi_1,\cdots,\pi_7]$ numerically.

The capacity of the non-uniform queue-based channel with $K=3$ is

$$
\begin{aligned}
C_3 &= 1-h(Z_\infty) \\
&\triangleq 1-\lim_{n\to\infty}\frac{1}{n}H(Z_1,\cdots,Z_n) \\
&= 1-H(Z_4|Z_3,Z_2,Z_1) \\
&= 1-H(\underline{S}_2|\underline{S}_1) \\
&= 1-[-\sum_{i,j=0}^{7}\pi_i p_{ij}\log_2 p_{ij}],
\end{aligned}
\tag{3.30}
$$

where $p_{ij}$ is the $(i,j)^{th}$ entry of the transition matrix. Similarly, we can get the results for the cases $K=2$ and $K=4$.

The non-uniform queue-based channel has a similar burst frequency as the uniform queue-based channel. For example, for the case $K=2$, the burst frequency of the non-uniform queue-based channel is given by the following.

- For $1 \leq l \leq K-1$, where $K=2$,

$$
\begin{aligned}
& \Pr[B_n = l] \\
= \; & \Pr[Z_{n+1} = 1, \cdots, Z_{n+l-1} = 1, Z_{n+l} = 0 | Z_n = 0] \\
= \; & \frac{\Pr[Z_1 = 0, Z_2 = 0]}{\Pr[Z_1 = 0]} \\
= \; & \frac{\pi_0}{1-p}.
\end{aligned} \tag{3.31}
$$

- For $l = K$, where $K = 2$,

$$
\begin{aligned}
& \Pr[B_n = l] \\
= \; & \Pr[Z_{n+1} = 1, \cdots, Z_{n+K-1} = 1, Z_{n+K} = 0 | Z_n = 0] \\
= \; & \frac{\Pr[Z_1 = 0, Z_2 = 1, Z_3 = 0]}{\Pr[Z_1 = 0]} \\
= \; & \frac{\Pr[Z_1 = 0, Z_2 = 1] \Pr[Z_3 = 0 | Z_2 = 1, Z_1 = 0]}{1-p} \\
= \; & \frac{\pi_2}{1-p} \cdot [\varepsilon(1-q_1) + (1-\varepsilon)(1-p)].
\end{aligned} \tag{3.32}
$$

- For $l \geq K+1$, where $K = 2$,

$$
\begin{aligned}
& \Pr[B_n = l] \\
= \; & \Pr[Z_{n+1} = 1, \cdots, Z_{n+l-1} = 1, Z_{n+l} = 0 | Z_n = 0] \\
= \; & \frac{\Pr[Z_1 = 0, Z_2 = 1, \cdots, Z_l = 1, Z_{l+1} = 0]}{\Pr[Z_1 = 0]} \\
= \; & \frac{\Pr[Z_1 = 0, Z_2 = 1, Z_3 = 1, \cdots, Z_l = 1, Z_{l+1} = 0]}{1-p} \\
= \; & \frac{\Pr[Z_1 = 0, Z_2 = 1]}{1-p} \cdot \Pr[Z_3 = 1 | Z_2 = 1, Z_1 = 0] \\
& \cdot \left\{ \prod_{i=4}^{l} \Pr[Z_i = 1 | Z_{i-1} = 1, Z_{i-2} = 1] \right\} \Pr[Z_{l+1} = 0 | Z_l = 1, Z_{l-1} = 1] \\
= \; & \frac{\pi_2}{1-p} \cdot [\varepsilon q_1 + (1-\varepsilon)p] \cdot [\varepsilon + (1-\varepsilon)p]^{l-3} \cdot [(1-\varepsilon)(1-p)].
\end{aligned} \tag{3.33}
$$

# Chapter 4

# Comparisons with Other Channels with Memory

In this chapter, we compare the uniform queue-based channel with the finite-memory contagion channel and the Gilbert-Elliott channel in terms of capacity, burst frequency and average burst duration. Furthermore, we study the capacity of the non-uniform queue-based channel with memory and compare it with that of the uniform queue-based channel with memory, and we compare the burst frequency of the non-uniform queue-based channel with memory with that of the uniform queue-based channel with memory and the Gilbert-Elliott channel.

In Section 4.1, we first rewrite the block transition probability in terms of the channel memory, $BER$ and correlation coefficient for the finite-memory contagion channel and the uniform queue-based channel with memory. We suprisingly find that the two channels have the *same* block transition probability for the same memory, $BER$ and correlation coefficient value. In Section 4.2, we rewrite the capacity in terms of the channel memory, $BER$ and correlation coefficient for the uniform queue-based channel with memory (which is the same as the capacity of the finite-memory contagion channel) and the Gilbert-Elliott channel. This gives us expressions for channel capacity upon which the plots in Sections 4.3 and 4.4 are based. In Section 4.3, we provide plots of channel capacity vs. memory for various $BER$ and correlation combinations for all three channels. In Section 4.4, we provide plots of channel capacity vs. $BER$ for various memory and correlation combinations for all three channels. In Section 4.5, we provide expressions for the burst frequency and the average burst duration for the finite-memory contagion channel and the Gilbert-Elliott channel, and provide plots of burst frequency vs. burst length for all three channel models. In Section 4.6, we compare the capacity of the non-uniform queue-based channel with memory with the capacity of the uniform queue-based channel with memory. In Section 4.7, we compare the burst frequency of the non-uniform queue-based channel with memory with the burst frequency of the uniform queue-based channel with memory and the Gilbert-Elliott channel. Finally, we conclude the chapter with a discussion in Section 4.8.

## 4.1 Block Transition Probability in Terms of Memory, $BER$ and Correlation Coefficient

### 4.1.1 The finite-memory contagion channel

For the finite-memory contagion channel, we know that the noise process is stationary and hence identically distributed [2]. We get the bit error rate ($BER$)

$$BER = \Pr[Z_i = 1] = \rho, \tag{4.1}$$

and the correlation coefficient (using $Cor$ instead of $Cor(Z_i, Z_{i+1})$ for brevity)

$$Cor = \frac{Cov(Z_i, Z_{i+1})}{\sqrt{Var(Z_i)Var(Z_{i+1})}} = \frac{\delta}{1+\delta} > 0, \tag{4.2}$$

for $i = 1, 2, 3, \cdots$. From (4.1) and (4.2) we can express $\rho$ and $\delta$ in terms of $BER$ and $Cor$.

$$\rho = BER, \tag{4.3}$$

and

$$\delta = \frac{Cor}{1 - Cor}. \tag{4.4}$$

Rewriting (2.1) and (2.2), we get the following in terms of $BER$, $Cor$ and $M$.

- For block length $n \leq M$, the block transition probability of this channel is

$$P(\underline{Y} = \underline{y} | \underline{X} = \underline{x}) = \frac{\prod_{s=0}^{d-1}(BER + s\frac{Cor}{1-Cor}) \prod_{j=0}^{n-1-d}((1 - BER) + j\frac{Cor}{1-Cor})}{\prod_{l=1}^{n-1}(1 + l\frac{Cor}{1-Cor})}, \tag{4.5}$$

  where $\prod_{i=0}^{a}(\cdot) \stackrel{\triangle}{=} 1$, if $a < 0$ and $d = d(\underline{y}, \underline{x})$=weight($\underline{z} = \underline{y} \oplus \underline{x}$).

- For $n \geq M + 1$,the block transition probability of this channel is

$$\begin{aligned} & P_M(\underline{Y} = \underline{y} | \underline{X} = \underline{x}) \\ = \ & P(\underline{Z} = \underline{e}) \\ = \ & L \prod_{i=M+1}^{n} \left[ \frac{BER + \lambda_{i-1}\frac{Cor}{1-Cor}}{1 + M\frac{Cor}{1-Cor}} \right]^{e_i} \left[ \frac{(1 - BER) + (M - \lambda_{i-1})\frac{Cor}{1-Cor}}{1 + M\frac{Cor}{1-Cor}} \right]^{1-e_i}, \end{aligned} \tag{4.6}$$

  where

$$L = \frac{\prod_{s=0}^{k-1}(BER + s\frac{Cor}{1-Cor}) \prod_{j=0}^{M-1-k}((1 - BER) + j\frac{Cor}{1-Cor})}{\prod_{l=1}^{M-1}(1 + l\frac{Cor}{1-Cor})},$$

  $\prod_{i=0}^{a}(\cdot) \stackrel{\triangle}{=} 1$, if $a < 0$, $e_i = x_i \oplus y_i$, $k = e_1 + \cdots + e_M$ and $\lambda_{i-1} = e_{i-1} + \cdots + e_{i-M}$.

### 4.1.2 The uniform queue-based channel with memory

From the previous chapter, we know that the noise process of the queue-based channel is stationary; hence it is identically distributed. We first calculate the $BER$ and the correlation coefficient.

$$BER = \Pr[Z_i = 1] = \Pr[Z_1 = 1] = p \tag{4.7}$$

from Section 3.4.

Also

$$Cor = \frac{Cov(Z_i, Z_{i+1})}{\sqrt{Var(Z_i)Var(Z_{i+1})}} = \frac{Cov(Z_2, Z_1)}{Var(Z_1)}.$$

We have

$$Cov(Z_2, Z_1) = E[Z_2 Z_1] - E[Z_2]E[Z_1] = \Pr[Z_2 = 1, Z_1 = 1] - p^2,$$

and

$$Var(Z_1) = E[Z_1^2] - E[Z_1]^2 = \Pr[Z_1 = 1] - p^2.$$

Using (3.16) with $n = 2$ and $d = 2$, we get

$$Cov(Z_2, Z_1) = \frac{\varepsilon\frac{1}{K}}{1 - \varepsilon\frac{K-1}{K}} p(1-p).$$

Since $\Pr[Z_1 = 1] = p$, we have

$$Var(Z_1) = p - p^2 = p(1-p).$$

Thus

$$Cor = \frac{\frac{\varepsilon\frac{1}{K}}{1-\varepsilon\frac{K-1}{K}} p(1-p)}{p(1-p)} = \frac{\varepsilon\frac{1}{K}}{1 - \varepsilon\frac{K-1}{K}}. \tag{4.8}$$

We then can deduce the parameters $p$ and $\varepsilon$ from (4.7) and (4.8):

$$p = BER, \tag{4.9}$$

and

$$\varepsilon = \frac{K \cdot Cor}{(K-1) \cdot Cor + 1}. \tag{4.10}$$

Rewriting (3.16) and (3.17), we get the following in terms of $BER$, $Cor$ and $K$.

- For block length $n \le K$, the block transition probability of this channel is

$$
\begin{aligned}
& \Pr(\underline{Y}=\underline{y}|\underline{X}=\underline{x}) \\
= \; & \Pr(\underline{Z}=\underline{a}) \\
= \; & \frac{1}{\prod_{l=K-n+1}^{K}\left(1-\frac{K\cdot Cor}{(K-1)\cdot Cor+1}\frac{l}{K}\right)} \\
& \cdot \prod_{s=0}^{n-d-1}\left[\frac{K\cdot Cor}{(K-1)\cdot Cor+1}\cdot\frac{s}{K}+\left(1-\frac{K\cdot Cor}{(K-1)\cdot Cor+1}\right)(1-BER)\right] \\
& \cdot \prod_{t=0}^{d-1}\left[\frac{K\cdot Cor}{(K-1)\cdot Cor+1}\cdot\frac{t}{K}+\left(1-\frac{K\cdot Cor}{(K-1)\cdot Cor+1}\right)BER\right] \\
= \; & \frac{\prod_{s=0}^{n-d-1}[s\frac{Cor}{1-Cor}+(1-BER)]\prod_{t=0}^{d-1}[t\frac{Cor}{1-Cor}+BER]}{\prod_{l=K-n+1}^{K}[1+(K-l)\frac{Cor}{1-Cor}]} \\
= \; & \frac{\prod_{s=0}^{n-d-1}[s\frac{Cor}{1-Cor}+(1-BER)]\prod_{t=0}^{d-1}[t\frac{Cor}{1-Cor}+BER]}{\prod_{r=1}^{n-1}[1+r\frac{Cor}{1-Cor}]},
\end{aligned}
\tag{4.11}
$$

where $d$ is the number of "ones" in block $\underline{a}$, and $\prod_{i=0}^{a}(\cdot)\stackrel{\triangle}{=}1$, if $a<0$.

- For block length $n \ge K+1$, the block transition probability of this channel is

$$
\begin{aligned}
& \Pr(\underline{Y}=\underline{y}|\underline{X}=\underline{x}) \\
= \; & \Pr(\underline{Z}=\underline{a}) \\
= \; & L\left\{\prod_{i=K+1}^{n}\left[\frac{K\cdot Cor}{(K-1)\cdot Cor+1}\cdot\frac{\lambda_{i-1}}{K}+\left(1-\frac{K\cdot Cor}{(K-1)\cdot Cor+1}\right)BER\right]^{a_i}\right. \\
& \left.\cdot\left[\frac{K\cdot Cor}{(K-1)\cdot Cor+1}\cdot\frac{K-\lambda_{i-1}}{K}+\left(1-\frac{K\cdot Cor}{(K-1)\cdot Cor+1}\right)(1-BER)\right]^{1-a_i}\right\} \\
= \; & L\prod_{i=K+1}^{n}\left[\frac{BER+\lambda_{i-1}\frac{Cor}{1-Cor}}{1+K\frac{Cor}{1-Cor}}\right]^{a_i}\left[\frac{(1-BER)+(K-\lambda_{i-1})\frac{Cor}{1-Cor}}{1+K\frac{Cor}{1-Cor}}\right]^{1-a_i},
\end{aligned}
\tag{4.12}
$$

where

$$
L = \frac{1}{\prod_{t=1}^{K}\left(1-\frac{t\cdot Cor}{(K-1)\cdot Cor+1}\right)}
$$

$$
\begin{aligned}
&\cdot \prod_{j=0}^{K-1-\lambda_K} \left[\frac{j \cdot Cor}{(K-1)\cdot Cor+1} + \left(1 - \frac{K\cdot Cor}{(K-1)\cdot Cor+1}\right)(1-BER)\right] \\
&\cdot \prod_{l=0}^{\lambda_K-1} \left[\frac{l \cdot Cor}{(K-1)\cdot Cor+1} + \left(1 - \frac{K\cdot Cor}{(K-1)\cdot Cor+1}\right)BER\right] \\
= \quad & \frac{\prod_{j=0}^{K-1-\lambda_K}[j\frac{Cor}{1-Cor} + (1-BER)]\prod_{l=0}^{\lambda_K-1}[l\frac{Cor}{1-Cor}+BER]}{\prod_{t=1}^{K}[1+(K-t)\frac{Cor}{1-Cor}]} \\
= \quad & \frac{\prod_{j=0}^{K-1-\lambda_K}[j\frac{Cor}{1-Cor} + (1-BER)]\prod_{l=0}^{\lambda_K-1}[l\frac{Cor}{1-Cor}+BER]}{\prod_{r=1}^{K-1}[1+r\frac{Cor}{1-Cor}]},
\end{aligned}
\tag{4.13}
$$

$\prod_{i=0}^{a}(\cdot)\overset{\triangle}{=}1$, if $a<0$, and $\lambda_{i-1} = a_{i-1} + \cdots + a_{i-K}$, $i = K+1, \cdots, n$.

Comparing (4.11) with (4.5) and (4.12) with (4.6), we can surprisingly observe that the finite-memory contagion channel and the uniform queue-based channel with memory have *identical* block transition probabilities for the same memory, $BER$ and correlation coefficient value.

### 4.1.3 The non-uniform queue-based channel with memory

For the non-uniform queue-based channel, we get the $BER$ and the correlation coefficient with memory $K = 3$ as follows.

$$
BER = \Pr[Z_i = 1] = \Pr[Z_1 = 1] = \pi_4 + \pi_5 + \pi_6 + \pi_7 = p, \tag{4.14}
$$

and

$$
\begin{aligned}
Cor &= \frac{Cov(Z_i, Z_{i+1})}{\sqrt{Var(Z_i)Var(Z_{i+1})}} \\
&= \frac{E[Z_2Z_1] - E[Z_2]E[Z_1]}{Var(Z_1)} \\
&= \frac{\Pr[Z_2 = 1, Z_1 = 1] - E[Z_1]^2}{E[Z_1^2] - E[Z_1]^2} \\
&= \frac{(\pi_6+\pi_7) - \Pr[Z_1 = 1]^2}{\Pr[Z_1] - \Pr[Z_1 = 1]^2} \\
&= \frac{(\pi_6+\pi_7) - p^2}{p - p^2}
\end{aligned}
\tag{4.15}
$$

Similarly, we can get the results for the cases $K = 2$ and $K = 4$.

## 4.2 Capacity in Terms of Memory, $BER$ and Correlation Coefficient

### 4.2.1 The uniform queue-based channel with memory

For the uniform queue-based channel, rewriting (3.18), we get

$$C_K = 1 - \sum_{\omega(i)=0}^{K} \binom{K}{\omega(i)} L_{\omega(i)} h_b \left( \frac{\omega(i) \cdot Cor}{(K-1) \cdot Cor + 1} + \frac{(1 - Cor)BER}{(K-1) \cdot Cor + 1} \right), \qquad (4.16)$$

where

$$\begin{aligned} L_{\omega(i)} &= \frac{\prod_{j=0}^{K-1-\omega(i)} \left[ \frac{j \cdot Cor}{(K-1) \cdot Cor+1} + \frac{1-Cor}{(K-1) \cdot Cor+1}(1 - BER) \right]}{\prod_{m=1}^{K} \left[ \frac{(1-Cor)+(K-m)Cor}{(K-1) \cdot Cor+1} \right]} \\ &\quad \cdot \prod_{l=0}^{\omega(i)-1} \left[ \frac{l \cdot Cor}{(K-1) \cdot Cor + 1} + \frac{1 - Cor}{(K-1) \cdot Cor + 1} BER \right], \end{aligned}$$

and $\prod_{i=0}^{a}(\cdot) \triangleq 1$, if $a < 0$.

Note that comparing with (2.5), for the same $BER$, memory and correlation coefficient, the capacity of the uniform queue-based channel is identical to that of the finite-memory contagion channel.

### 4.2.2 The Gilbert-Elliott channel

For the Gilbert-Elliott channel, there are four parameters: $g$, $b$, $p_G$, $p_B$. Since the Gilbert-Elliott channel is stationary, we have

$$BER = \Pr[Z_i = 1] = \Pr[Z_1 = 1] = \boldsymbol{\pi}^T \mathbf{P}(1) \mathbf{1}. \qquad (4.17)$$

Substituting matrix (2.46) and vectors (2.47) and (2.48) into (4.17), we get

$$BER = \pi_0 \cdot p_G + \pi_1 \cdot p_B = \frac{g}{b+g} p_G + \frac{b}{b+g} p_B = \frac{\rho p_G + p_B}{\rho + 1}. \tag{4.18}$$

Now we calculate the correlation coefficient.

$$Cor = \frac{Cov(Z_i, Z_{i+1})}{\sqrt{Var(Z_i)Var(Z_{i+1})}} = \frac{Cov(Z_1, Z_2)}{\sqrt{Var(Z_1)Var(Z_2)}} = \frac{Cov(Z_1, Z_2)}{Var(Z_1)}, \tag{4.19}$$

where

$$\begin{aligned} Cov(Z_1, Z_2) &= E[Z_2 Z_1] - E[Z_2]E[Z_1] \\ &= \Pr[Z_2 = 1, Z_1 = 1] - BER^2 \\ &= \boldsymbol{\pi}^T \mathbf{P}(1)\mathbf{P}(1)\mathbf{1} - BER^2. \end{aligned} \tag{4.20}$$

Substituting matrix (2.44) and matrix (2.46) and vectors (2.47) and (2.48) into (4.20), we get

$$\begin{aligned} Cov(Z_1, Z_2) &= (1 - g - b)(BER - p_G)(p_B - BER) \\ &= \mu(BER - p_G)(p_B - BER), \end{aligned} \tag{4.21}$$

and

$$Var(Z_1) = E[Z_1^2] - E[Z_1]^2 = BER - BER^2 = BER(1 - BER). \tag{4.22}$$

Substituting (4.21) and (4.22) into (4.19), we get

$$Cor = \frac{\mu(BER - p_G)(p_B - BER)}{BER(1 - BER)}. \tag{4.23}$$

We next deduce the needed parameters from (4.18) and (4.23). We divide this problem into two parts.

- Fix the values of $p_G$ and $p_B$ and deduce $g$ and $b$:

$$\rho = \frac{p_B - BER}{BER - p_G} \tag{4.24}$$

and

$$\mu = \frac{BER \cdot (1 - BER) \cdot Cor}{(BER - p_G)(p_B - BER)} \tag{4.25}$$

.

Using $\rho = g/b$ and $\mu = 1 - g - b = 1 - \rho \cdot b - b$, we get

$$b = (1-\mu)/(1+\rho) = \frac{1 - \frac{BER(1-BER)\cdot Cor}{(BER-p_G)(p_B-BER)}}{1 + \frac{p_B - BER}{BER - p_G}}, \tag{4.26}$$

and

$$g = \rho \cdot b = \frac{1-\mu}{1+\frac{1}{\rho}} = \frac{1 - \frac{BER(1-BER)\cdot Cor}{(BER-p_G)(p_B-BER)}}{1 + \frac{BER - p_G}{p_B - BER}}. \tag{4.27}$$

- Fix the values of $g$ and $b$ and deduce $p_G$ and $p_B$:

$$p_G = BER - \sqrt{\frac{b \cdot BER(1 - BER) \cdot Cor}{g(1 - g - b)}} \tag{4.28}$$

and

$$p_B = BER + \sqrt{\frac{g \cdot BER(1 - BER) \cdot Cor}{b(1 - g - b)}}. \tag{4.29}$$

In our computations, we will fix the values of the $BER$, the correlation coefficient, $p_G = 0.00002$ and $p_B = 0.92$ to calculate the capacity through (2.38) for $l = 9$. For this choice of $l$, the value of $C$ is accurate up to 6 decimal digits.

## 4.3 Capacity vs. Memory Plots

The following are the figures of capacity vs. memory for different $BER$s and correlation coefficients for the uniform queue-based channel with memory, the finite-memory contagion channel and the Gilbert-Elliott channel. Note that the capacity of the Gilbert-Elliott channel does not vary with memory.

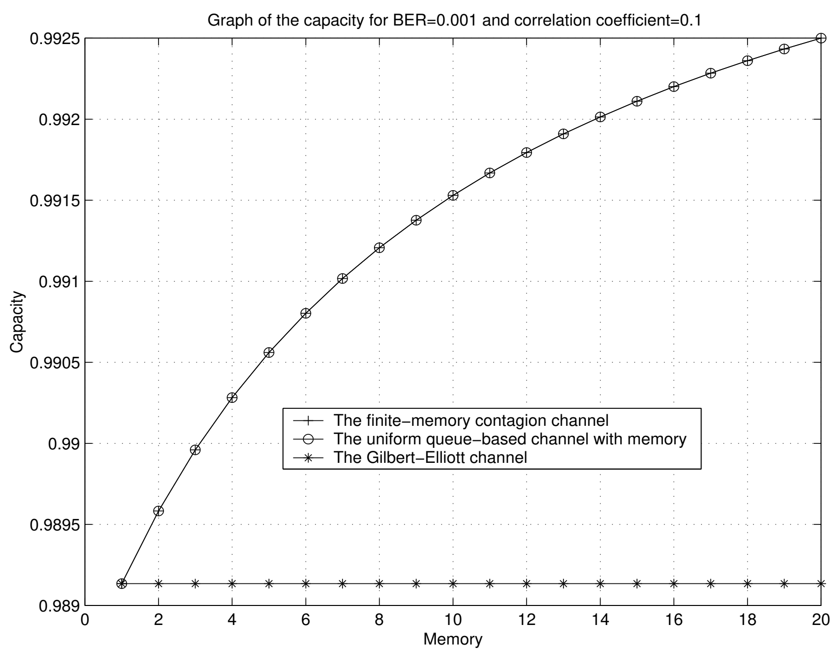



Figure 4.1: Graph of capacity vs. memory for $BER$=0.001 and $Cor$=0.1.

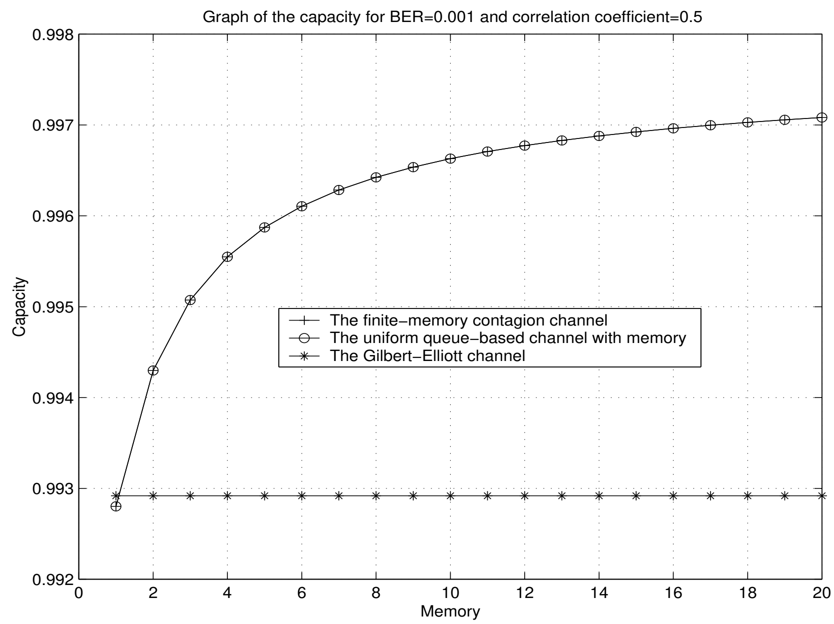



Figure 4.2: Graph of capacity vs. memory for $BER$=0.001 and $Cor$=0.5.

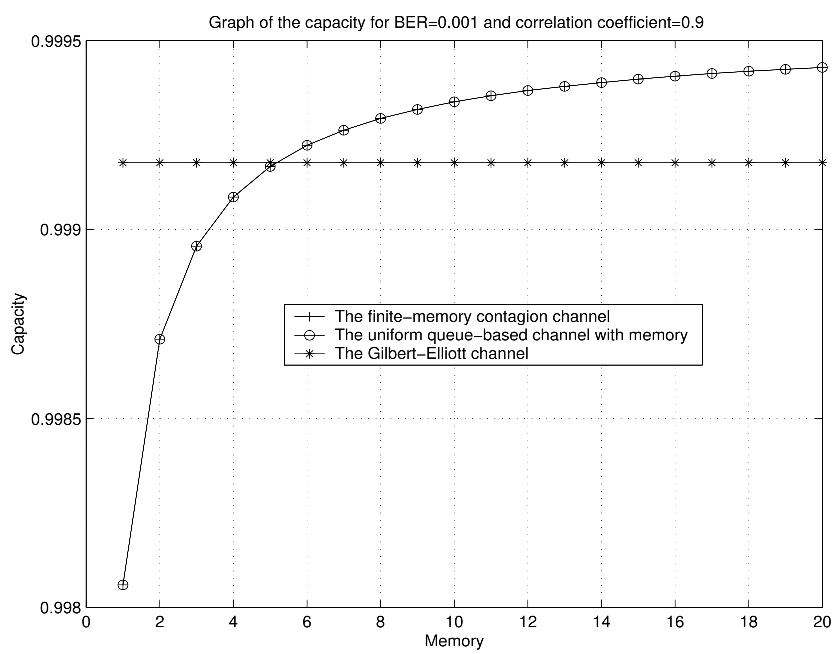



Figure 4.3: Graph of capacity vs. memory for $BER$=0.001 and $Cor$=0.9.

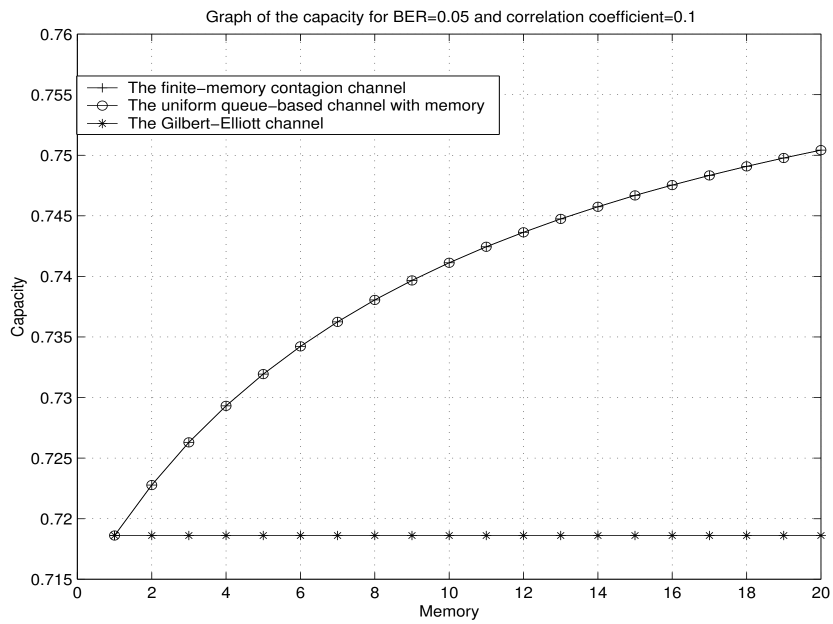



Figure 4.4: Graph of capacity vs. memory for $BER$=0.05 and $Cor$=0.1.

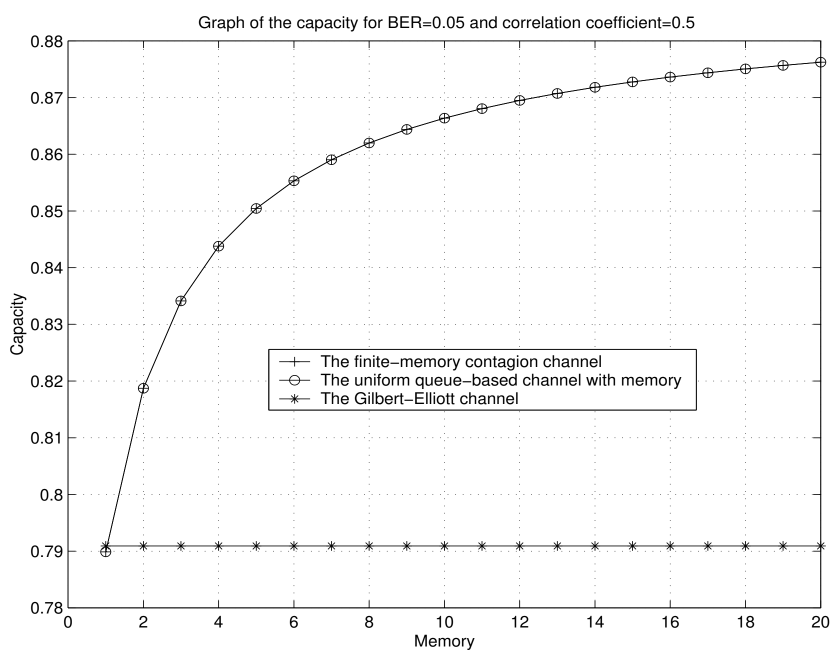



Figure 4.5: Graph of capacity vs. memory for $BER$=0.05 and $Cor$=0.5.

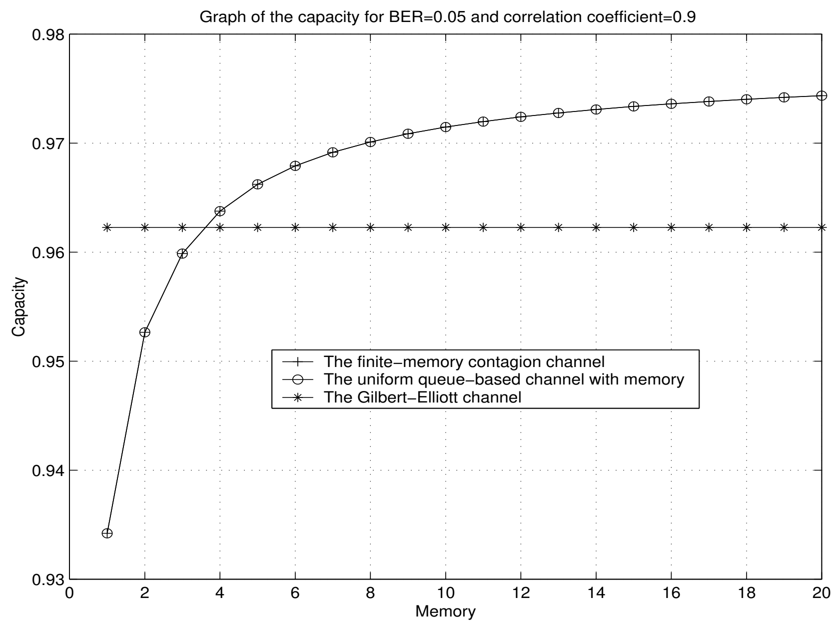



Figure 4.6: Graph of capacity vs. memory for $BER$=0.05 and $Cor$=0.9.

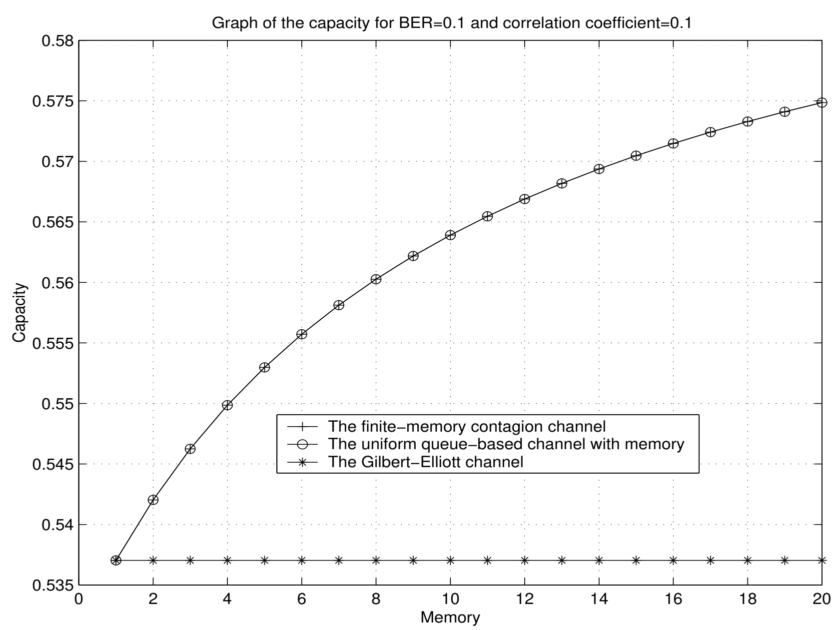



Figure 4.7: Graph of capacity vs. memory for $BER$=0.1 and $Cor$=0.1.

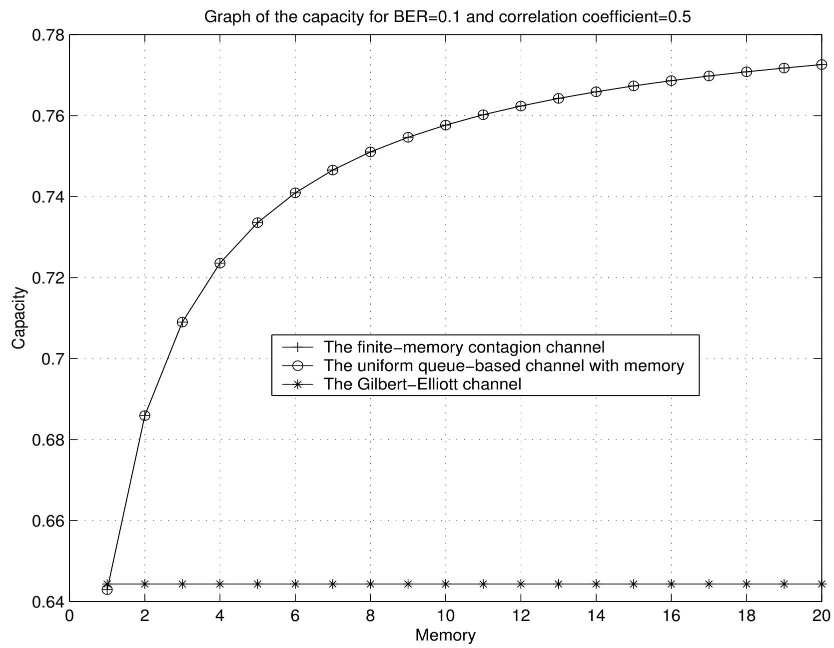



Figure 4.8: Graph of capacity vs. memory for $BER$=0.1 and $Cor$=0.5.

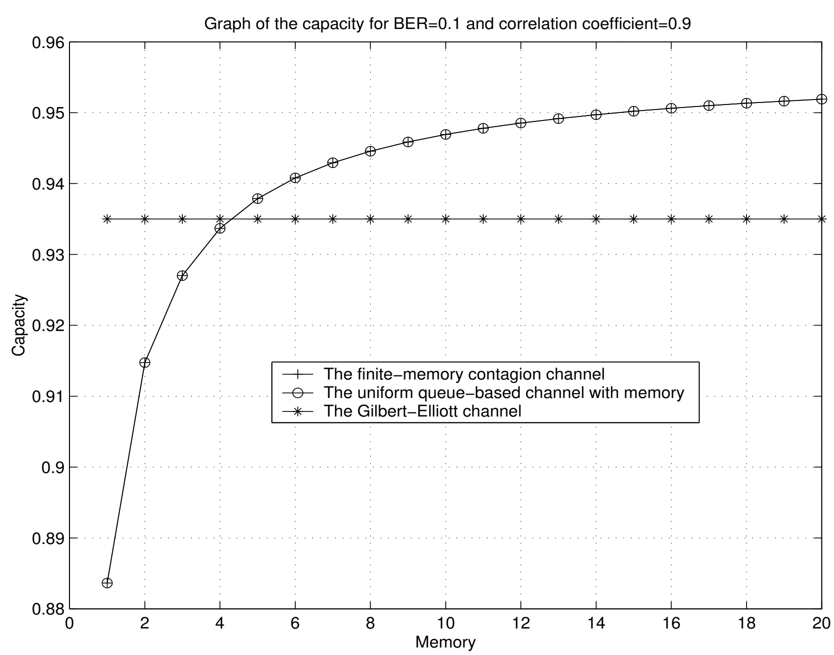



Figure 4.9: Graph of capacity vs. memory for $BER$=0.1 and $Cor$=0.9.

## 4.4 Capacity vs. $BER$ Plots

The following are the figures of capacity vs. $BER$ for different memory values and correlation coefficients for the uniform queue-based channel with memory, the finite-memory contagion channel and the Gilbert-Elliott channel.

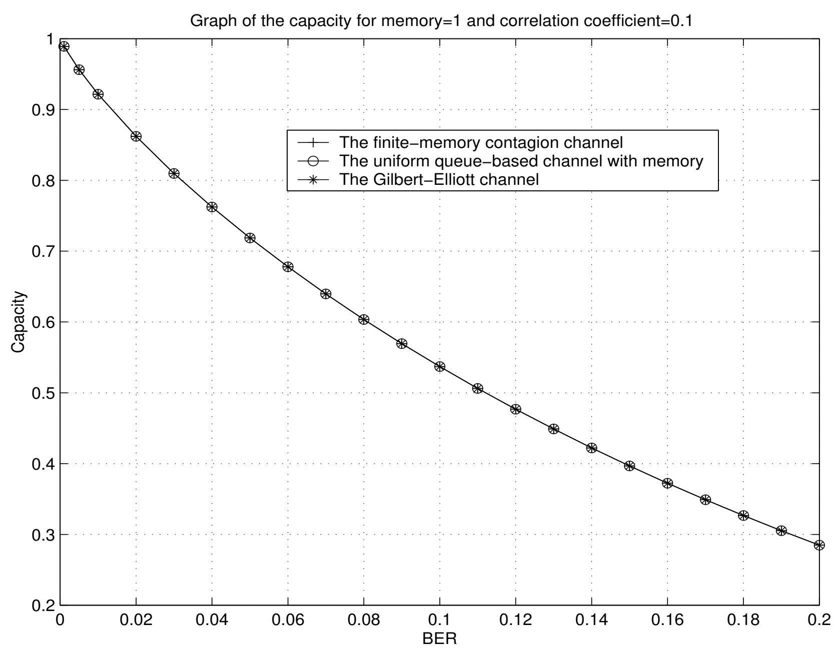



Figure 4.10: Graph of capacity vs. $BER$ for memory=1 and $Cor$=0.1.

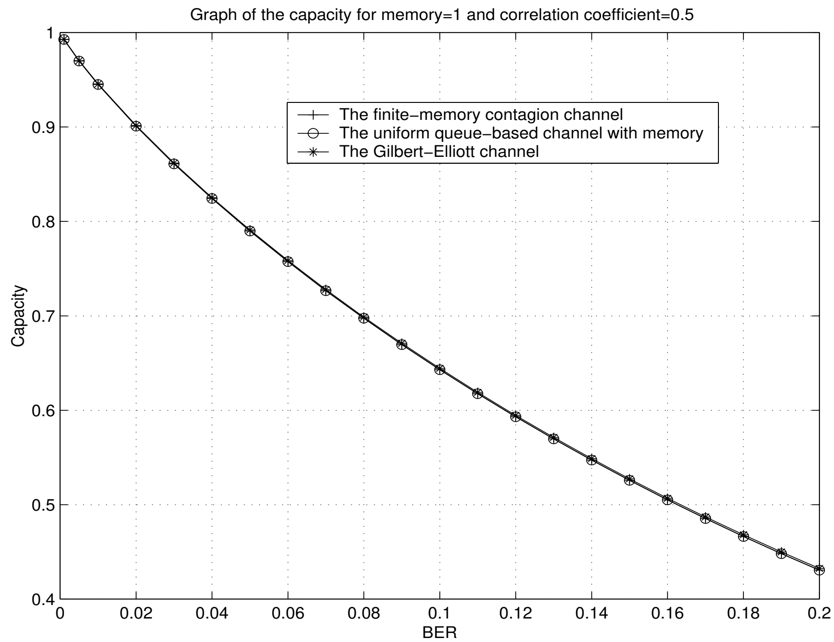



Figure 4.11: Graph of capacity vs. $BER$ for memory=1 and $Cor$=0.5.

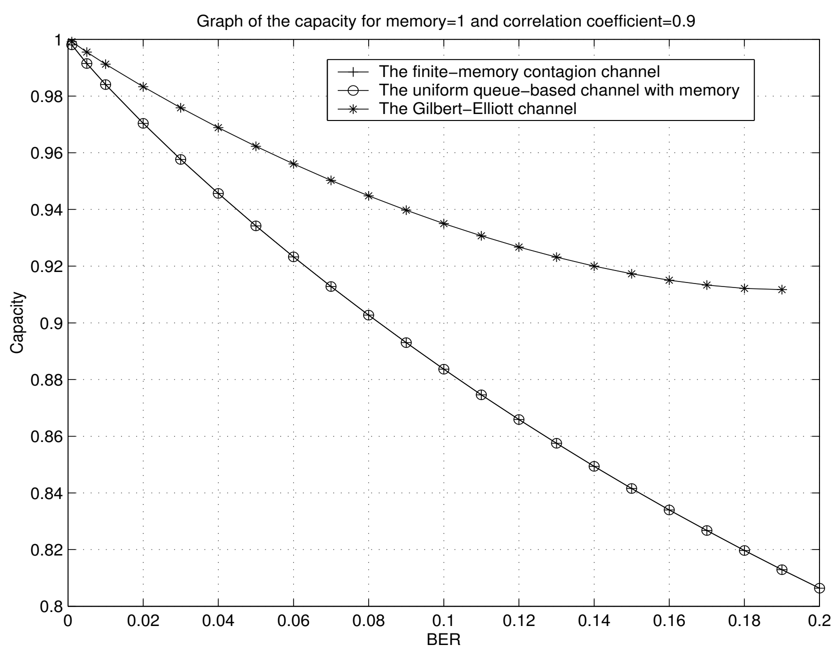



Figure 4.12: Graph of capacity vs. $BER$ for memory=1 and $Cor$=0.9.

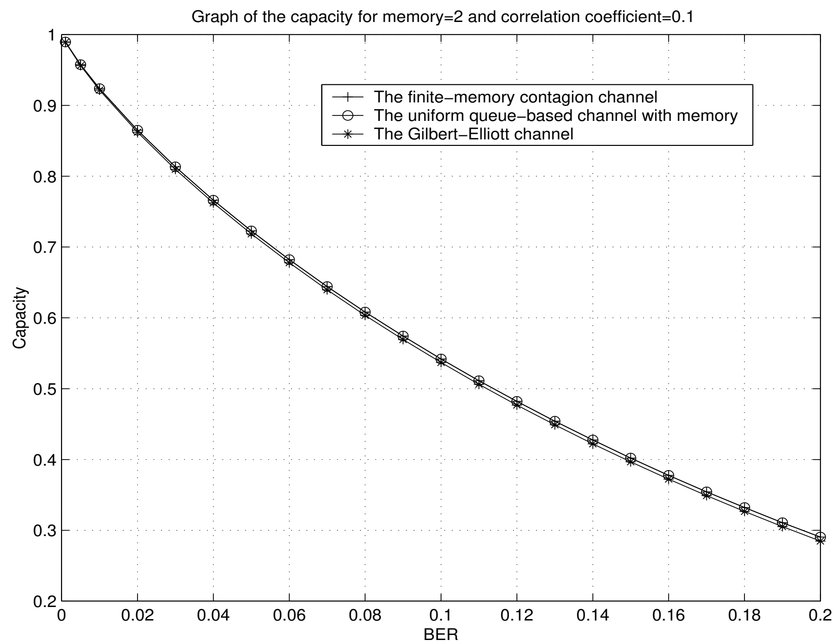



Figure 4.13: Graph of capacity vs. $BER$ for memory=2 and $Cor$=0.1.

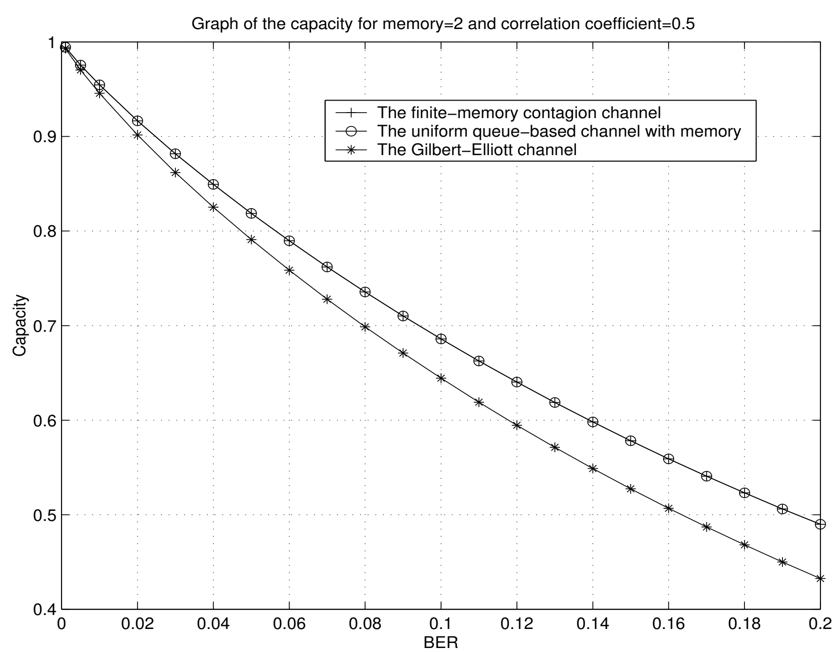



Figure 4.14: Graph of capacity vs. $BER$ for memory=2 and $Cor$=0.5.

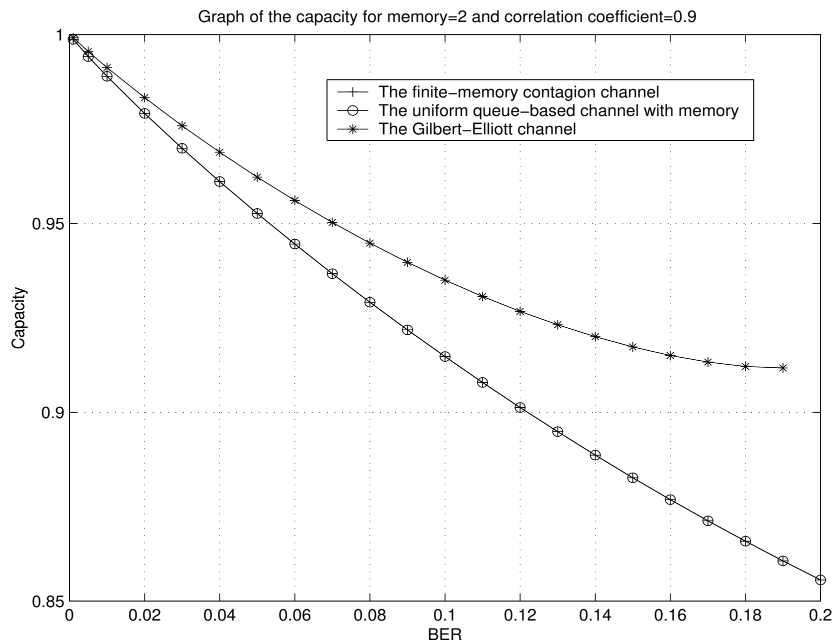



Figure 4.15: Graph of capacity vs. $BER$ for memory=2 and $Cor$=0.9.

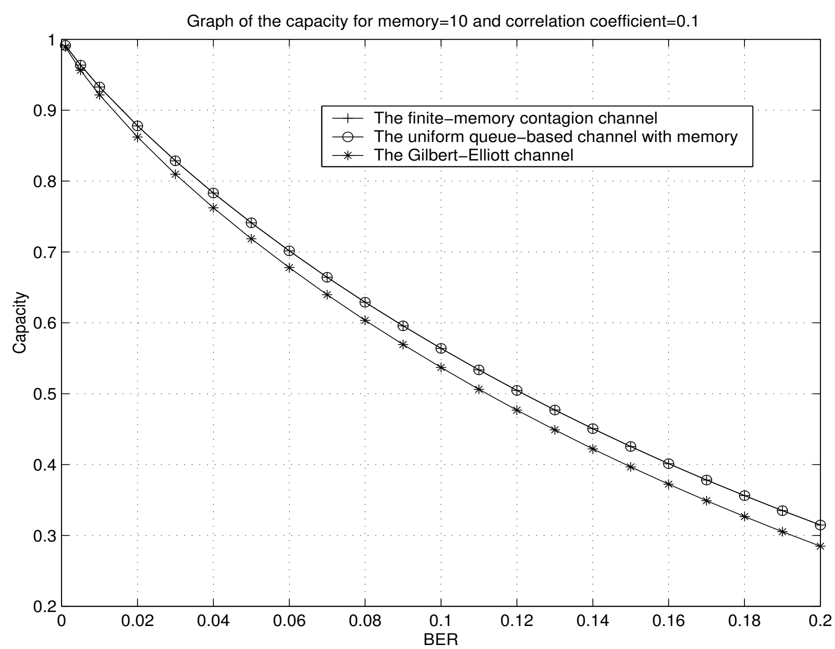



Figure 4.16: Graph of capacity vs. $BER$ for memory=10 and $Cor$=0.1.

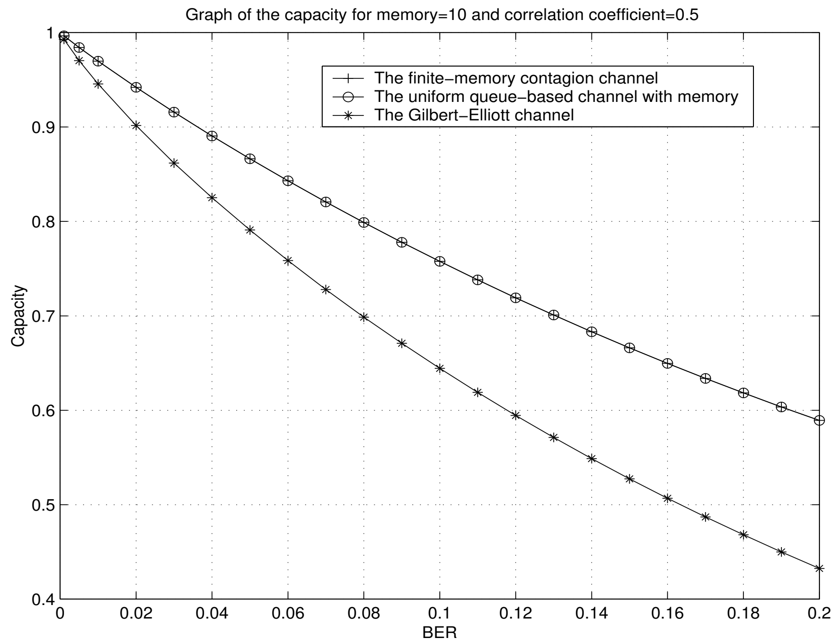



Figure 4.17: Graph of capacity vs. $BER$ for memory=10 and $Cor$=0.5.

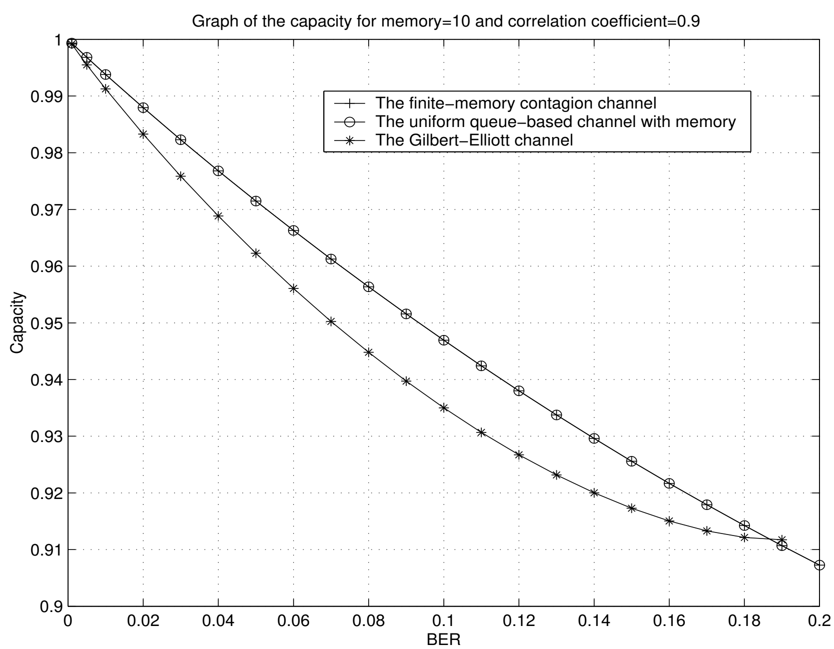



Figure 4.18: Graph of capacity vs. $BER$ for memory=10 and $Cor$=0.9.

## 4.5 Burst Frequency vs. Burst Length

In Section 4.5.1 and 4.5.2 we provide expressions for the burst frequencies and average burst durations for the uniform queue-based channel (which is identical to the finite-memory contagion channel) and the Gilbert-Elliot channel. In section 4.5.3, plots of the burst frequency vs. burst length for all three channels are provided.

### 4.5.1 Burst frequency and average burst duration for the uniform queue-based channel with memory

Using (4.9) and (4.10), rewriting (3.26) to (3.28), we get the following in terms of $BER$, $Cor$ and $K$.

- For $1 \leq l \leq K-1$, where $K > 1$,

$$\Pr[B_n = l] = \frac{\prod_{j=0}^{l-2}(BER + j\frac{Cor}{1-Cor}) \cdot (1 - BER + \frac{Cor}{1-Cor})}{\prod_{i=1}^{l}(1 + i\frac{Cor}{1-Cor})}, \tag{4.30}$$

  where $\prod_{i=0}^{a}(\cdot)\stackrel{\triangle}{=}1$, if $a < 0$.

- For $l = K$,

$$\Pr[B_n = K] = \frac{\prod_{j=0}^{K-2}(BER + j\frac{Cor}{1-Cor})}{\prod_{i=1}^{K-1}(1 + i\frac{Cor}{1-Cor})} \cdot \left(\frac{1 - BER + \frac{Cor}{1-Cor}}{1 + K\frac{Cor}{1-Cor}}\right). \tag{4.31}$$

- For $l \geq K+1$,

$$\begin{aligned} \Pr[B_n = l] \quad = \quad & \frac{\prod_{j=0}^{K-2}(BER + j\frac{Cor}{1-Cor})}{\prod_{i=1}^{K-1}(1 + i\frac{Cor}{1-Cor})} \cdot \left[\frac{BER + (K-1)\frac{Cor}{1-Cor}}{1 + K\frac{Cor}{1-Cor}}\right] \\ & \cdot \left[\frac{BER + K\frac{Cor}{1-Cor}}{1 + K\frac{Cor}{1-Cor}}\right]^{l-K-1} \cdot \left(\frac{1 - BER}{1 + K\frac{Cor}{1-Cor}}\right). \end{aligned} \tag{4.32}$$

The average burst duration is

$$E(B_n) = \frac{1}{P(0)} = \frac{1}{\Pr[Z_n = 0]} = \frac{1}{1 - BER}. \tag{4.33}$$

### 4.5.2 Burst frequency and average burst duration for the Gilbert-Elliott channel

The burst frequencies for the Gilbert-Elliott channel are

$$\Pr[B_n = l] = \frac{\boldsymbol{\pi}^T \mathbf{P}(0)[\mathbf{P}(1)]^{l-1}\mathbf{P}(0)\mathbf{1}}{\boldsymbol{\pi}^T \mathbf{P}(0)\mathbf{1}}, \tag{4.34}$$

where $\boldsymbol{\pi}$, $\mathbf{P}(0)$, $\mathbf{P}(1)$ and $\mathbf{1}$ are defined in (2.45) - (2.48).

The average burst duration is

$$E(B_n) = \frac{1}{P(0)} = \frac{1}{\Pr[Z_n = 0]} = \frac{1}{1 - BER}. \tag{4.35}$$

### 4.5.3 Burst frequency vs. burst length plots

The following are the figures of burst frequency vs. burst length for different memory values, $BER$s and correlation coefficients for the uniform queue-based channel with memory, the finite-memory contagion channel and the Gilbert-Elliott channel.

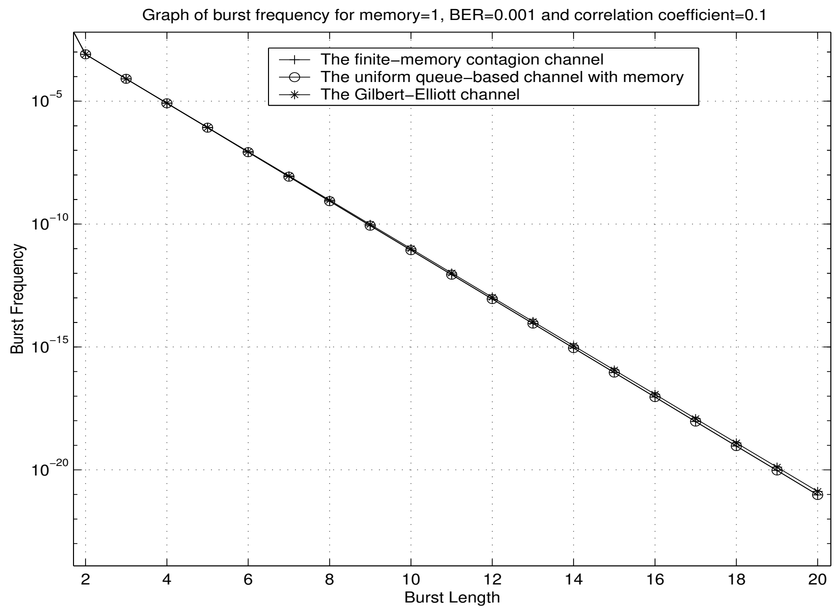



Figure 4.19: Graph of burst frequency vs. burst length for memory=1, $BER$=0.001 and $Cor$=0.1.

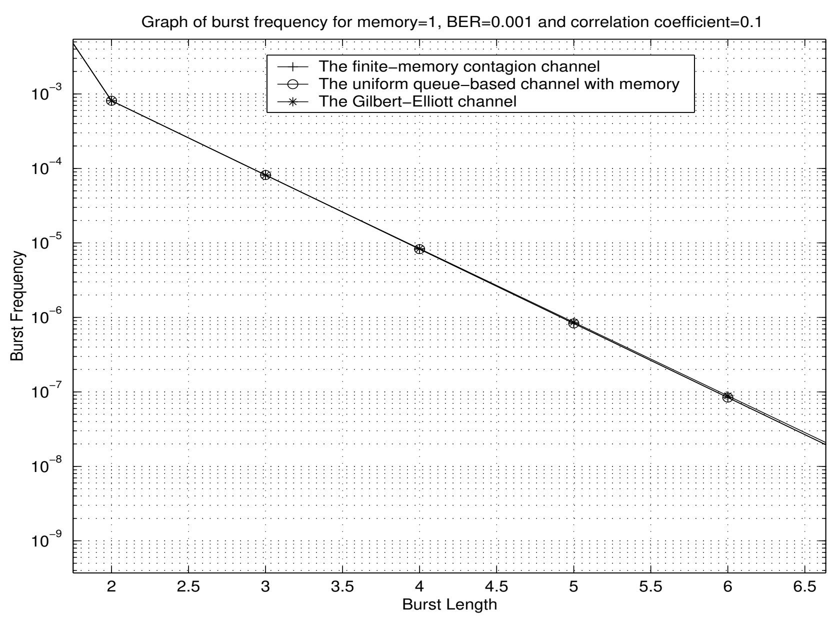



Figure 4.20: Zoom graph of burst frequency vs. burst length for memory=1, $BER$=0.001 and $Cor$=0.1.

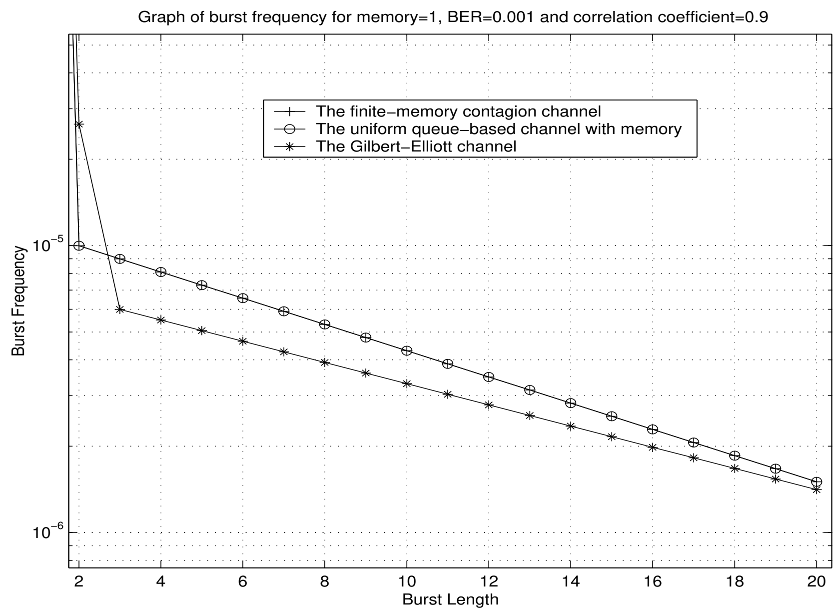



Figure 4.21: Graph of burst frequency vs. burst length for memory=1, $BER$=0.001 and $Cor$=0.9.

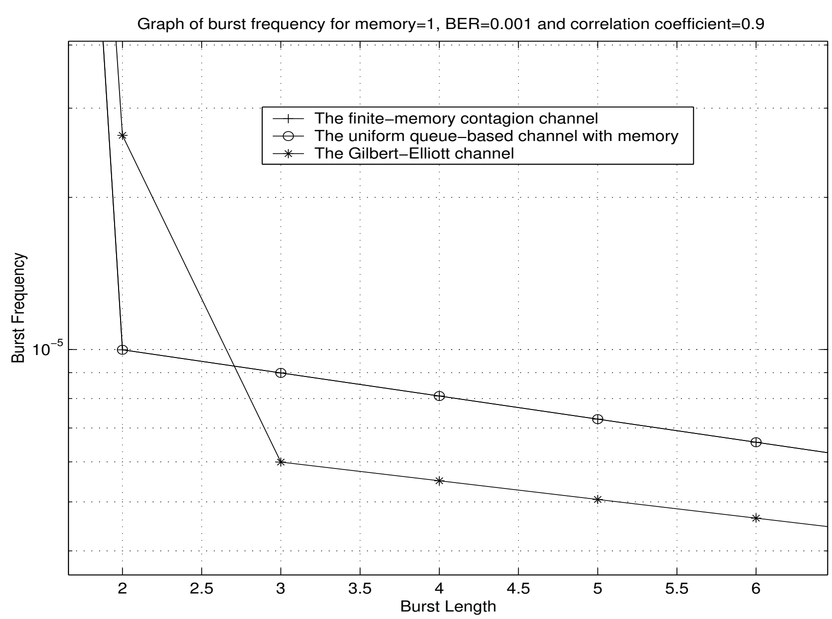



Figure 4.22: Zoom graph of burst frequency vs. burst length for memory=1, $BER$=0.001 and $Cor$=0.9.

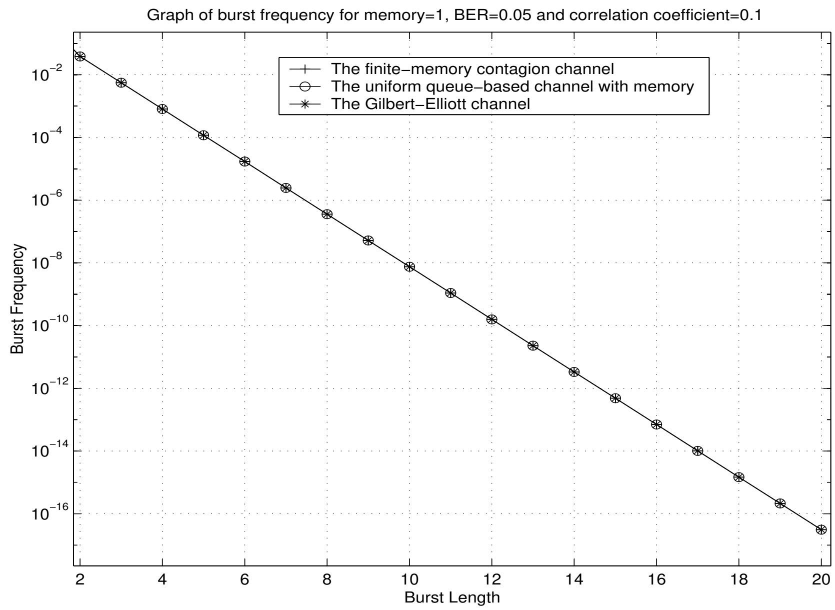



Figure 4.23: Graph of burst frequency vs. burst length for memory=1, $BER$=0.05 and $Cor$=0.1.

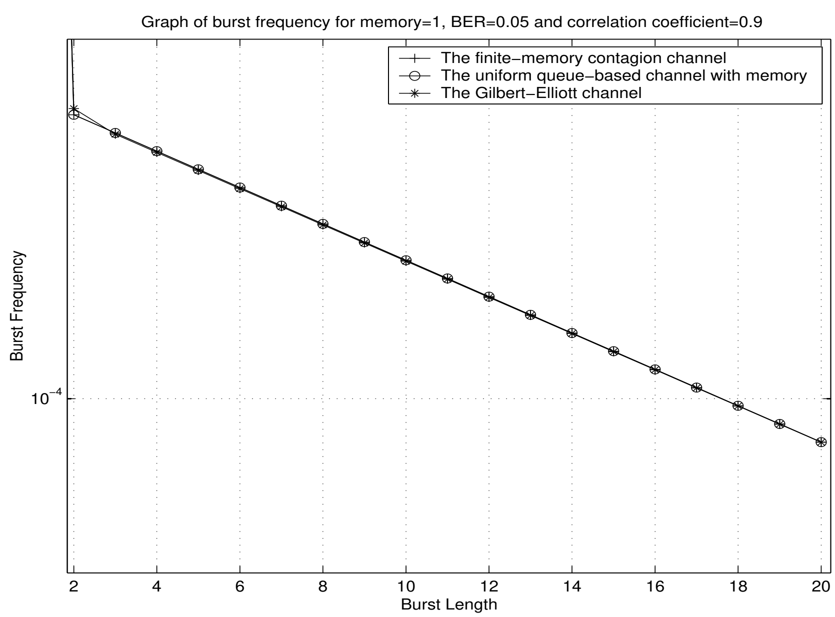



Figure 4.24: Graph of burst frequency vs. burst length for memory=1, $BER$=0.05 and $Cor$=0.9.

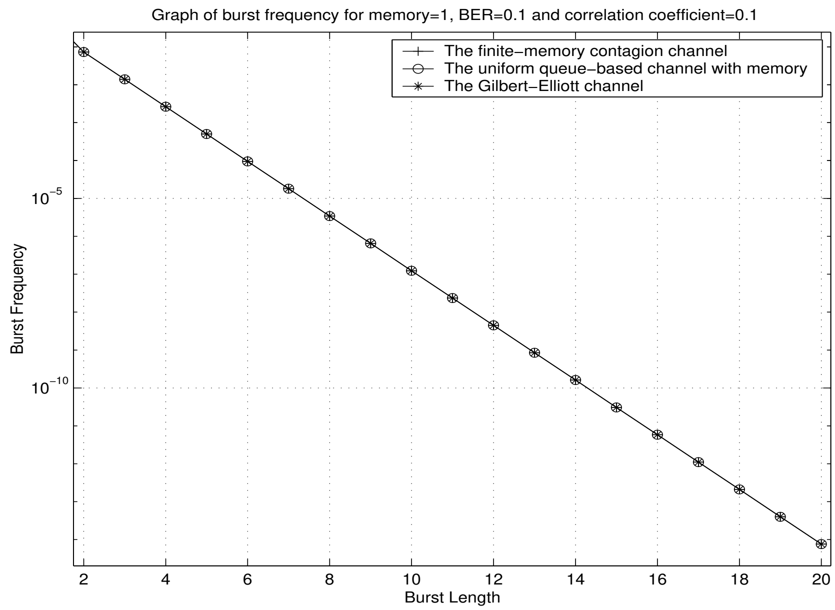



Figure 4.25: Graph of burst frequency vs. burst length for memory=1, $BER$=0.1 and $Cor$=0.1.

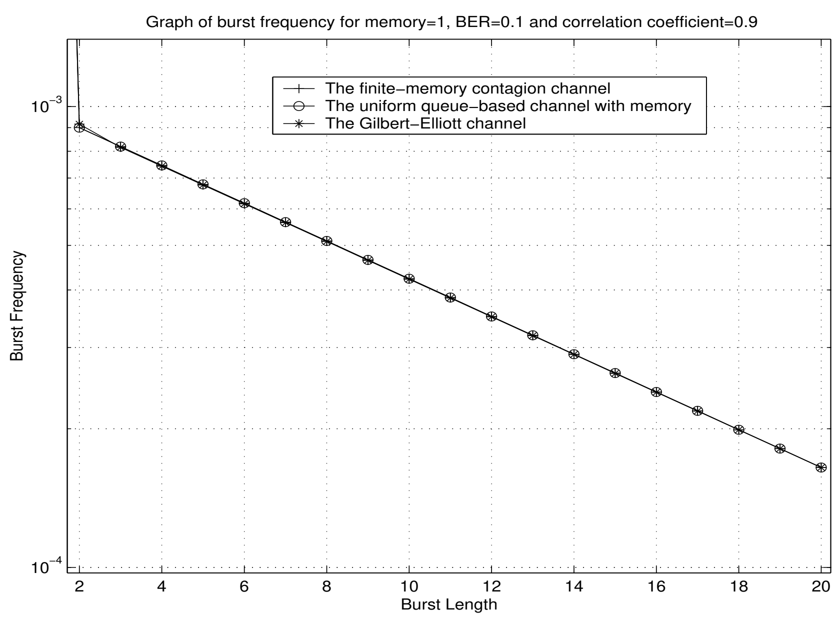



Figure 4.26: Graph of burst frequency vs. burst length for memory=1, $BER$=0.1 and $Cor$=0.9.

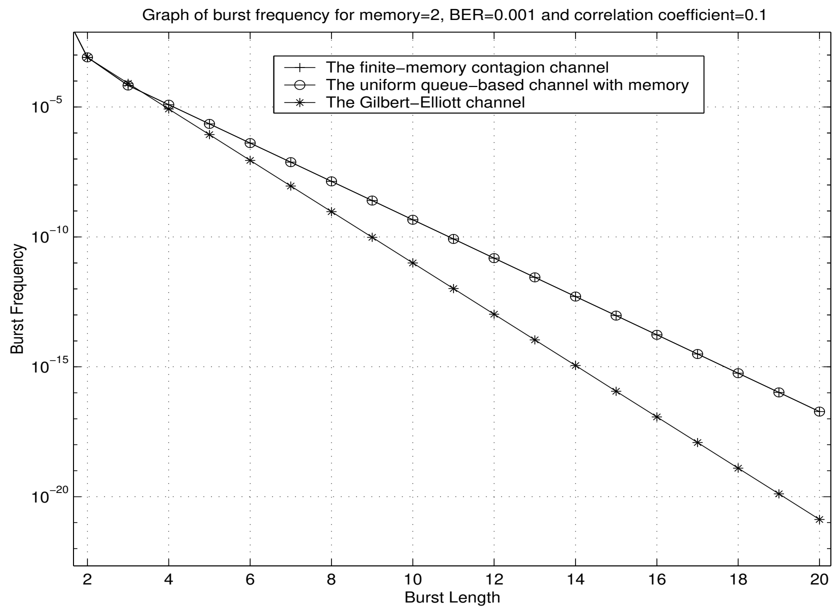



Figure 4.27: Graph of burst frequency vs. burst length for memory=2, $BER$=0.001 and $Cor$=0.1.

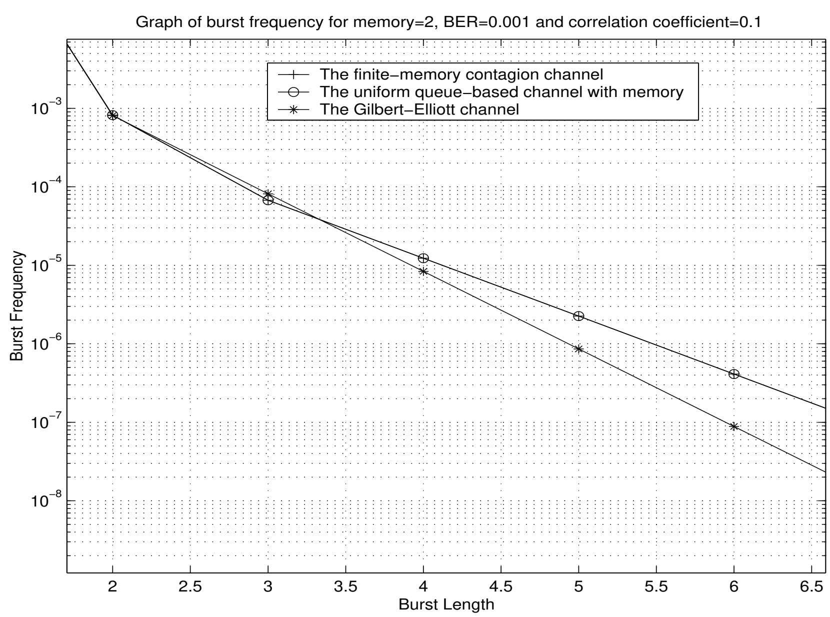



Figure 4.28: Zoom graph of burst frequency vs. burst length for memory=2, $BER$=0.001 and $Cor$=0.1.

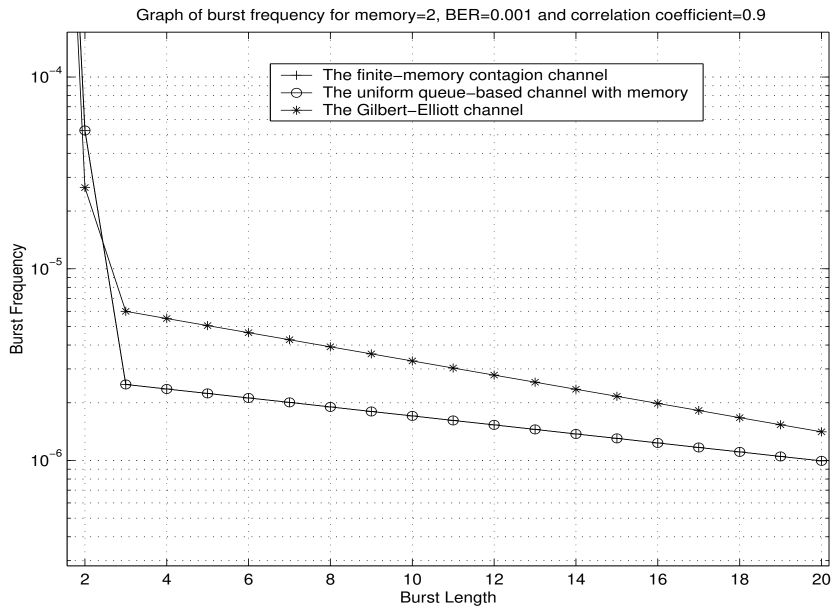



Figure 4.29: Graph of burst frequency vs. burst length for memory=2, $BER$=0.001 and $Cor$=0.9.

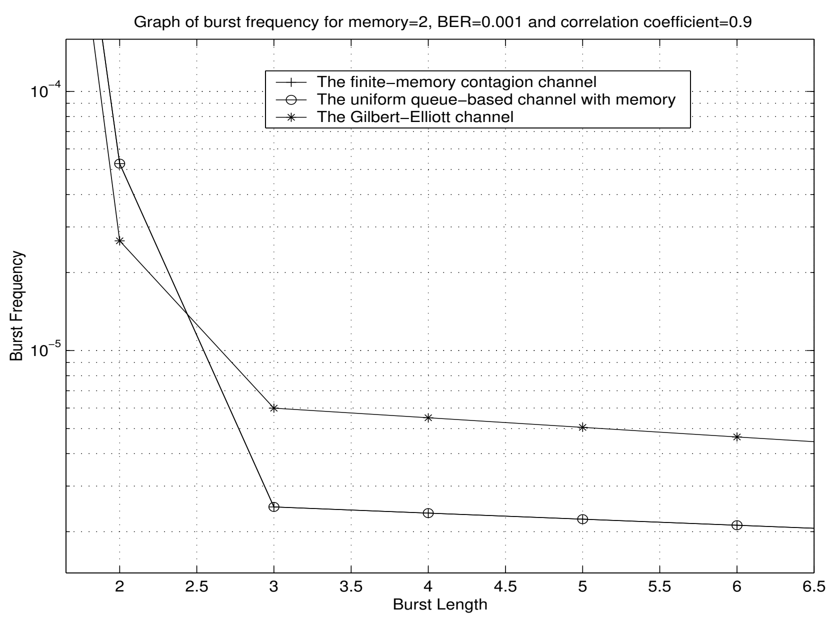



Figure 4.30: Zoom graph of burst frequency vs. burst length for memory=2, $BER$=0.001 and $Cor$=0.9.

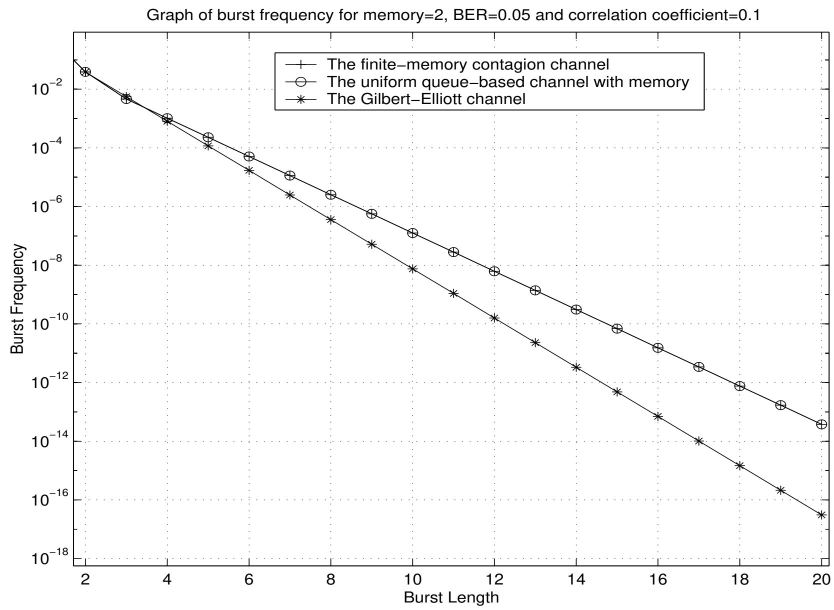



Figure 4.31: Graph of burst frequency vs. burst length for memory=2, $BER$=0.05 and $Cor$=0.1.

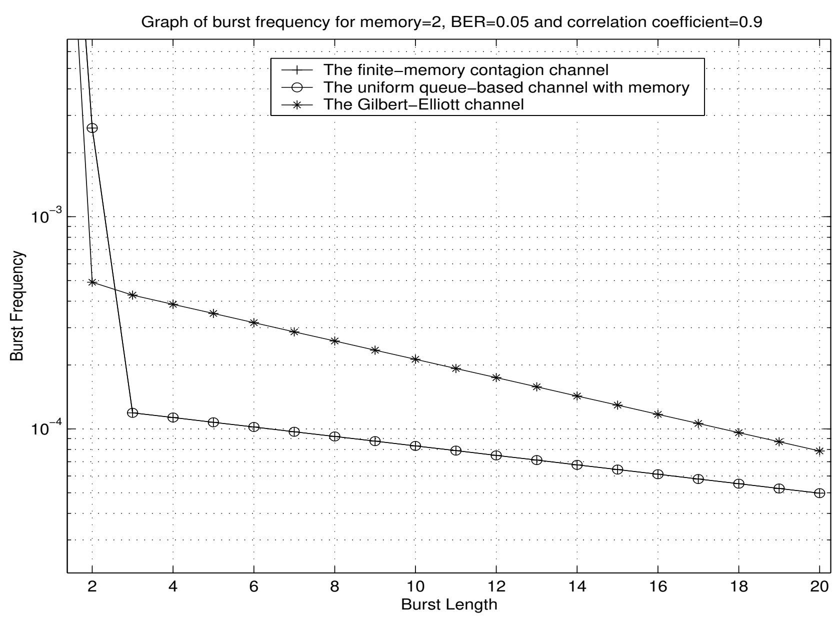



Figure 4.32: Graph of burst frequency vs. burst length for memory=2, $BER$=0.05 and $Cor$=0.9.

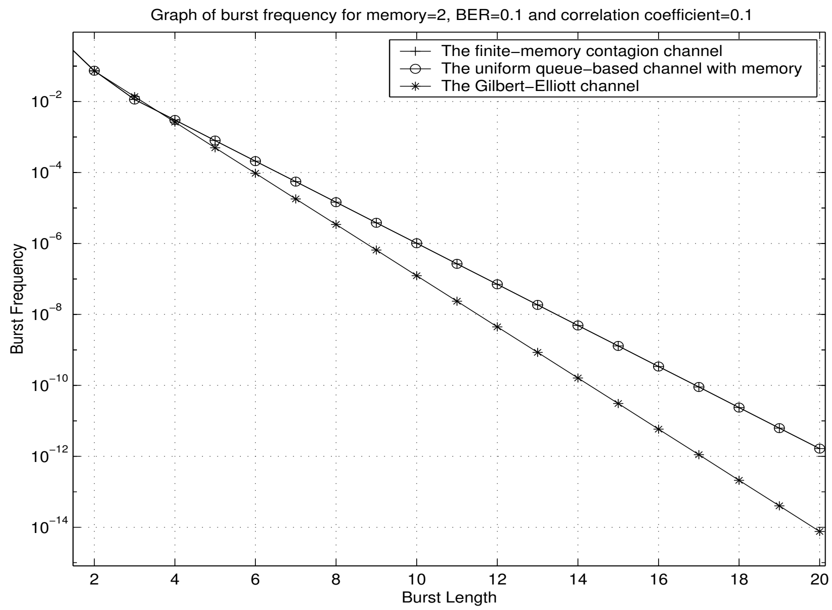



Figure 4.33: Graph of burst frequency vs. burst length for memory=2, $BER$=0.1 and $Cor$=0.1.

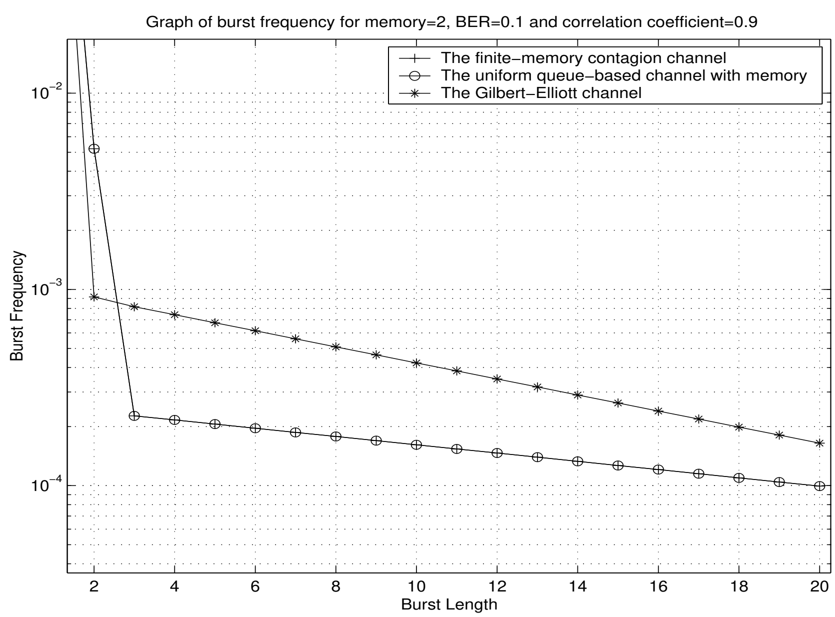



Figure 4.34: Graph of burst frequency vs. burst length for memory=2, $BER$=0.1 and $Cor$=0.9.

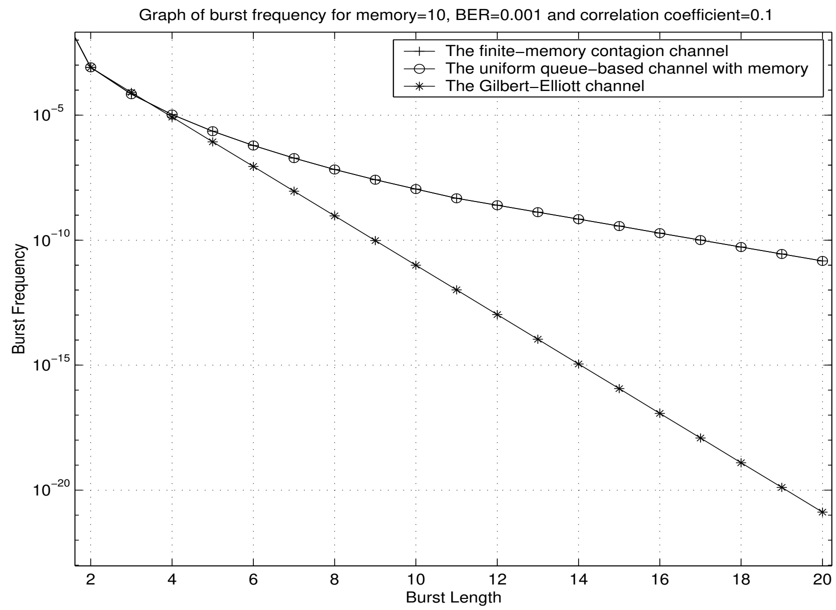



Figure 4.35: Graph of burst frequency vs. burst length for memory=10, $BER$=0.001 and $Cor$=0.1.

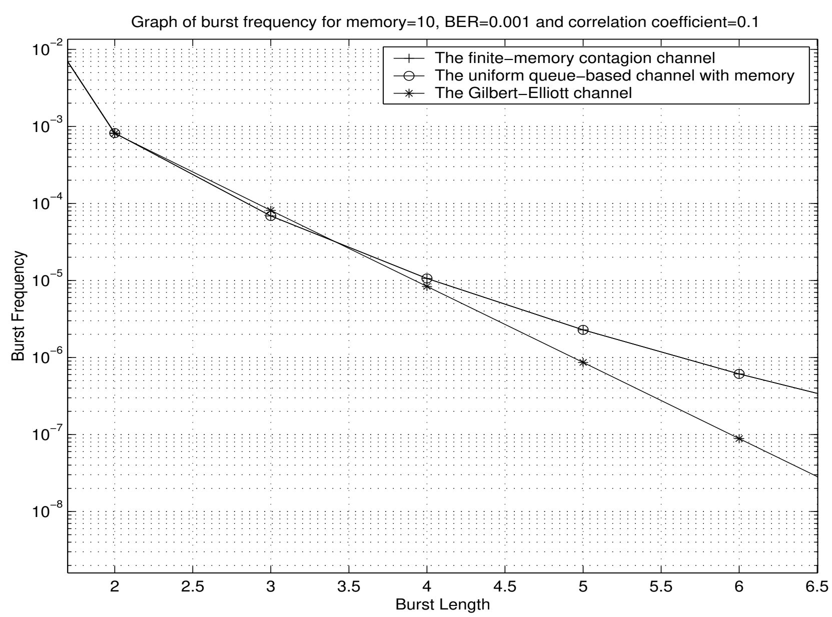



Figure 4.36: Zoom graph of burst frequency vs. burst length for memory=10, $BER$=0.001 and $Cor$=0.1.

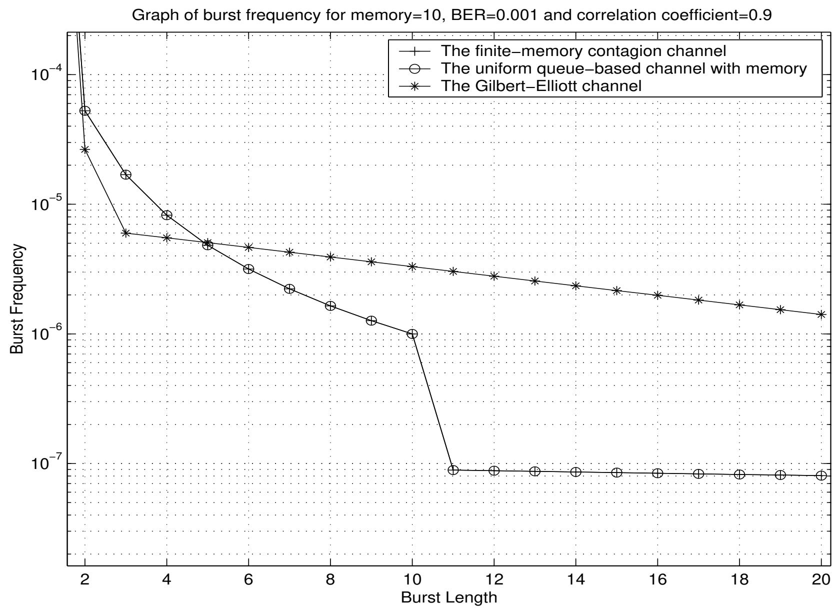



Figure 4.37: Graph of burst frequency vs. burst length for memory=10, $BER$=0.001 and $Cor$=0.9.

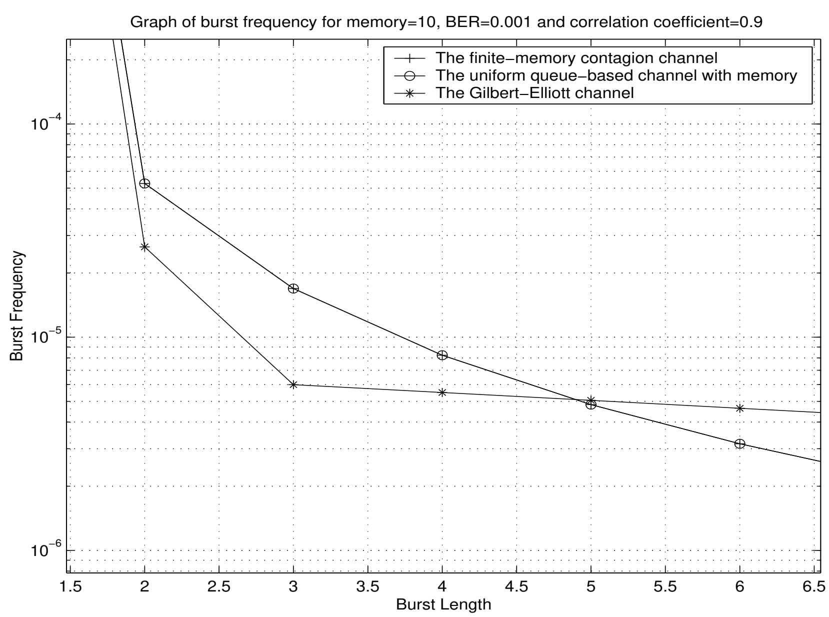



Figure 4.38: Zoom graph of burst frequency vs. burst length for memory=10, $BER$=0.001 and $Cor$=0.9.

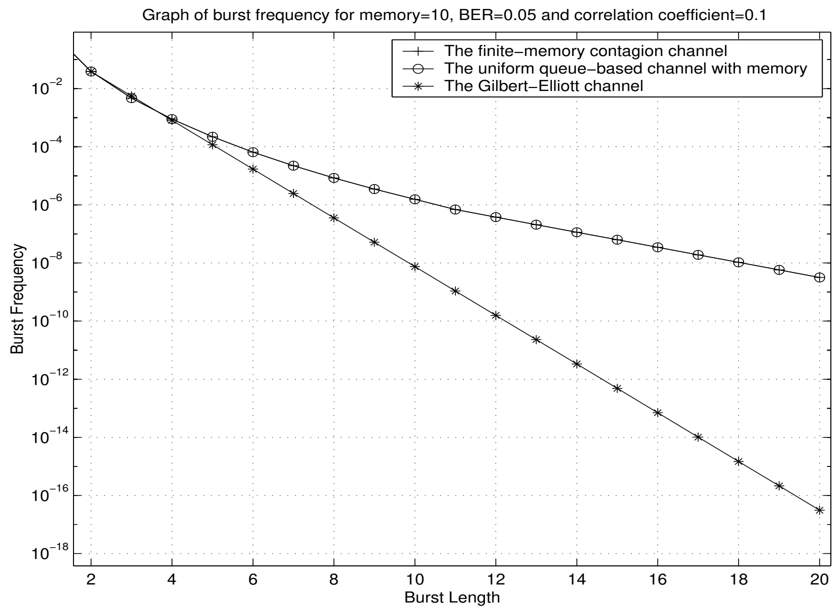



Figure 4.39: Graph of burst frequency vs. burst length for memory=10, $BER$=0.05 and $Cor$=0.1.

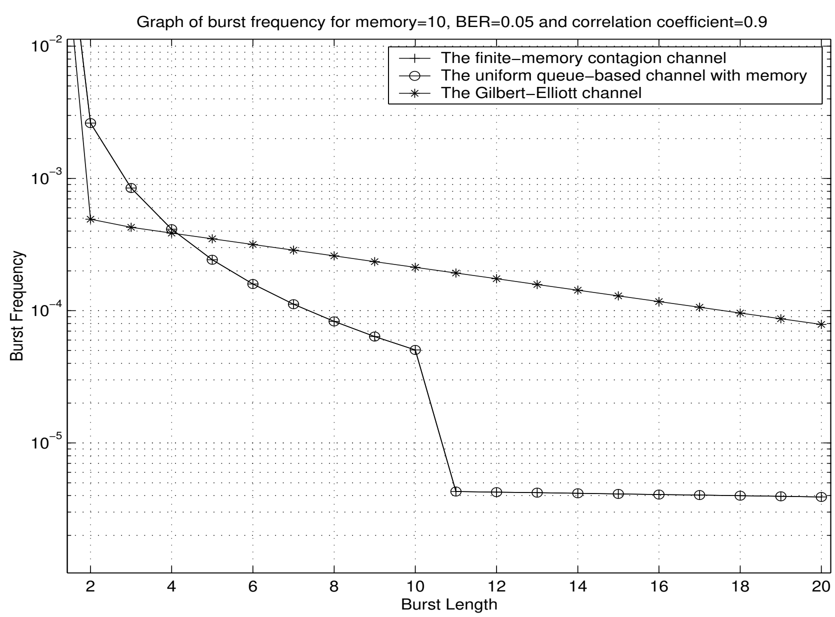



Figure 4.40: Graph of burst frequency vs. burst length for memory=10, $BER$=0.05 and $Cor$=0.9.

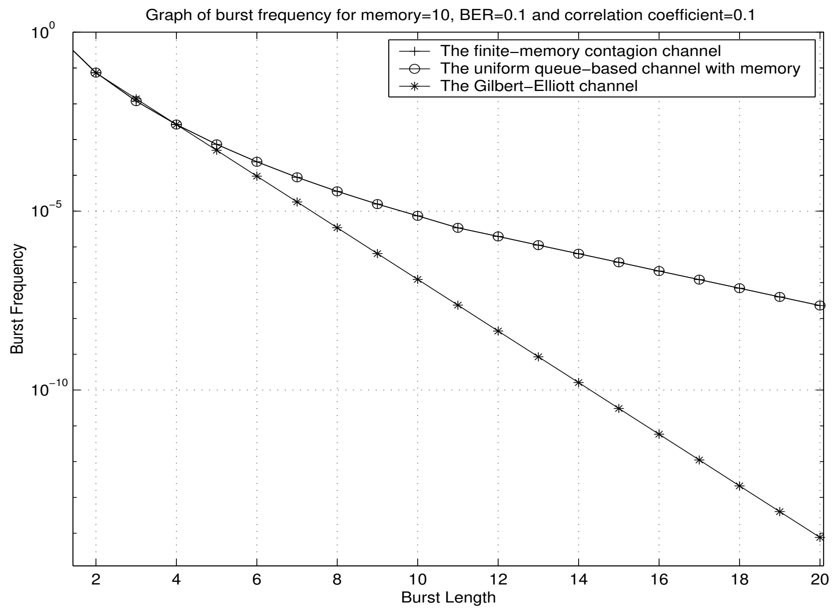



Figure 4.41: Graph of burst frequency vs. burst length for memory=10, $BER$=0.1 and $Cor$=0.1.

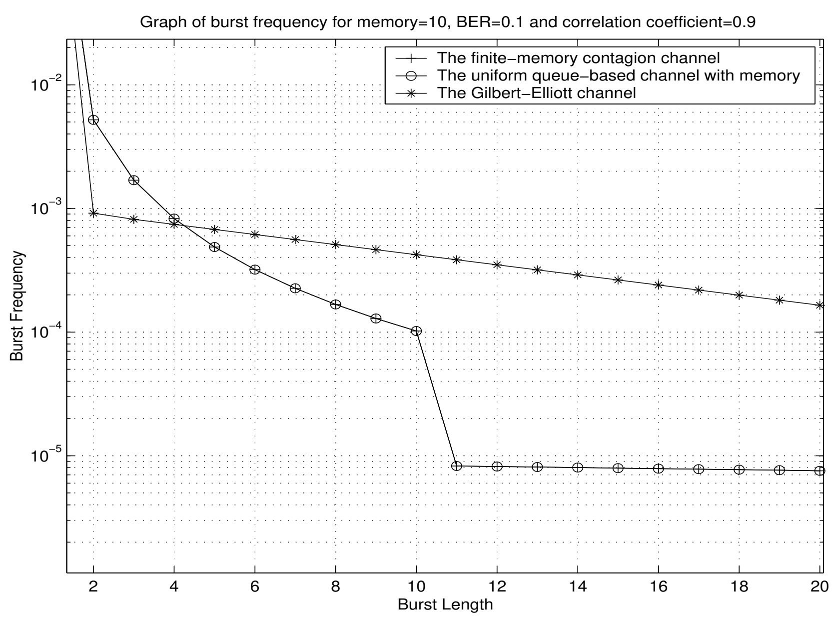



Figure 4.42: Graph of burst frequency vs. burst length for memory=10, $BER$=0.1 and $Cor$=0.9.

## 4.6 Capacity of Non-Uniform Queue-Based Channel with Memory

The following are the figures of capacity vs. $q_1$ for different memory values, $p = 0.1$ and $\varepsilon = 0.5$ for the non-uniform queue-based channel with memory.

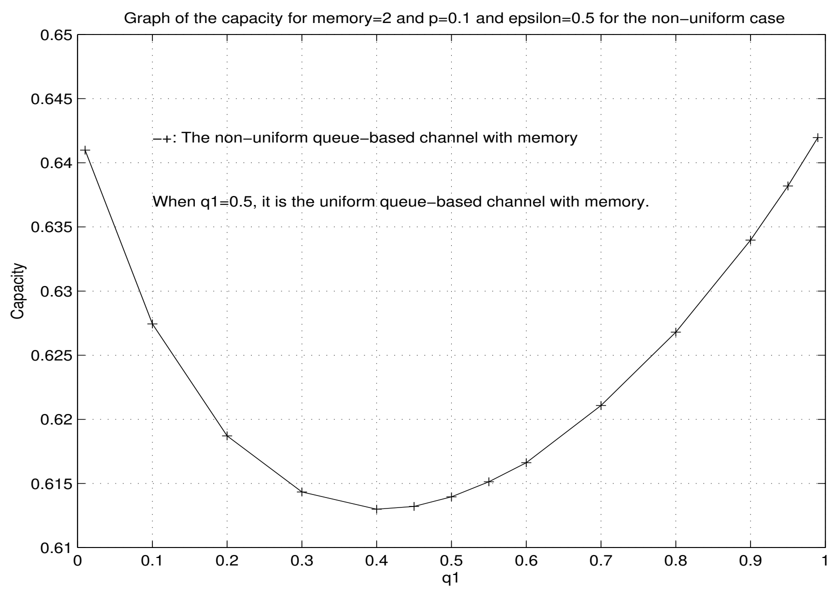



Figure 4.43: Graph of capacity for the non-uniform queue-based channel with memory=2.

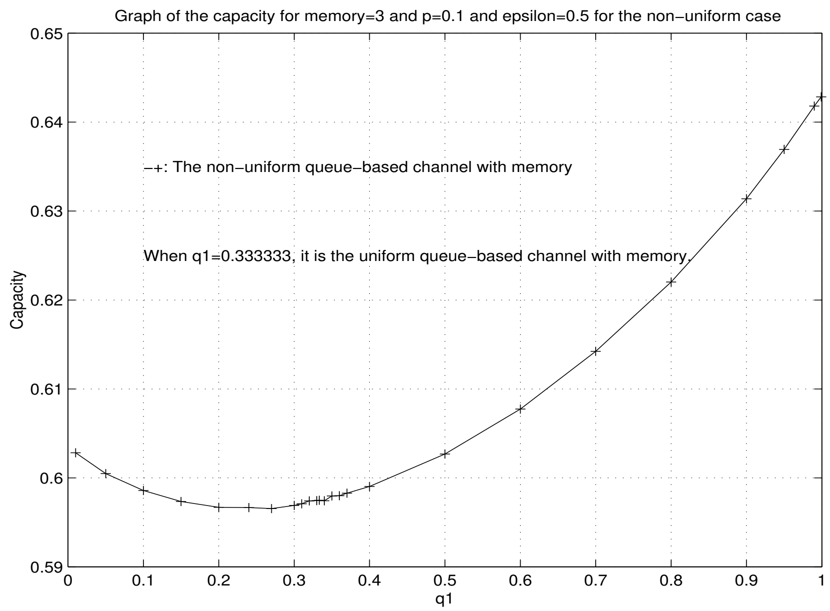



Figure 4.44: Graph of capacity for the non-uniform queue-based channel with memory=3.

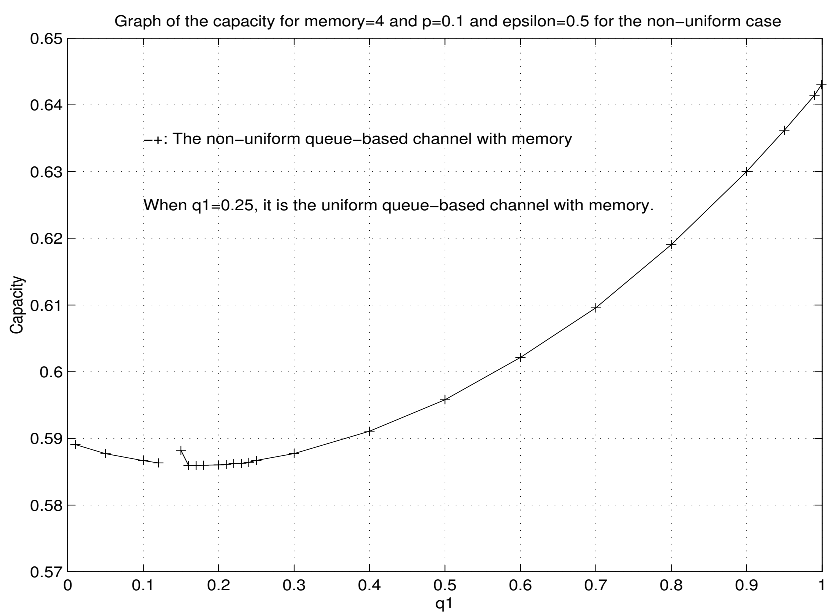



Figure 4.45: Graph of capacity for the non-uniform queue-based channel with memory=4.

Next, we compare the capacity for uniform and non-uniform queue-based channel given the same correlation coefficient for different memory values.

| $q_1$ | correlation coefficient | capacity for non-uniform | capacity for uniform |
|---|---|---|---|
| 0.01 | 0.01 | 0.640983 | 0.531144 |
| 0.1 | 0.0911111 | 0.627438 | 0.540342 |
| 0.3 | 0.231111 | 0.614337 | 0.577149 |
| 0.5 | 0.333333 | 0.613961 | 0.613961 |
| 0.6 | 0.375344 | 0.61662 | 0.630839 |
| 0.7 | 0.412222 | 0.62108 | 0.646386 |
| 0.8 | 0.444444 | 0.626793 | 0.660498 |
| 0.9 | 0.473333 | 0.633968 | 0.673552 |
| 0.99 | 0.497778 | 0.64196 | 0.684888 |

Table 4.1: Table of capacity for the non-uniform and uniform queue-based channel with memory=2 and $BER$=0.1.

| $q_1$ | correlation coefficient | capacity for non-uniform | capacity for uniform |
|---|---|---|---|
| 0.01 | 0.0956556 | 0.602821 | 0.545132 |
| 0.1 | 0.141111 | 0.59858 | 0.55813 |
| 0.2 | 0.188889 | 0.596694 | 0.574272 |
| 0.333333 | 0.25 | 0.597483 | 0.597483 |
| 0.5 | 0.3211 | 0.602684 | 0.627007 |
| 0.6 | 0.3601 | 0.607736 | 0.644062 |
| 0.7 | 0.397878 | 0.614227 | 0.66106 |
| 0.8 | 0.433444 | 0.622028 | 0.677447 |
| 0.9 | 0.468678 | 0.631387 | 0.694021 |
| 0.95 | 0.484444 | 0.636936 | 0.701543 |
| 0.99 | 0.496767 | 0.641804 | 0.707466 |

Table 4.2: Table of capacity for the non-uniform and uniform queue-based channel with memory=3 and $BER$=0.1.

| $q_1$ | correlation coefficient | capacity for non-uniform | capacity for uniform |
|---|---|---|---|
| 0.01 | 0.0876667 | 0.58906 | 0.546145 |
| 0.1 | 0.1302 | 0.586663 | 0.559906 |
| 0.2 | 0.178567 | 0.58601 | 0.57807 |
| 0.25 | 0.2 | 0.586703 | 0.586703 |
| 0.4 | 0.265656 | 0.591077 | 0.614629 |
| 0.5 | 0.308778 | 0.595794 | 0.633826 |
| 0.6 | 0.348989 | 0.602134 | 0.652154 |
| 0.7 | 0.388689 | 0.60957 | 0.670568 |
| 0.8 | 0.426767 | 0.619047 | 0.688483 |
| 0.9 | 0.463533 | 0.630005 | 0.705992 |
| 0.95 | 0.482322 | 0.636202 | 0.715016 |
| 0.99 | 0.496567 | 0.641435 | 0.721892 |

Table 4.3: Table of capacity for the non-uniform and uniform queue-based channel with memory=4 and $BER$=0.1.

## 4.7 Burst Frequency of Non-Uniform Queue-Based Channel with Memory

The following are the figures of burst frequency (see (3.31) to (3.33)) vs. burst length for different $BER$ and correlation coefficient values, $q_1 = 0.9$ for the non-uniform queue-based channel with memory.

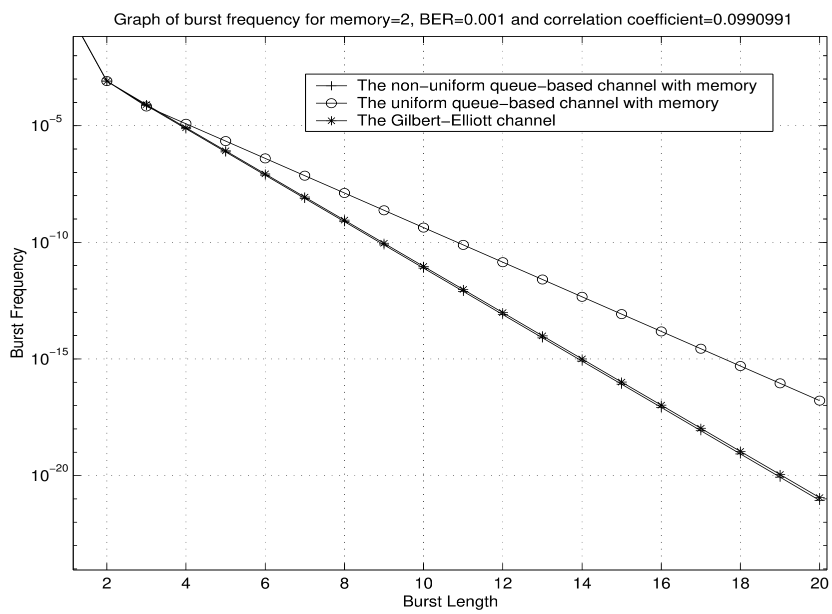



Figure 4.46: Graph of burst frequency vs. burst length for the non-uniform queue-based channel with memory=2, $BER$=0.001 and $Cor$=0.0990991.

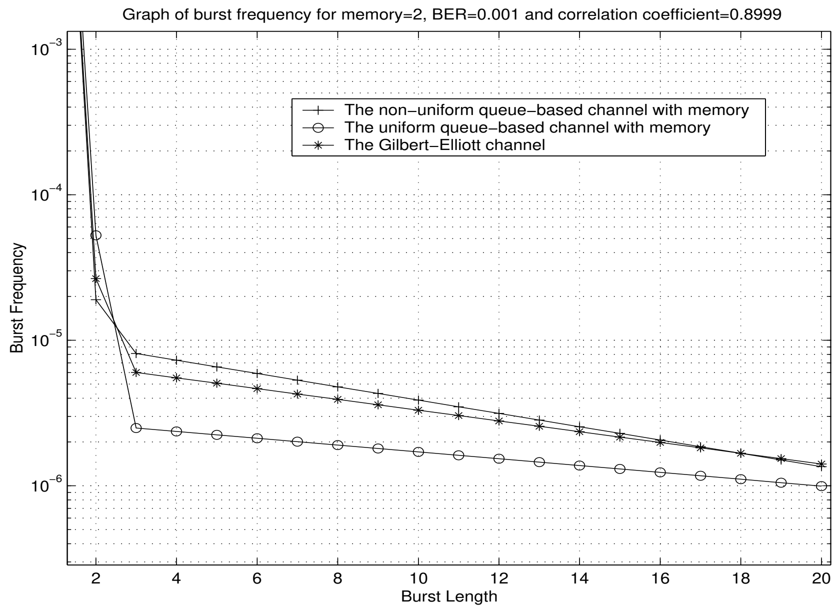



Figure 4.47: Graph of burst frequency vs. burst length for the non-uniform queue-based channel with memory=2, $BER$=0.001 and $Cor$=0.8999.

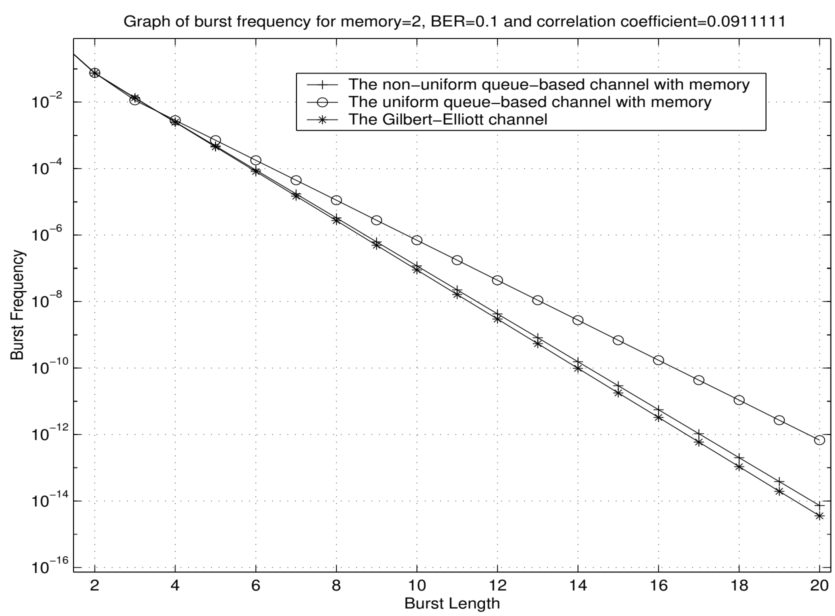



Figure 4.48: Graph of burst frequency vs. burst length for the non-uniform queue-based channel with memory=2, $BER$=0.1 and $Cor$=0.0911111.

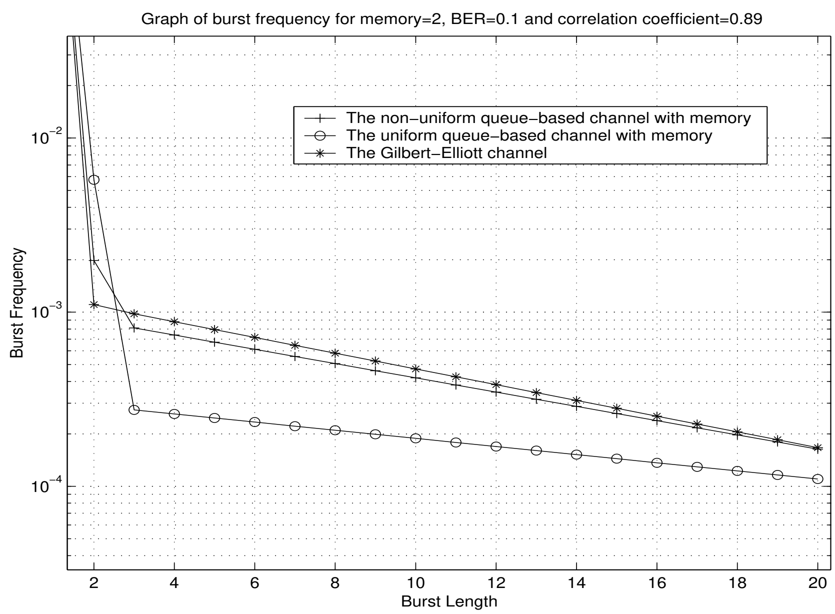



Figure 4.49: Graph of burst frequency vs. burst length for the non-uniform queue-based channel with memory=2, $BER$=0.1 and $Cor$=0.89.

## 4.8 Discussion

We have shown that the finite-memory contagion channel and the uniform queue-based channel with memory are surprisingly identical; i.e., they have the same block transition probability for the same memory, $BER$ and correlation coefficient. Therefore the two channels have identical capacities and burst frequencies under the above conditions.

Capacity increases as $BER$ decreases and as the correlation coefficient increases. For the uniform queue-based channel and the finite-memory contagion channel, capacity also increases with memory. When $Cor = 0.1$, the capacities of the uniform queue-based channel and the finite-memory contagion channel are always larger than that of the Gilbert-Elliott channel for any memory values. But, as the correlation coefficient increases, the capacity of the Gilbert-Elliott channel grows faster. When $Cor = 0.9$, the uniform queue-based channel and the finite-memory contagion channel have lower capacities than the Gilbert-Elliott channel at smaller memory values and have higher capacities at larger memory values.

It is clear in Fig. 4.10 and from Fig. 4.19 to Fig. 4.26 that the three channels have almost equal capacities and burst frequencies when memory=1. This means that in these cases we can replace the Gilbert-Elliott channel with the (less complex) uniform queue-based channel (or the finite-memory contagion channel) if our target is to achieve an error burst behavior and capacity that are close to those of the Gilbert-Elliott channel.

When memory=2,10, the uniform queue-based channel with memory and the finite-memory contagion channel have higher burst frequencies than the Gilbert-Elliott channel for low correlation coefficients. The former two have lower burst frequency than the latter for high correlation coefficients and burst length$=5, \cdots, 20$.

We also notice that the burst frequencies of the uniform queue-based channel with memory and the finite-memory contagion channel increase with memory for low correlations and decrease with memory for high correlations.

Since the noise processes of the three channels are all stationary, from (4.33) and (4.35), we arrive at the conclusion that the average burst duration of the three channels are equal given the same $BER$; i.e., $E(B_n) = \frac{1}{1-BER}$.

From Table 4.1 to Table 4.3, we conclude that the capacity values for the non-uniform case are larger than those for the uniform case when $q_1 < \frac{1}{K}$ and smaller than those for the uniform case when $q_1 > \frac{1}{K}$. When $q_1 \rightarrow 1$, the non-uniform case with memory $K$ converges to the uniform case with memory=1.

The non-uniform queue-based channel has lower burst frequencies than the uniform queue-based channel for low correlation coefficients (see Fig. 4.46 and Fig. 4.48). But it has higher burst frequencies for high correlation coefficients and burst length$= 3, \cdots, 20$ (see Fig. 4.47 and Fig. 4.49). But the burst frequencies of the non-uniform queue-based channel decreases faster than those of the uniform queue-based channel; thus the former eventually has lower burst frequency when memory is big enough. We notice that the non-uniform queue-based channel has similar burst frequency as the Gilbert-Elliott channel. This is because the non-uniform queue-based channel converges to the uniform case with memory=1 when $q_1 \rightarrow 1$ (we chose $q_1 = 0.9$).

# Chapter 5

# Summary

In this project we considered a binary channel with additive bursty noise based on a finite queue. We considered the channel in two cases: a uniform queue-based channel with memory where we experiment on the cells of the queue with equal probability, and a non-uniform queue-based channel with memory where we experiment on the cells of the queue with different probabilities. The channel noise process is stationary and ergodic. The state of the channel is a one-step Markov process with $2^K$ states. We derived the stationary noise distribution, the capacity, and the gap and the burst distributions of the channel in terms of $K$ for the uniform case. For the non-uniform case, we numerically calculated the stationary distributions for memory=2,3,4 using a matlab program. Then we obtained the $BER$, the correlation coefficient, capacity, and burst frequency in terms of the stationary distribution and got the numerical results using C++ programs.

Finally, we compared the capacity, burst frequency and average burst duration of the uniform and the non-uniform queue-based channel with those of the finite-memory contagion channel and the Gilbert-Elliott channel. From the discussion in the previous chapter, we concluded the following.

- The uniform queue-based channel and the finite-memory contagion channel have identical block transition probabilities for the same memory, $BER$ and correlation coefficient. Therefore the two channels have identical capacity and burst frequency under the above conditions. It seems we can model the same noise process through two different channel models: one based on a queue and the other based on a contagion urn scheme.

- The uniform queue-based channel, the finite-memory contagion channel, and the Gilbert-Elliott channel have almost identical capacities and burst frequency when memory=1. This means that we possibly can replace the Gilbert-Elliott channel with the less-complex uniform queue-based channel with memory=1 (or the finite-memory contagion channel) if our objective is to achieve an error burst behavior and capacity that are close to those of the Gilbert-Elliott channel.

- Since the noise processes of the three channels are all stationary, we arrive at the conclusion that the average burst duration of the three channels are equal given equal $BER$; i.e., $E(B_n) = \frac{1}{1-BER}$.

- The capacity values for the non-uniform case are larger than those for the uniform case when $q_1 < \frac{1}{K}$ and smaller than those for the uniform case when $q_1 > \frac{1}{K}$. When $q_1 \to 1$, the non-uniform case converges to the uniform case with memory=1.

- The non-uniform queue-based channel has lower burst frequencies than the uniform queue-based channel for low correlation coefficients. At high correlation coefficit-nts, the former has higher burst frequencies than the latter and has lower burst frequency when memory is big enough.

# Bibliography


[1] J. A. Adoul, B. D. Fritchman and L. N. Kanal, "A Critical Statistic for Channels with Memory," *IEEE Trans. Inform. Theory*, Vol. IT-18, No. 1, pp. 133–141, January 1972.

[2] F. Alajaji and T. Fuja, "A Communication Channel Modeled on Contagion," *IEEE Trans. Inform. Theory*, Vol. 40, No. 6, pp. 2035–2041, November 1994.

[3] E. O. Elliott, "Estimates of Error Rates for Codes on Burst-Noise Channel," *Bell Syst. Tech. J.*, Vol. 42, pp. 1977-1997, Sept. 1963.

[4] W. Feller, *An Introduction to Probability Theory and its Applications*, 2nd ed. New York: Wiley, 1971.

[5] R. G. Gallager, *Information Theory and Reliable Communication*, New York: Wiley, 1968.

[6] E. N. Gilbert, "Capacity of a Burst-Noise Channel," *Bell Syst. Tech. J.*, Vol. 39, pp. 1253-1266, Sept. 1960.

[7] L. N. Kanal and A. R. K. Sastry, "Models for Channels with Memory and their Applications to Error Control," *Proceedings of the IEEE* , Vol. 66, No. 7, pp. 724-744, July 1978.

[8] M. Mushkin and I. Bar-David, "Capacity and Coding for the Gilbert-Elliott Channel," *IEEE Trans. Inform. Theory*, Vol. 35, No. 6, pp. 1277-1290, November 1989.

[9] K. Petersen, *Ergodic Theory*, Cambridge Univ. Press, 1983.

[10] C. J. L. Pimentel, "Enumeration Techniques for Finite State Channels," Ph.D. Thesis, Univ. of Waterloo, 1996.