A Binary Burst-Noise Communication Channel Modeled by a Finite Queue: Information Theoretic Properties and Applications to Wireless Communications

by

Libo Zhong

A thesis submitted to the

Department of Mathematics and Statistics

in conformity with the requirements for

the degree of Doctor of Philosophy

Queen's University

Kingston, Ontario, Canada

June, 2005

Abstract

A model for a binary additive noise communication channel with memory is introduced. The motivation of studying such a model is the fact that most real-world channels have memory and most current communication systems do not exploit this memory due (in part) to the lack of mathematically tractable models for such channels. The channel noise process, which is generated according to a ball sampling mechanism involving a queue of finite length M, is a stationary ergodic Mth order Markov source. The channel properties are analyzed and several of its statistical and information theoretical quantities (e.g., block transition distribution, autocorrelation function (ACF), capacity, and reliability function) are derived in either closed or easily computable form in terms of its four parameters. The capacity of the queue-based channel (QBC) is also analytically and numerically compared for a variety of channel conditions with the capacity of other binary models with memory, such as the well-known Gilbert-Elliott channel (GEC), the Fritchman channel and the finite-memory contagion channel.

We next investigate the modeling of the traditional GEC using this queue-based channel (QBC) model. The QBC parameters are estimated by minimizing the Kullback-Leibler divergence rate (KLDR) between the probability of noise sequences generated by the GEC and the QBC, while maintaining identical bit error rates (BER) and

correlation coefficients (Cor). The accuracy of fitting the GEC via the QBC is evaluated in terms of ACF, channel capacity and reliability function. Numerical results indicate that the QBC provides a good approximation of the GEC for various channel conditions.

Finally, we study the modeling of a family of hard-decision frequency-shift keying demodulated correlated Rician fading channels using the QBC model. As in the case of fitting the GEC, the QBC parameters are estimated by minimizing the KLDR between the distributions of error sequences generated by the QBC and the fading channels, and the modeling accuracy is evaluated in terms of channel capacity and ACF. Numerical results indicate that the QBC provides a very good approximation of the fading channels for a wide range of channel conditions; it thus offers an interesting alternative to the GEC and general finite-state Markov models recently studied by Pimentel *et. al.* while remaining mathematically tractable.

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 work.

I also would like to gratefully acknowledge the financial support received from the Natural Sciences and Engineering Research Council (NSERC) of Canada and the Premier's Research Excellence Award (PREA) of Ontario, the Ontario Graduate Scholarship (OGS), the Ontario Graduate Scholarship in Science and Technology (OGSST), the Department of Mathematics and Statistics, and the School of Graduate Studies of Queen's University.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

In recent years, there has been an increasing interest in transmitting voice, data, image and video signals over wireless communication channels. However, the quality of wireless communication systems is subject to a number of sources of degradation because wireless channels undergo a variety of time-varying channel impairments caused by propagation loss, shadowing, multipath fading, and thermal noise. In particular, it is important to understand the deleterious effects of fading on wireless transmission. A common feature of many fading channels is that they experience symbol errors during transmission in a bursty fashion [31] indicating a significant degree of correlation or dependence in the error or noise process.

In the presence of error bursts, interleaving is usually applied to destroy or mitigate the memory because most error control systems are designed under the assumption of

independent and identically distributed (iid) error processes. With perfect interleaving, it is possible to model the fading channels as binary symmetric channels (BSC). However, the use of interleaving introduces extra delay and complexity, and perfect interleavers do not exist in any practical system. In real-time personal communication systems, data is transmitted in short blocks and fairly strict delay constraints must be obeyed (e.g., see [68]). Non-interleaved or finite-interleaved packet transmission over fading channels has received significant recent attention [7, 42].

Therefore, in this thesis we start with the premise that the inherent memory of fading communication channels cannot be neglected. Actually, an advantageous feature of memory is that the channel quality in the near future can be forecast based on the knowledge of previous channel conditions due to the dependence of errors and hence greater capacity is achievable by receivers [37]. In order to obtain highly reliable data transmissions over channels with memory, we should take advantage of channel memory to design effective error control coding strategies. For this reason, it is critical to fully understand the error statistics of such channels. This is achieved via channel modeling, where the primary objective is to provide a model whose properties are both complex enough to closely capture the real channel statistical characteristics, yet simple enough to allow mathematically tractable system analysis. In this thesis, we present a binary additive channel model with memory based on a finite queue that is tractable (e.g., it has an explicit closed-form expression for channel capacity) and

reliably describes a family of wireless communication channels.

The first subject of the thesis is the investigation of a stationary ergodic Mth order Markov model for a binary additive noise communication channel with memory based on a finite queue of length M. The second subject is to approximate the well-known Gilbert-Elliott channel via the queue-based channel using parameter estimation and optimization. The third subject is the investigation of the modeling of a family of hard-decision frequency-shift keying demodulated correlated Rician fading channels via the queue-based model. In this chapter, we present the literature review of the articles upon which our research is based. We then specify our main research contributions. Finally, we outline the remainder of the thesis.

1.1 Literature Review

During the past several decades, a variety of channel models have been proposed and studied for wireless channels [32, 33, 36, 39, 43, 46, 53, 54, 55, 58]. A finite-state Markov channel (FSMC) is a discrete valued channel with a finite set of possible states whose transition is governed by an underlying Markov chain and with a probability assignment that is independent of time [23, 31, 34]. FSMC's have been widely adopted for the description of the correlation structures and success/failure processes of wireless channels with bursty behavior [62] because they are efficient in quick sim-

ulations, system performance evaluations and protocol investigations.

The most commonly used models for representing the "discretized" version (under hard-decision demodulation) of binary-input fading channels with memory are the Gilbert-Elliott channel (GEC) [25, 20] and the Fritchman channel (FC) [22]. These models have been partly adopted for historical reasons since they were introduced in the 1960s and were the first available models for channels with memory. They were for example employed to model high-frequency channels [61], mobile radio channels [1, 11, 57], low earth orbit satellite channels [13] and magnetic tape recorders [18]. The GEC model also has been used to evaluate the performance of coding and decoding schemes over bursty channels [48, 51, 68]. A generalized GEC was studied in [35].

Including the above works, many FSMC models have been proposed to fit realistic wireless channels. In [66], Wang and Moayeri proposed an FSMC based on the partitioning of the received signal-to-noise ratio (SNR) into a finite number of states to model Rayleigh fading channels. The same approach was also presented independently in [7] and used in [60, 69]. The model proposed in [66] attracted much attention because it has a good balance between accuracy and complexity. It was applied to the evaluation of system-related channel properties (such as the correlation properties of the fading mobile radio channel) in [77, 78] by modeling the channel as a first-order Markov process whose transition probabilities are a function of the channel characteristics. In [12], an analytical model was used to evaluate the effect of mobile

velocity on the performance of a communication system operating in a multi-path fading channel.

There are other papers assessing first-order FSMC models to represent the quantized Rayleigh fading channel [4, 59, 67]. First-order FSMC's were also used to simulate Rician fading channels in [5] and to simulate diversity Nakagami fading channels in [30]. Based on information metrics, statistical methods have been proposed to estimate how reliable such models are in applications (for example in communication protocol evaluation). Wang and Chang used an information theoretic criterion to show that the first-order Markov process is a good approximation to model the envelope of a Rayleigh channel in [67]. In [60], the limitations $\longrightarrow$ " of this criterion and the applicability of the first-order assumption were discussed. Several variations of the FSMC of unspecified order were examined, for example, in [6] for flat fading.

FSMC's are often generated via finite-state hidden Markov models (HMM's).[1] General HMM's were studied in [63] to model flat fading channels. Due to their HMM structure, such channels can be difficult to mathematically analyze (e.g., they do not admit an exact closed-form expression for their capacity and their block transition distribution is not transparently expressed in terms of the channel parameters), particularly when incorporated within an overall source and/or channel coded system.

[1] A description of other lesser known, but related, finite or infinite state HMM based channel models is provided in [31].

This may partly explain why, to date, few coding techniques that effectively exploit the noise memory have been successfully constructed for such HMM-based channel models and for channels with memory in general. Instead, most current wireless communication systems are designed for memoryless channels and employ channel block interleaving in an attempt to disperse the channel memory and render the channel "locally memoryless" – i.e., use a long interleaving span to spread the error bursts over several codewords so that the noise appears random within each codeword [37]. However, the use of interleaving results in increased complexity and delay. More importantly, the failure to exploit the channel's memory leads to a waste of channel capacity since it is well known that *memory increases capacity*[2] for a wide class of channels (the class of information stable channels [17, 2]). It is therefore vital to construct channel models which can well represent the behavior of real-world channels while remaining analytically tractable for design purposes.

In [2, Section VI], Alajaji and Fuja proposed a simple binary additive noise channel with memory, referred to as the finite memory contagion channel (FMCC), where the noise process is generated via a finite-memory version of Polya's urn scheme for the spread of a contagious disease through a population [45]. In such a channel, every error (or "infection", if we use the contagion interpretation) effectively increases the

[2]In other words, the capacity of the "equivalent" memoryless channel achieved by ideal interleaving (with infinite interleaving span) is smaller than the capacity of the original channel (e.g., see [37]).

probability of future errors ([45], [21, p. 57]), and hence may lead to a clustering or burst of errors (i.e., an "epidemic" in the population). The resulting channel has a stationary ergodic Mth order Markov noise source and is fully described by only three parameters. Furthermore, it admits single-letter analytical expressions for its block transition distribution and capacity, which is an attractive feature for mathematical analysis. This model was adopted in several joint source-channel coding studies (e.g., [3, 9, 10, 28, 29, 41, 56]) where the channel statistics are incorporated into the system design in order to exploit the noise memory. The construction of low density parity check (LDPC) codes for this channel was also recently investigated in [38].

It is also important to point out that Pimentel *et. al.* recently showed in a numerical study [44] that the class of binary channel models with additive Mth order Markov noise (to which both the FMCC model and the channel model studied in this thesis belong) is a good approximation, in terms of the autocorrelation function (ACF) and variational distance, to the family of hard-decision frequency-shift keying demodulated correlated Rayleigh and Rician fading channels for a good range of fading environments, particularly for medium and fast fading rates. Note however, that the general Kth order Markov noise channels considered in [44] have a complexity (number of parameters) that grows exponentially with M, rendering it impractical for the modeling of channels with large memory such as very slow Rayleigh fading channels (e.g., see [44, Fig. 11]).

1.2 Contribution

In this thesis, we introduce a new binary additive noise channel based on a finite queue of length M. The proposed queue-based channel (QBC) model features an Mth order Markov noise source that is fully characterized by four parameters (M, p, ϵ, α), making it more sophisticated than the FMCC for channel modeling (as it has an additional degree of freedom) while remaining mathematically tractable. Our QBC model does not suffer from the limitation of complexity as it is fully described by only four parameters and it can accommodate very large values of the memory M. It also enjoys a transparent formula for its n-fold statistics and a closed form formula for its capacity, which are appealing features as they provide powerful analytical tools for code design and system analysis.

The contributions of this thesis (parts of which appeared in [70]-[76]) are as follows.

- Investigation of the statistical properties of the QBC, such as state transition matrix, stationary distribution, and proof of a recursion property of the stationary distribution.

- Closed form formulas for the block transition probability and capacity and a recursive expression for the ACF in terms of the four parameters of the QBC.

- Proofs of two important theorems on the QBC capacity as a function of the parameters α and M.

- Analytical and numerical comparison studies of the QBC, the FMCC, the GEC and the Fritchman channel in terms of capacity.

- Statistical approximation of the GEC via the QBC model using parameter estimation and optimization. The accuracy of the approximation is measured in terms of channel capacity, ACF and channel reliability function.

- Approximating the class of hard-decision demodulated Rician fading channels with memory, referred to as the discrete channel with Clarke's autocorrelation (DCCA) model [44] via the QBC.

1.3 Thesis Overview

The rest of this thesis is organized as follows.

In Chapter 2, we review some useful properties and results regarding discrete stochastic processes and discrete Markov chains. We next present information theoretic quantities, such as capacity and reliability function, of discrete communication channels. We also discribe three previous discrete-time binary additive channel models with memory: the GEC, the Fritchman channel and FMCC.

In Chapter 3, the QBC is introduced and its statistical (noise stationary distribution and block transition probability) and information theoretic (capacity and reliability function) quantities are investigated. The channel is also studied in the

special case when the queue cells are operated on uniformly; the resulting channel is called the uniform queue-based channel (UQBC). The QBC is next compared analytically and numerically in terms of channel capacity with the FMMC, the GEC and a particular class of the Fritchman channel.

In Chapter 4, we study the approximation of the GEC via our QBC model. For a given GEC, we construct its "closest" QBC peer in the Kullback-Leibler distance sense; i.e., we estimate the QBC parameters by minimizing the Kullback-Leibler divergence rate (KDLR) between the block transition probabilities of both channels, under the constraint of maintaining identical bit error rates (BER) and correlation coefficients (Cor) (between two consecutive noise samples). We then evaluate the accuracy of the fit between the QBC and the GEC in terms of channel capacity, ACF and channel reliability function (or error exponent). Numerical results show that the QBC provides a good approximation of the GEC for a broad range of channel conditions, and it thus offers an attractive alternative to the GEC.

In Chapter 5, we study the problem of approximating the class of discretized correlated Rician fading channels (the discrete channels with Clarke's autocorrelation (DCCA) model [44]) via the QBC. For a given DCCA, we construct a QBC whose noise process is statistically "close" in the Kullback-Leibler sense to the error or noise process generated by the DCCA under the constraint of maintaining identical BER and Cor. We evaluate the accuracy of the fit between the QBC and the DCCA in

terms of ACF and channel capacity. Numerical results show that the QBC provides a very good approximation of the DCCA in terms of channel capacity and ACF for a wide range of fading rates (including slow fading rates), making the QBC a better fit than the GEC and Markov models studied in [44].

In Chapter 6, we conclude with a summary along with directions for future work.

Chapter 2

Preliminaries

We begin with some useful definitions and important properties about discrete Markov processes and communication channels that can be found in textbooks on information theory such as [15, 23, 26].

2.1 Discrete Markov Processes

Definition 2.1 *A stochastic process* $\{Z_1, Z_2, \cdots\}$ *with finite-alphabet* $\mathcal{K} = \{0, 1, 2, \cdots, L-1\}$ *is said to be* stationary *if the joint distribution of any subset of the sequence of random variables is invariant with respect to time shifts; i.e.,*

$$\Pr\{Z_1 = z_1, Z_2 = z_2, \cdots, Z_n = z_n\} = \Pr\{Z_{1+\tau} = z_1, Z_{2+\tau} = z_2, \cdots, Z_{n+\tau} = z_n\},$$

for every time shift τ *and for all* $z_1, \cdots, z_n \in \mathcal{K}$.

Throughout this work, we consider the binary case, i.e., $\mathcal{K} = \{0, 1\}$. In practice, many sources can be well modeled via stationary processes.

Definition 2.2 *A discrete process* $\{Z_1, Z_2, \cdots\}$ *with finite-alphabet* $\mathcal{K}$ *is said to be a* Markov chain *or* Markov process *if for* $n = 1, 2, \cdots$

$$\begin{aligned} & \Pr\{Z_n = z_n | Z_{n-1} = z_{n-1}, Z_{n-2} = z_{n-2}, \cdots, Z_1 = z_1\} \\ = \; & \Pr\{Z_n = z_n | Z_{n-1} = z_{n-1}\}, \end{aligned}$$

for all $z_1, \cdots, z_n \in \mathcal{K}$. *In this case,*

$$\Pr\{Z^n = z^n\} = \Pr\{Z_1 = z_1\} \prod_{i=2}^{n} \Pr\{Z_i = z_i | Z_{i-1} = z_{i-1}\},$$

where $z^n \stackrel{\triangle}{=} (z_1, z_2, \cdots, z_n)$.

Furthermore, a process is a Markov process of order M *(or memory* M*), where* $M > 0$ *is fixed, if*

$$\begin{aligned} & \Pr\{Z_n = z_n | Z_{n-1} = z_{n-1}, Z_{n-2} = z_{n-2}, \cdots, Z_1 = z_1\} \\ = \; & \Pr\{Z_n = z_n | Z_{n-1} = z_{n-1}, Z_{n-2} = z_{n-2}, \cdots, Z_{n-M} = z_{n-M}\}, \end{aligned}$$

for $n > M$ *and for all* $z_1, \cdots, z_n \in \mathcal{K}$.

Define $\{\underline{S}_n\}$ as the process obtained by M*-step blocking* the Markov process $\{Z_n\}$; i.e.,

$$\underline{S}_n \stackrel{\triangle}{=} (Z_n, Z_{n+1}, \cdots, Z_{n+M-1}).$$

Then

$$\begin{aligned} & \Pr\{\underline{S}_n = \underline{s}_n | \underline{S}_{n-1} = \underline{s}_{n-1}, \cdots, \underline{S}_1 = \underline{s}_1\} \\ = \; & \Pr\{Z_{n+M-1} = z_{n+M-1} | Z_{n+M-2} = z_{n+M-2}, \cdots, Z_{n-1} = z_{n-1}\} \\ = \; & \Pr\{\underline{S}_n = \underline{s}_n | \underline{S}_{n-1} = \underline{s}_{n-1}\}, \end{aligned}$$

and hence, $\{\underline{S}_n\}$ is a Markov process with $|\mathcal{K}|^M$ states.

Definition 2.3 *A Markov process is said to be* time-invariant *(or* homogeneous*) if the conditional probability* $\Pr\{Z_n = z_n | Z_{n-1} = z_{n-1}\}$ *does not depend on* n*, i.e.,*

$$\Pr\{Z_n = a | Z_{n-1} = b\} = \Pr\{Z_2 = a | Z_1 = b\},$$

for $n > 1$ *and for all* $a, b \in \mathcal{K}$*.*

For a Markov process $\{Z_n\}_{n=1}^{\infty}$, Z_n is called the *state* at time n. A homogeneous Markov process is characterized by its initial state distribution (i.e., the distribution of Z_1) and a *probability transition matrix* $\boldsymbol{Q}$ whose (i,j)th entry p_{ij} is the conditional probability of a transition from state i to state j, i.e.,

$$p_{ij} \stackrel{\triangle}{=} \Pr\{Z_n = j | Z_{n-1} = i\}.$$

Furthermore, a Markov process is *irreducible* if any state can reach any other state in a finite number of steps with positive probability.

Definition 2.4 *For a Markov process* $\{Z_n\}$*, a distribution on the states such that the distribution at time* $n+1$ *is the same as the distribution at time* n *is called a* stationary distribution *and is denoted by* $\boldsymbol{\pi} \stackrel{\triangle}{=} (\pi_0;\, \pi_1; \cdots;\, \pi_{|\mathcal{K}|-1})$.

Remark: For a finite-alphabet Markov process with probability transition matrix $\boldsymbol{Q}$, its stationary distribution $\boldsymbol{\pi}$ always exists [23, p. 110] and can be obtained by solving $\boldsymbol{\pi} = \boldsymbol{\pi}\boldsymbol{Q}$. Furthermore, if the initial state of a homogeneous Markov process is drawn according to the stationary distribution $\boldsymbol{\pi}$, then the Markov process is a stationary process.

Definition 2.5 *The* entropy rate *of a stochastic process* $\{Z_1, Z_2, \cdots\}$ *is defined by*

$$\mathcal{H}(Z) = \lim_{n\to\infty} \frac{1}{n} H(Z_1, Z_2, \cdots, Z_n)$$

when the limit exists, where

$$H(Z_1, Z_2, \cdots, Z_n) \stackrel{\triangle}{=} -\sum_{z_1,\cdots,z_n} \Pr\{Z_1 = z_1, \cdots, Z_n = z_n\} \log \Pr\{Z_1 = z_1, \cdots, Z_n = z_n\}.$$

Definition 2.6 *We can also define a related quantity for entropy rate:*

$$\mathcal{H}'(Z) = \lim_{n\to\infty} H(Z_n | Z_{n-1}, Z_{n-2}, \cdots, Z_1)$$

when the limit exists, where

$$\begin{aligned} & H(Z_n | Z_{n-1}, \cdots, Z_1) \\ \stackrel{\triangle}{=} \; & -\sum_{z_1,\cdots,z_n} \Pr\{Z_1 = z_1, \cdots, Z_n = z_n\} \log \Pr\{Z_n = z_n | Z_{n-1} = z_{n-1}, \cdots, Z_1 = z_1\}. \end{aligned}$$

For a stationary process, $H(Z_n|Z_{n-1}, Z_{n-2}, \cdots, Z_1)$ and $\frac{1}{n}H(Z_1, Z_2, \cdots Z_n)$ are decreasing in n (see [15], Theorem 4.2.2 and Problem 4.a on P. 74).

Proposition 2.1 **[15]** *For a stationary source, $\mathcal{H}(Z)$ and $\mathcal{H}'(Z)$ exist and are equal.*

Proposition 2.2 *Let $\{Z_1, Z_2, \cdots\}$ be a stationary Markov process with stationary distribution $\boldsymbol{\pi}$ and transition matrix $\boldsymbol{Q}$. Then its entropy rate is given by*

$$\mathcal{H}(Z) = H(Z_2|Z_1) = -\sum_{i,j\in\mathcal{K}} \pi_i p_{ij} \log p_{ij},$$

where p_{ij} is the (i,j)th entry of $\boldsymbol{Q}$. The same result also holds for irreducible (not necessarily stationary) homogeneous Markov processes.

Proposition 2.3 **[52]** *A stationary (finite-state) Markov process is ergodic if and only if it is irreducible.*

Proposition 2.4 **[26]** *Let p and q be two possible distributions for a source $\{Z_n\}$ with n-dimensional distributions $p^{(n)}$ and $q^{(n)}$, respectively. If $\{Z_n\}$ is stationary under p and homogeneous Mth order Markov under q, then the* Kullback-Leibler divergence rate *between p and q is given by*

$$\lim_{n\to\infty} \frac{1}{n} D(p^{(n)} \| q^{(n)}) \quad = \quad -\mathcal{H}_p(Z) - E_p[\log q\{Z_{M+1}|Z^M\}],$$

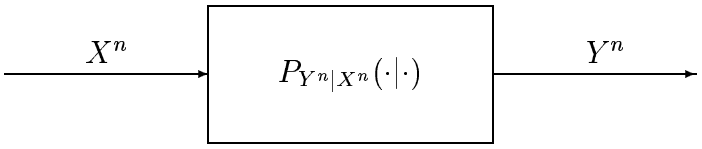
where $\mathcal{H}_p(Z)$ is the entropy rate of p which exists by Proposition 2.1. $E_p[X]$ denotes the expectation with respect to p and $q\{Z_{M+1}|Z^M\}$ denotes the conditional probability given M previous symbols.

2.2 Discrete Communication Channels: Information Theoretic Properties

Most of the following can be found in [23, 50].

Definition 2.7 *A* discrete communication channel *is a system with:*

- *A finite input alphabet* $\mathcal{X}$*;*
- *A finite output alphabet* $\mathcal{Y}$*;*
- *Transition probabilities:* $\{P_{Y^n|X^n}(y^n \mid x^n)\}_{n=1}^{\infty}$ *of receiving the n-tuple* $Y^n = y^n$ *at the channel output given that* $X^n = x^n$ *was sent, as illustrated below.*



Definition 2.8 *A* discrete memoryless channel *(DMC) satisfies*

$$P_{Y^n|X^n}(y^n \mid x^n) = \prod_{i=1}^{n} P_{Y_i|X_i}(y_i \mid x_i).$$

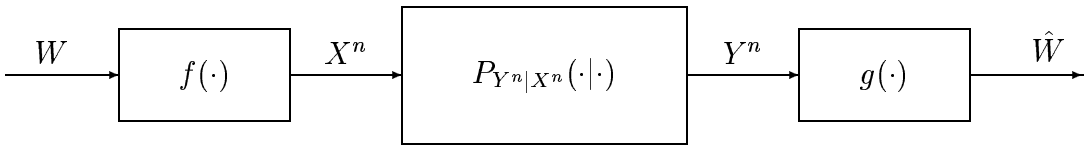
The DMC is uniquely determined by the channel transition matrix $\boldsymbol{Q} = [P(y|x)]$ for $x \in \mathcal{X}$ and $y \in \mathcal{Y}$. Note that for channels with memory the above property does *not* hold.

Definition 2.9 *A* code $\mathcal{C}$ *with blocklength* n *and size* K *for a discrete channel* $\{P_{Y^n|X^n}(y^n \mid x^n)\}$ *consists of an encoder-decoder pair* (f, g)*:*

$$f : \{1, 2, \ldots, K\} \to \mathcal{X}^n,$$

and

$$g : \mathcal{Y}^n \to \{1, 2, \ldots, K\}.$$



- Encoder*: encodes message* $W \in \{1, 2, \ldots, K\}$ *with codeword*

$$f(W) = X^n = (X_1, X_2, \ldots, X_n).$$

- Decoder*: for received* $Y^n = (Y_1, Y_2, \ldots, Y_n)$*, chooses as estimate of the message* $\hat{W} = g(Y^n)$.

- Code rate *is:* $R(\mathcal{C}) = (1/n) \log_2 K$ *bits/channel use.*

We assume that the message W *is uniform over* $\{1, 2, \ldots, K\}$*. Then the code's* probability of decoding error *is*

$$P_e(\mathcal{C}) = \Pr\{\hat{W} \neq W\} = \frac{1}{K} \sum_{w=1}^{K} \Pr\{Y^n \notin \mathcal{B}_w \mid w \text{ sent}\},$$

where

$$\mathcal{B}_w = \{y^n \in \mathcal{Y}^n : g(y^n) = w\},$$

and

$$\Pr\{Y^n \notin \mathcal{B}_w \mid w \text{ sent}\} = \sum_{y^n \notin \mathcal{B}_w} P_{Y^n|X^n}(y^n \mid f(w)).$$

Definition 2.10 *Rate R is* achievable *if there exists a sequence of codes with blocklength n and rate no less than R such that*

$$\lim_{n\to\infty} P_e(\mathcal{C}) = 0.$$

For any achievable rate R, it is possible to construct a code that conveys R bits of information with every use of the channel and yields arbitrarily small probability of error for sufficiently large blocklength. The channel capacity C is defined as the supremum of all achievable code rates.

Definition 2.11 *Fix $R \geq 0$ and let $P_e^*(R,n)$ denoted the smallest possible error probability of any block code $\mathcal{C}$ for a given channel with blocklength n and rate no less than R:*

$$P_e^*(R,n) \stackrel{\triangle}{=} \inf_{\mathcal{C}\subset\mathcal{X}^n:\ R(\mathcal{C})\geq R} P_e(\mathcal{C}).$$

Then the channel reliability function *is defined as the rate of (asymptotic) exponential decay of $P_e^*(R,n)$:*

$$E(R) \stackrel{\triangle}{=} \lim_{n\to\infty} -\frac{1}{n}\log_2 P_e^*(R,n),$$

assuming the limit exists. If not, then lower and upper bounds are given by

$$\liminf_{n\to\infty} -\frac{1}{n}\log_2 P_e^*(R,n) \text{ and } \limsup_{n\to\infty} -\frac{1}{n}\log_2 P_e^*(R,n).$$

It is shown via the Channel Coding Theorem that the capacity of a "wide" class of channels with memory considered in Information Theory admits the following expression:

$$C = \lim_{n\to\infty} \sup_{X^n} \frac{1}{n} I(X^n; Y^n),$$

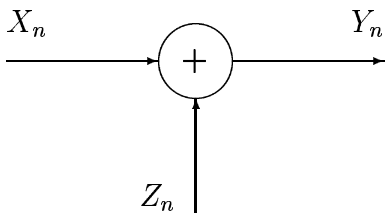
where $I(X^n; Y^n)$ is the *mutual information* between X^n and Y^n

$$I(X^n; Y^n) \stackrel{\triangle}{=} \sum_{y^n \in \mathcal{Y}^n} \sum_{x^n \in \mathcal{X}^n} P_{X^n Y^n}(x^n, y^n) \log \frac{P_{X^n Y^n}(x^n, y^n)}{P_{X^n}(x^n) P_{Y^n}(y^n)}.$$

By "wide" class of channels, we mean *information stable* channels [64]– i.e. the channel input X^n that maximizes $I(X^n; Y^n)$ and its output behave ergodically (e.g., the asymptotic equipartition property (AEP) holds). In particular, for a DMC:

$$C = \sup_X I(X; Y).$$

A formula for the capacity of *arbitrary* (non-information stable, non-stationary, etc) channels was given by Verdú/Han [64], using generalized information measures and information spectrum techniques.



$$Y_n = X_n \oplus Z_n$$

Figure 2.1: Discrete-time, binary, additive noise communication channel.

Hereafter, a *discrete-time binary additive noise communication channel* (as shown in Fig. 2.1) refers to a channel with common input, noise and output alphabet $\mathcal{X} = \mathcal{Z} = \mathcal{Y} = \{0, 1\}$ described by $Y_n = X_n \oplus Z_n$, for $n = 1, 2, 3, \cdots$, where $\oplus$ denotes addition modulo 2, and where X_n, Z_n, and Y_n denote, respectively, the channel input, noise, and output at time n. Hence a transmission error occurs at time n whenever $Z_n = 1$. It is assumed that the input and noise sequences are independent of each other. In this thesis, a given noise process $\{Z_n\}_{n=1}^{\infty}$ will be generated according to one of the models introduced in Section 2.3 (the GEC, the FC and the FMCC) and our proposed queue-based channel model (QBC).

Example: Assume that $\{Z_n\}_{n=1}^{\infty}$ in the above channel is stationary and ergodic. Then this channel with memory is information stable [64] and its capacity is given by

$$C = \lim_{n\to\infty} \sup_{X^n} \frac{1}{n} I(X^n; Y^n) = 1 - \mathcal{H}(Z), \tag{2.1}$$

where $\mathcal{H}(Z)$ is the entropy rate of $\{Z_n\}_{n=1}^{\infty}$.

2.3 Previous Binary Channels with Memory

2.3.1 Gilbert-Elliott Channel

The GEC model belongs to the family of finite-state channels, which is thoroughly studied in [23, pp. 97-111]. It is driven by an underlying stationary ergodic Markov chain with two states: a good state and a bad state, denoted by G and B, respectively

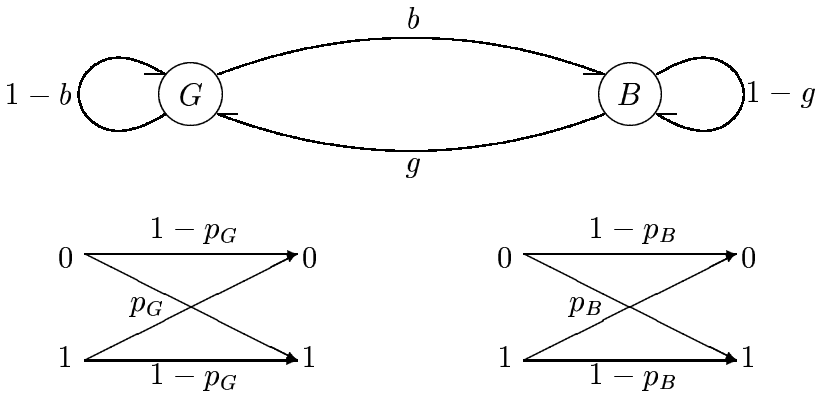


Figure 2.2: The Gilbert-Elliott channel model.

(see Fig. 2.2). In a fixed state, the channel behaves like a binary symmetric channel (BSC). The GEC is thus a time-varying BSC, where p_G and p_B are the crossover probabilities in the good and bad states, respectively (the Gilbert Channel (GC) [25] is obtained when $p_G = 0$, i.e., it behaves like a noiseless BSC in the good state). After every channel transmission, the Markov chain makes a state transition according to the transition probability matrix

$$\boldsymbol{P} = \begin{bmatrix} 1-b & b \\ g & 1-g \end{bmatrix},$$

where $0 < b < 1$ and $0 < g < 1$.

The GEC is a binary additive channel: $Y_n = X_n \oplus Z_n$, where the noise $\{Z_n\}$ is a stationary ergodic *hidden* Markov source (infinite memory). A useful approach for calculating the probability of an error or noise sequence for the GEC is discussed in [43]. By the law of total probability, the probability of a noise sequence of length n,

$z^n = (z_1, z_2, \cdots, z_n)$, may be expressed as

$$\Pr\{Z^n = z^n\} = \boldsymbol{\pi}^T \left(\prod_{k=1}^{n} \boldsymbol{P}(z_k) \right) \mathbf{1}, \tag{2.2}$$

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

$$\mathbf{P}(0) = \begin{bmatrix} (1-b)(1-p_G) & b(1-p_B) \\ g(1-p_G) & (1-g)(1-p_B) \end{bmatrix}, \tag{2.3}$$

$$\mathbf{P}(1) = \begin{bmatrix} (1-b)p_G & bp_B \\ gp_G & (1-g)p_B \end{bmatrix}. \tag{2.4}$$

$\mathbf{1}$ is the 2-dimensional vector with all ones and the vector $\boldsymbol{\pi}$ indicates the stationary distribution vector of the underlying Markov chain,

$$\boldsymbol{\pi} = \begin{bmatrix} \pi_0 \\ \pi_1 \end{bmatrix} = \begin{bmatrix} \frac{g}{b+g} \\ \frac{b}{b+g} \end{bmatrix}. \tag{2.5}$$

2.3.2 Fritchman Channel

In 1967, Fritchman proposed a class of models by partitioning the state space $\Omega = \{0, 1, \cdots, N-1\}$ of an N-state stationary ergodic Markov chain into two groups of states $\Lambda_0 = \{0, 1, \cdots, K-1\}$ and $\Lambda_1 = \{K, K+1, \cdots, N-1\}$ resulting in the so-called $(K, N-K)$ Fritchman channel (FC) [22]. Corresponding to this partition, the state transition probability matrix $\boldsymbol{P}$ and the stationary vector $\boldsymbol{\pi}$ can be written

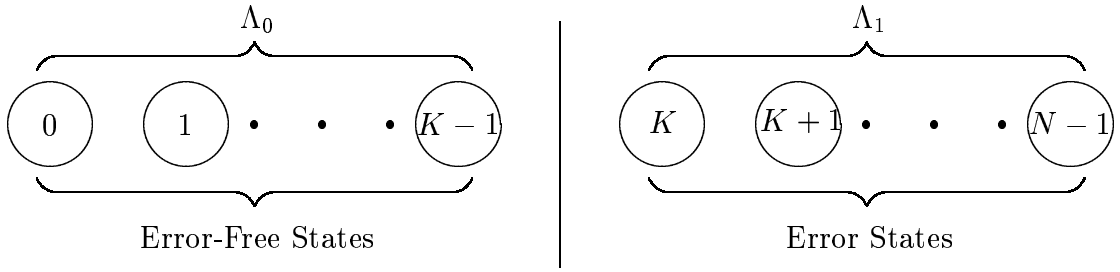


Figure 2.3: Partitioning of the state space of the Fritchman channel model.

in the block form

$$\boldsymbol{P} = \begin{bmatrix} \boldsymbol{P}_{00} & \boldsymbol{P}_{01} \\ \boldsymbol{P}_{10} & \boldsymbol{P}_{11} \end{bmatrix},$$

and $\boldsymbol{\pi} = [\boldsymbol{\pi}_0 \ \boldsymbol{\pi}_1]$, where the submatrix $\boldsymbol{P}_{ij}$ contains the transition probabilities from the set Λ_i to Λ_j. The noise process at the nth time interval Z_n is generated by a deterministic function of the state S_n at the nth time interval

$$Z_n = \begin{cases} 0, & \text{for } S_n \in \Lambda_0, \\ 1, & \text{for } S_n \in \Lambda_1. \end{cases}$$

In [43], Pimentel and Blake represent the probability of a noise sequence of length n by

$$\Pr\{Z^n = z^n\} = \boldsymbol{\pi}_{z_1}^T \left(\prod_{l=1}^{n-1} \boldsymbol{P}_{z_l z_{l+1}} \right) \mathbf{1}. \tag{2.6}$$

2.3.3 Finite-Memory Contagion Channel

The noise process of the FMCC [2] is generated 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 jth draw, $j = 1, 2, \cdots$, we select a ball from the urn and replace it with $1 + \Delta$ balls of the same color ($\Delta > 0$); then M draws later - after the $(j + M)$th draw - we retrieve from the urn Δ 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 $\{Z_i\}_{i=1}^{\infty}$ corresponds to the outcomes of the draws from the urn, where

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

It can be shown that the noise process $\{Z_i\}_{i=M+1}^{\infty}$ is a stationary ergodic Markov source of order M [2]. For an input block $X^n = (X_1, X_2, \cdots, X_n)$ and an output block $Y^n = (Y_1, Y_2, \cdots, Y_n)$, the block transition probability

$$\Pr\{Y^n = y^n \mid X^n = x^n\} = \Pr\{Z^n = z^n\},$$

where $z_i = y_i \oplus x_i$ for $i = 1, \cdots, n$, can expressed as follows [2]:

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

$$\Pr\{Z^n = z^n\} = \frac{\rho(\rho+\delta)\cdots[\rho+(d-1)\delta]\sigma(\sigma+\delta)\cdots[\sigma+(n-d-1)\delta]}{(1+\delta)(1+2\delta)\cdots[1+(n-1)\delta]}, \quad (2.7)$$

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

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

$$\Pr\{Z^n = z^n\} = L \prod_{i=M+1}^{n} \left[\frac{\rho + \lambda_{i-1}\delta}{1+M\delta}\right]^{z_i} \left[\frac{\sigma + (M-\lambda_{i-1})\delta}{1+M\delta}\right]^{1-z_i}, \qquad (2.8)$$

  where

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

  $\prod_{j=0}^{a}(\cdot) \stackrel{\triangle}{=} 1$, if $a < 0$, $z_i = x_i \oplus y_i$, and $\lambda_{i-1} = z_{i-1} + \cdots + z_{i-M}$ for $i = M+1, \cdots, n$.

Theorem 2.1 [2] *The resulting channel noise is stationary ergodic Markov with memory M, and the channel capacity increases with the memory M (for fixed BER and correlation) and is given by*

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

where

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

$\rho = R/T$ is the channel BER, and $\delta = \Delta/T$ is a correlation parameter.

Notice that the FMCC is fully described by three parameters: the BER ρ, M and δ. When $\delta = 0$, FMCC $\Longleftrightarrow$ BSC(ρ), and $C_{\text{FMCC}}^{(M)} > C_{\text{BSC}}$ for $\delta > 0$, which means that ideal channel interleaving is *bad* (as expected).

Chapter 3

Queue-Based Channel with Memory

In this chapter, we first present the queue-based binary channel with memory whose additive noise process is generated according to a sampling mechanism involving the following two parcels.

- **Parcel 1** is a queue of length M as shown in Fig. 3.1, that contains initially M balls, either red or black.

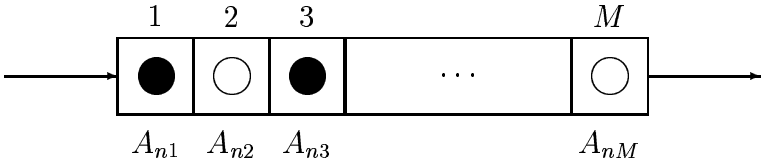


Figure 3.1: A queue of length M.

The random variables A_{nk} (n is a time index referring to the nth experiment, $n \geq 1$; k represents the position in the queue, $k = 1, 2, \cdots, M$) are defined by:

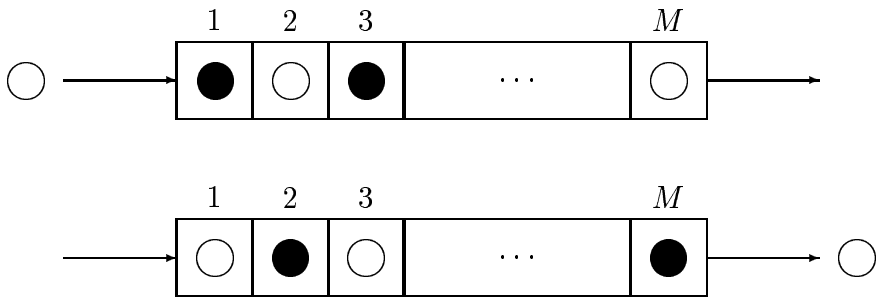
$$A_{nk} = \begin{cases} 1, & \text{if the } k\text{th cell contains a red ball,} \\ 0, & \text{if the } k\text{th cell contains a black ball.} \end{cases}$$

- **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 $p \in (0,1)$, $p \ll 1/2$.

We assume that the probability of selecting parcel 1 (the queue) is ε, while the probability of selecting parcel 2 (the urn) is $1-\varepsilon$ and $\varepsilon \in [0,1)$. Notice that the channel is actually a memoryless binary symmetric channel (BSC) with crossover probability p when $\varepsilon = 0$, in which case we experiment on the urn only.

The noise process $\{Z_n\}_{n=1}^{\infty}$ is generated according to the following procedure. By flipping a biased coin (with Pr(Head)=ε), we select one of the two parcels (select the queue if Heads and the urn if Tails). If parcel 2 (the urn) is selected, a pointer randomly points at a ball, and identifies its color. If parcel 1 (the queue) is selected, the procedure is determined by the length of the queue. If $M \geq 2$, a pointer points at the ball in cell k with probability $1/(M-1+\alpha)$, for $k = 1, 2, \cdots, M-1$ and $\alpha \geq 0$, and points at the ball in cell M with probability $\alpha/(M-1+\alpha)$, and identifies its color. If $M = 1$, a pointer points at the ball in the only cell of the queue with probability 1; in this case, we set $\alpha = 1$. If the selected ball from either parcel is

red (respectively black), we introduce a red (respectively black) ball in cell 1 of the queue, pushing the last ball in cell M out.



The noise process $\{Z_n\}_{n=1}^{\infty}$ is then modeled as follows:

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

We define the state of the channel to be $\underline{S}_n \stackrel{\triangle}{=} (A_{n1}, A_{n2}, \cdots, A_{nM})$, the binary $M-$tuple in the queue after the nth experiment is completed. Note that, in terms of the noise process, the channel state at time n can be written as $\underline{S}_n = (Z_n, Z_{n-1}, \cdots, Z_{n-M+1})$, for $n \geq M$. If $\varepsilon = 1$ (i.e., we always choose the queue), the channel state at time n can be either all 1s or all 0s for n sufficiently large since the two states are absorbing states. In this case the process $\{\underline{S}_n\}$ is reducible; hence it is non-ergodic.

3.1 Statistical and Information Theoretical Properties of the QBC

We next investigate the properties of the binary noise process $\{Z_n\}_{n=1}^{\infty}$. We first observe that, for $n \geq M+1$,

$$\begin{aligned}
&\Pr^{(M)}\{Z_n = 1 \mid Z_{n-1} = z_{n-1}, \cdots, Z_1 = z_1\} \\
&= \varepsilon \left(z_{n-1} \cdot \frac{1}{M-1+\alpha} + \cdots + z_{n-M+1} \cdot \frac{1}{M-1+\alpha} + z_{n-M} \cdot \frac{\alpha}{M-1+\alpha} \right) + (1-\varepsilon)p \\
&= \Pr\{Z_n = 1 \mid Z_{n-1} = z_{n-1}, \cdots, Z_{n-M} = z_{n-M}\}, \qquad (3.1)
\end{aligned}$$

where $z_l \in \{0,1\}$, $l = 1, \cdots, n-1$. Hence $\{Z_n\}_{n=1}^{\infty}$ is a homogeneous Markov process of order M.

3.1.1 Stationary Distribution

Throughout this work, we consider the case where the initial distribution of the channel state $\{\underline{S}_n\}_{n=1}^{\infty}$ is drawn according to its stationary distribution; hence the noise process $\{Z_n\}_{n=1}^{\infty}$ is stationary. In this section, we first give two examples of first-order and second-order QBC models for simplicity. Later we generalize the results for the QBC of arbitrary order M.

Example 1:

For $M = 1$, the noise process $\{Z_n\}_{n=1}^{\infty}$ is a simple stationary Markov process. The channel state process $\{\underline{S}_n\}_{n=1}^{\infty}$ is a homogeneous Markov process with 2 states, and its transition matrix is given by

$$\boldsymbol{Q}_{\mathrm{QBC}}^{(1)} = \begin{bmatrix} p_{00} & p_{01} \\ p_{10} & p_{11} \end{bmatrix} = \begin{bmatrix} \varepsilon + (1-\varepsilon)(1-p) & (1-\varepsilon)p \\ (1-\varepsilon)(1-p) & \varepsilon + (1-\varepsilon)p \end{bmatrix}.$$

We can observe from the transition matrix $\boldsymbol{Q}_{\mathrm{QBC}}^{(1)}$ 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). Consequently, the stationary distribution $\boldsymbol{\pi}^{(1)} \stackrel{\triangle}{=} (\pi_0^{(1)};\ \pi_1^{(1)})$ of the process exists and is unique. Solving $\boldsymbol{\pi}^{(1)} = \boldsymbol{\pi}^{(1)}\boldsymbol{Q}_{\mathrm{QBC}}^{(1)}$, the stationary distribution of the process is given by $\boldsymbol{\pi}^{(1)} = (1-p;\ p)$.

Example 2:

For $M = 2$, the noise process $\{Z_n\}_{n=1}^{\infty}$ is a 2nd order Markov process. The channel state process $\{\underline{S}_n\}_{n=1}^{\infty}$ 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).

Since $\{\underline{Z}_n\}_{n=1}^{\infty}$ is stationary, it is also 2-step block stationary, which implies that

$\{\underline{S}_n\}_{n=1}^{\infty}$ is a homogeneous Markov process with stationary distribution

$$\boldsymbol{\pi}^{(2)} \triangleq (\pi_0^{(2)};\ \pi_1^{(2)};\ \pi_2^{(2)};\ \pi_3^{(2)}).$$

If $p_{ij}^{(2)}$ denotes the transition probability that $\underline{S}_n$ goes from state i to state j, $i, j = 0, 1, 2, 3$, the transition matrix of the process $\{\underline{S}_n\}_{n=1}^{\infty}$ can be written as

$$\boldsymbol{Q}_{\mathrm{QBC}}^{(2)} = \begin{bmatrix} p_{00} & p_{01} & p_{02} & p_{03} \\ p_{10} & p_{11} & p_{12} & p_{13} \\ p_{20} & p_{21} & p_{22} & p_{23} \\ p_{30} & p_{31} & p_{32} & p_{33} \end{bmatrix}$$

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

From the transition matrix $\boldsymbol{Q}_{\mathrm{QBC}}^{(2)}$, 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). Solving $\boldsymbol{\pi}^{(2)} = \boldsymbol{\pi}^{(2)}\boldsymbol{Q}_{\mathrm{QBC}}^{(2)}$, the stationary distribution of the state process is given by

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

General case

We now analyze the noise process $\{Z_n\}_{n=1}^{\infty}$ for the general case as a function of M. The noise process, as we showed earlier, is an Mth order stationary Markov process. The channel state process $\{\underline{S}_n\}_{n=1}^{\infty}$ is a homogeneous Markov process with stationary distribution

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

where state i gives the decimal representation of the corresponding binary M-tuple.

If $p_{ij}^{(M)}$ denotes the transition probability that $\underline{S}_n$ goes from state i to state j, $i, j = 0, 1, \cdots, 2^M - 1$, the transition matrix of the process $\{\underline{S}_n\}_{n=1}^{\infty}$ can be written as $\boldsymbol{Q}_{\text{QBC}}^{(M)} = \left[p_{ij}^{(M)}\right]$ with

$$p_{ij}^{(M)} = \begin{cases} \left(M - \omega_i^{(M)} - 1 + \alpha\right) \frac{\varepsilon}{M-1+\alpha} + (1-\varepsilon)(1-p), & \text{if } j = \frac{i}{2}, \text{ and } i \text{ is even}, \\ \left(M - \omega_i^{(M)}\right) \frac{\varepsilon}{M-1+\alpha} + (1-\varepsilon)(1-p), & \text{if } j = \lfloor \frac{i}{2} \rfloor, \text{ and } i \text{ is odd}, \\ \omega_i^{(M)} \frac{\varepsilon}{M-1+\alpha} + (1-\varepsilon)p, & \text{if } j = \frac{i+2^M}{2}, \text{ and } i \text{ is even}, \\ \left(\omega_i^{(M)} - 1 + \alpha\right) \frac{\varepsilon}{M-1+\alpha} + (1-\varepsilon)p, & \text{if } j = \lfloor \frac{i+2^M}{2} \rfloor, \text{ and } i \text{ is odd}, \\ 0, & \text{otherwise}; \end{cases} \tag{3.2}$$

where $\omega_i^{(M)}$ is the number of "ones" in the M-bit binary representation of the decimal integer i. We next derive the analytic expression for $\pi_i^{(M)}$.

Lemma 3.1 *The stationary distribution $\boldsymbol{\pi}^{(M)}$ of the QBC Markov noise process is*

given by

$$\pi_i^{(M)} = \frac{\prod_{j=0}^{M-1-\omega_i^{(M)}} \left[j\frac{\varepsilon}{M-1+\alpha} + (1-\varepsilon)(1-p) \right] \prod_{j=0}^{\omega_i^{(M)}-1} \left[j\frac{\varepsilon}{M-1+\alpha} + (1-\varepsilon)p \right]}{\prod_{j=0}^{M-1} \left[1 - (\alpha+j)\frac{\varepsilon}{M-1+\alpha} \right]}, \quad (3.3)$$

for $i = 0, 1, 2, \cdots, 2^M - 1$, *where* $\prod_{j=0}^{a}(\cdot) \stackrel{\triangle}{=} 1$ *if* $a < 0$.

Proof If two states have the same binary representation except in the last bit, then the weights of the two states are different by 1 (suppose one is $\omega_i^{(M)}$, which is the number of "ones" in state i (the weight of state i), the other is $\omega_{i+1}^{(M)} = \omega_i^{(M)} + 1$ when i is even). Solving $\boldsymbol{\pi}^{(M)} = \boldsymbol{\pi}^{(M)} \boldsymbol{Q}_{\mathrm{QBC}}^{(M)}$ is equivalent to verifying that (3.3) satisfies the following:

$$\begin{aligned} \pi_{\frac{i}{2}}^{(M)} &= \pi_i^{(M)} \left[\frac{\left(M - \omega_i^{(M)} - 1 + \alpha \right) \varepsilon}{M - 1 + \alpha} + (1-\varepsilon)(1-p) \right] \\ &\quad + \pi_{i+1}^{(M)} \left[\frac{\left(M - \omega_i^{(M)} - 1 \right) \varepsilon}{M - 1 + \alpha} + (1-\varepsilon)(1-p) \right], \end{aligned} \quad (3.4)$$

and

$$\pi_{\frac{i+2^M}{2}}^{(M)} = \pi_i^{(M)} \left[\frac{\omega_i^{(M)} \varepsilon}{M - 1 + \alpha} + (1-\varepsilon)p \right] + \pi_{i+1}^{(M)} \left[\frac{\left(\omega_i^{(M)} + \alpha \right) \varepsilon}{M - 1 + \alpha} + (1-\varepsilon)p \right], \quad (3.5)$$

which are directly obtained from the transition probabilities (3.2). First we show that (3.4) is satisfied. State $\frac{i}{2}$ is reached from state i or state $i+1$ by transmitting a 0; thus $\pi_{\frac{i}{2}}^{(M)} = \pi_i^{(M)}$. Moreover,

$$\pi_{i+1}^{(M)} = \pi_i^{(M)} \frac{\omega_i^{(M)} \frac{\varepsilon}{M-1+\alpha} + (1-\varepsilon)p}{(M - \omega_i^{(M)} - 1)\frac{\varepsilon}{M-1+\alpha} + (1-\varepsilon)(1-p)}. \quad (3.6)$$

Therefore, (3.4) is equivalent to

$$1 = \frac{\left(M - \omega_i^{(M)} - 1 + \alpha\right)\varepsilon}{M - 1 + \alpha} + (1-\varepsilon)(1-p) + \omega_i^{(M)} \frac{\varepsilon}{M - 1 + \alpha} + (1-\varepsilon)p,$$

which is easily seen to hold. Next, we show that (3.5) is satisfied. State $\frac{i+2^M}{2}$ is reached from state i or state $i+1$ by transmitting a 1; thus $\pi^{(M)}_{\frac{i+2^M}{2}} = \pi^{(M)}_{i+1}$. Using (3.6), we see that (3.5) is equivalent to

$$1 = (M - \omega_i^{(M)} - 1)\frac{\varepsilon}{M - 1 + \alpha} + (1-\varepsilon)(1-p) + \frac{\left(\omega_i^{(M)} + \alpha\right)\varepsilon}{M - 1 + \alpha} + (1-\varepsilon)p,$$

which again is easily seen to hold. $\sum_{i=0}^{2^M-1} \pi_i^{(M)} = 1$ is satisfied by taking pairwise summation of $\pi_i^{(M)}$, $i = 0, 1, \cdots, 2^M - 1$, taking out the common factor, and cancelling another common factor in the numerator and denominator. Thus, it is verified that (3.3) is the stationary distribution of the QBC. □

Observation: We note that the stationary distribution (3.3) is identical to the stationary distribution of the channel state for the FMCC ([2], Sec.VI) if we set the parameters ρ and δ of the FMCC to $\rho = p$ and $\delta = \varepsilon/[(1-\varepsilon)(M-1+\alpha)]$. Thus, the set of possible stationary distributions for the QBC channel state is the same as that for the FMCC channel state. However, for a given M, bit error rate (BER) and correlation coefficient (Cor), the parameters of the FMCC are determined while we may still vary the parameters ε and α for the QBC to obtain different transition probabilities in (3.2) and hence different noise processes with the same stationary distribution but different entropies and ACFs.

3.1.2 Block Transition Probability

For a given input block $X^n = (X_1, \cdots, X_n)$ and a given output block $Y^n = (Y_1, \cdots, Y_n)$, where n is the blocklength, it can be shown using the Markovian property of the noise and state sources that the block transition probability of the resulting binary channel is

$$\Pr^{(M)}\{Y^n = y^n | X^n = x^n\} = \Pr^{(M)}\{Z^n = z^n\},$$

where $z_i = x_i \oplus y_i$, for $i = 1, 2, \cdots, n$, and the noise n-fold distribution is as follows.

- For blocklength $n \leq M$,

$$\begin{aligned}
&\Pr^{(M)}\{Z^n = z^n\} \\
&= \frac{\prod_{j=0}^{n-d_1^n-1}\left[j\frac{\varepsilon}{M-1+\alpha} + (1-\varepsilon)(1-p)\right]\prod_{j=0}^{d_1^n-1}\left[j\frac{\varepsilon}{M-1+\alpha} + (1-\varepsilon)p\right]}{\prod_{j=M-n}^{M-1}\left[1-(\alpha+j)\frac{\varepsilon}{M-1+\alpha}\right]},
\end{aligned} \tag{3.7}$$

  where $d_a^b = z_b + z_{b-1} + \cdots + z_a$ ($d_a^b = 0$ if $a > b$), and $\prod_{j=0}^{a}(\cdot) \stackrel{\triangle}{=} 1$ if $a < 0$.

- For blocklength $n \geq M + 1$,

$$\begin{aligned}
&\Pr^{(M)}\{Z^n = z^n\} = L^{(M)} \prod_{i=M+1}^{n}\left[\left(d_{i-M+1}^{i-1} + \alpha z_{i-M}\right)\frac{\varepsilon}{M-1+\alpha} + (1-\varepsilon)p\right]^{z_i} \\
&\left\{\left[\left(M-1-d_{i-M+1}^{i-1}\right) + \alpha(1-z_{i-M})\right]\frac{\varepsilon}{M-1+\alpha} + (1-\varepsilon)(1-p)\right\}^{1-z_i},
\end{aligned} \tag{3.8}$$

  where

$$L^{(M)} = \frac{\prod_{j=0}^{M-1-d_1^M}\left[j\frac{\varepsilon}{M-1+\alpha} + (1-\varepsilon)(1-p)\right]\prod_{j=0}^{d_1^M-1}\left[j\frac{\varepsilon}{M-1+\alpha} + (1-\varepsilon)p\right]}{\prod_{j=0}^{M-1}\left[1-(\alpha+j)\frac{\varepsilon}{M-1+\alpha}\right]}.$$

The QBC noise process is stationary; hence it is identically distributed. The channel BER and Cor are next determined.

$$\text{BER} = \Pr\{Z_i = 1\} = \Pr\{Z_1 = 1\} = p, \tag{3.9}$$

and

$$\text{Cor} = \frac{Cov(Z_i, Z_{i+1})}{\sqrt{Var(Z_i)Var(Z_{i+1})}} = \frac{Cov(Z_2, Z_1)}{Var(Z_1)}, \tag{3.10}$$

where $Cov(Z_2, Z_1) = E[Z_2Z_1] - E[Z_2]E[Z_1] = \Pr\{Z_2 = 1, Z_1 = 1\} - p^2$ and $Var(Z_1) = E[Z_1^2] - E[Z_1]^2 = \Pr\{Z_1 = 1\} - p^2 = p - p^2$. To obtain $Cov(Z_2, Z_1)$, we use (3.8) if $M = 1$ (with $n = 2$ and $\alpha = 1$):

$$\Pr\{Z_2 = 1, Z_1 = 1\} = \frac{(1-\varepsilon)p}{1-\varepsilon}\left[\varepsilon + (1-\varepsilon)p\right] = p\left[\varepsilon + (1-\varepsilon)p\right].$$

Thus, for $M = 1$, we get

$$Cov(Z_2, Z_1) = p\left[\varepsilon + (1-\varepsilon)p\right] - p^2 = p(1-p)\varepsilon.$$

When $M \geq 2$, we use (3.7) (with $n = 2$) to obtain

$$\Pr\{Z_2 = 1, Z_1 = 1\} = \frac{p\left[\frac{\varepsilon}{M-1+\alpha} + (1-\varepsilon)p\right]}{1 - (M-2+\alpha)\frac{\varepsilon}{M-1+\alpha}}.$$

Thus, for $M \geq 2$, we obtain

$$\begin{aligned} Cov(Z_2, Z_1) &= p\frac{\frac{\varepsilon}{M-1+\alpha} + (1-\varepsilon)p}{1 - (M-2+\alpha)\frac{\varepsilon}{M-1+\alpha}} - p^2 \\ &= p(1-p)\frac{\frac{\varepsilon}{M-1+\alpha}}{1 - (M-2+\alpha)\frac{\varepsilon}{M-1+\alpha}}. \end{aligned} \tag{3.11}$$

Note that, since $\alpha = 1$ when $M = 1$, (3.11) also holds for $M = 1$. Thus, for $M \geq 1$ we have

$$\text{Cor} = \frac{Cov(Z_2, Z_1)}{Var(Z_1)} = \frac{\frac{\varepsilon}{M-1+\alpha}}{1-(M-2+\alpha)\frac{\varepsilon}{M-1+\alpha}}. \tag{3.12}$$

Remark: For fixed $M \geq 1$ and $\alpha \geq 0$, $\text{Cor} = 0$ if and only if $\varepsilon = 0$. Thus, the QBC noise process is iid if and only if $\text{Cor} = 0$. This property does not hold in general for stationary binary sources ($\text{Cor} = 0$ only implies pairwise independence of consecutive bits). Finally, note that $\text{Cor} \geq 0$.

From (3.9) and (3.12), the parameters p and ε can be expressed in terms of BER, Cor, M and α as follows:

$$p = \text{BER}, \tag{3.13}$$

and

$$\varepsilon = \frac{(M-1+\alpha)\text{Cor}}{1+(M-2+\alpha)\text{Cor}}. \tag{3.14}$$

Similarly, since $\frac{\varepsilon}{1-\varepsilon} = \frac{(M-1+\alpha)\text{Cor}}{1-\text{Cor}}$, (3.7) and (3.8) can be written in terms of BER, Cor, M and α as follows:

- For blocklength $n \leq M$,

$$\Pr^{(M)}\{Z^n = z^n\} = \frac{\prod_{j=0}^{n-d_1^n-1}[j\frac{\text{Cor}}{1-\text{Cor}} + (1-\text{BER})]\prod_{j=0}^{d_1^n-1}[j\frac{\text{Cor}}{1-\text{Cor}} + \text{BER}]}{\prod_{j=0}^{n-1}[1+j\frac{\text{Cor}}{1-\text{Cor}}]}. \tag{3.15}$$

- For blocklength $n \geq M+1$,

$$\Pr^{(M)}\{Z^n = z^n\} = L^{(M)} \prod_{i=M+1}^{n} \left\{ \frac{\left(d_{i-M+1}^{i-1} + \alpha z_{i-M}\right) \frac{\mathrm{Cor}}{1-\mathrm{Cor}} + \mathrm{BER}}{1 + (M-1+\alpha)\frac{\mathrm{Cor}}{1-\mathrm{Cor}}} \right\}^{z_i} \left\{ \frac{\left[M-1-d_{i-M+1}^{i-1} + \alpha(1-z_{i-M})\right] \frac{\mathrm{Cor}}{1-\mathrm{Cor}} + (1-\mathrm{BER})}{1 + (M-1+\alpha)\frac{\mathrm{Cor}}{1-\mathrm{Cor}}} \right\}^{1-z_i}, \quad (3.16)$$

  where

$$L^{(M)} = \frac{\prod_{j=0}^{M-1-d_1^M} \left[j\frac{\mathrm{Cor}}{1-\mathrm{Cor}} + (1-\mathrm{BER})\right] \prod_{j=0}^{d_1^M-1} \left(j\frac{\mathrm{Cor}}{1-\mathrm{Cor}} + \mathrm{BER}\right)}{\prod_{j=0}^{M-1}\left(1 + j\frac{\mathrm{Cor}}{1-\mathrm{Cor}}\right)}. \quad (3.17)$$

The stationary distribution (3.3) can be expressed in terms of M, BER and Cor as follows.

$$\pi_i^{(M)} = \frac{\prod_{j=0}^{M-\omega_i^{(M)}-1} \left[j\frac{\mathrm{Cor}}{1-\mathrm{Cor}} + (1-\mathrm{BER})\right] \prod_{j=0}^{\omega_i^{(M)}-1} \left(j\frac{\mathrm{Cor}}{1-\mathrm{Cor}} + \mathrm{BER}\right)}{\prod_{j=0}^{M-1}\left(1 + j\frac{\mathrm{Cor}}{1-\mathrm{Cor}}\right)}, \quad (3.18)$$

for $i = 0, 1, 2, \cdots, 2^M - 1$.

In [29], the authors study the joint source-channel coding problem of designing an index assignment based on the Hadamard transform for the robust transmission of vector quantization indices over the FMCC. It is noted that the block distribution of the FMCC noise obeys well structured recursions, hence simplifying the Hadamard transform analysis and the evaluation of the channel distortion. As a result an efficient index assignment method with appealing robustness properties is constructed. We herein present a simple recursion property on the stationary distribution of the QBC.

Lemma 3.2 *For fixed BER and Cor, the stationary distribution* $\pi_i^{(M)}$ *obeys the following recursion:*

$$\pi_i^{(M)} = \pi_{2i}^{(M+1)} + \pi_{2i+1}^{(M+1)}, \quad \textit{for } i = 0, 1, 2, \cdots, 2^M - 1. \tag{3.19}$$

Proof We notice that

$$\omega_s^{(M+1)} = \begin{cases} \omega_i^{(M)}, & \text{if } s = 2i, \\ \omega_i^{(M)} + 1, & \text{if } s = 2i+1, \end{cases}$$

for $s = 0, 1, \cdots, 2^{M+1} - 1$, where $\omega_s^{(M+1)}$ is the number of "ones" in the $(M+1)$-bit binary representation of the decimal integer s. Thus, from (3.18),

$$\pi_{2i}^{(M+1)} = \frac{\prod_{j=0}^{M-\omega_i^{(M)}} \left[j\frac{\text{Cor}}{1-\text{Cor}} + (1-\text{BER}) \right] \prod_{j=0}^{\omega_i^{(M)}-1} \left(j\frac{\text{Cor}}{1-\text{Cor}} + \text{BER} \right)}{\prod_{j=1}^{M} \left(1 + j\frac{\text{Cor}}{1-\text{Cor}} \right)}, \tag{3.20}$$

and

$$\pi_{2i+1}^{(M+1)} = \frac{\prod_{j=0}^{M-\omega_i^{(M)}-1} \left[j\frac{\text{Cor}}{1-\text{Cor}} + (1-\text{BER}) \right] \prod_{j=0}^{\omega_i^{(M)}} \left(j\frac{\text{Cor}}{1-\text{Cor}} + \text{BER} \right)}{\prod_{j=1}^{M} \left(1 + j\frac{\text{Cor}}{1-\text{Cor}} \right)}. \tag{3.21}$$

Summing (3.20) and (3.21) we get

$$\begin{aligned} & \pi_{2i}^{(M+1)} + \pi_{2i+1}^{(M+1)} \\ = \; & \frac{\prod_{j=0}^{M-\omega_i^{(M)}-1} \left[j\frac{\text{Cor}}{1-\text{Cor}} + (1-\text{BER}) \right] \prod_{j=0}^{\omega_i^{(M)}-1} \left(j\frac{\text{Cor}}{1-\text{Cor}} + \text{BER} \right)}{\prod_{j=0}^{M} \left(1 + j\frac{\text{Cor}}{1-\text{Cor}} \right)} \\ & \times \left[\left(M - \omega_i^{(M)} \right) \frac{\text{Cor}}{1-\text{Cor}} + (1-\text{BER}) + \omega_i^{(M)} \frac{\text{Cor}}{1-\text{Cor}} + \text{BER} \right] \\ = \; & \frac{\prod_{j=0}^{M-\omega_i^{(M)}-1} \left[j\frac{\text{Cor}}{1-\text{Cor}} + (1-\text{BER}) \right] \prod_{j=0}^{\omega_i^{(M)}-1} \left(j\frac{\text{Cor}}{1-\text{Cor}} + \text{BER} \right)}{\prod_{j=0}^{M-1} \left(1 + j\frac{\text{Cor}}{1-\text{Cor}} \right)}. \end{aligned} \tag{3.22}$$

By comparing (3.22) and (3.18), we get

$$\pi_i^{(M)} = \pi_{2i}^{(M+1)} + \pi_{2i+1}^{(M+1)}. \tag{3.23}$$

□

3.1.3 Autocorrelation Function

The ACF of a binary stationary process $\{Z_n\}_{n=1}^{\infty}$ is defined by

$$R[m] = E[Z_i Z_{i+m}] = \Pr\{Z_i = 1, Z_{i+m} = 1\}.$$

Lemma 3.3 *The ACF of the QBC satisfies the following.*

$$R[m] = \begin{cases} p & \text{if } m = 0; \\ \frac{\frac{\varepsilon}{M-1+\alpha} + (1-\varepsilon)p}{1 - \frac{M-2+\alpha}{M-1+\alpha}\varepsilon} p & \text{if } 1 \le m \le M-1; \\ (1-\varepsilon)p^2 + \frac{\varepsilon}{M-1+\alpha}\left(\sum_{i=m-M+1}^{m-1} R[i] + \alpha R[m-M]\right) & \text{if } m \ge M. \end{cases}$$

Proof The autocorrelation function (ACF) of a binary stationary process $\{Z_n\}$ is given by:

$$\begin{aligned} R[m] &= E[Z_i Z_{i+m}] = \Pr\{Z_i = 1, Z_{i+m} = 1\} \\ &= \sum_{z_{i+1}=0}^{1} \cdots \sum_{z_{i+m-1}=0}^{1} \Pr\{Z_i = 1, Z_{i+1} = z_{i+1}, \cdots, Z_{i+m-1} = z_{i+m-1}, Z_{i+m} = 1\}, \end{aligned}$$

Using (3.7) and (3.8), the ACF of the QBC is expressed as follows.

- If $m = 0$, we have

$$R[0] = \Pr\{Z_i = 1\} = p.$$

- If $m \leq M-1$, we have

$$\begin{aligned} R[m] &= \Pr\{Z_i = 1, Z_{i+1} = 0, \cdots, Z_{i+m-2} = 0, Z_{i+m-1} = 0, Z_{i+m} = 1\} \\ &+ \Pr\{Z_i = 1, Z_{i+1} = 0, \cdots, Z_{i+m-2} = 0, Z_{i+m-1} = 1, Z_{i+m} = 1\} \\ &+ \cdots\cdots \\ &+ \Pr\{Z_i = 1, Z_{i+1} = 1, \cdots, Z_{i+m-2} = 1, Z_{i+m-1} = 1, Z_{i+m} = 1\}. \end{aligned}$$

We notice that, if two binary error sequences of length $m+1$ are different in only one digit, then the weights of the two error sequences are different by 1 (suppose one is ω, which is the number of "ones", the other is $\omega+1$). Thus the formulas of the block transition probabilities of the two error sequences are different by only one term; i.e., one includes $[\varepsilon\frac{m-\omega}{M-1+\alpha} + (1-\varepsilon)(1-p)]$ and the other includes $[\varepsilon\frac{\omega}{M-1+\alpha} + (1-\varepsilon)p]$. We can calculate the sum of these two terms to get $(1-\varepsilon\frac{M-m-1+\alpha}{M-1+\alpha})$, which can be cancelled by $(1-\varepsilon\frac{M-m-1+\alpha}{M-1+\alpha})$ in the denominator. Then the sum of the two stationary probabilities is

$$\frac{\prod_{j=0}^{m-\omega-1}\left[j\frac{\varepsilon}{M-1+\alpha} + (1-\varepsilon)(1-p)\right]\prod_{j=0}^{\omega-1}\left[j\frac{\varepsilon}{M-1+\alpha} + (1-\varepsilon)p\right]}{\prod_{j=M-m}^{M-1}\left[1-(\alpha+j)\frac{\varepsilon}{M-1+\alpha}\right]}.$$

Taking pairwise summations of 2^{m-1} terms, the total number of terms decreases to 2^{m-2}. By repeating these operations of 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:

$$\frac{\frac{\varepsilon}{M-1+\alpha} + (1-\varepsilon)p}{1-\frac{M-2+\alpha}{M-1+\alpha}\varepsilon}p.$$

- If $m \geq M$, we have

$$
\begin{aligned}
R[m] &= \sum_{z_{i+1}=0}^{1} \cdots \sum_{z_{i+m-1}=0}^{1} \Pr\{Z_i^{i+m-M-1} = z_i^{i+m-M-1}\} \\
&\quad \times \prod_{j=i+m-M}^{i+m-1} \Pr\{Z_j = z_j | Z_{j-1} = z_{j-1}, \cdots, Z_{j-M} = z_{j-M}\} \\
&\quad \times \Pr\{Z_{i+m} = 1 | Z_{i+m-1} = z_{i+m-1}, \cdots, Z_{i+m-M} = z_{i+m-M}\} \\
&= \sum_{z_{i+1}=0}^{1} \cdots \sum_{z_{i+m-2}=0}^{1} \Pr\{Z_i^{i+m-M-1} = z_i^{i+m-M-1}\} \\
&\quad \times \prod_{j=i+m-M}^{i+m-2} \Pr\{Z_j = z_j | Z_{j-1} = z_{j-1}, \cdots, Z_{j-M} = z_{j-M}\} \\
&\quad \times \Big[\Pr\{Z_{i+m} = 1 | Z_{i+m-1} = 1, \cdots, Z_{i+m-M} = z_{i+m-M}\} \\
&\qquad \times \Pr\{Z_{i+m-1} = 1 | Z_{i+m-2} = z_{i+m-2}, \cdots, Z_{i+m-M-1} = z_{i+m-M-1}\} \\
&\qquad + \Pr\{Z_{i+m} = 1 | Z_{i+m-1} = 0, \cdots, Z_{i+m-M} = z_{i+m-M}\} \\
&\qquad \times \Pr\{Z_{i+m-1} = 0 | Z_{i+m-2} = z_{i+m-2}, \cdots, Z_{i+m-M-1} = z_{i+m-M-1}\} \Big],
\end{aligned}
$$

where $Z_i^{i+m-M-1}$ denotes $(Z_i = 1, \cdots, Z_{i+m-M-1})$. Notice that the weight of $(Z_{i+m-1} = 1, \cdots, Z_{i+m-M})$ and $(Z_{i+m-1} = 0, \cdots, Z_{i+m-M})$ are different by 1, and

$$
\begin{aligned}
&\Pr\{Z_{i+m-1} = 0 | Z_{i+m-2} = z_{i+m-2}, \cdots, Z_{i+m-M-1} = z_{i+m-M-1}\} \\
= \;& 1 - \Pr\{Z_{i+m-1} = 1 | Z_{i+m-2} = z_{i+m-2}, \cdots, Z_{i+m-M-1} = z_{i+m-M-1}\}.
\end{aligned}
$$

We obtain

$$
R[m] = \sum_{z_{i+1}=0}^{1} \cdots \sum_{z_{i+m-2}=0}^{1} \Pr\{Z_i^{i+m-M-1} = z_i^{i+m-M-1}\}
$$

$$
\begin{aligned}
&\times \prod_{j=i+m-M}^{i+m-2} \Pr\{Z_j = z_j | Z_{j-1} = z_{j-1}, \cdots, Z_{j-M} = z_{j-M}\} \\
&\times \Big[\Pr\{Z_{i+m-1} = 1 | Z_{i+m-2} = z_{i+m-2}, \cdots, Z_{i+m-M-1} = z_{i+m-M-1}\} \\
&\quad \times \frac{\varepsilon}{M-1+\alpha} + \Pr\{Z_{i+m} = 1 | Z_{i+m-1} = 0, \cdots, Z_{i+m-M}\} \Big] \\
= &\sum_{z_{i+1}=0}^{1} \cdots \sum_{z_{i+m-2}=0}^{1} \Pr\{Z_i^{i+m-M-1} = z_i^{i+m-M-1}\} \\
&\times \prod_{j=i+m-M}^{i+m-2} \Pr(Z_j = z_j | Z_{j-1} = z_{j-1}, \cdots, Z_{j-M} = z_{j-M}\} \\
&\times \Pr\{Z_{i+m} = 1 | Z_{i+m-1} = 0, \cdots, Z_{i+m-M} = z_{i+m-M}\} \\
&+ \frac{\varepsilon}{M-1+\alpha} \sum_{z_{i+1}=0}^{1} \cdots \sum_{z_{i+m-2}=0}^{1} \Pr\{Z_i^{i+m-M-1} = z_i^{i+m-M-1}\} \\
&\times \prod_{j=i+m-M}^{i+m-2} \Pr\{Z_j = z_j | Z_{j-1} = z_{j-1}, \cdots, Z_{j-M} = z_{j-M}\} \\
&\times \Pr(Z_{i+m-1} = 1 | Z_{i+m-2} = z_{i+m-2}, \cdots, Z_{i+m-M-1} = z_{i+m-M-1}\}.
\end{aligned}
$$

The second term can be expressed by $\frac{\varepsilon}{M-1+\alpha} R[m-1]$. By repeating the operations of taking pairwise summation in the first term, we get

$$
\begin{aligned}
R[m] \;=\; &\sum_{z_{i+1}=0}^{1} \cdots \sum_{z_{i+m-M-1}=0}^{1} \Pr\{Z_i^{i+m-M-1} = z_i^{i+m-M-1}\} \\
&\times \Pr\{Z_{i+m} = 1 | Z_{i+m-1} = 0, \cdots, Z_{i+m-M} = 0\} \\
&+ \frac{\varepsilon}{M-1+\alpha} \left(\sum_{i=m-M+1}^{m-1} R[i] + \alpha R[m-M] \right).
\end{aligned}
$$

The final recursion expression for ACF is shown as follows.

$$
R[m] = (1-\varepsilon)p^2 + \frac{\varepsilon}{M-1+\alpha} \left(\sum_{i=m-M+1}^{m-1} R[i] + \alpha R[m-M] \right).
$$

□

Lemma 3.4 *If $Cor > 0$, then $R[m]$ is strictly decreasing for $m > M$ and*

$$\lim_{m\to\infty} R[m] = p^2;$$

thus, Z_i and Z_{i+m} are asymptotically independent (since they are binary valued).

Proof The proof is by induction. First, we prove that the initial condition $R[M] > R[M+1]$ holds; then we prove that $R[j] > R[j+1]$ holds if $R[m]$ is strictly decreasing for $M < m < j$.

Initial condition:

$$\begin{aligned} R[M+1] - R[M] &= (1-\varepsilon)p^2 + \frac{\varepsilon}{M-1+\alpha}\left(\sum_{i=2}^{M} R[i] + \alpha R[1]\right) \\ &\quad -(1-\varepsilon)p^2 - \frac{\varepsilon}{M-1+\alpha}\left(\sum_{i=1}^{M-1} R[i] + \alpha R[0]\right) \\ &= \frac{\varepsilon}{M-1+\alpha}\left(R[M] + \alpha R[1] - R[1] - \alpha R[0]\right). \end{aligned}$$

Since $R[i] = R[1]$ for $1 \le i \le M-1$, we obtain

$$\begin{aligned} & R[M+1] - R[M] \\ = \; & \frac{\varepsilon}{M-1+\alpha}\Big\{(1-\varepsilon)p^2 + \frac{\varepsilon}{M-1+\alpha}\Big((M-1)R[1] + \alpha R[0]\Big) \\ & +\alpha R[1] - R[1] - \alpha R[0]\Big\} \\ = \; & \frac{\varepsilon}{M-1+\alpha}\Big\{(1-\varepsilon)p^2 + \frac{\varepsilon}{M-1+\alpha}\Big((M-1+\alpha)R[1] + \alpha(R[0] - R[1])\Big) \\ & +\alpha R[1] - R[1] - \alpha R[0]\Big\} \\ = \; & \frac{\varepsilon}{M-1+\alpha}\Big\{(1-\varepsilon)p^2 + \varepsilon R[1] + \frac{\varepsilon\alpha(R[0] - R[1])}{M-1+\alpha} + \alpha R[1] - R[1] - \alpha R[0]\Big\} \end{aligned}$$

$$
\begin{aligned}
&= \frac{\varepsilon}{M-1+\alpha}\Big\{(1-\varepsilon)p^2-(1-\varepsilon)R[1]-\Big(1-\frac{\varepsilon}{M-1+\alpha}\Big)\alpha R[0] \\
&\quad +\Big(1-\frac{\varepsilon}{M-1+\alpha}\Big)\alpha R[1]\Big\} \\
&= \frac{\varepsilon}{M-1+\alpha}\Big\{(1-\varepsilon)(p^2-R[1])+\alpha\Big(1-\frac{\varepsilon}{M-1+\alpha}\Big)(R[1]-R[0])\Big\}.
\end{aligned}
$$

Since $p^2 < R[1]$ (since Cor > 0) and $R[1] < R[0]$ for $0 < p \leq 1$, we have that $R[M+1] < R[M]$.

Inductive Step: Assume $R[m]$ is strictly decreasing for $M < m < j$, then

$$
\begin{aligned}
& R[j+1]-R[j] \\
&= (1-\varepsilon)p^2+\frac{\varepsilon}{M-1+\alpha}\left(\sum_{i=j-M+2}^{j}R[i]+\alpha R[j-M+1]\right) \\
&\quad -(1-\varepsilon)p^2-\frac{\varepsilon}{M-1+\alpha}\left(\sum_{i=j-M+1}^{j-1}R[i]+\alpha R[j-M]\right) \\
&= \frac{\varepsilon}{M-1+\alpha}\Big(R[j]+\alpha R[j-M+1]-R[j-M+1]-\alpha R[j-M]\Big).
\end{aligned}
$$

We have $R[j+1] < R[j]$ based on the assumption of $R[m]$ is strictly decreasing for $M < m < j$.

Therefore $R[m]$ is strictly decreasing for $m > M$ and converges to a positive value:

$$
\begin{aligned}
\lim_{m\to\infty}R[m] &= (1-\varepsilon)p^2+\frac{\varepsilon}{M-1+\alpha}\Big\{(M-1)\lim_{m\to\infty}R[m]+\alpha\lim_{m\to\infty}R[m-M]\Big\} \\
&= (1-\varepsilon)p^2+\frac{\varepsilon}{M-1+\alpha}\Big\{(M-1+\alpha)\lim_{m\to\infty}R[m]\Big\}.
\end{aligned}
$$

Therefore

$$\lim_{m\to\infty}R[m]=p^2.$$

□

3.1.4 Channel Capacity

The information capacity for channels with memory is defined by Shannon's familiar expression [26, pp. 287], [64]

$$C = \lim_{n\to\infty} \sup_{X^n} \frac{1}{n} I(X^n; Y^n), \tag{3.24}$$

where $I(X^n; Y^n)$ denotes the block mutual information [15, pp. 18] between input $X^n = (X_1, \cdots, X_n)$ and output $Y^n = (Y_1, \cdots, Y_n)$ and where the supremum is taken over all possible inputs X^n. As noted in Chapter 2, for the wide class of channels with memory that are information stable [64] (e.g., the input process that maximizes the block mutual information and its corresponding output process form a stationary ergodic joint input-output process) the information capacity shown above has an important operational meaning as established by Shannon [50], since it represents the largest rate at which information can be transmitted over the channel via a channel code and recovered at the receiver with asymptotically vanishing probability of error (as the code blocklength approaches infinity). It is thus clear that C is a key quantity in the investigation of communication channels.

Since the QBC is a channel with additive stationary ergodic noise, it is information stable, and its (operational) capacity, $C_{\text{QBC}}^{(M)}$, is hence given by (3.24). Due to the channel's symmetry, it can be shown that input n-tuples X^n that are uniformly

distributed over $\{0,1\}^n$ maximize $I(X^n;Y^n)$ in (3.24). Thus,

$$C_{\mathrm{QBC}}^{(M)} \;=\; 1-\mathcal{H}^{(M)}(Z), \tag{3.25}$$

where $\mathcal{H}^{(M)}(Z)$ is the entropy rate of the Mth order Markov noise process. It can be obtained using (3.2) and (3.3), as follows.

$$\begin{aligned}
\mathcal{H}^{(M)}(Z) &\stackrel{\triangle}{=} \lim_{n\to\infty}\frac{1}{n}H^{(M)}(Z_1,\cdots,Z_n)\\
&= H^{(M)}(Z_{M+1}\mid Z_M,Z_{M-1},\cdots,Z_1)\\
&= H^{(M)}(\underline{S}_2\mid\underline{S}_1)\\
&= -\sum_{i,j=0}^{2^M-1}\pi_i^{(M)}p_{ij}^{(M)}\log_2 p_{ij}^{(M)}\\
&= \sum_{\omega=0}^{M-1}\binom{M-1}{\omega}L_\omega^{(M)}h_b\left[\omega\frac{\varepsilon}{M-1+\alpha}+(1-\varepsilon)p\right]\\
&+ \sum_{\omega=1}^{M}\binom{M-1}{\omega-1}L_\omega^{(M)}h_b\left[(\omega-1+\alpha)\frac{\varepsilon}{M-1+\alpha}+(1-\varepsilon)p\right],
\end{aligned}\tag{3.26}$$

where

$$L_\omega^{(M)}=\frac{\prod_{j=0}^{M-1-\omega}\left[j\frac{\varepsilon}{M-1+\alpha}+(1-\varepsilon)(1-p)\right]\prod_{j=0}^{\omega-1}\left[j\frac{\varepsilon}{M-1+\alpha}+(1-\varepsilon)p\right]}{\prod_{j=0}^{M-1}\left[1-(\alpha+j)\frac{\varepsilon}{M-1+\alpha}\right]},$$

$\prod_{j=0}^{a}(\cdot)\stackrel{\triangle}{=}1$, if $a<0$, $\binom{a}{b}\stackrel{\triangle}{=}1$, if $a=0$, and $h_b(\cdot)$ is the binary entropy function: $h_b(g)=-g\log_2 g-(1-g)\log_2(1-g)$. It is clear that $C_{\mathrm{QBC}}^{(M)}$ is positive since $\mathcal{H}^{(M)}(Z)<1$ for fixed M, ε, p and α. Using (3.25) and (3.26), we obtain the following expression for $C_{\mathrm{QBC}}^{(M)}$.

$$\begin{aligned}
C_{\mathrm{QBC}}^{(M)} &= 1-\sum_{\omega=0}^{M-1}\binom{M-1}{\omega}L_\omega^{(M)}h_b\left[\omega\frac{\varepsilon}{M-1+\alpha}+(1-\varepsilon)p\right]\\
&\quad-\sum_{\omega=1}^{M}\binom{M-1}{\omega-1}L_\omega^{(M)}h_b\left[(\omega+\alpha-1)\frac{\varepsilon}{M-1+\alpha}+(1-\varepsilon)p\right].
\end{aligned}\tag{3.27}$$

In terms of the channel parameters M, BER, Cor and α, the capacity in (3.27) can be written as

$$\begin{aligned} C_{\mathrm{QBC}}^{(M)} &= 1 - \sum_{\omega=0}^{M-1} \binom{M-1}{\omega} L_{\omega}^{(M)} h_b \left[\frac{\omega \frac{\mathrm{Cor}}{1-\mathrm{Cor}} + \mathrm{BER}}{1 + (M-1+\alpha)\frac{\mathrm{Cor}}{1-\mathrm{Cor}}} \right] \\ &\quad - \sum_{\omega=1}^{M} \binom{M-1}{\omega-1} L_{\omega}^{(M)} h_b \left[\frac{(\omega-1+\alpha)\frac{\mathrm{Cor}}{1-\mathrm{Cor}} + \mathrm{BER}}{1 + (M-1+\alpha)\frac{\mathrm{Cor}}{1-\mathrm{Cor}}} \right], \end{aligned} \tag{3.28}$$

where

$$L_{\omega}^{(M)} = \frac{\prod_{j=0}^{M-1-\omega} \left[j\frac{\mathrm{Cor}}{1-\mathrm{Cor}} + (1-\mathrm{BER}) \right] \prod_{j=0}^{\omega-1} \left(j\frac{\mathrm{Cor}}{1-\mathrm{Cor}} + \mathrm{BER} \right)}{\prod_{j=0}^{M-1} \left(1 + j\frac{\mathrm{Cor}}{1-\mathrm{Cor}} \right)}, \tag{3.29}$$

which is not a function of α .

When $\varepsilon = 0$ (or $\mathrm{Cor} = 0$), the channel is a BSC and the channel capacity is

$$C_{\mathrm{QBC}}^{(M)} = 1 - h_b(p) = 1 - h_b(\mathrm{BER}). \tag{3.30}$$

Theorem 3.1 *The capacity $C_{\mathrm{QBC}}^{(M)}$ of the QBC strictly increases with α for fixed $M \geq 2$, BER and Cor $\in (0,1)$.*

Proof Notice that for various α, ε has to change to keep Cor fixed from (3.12). Rewriting (3.28), we get

$$C_{\mathrm{QBC}}^{(M)} = 1 - \sum_{\omega=0}^{M-1} \binom{M-1}{\omega} \left\{ L_{\omega}^{(M)} h_b \left[f(\alpha) \right] + L_{\omega+1}^{(M)} h_b \left[g(\alpha) \right] \right\}, \tag{3.31}$$

where $L_{\omega}^{(M)}$ is expressed in (3.29), $h_b(\cdot)$ is the binary entropy function,

$$f(\alpha) = \frac{\omega \frac{\mathrm{Cor}}{1-\mathrm{Cor}} + \mathrm{BER}}{1 + (M-1+\alpha)\frac{\mathrm{Cor}}{1-\mathrm{Cor}}},$$

and

$$g(\alpha) = \frac{(\omega+\alpha)\frac{\mathrm{Cor}}{1-\mathrm{Cor}} + \mathrm{BER}}{1+(M-1+\alpha)\frac{\mathrm{Cor}}{1-\mathrm{Cor}}}.$$

Differentiating $C_{\mathrm{QBC}}^{(M)}$ over α for fixed M, BER and Cor yields

$$\frac{d\,C_{\mathrm{QBC}}^{(M)}}{d\,\alpha} = -\sum_{\omega=0}^{M-1}\binom{M-1}{\omega}\left[L_{\omega}^{(M)}\frac{d\,h_b\,(f)}{d\,f}\frac{d\,f}{d\,\alpha} + L_{\omega+1}^{(M)}\frac{d\,h_b\,(g)}{d\,g}\frac{d\,g}{d\,\alpha}\right]. \tag{3.32}$$

Since

$$\frac{d\,f}{d\,\alpha} = -\frac{\left(\omega\frac{\mathrm{Cor}}{1-\mathrm{Cor}} + \mathrm{BER}\right)\frac{\mathrm{Cor}}{1-\mathrm{Cor}}}{\left[1+(M-1+\alpha)\frac{\mathrm{Cor}}{1-\mathrm{Cor}}\right]^2},$$

$$\frac{d\,g}{d\,\alpha} = \frac{\left[(M-\omega-1)\,\frac{\mathrm{Cor}}{1-\mathrm{Cor}} + (1-\mathrm{BER})\right]\frac{\mathrm{Cor}}{1-\mathrm{Cor}}}{\left[1+(M-1+\alpha)\frac{\mathrm{Cor}}{1-\mathrm{Cor}}\right]^2},$$

$\frac{d\,h_b(f)}{d\,f} = \log_2\left(\frac{1-f}{f}\right)$, and $\frac{d\,h_b(g)}{d\,f} = \log_2\left(\frac{1-g}{g}\right)$, we obtain that

$$\begin{aligned}\frac{d\,C_{\mathrm{QBC}}^{(M)}}{d\,\alpha} \;=\; & -V^{(M)}\sum_{\omega=0}^{M-1}\left\{\binom{M-1}{\omega}L_{\omega}^{(M)}\left[-\left(\omega\frac{\mathrm{Cor}}{1-\mathrm{Cor}}+\mathrm{BER}\right)\right]\right. \\ & \times\log_2\left[\frac{(M-1+\alpha-\omega)\,\frac{\mathrm{Cor}}{1-\mathrm{Cor}}+(1-\mathrm{BER})}{\omega\frac{\mathrm{Cor}}{1-\mathrm{Cor}}+\mathrm{BER}}\right] \\ & +L_{\omega+1}^{(M)}\left[(M-\omega-1)\,\frac{\mathrm{Cor}}{1-\mathrm{Cor}}+(1-\mathrm{BER})\right] \\ & \left.\times\log_2\left[\frac{(M-\omega-1)\,\frac{\mathrm{Cor}}{1-\mathrm{Cor}}+(1-\mathrm{BER})}{(\omega+\alpha)\,\frac{\mathrm{Cor}}{1-\mathrm{Cor}}+\mathrm{BER}}\right]\right\}\end{aligned} \tag{3.33}$$

where

$$V^{(M)} = \frac{\frac{\mathrm{Cor}}{1-\mathrm{Cor}}}{\left[1+(M-1+\alpha)\frac{\mathrm{Cor}}{1-\mathrm{Cor}}\right]^2}.$$

Notice, from (3.29), that

$$\begin{aligned} W_\omega &\triangleq L_\omega^{(M)} \times \left(\omega \frac{\text{Cor}}{1-\text{Cor}} + \text{BER}\right) \\ &= L_{\omega+1}^{(M)} \times \left[(M-\omega-1)\,\frac{\text{Cor}}{1-\text{Cor}} + (1-\text{BER})\right]. \end{aligned} \tag{3.34}$$

Hence, (3.33) can be rewritten as

$$\begin{aligned} \frac{d\,C_{\text{QBC}}^{(M)}}{d\,\alpha} = -V^{(M)} \sum_{\omega=0}^{M-1} \Bigg\{ \binom{M-1}{\omega} W_\omega \\ \log_2 \left[\frac{\omega \frac{\text{Cor}}{1-\text{Cor}} + \text{BER}}{(\omega+\alpha)\,\frac{\text{Cor}}{1-\text{Cor}} + \text{BER}} \times \frac{(M-\omega-1)\,\frac{\text{Cor}}{1-\text{Cor}} + (1-\text{BER})}{(M-1+\alpha-\omega)\,\frac{\text{Cor}}{1-\text{Cor}} + (1-\text{BER})} \right] \Bigg\}. \end{aligned} \tag{3.35}$$

Since $\omega < (\omega+\alpha)$ and $(M-\omega-1) < (M-1+\alpha-\omega)\ \forall\ \alpha > 0$, we get

$$\log_2 \left[\frac{\omega \frac{\text{Cor}}{1-\text{Cor}} + \text{BER}}{(\omega+\alpha)\,\frac{\text{Cor}}{1-\text{Cor}} + \text{BER}} \times \frac{(M-\omega-1)\,\frac{\text{Cor}}{1-\text{Cor}} + (1-\text{BER})}{(M-1+\alpha-\omega)\,\frac{\text{Cor}}{1-\text{Cor}} + (1-\text{BER})} \right] < 0.$$

Since $V^{(M)} > 0$ and $W_\omega > 0$ for $\omega \in \{0, 1, \cdots, M-1\}$, we obtain that $\frac{d\,C_{\text{QBC}}^{(M)}}{d\,\alpha} > 0$ for $\alpha > 0$; i.e., the capacity of the QBC strictly increases with α for fixed $M \geq 2$, BER and Cor $\in (0,1)$. □

Observation: When $\alpha \to \infty$ we have from (3.28) that

$$h_b \left[\frac{\omega \frac{\text{Cor}}{1-\text{Cor}} + \text{BER}}{1 + (M-1+\alpha)\frac{\text{Cor}}{1-\text{Cor}}} \right] \to h_b(0) = 0, \ \text{ for } \omega \in \{0, \cdots, M-1\},$$

and

$$h_b \left[\frac{(\omega+\alpha-1)\frac{\text{Cor}}{1-\text{Cor}} + \text{BER}}{1 + (M-1+\alpha)\frac{\text{Cor}}{1-\text{Cor}}} \right] \to h_b(1) = 0, \ \text{ for } \omega \in \{1, \cdots, M\}.$$

Therefore, $C_{\text{QBC}}^{(M)} \to 1$ when $\alpha \to \infty$ for $M \geq 2$, BER $\in (0,1)$ and Cor $\in (0,1)$. Note that $\varepsilon \to 1$ as $\alpha \to \infty$ from (3.14). When $\varepsilon = 1$, however, the experiments always choose the queue; this results in a queue with all 0s or all 1s when the number of experiments is sufficiently large. Thus, the channel is non-ergodic when $\varepsilon = 1$, and is not the same as the limiting channel as $\varepsilon \to 1$.

3.1.5 Reliability Function

The coding theorem, given by Claude E. Shannon [50] in 1948, states that for a certain class of channels it is possible to allow an arbitrarily small probability of decoding error by choosing the blocklength n sufficiently large if the transmission rate R is below the channel capacity C. The channel reliability function or error exponent $E(R)$, which is a more comprehensive tool than channel capacity, is defined as the asymptotic exponent of the minimum error probability $P_e^*(R,n)$ over all codes with blocklength n and rate R (e.g., see [8, 23] and Definition 2.11). $E(R)$ is a non-negative and non-increasing function of R for all $0 \leq R < C$, and $E(R) = 0$ for $R \geq C$, where C is the channel capacity.

$E(R)$ is indeed a more comprehensive tool than channel capacity, although it is considerably more difficult to study for general channels with memory (including the GEC).[1] For our binary channels with additive stationary ergodic noise, two bounds

[1]Recall that even for discrete memoryless channels, $E(R)$ is not exactly known at low rates; so

for $E(R)$ are known: the random coding lower bound (RCLB) [23] and the sphere-packing upper bound (SPUB) [19].[2]

Proposition 3.1 (Random coding lower bound (RCLB) [23] and Sphere-packing upper bound (SPUB) [19]) *For a channel with additive stationary ergodic noise described by* $\{P_{Y^n|X^n} : \mathcal{X}^n \to \mathcal{Y}^n\}_{n=1}^{\infty}$,

$$E_r(R) \leq E(R) \leq E_s(R), \tag{3.36}$$

where

$$E_r(R) \stackrel{\triangle}{=} \sup_{0 \leq \rho \leq 1} \left[-\rho R + E_0^{(\infty)}(\rho)\right],$$

$$E_s(R) \stackrel{\triangle}{=} \sup_{\rho \geq 0} \left[-\rho R + E_0^{(\infty)}(\rho)\right],$$

$$E_0^{(\infty)}(\rho) \stackrel{\triangle}{=} \lim_{n \to \infty} \left\{E_0^{(n)}(\rho)\right\},$$

and

$$E_0^{(n)}(\rho) \stackrel{\triangle}{=} -\sup_{X^n} \frac{1}{n} \log_2 \sum_{y^n \in \mathcal{Y}^n} \left[\sum_{x^n \in \mathcal{X}^n} P_{X^n}(x^n) P_{Y^n|X^n}(y^n \mid x^n)^{\frac{1}{1+\rho}}\right]^{1+\rho}. \tag{3.37}$$

The above RCLB and SPUB for $E(R)$ are tight when $R \geq R_{cr}$, where R_{cr} is the critical rate given by

$$R_{cr} \stackrel{\triangle}{=} \left.\frac{\partial E_0^{(\infty)}(\rho)}{\partial \rho}\right|_{\rho=1}.$$

---

only upper and lower bounds to $E(R)$ can be examined (see for example [8, 16, 23, 65]).

[2] These bounds actually hold for more general discrete channels with memory.

Note also that for our binary additive noise channels, $E_0^{(n)}(\rho)$ is achieved by the uniform input distribution $P_{X^n}(x^n) = 2^{-n}$ for all $x^n \in \{0,1\}^n$, since the channels are symmetric. Furthermore, when the noise is a Markov source, $E_0^{(\infty)}(\rho)$ admits an expression as follows [23, 47].

Proposition 3.2 *For a channel with binary stationary ergodic Markov noise with transition matrix* $[p_{ij}]$ *and* n*-fold distribution* $p^{(n)}$*,*

$$E_0^{(\infty)}(\rho) = \rho - \rho \lim_{n\to\infty} \frac{1}{n} H_{\frac{1}{1+\rho}}(p^{(n)}) = \rho - (1+\rho)\log_2 \lambda(\rho),$$

where $\lim_{n\to\infty}(1/n)H_\alpha(p^{(n)})$ *is the* Rényi entropy rate [47] *of the Markov noise with parameter* α *(*$\alpha > 0$ *and* $\alpha \neq 1$*),*

$$H_\alpha(p^{(n)}) \triangleq \frac{1}{1-\alpha}\log\left\{\sum_{z^n}\left[p^{(n)}(z^n)\right]^\alpha\right\},$$

and $\lambda(\rho)$ *is the largest eigenvalue of the matrix* $[p_{ij}^{1/(1+\rho)}]$*.*

The above result also directly holds for Markov noise sources of order M (c.f., [47]). Hence, for the QBC (whose noise process is of memory M), the RCLB and SPUB bounds on $E(R)$ can be readily obtained and calculated using Propositions 3.1 and 3.2.

3.2 Uniform Queue-Based Channel with Memory

We next study a particular case of the QBC, the uniform queue-based channel (UQBC), by fixing $\alpha = 1$; i.e., we operate on the queue cells with equal probability $1/M$. The UQBC block transition or noise probability $\Pr^{(M)}\{Y^n = y^n | X^n = x^n\} = \Pr^{(M)}\{Z^n = z^n\}$, where $z_i = x_i \oplus y_i$, of the UQBC can be expressed in terms of M, BER and Cor from (3.15) and (3.16) as follows.

- For blocklength $n \leq M$,

$$\Pr^{(M)}\{Z^n = z^n\} = \frac{\prod_{j=0}^{n-d_1^n-1}[j\frac{\text{Cor}}{1-\text{Cor}} + (1-\text{BER})]\prod_{j=0}^{d_1^n-1}[j\frac{\text{Cor}}{1-\text{Cor}} + \text{BER}]}{\prod_{j=0}^{n-1}[1+j\frac{\text{Cor}}{1-\text{Cor}}]}. \quad (3.38)$$

- For blocklength $n \geq M+1$,

$$\begin{aligned}\Pr^{(M)}\{Z^n = z^n\} &= L^{(M)} \prod_{i=M+1}^{n} \left\{\frac{d_{i-M}^{i-1}\frac{\text{Cor}}{1-\text{Cor}} + \text{BER}}{1+M\frac{\text{Cor}}{1-\text{Cor}}}\right\}^{z_i} \\ &\quad \left\{\frac{\left(M - d_{i-M}^{i-1}\right)\frac{\text{Cor}}{1-\text{Cor}} + (1-\text{BER})}{1+M\frac{\text{Cor}}{1-\text{Cor}}}\right\}^{1-z_i}, \quad (3.39)\end{aligned}$$

  where $L^{(M)}$ is given by (3.17).

Similarly, from (3.28), the capacity of the UQBC in terms of M, BER, and Cor is given by

$$C_{\text{UQBC}}^{(M)} = 1 - \sum_{\omega=0}^{M} \binom{M}{\omega} L_\omega^{(M)} h_b\left(\frac{\omega\frac{\text{Cor}}{1-\text{Cor}} + \text{BER}}{1+M\frac{\text{Cor}}{1-\text{Cor}}}\right), \quad (3.40)$$

where $L_\omega^{(M)}$ is given by (3.29).

Lemma 3.5 *The UQBC with memory order M and the QBC with memory order $M+1$ and $\alpha = 0$ have identical block transition probability for fixed BER and Cor; therefore the two channels have identical capacity under the above conditions.*

Proof When $\alpha = 0$, we obtain, from (3.15) and (3.16),

- For blocklength $n \leq M$,

$$\begin{aligned} &\Pr^{(M+1)}\{Z^n = z^n\} \\ &= \frac{\prod_{j=0}^{n-d_1^n-1}[j\frac{\text{Cor}}{1-\text{Cor}} + (1-\text{BER})]\prod_{j=0}^{d_1^n-1}[j\frac{\text{Cor}}{1-\text{Cor}} + \text{BER}]}{\prod_{j=0}^{n-1}[1+j\frac{\text{Cor}}{1-\text{Cor}}]}. \end{aligned} \tag{3.41}$$

- For blocklength $n = M+1$,

$$\begin{aligned} &\Pr^{(M+1)}\{Z^n = z^n\} \stackrel{\triangle}{=} L^{(M+1)} \\ &= \frac{\prod_{j=0}^{M+1-d_1^{M+1}-1}\left[j\frac{\text{Cor}}{1-\text{Cor}} + (1-\text{BER})\right]\prod_{j=0}^{d_1^{M+1}-1}\left[j\frac{\text{Cor}}{1-\text{Cor}} + \text{BER}\right]}{\prod_{j=0}^{M}\left(1+j\frac{\text{Cor}}{1-\text{Cor}}\right)} \\ &= L^{(M)}\prod_{i=M+1}^{n}\left\{\frac{d_1^M\frac{\text{Cor}}{1-\text{Cor}} + \text{BER}}{1+M\frac{\text{Cor}}{1-\text{Cor}}}\right\}^{z_i}\left\{\frac{\left(M-d_1^M\right)\frac{\text{Cor}}{1-\text{Cor}} + (1-\text{BER})}{1+M\frac{\text{Cor}}{1-\text{Cor}}}\right\}^{1-z_i}. \end{aligned} \tag{3.42}$$

- For blocklength $n \geq M+2$,

$$\begin{aligned} &\Pr^{(M+1)}\{Z^n = z^n\} \\ &= L^{(M+1)}\prod_{i=M+2}^{n}\left\{\frac{d_{i-M}^{i-1}\frac{\text{Cor}}{1-\text{Cor}} + \text{BER}}{1+M\frac{\text{Cor}}{1-\text{Cor}}}\right\}^{z_i}\left\{\frac{\left(M-d_{i-M}^{i-1}\right)\frac{\text{Cor}}{1-\text{Cor}} + (1-\text{BER})}{1+M\frac{\text{Cor}}{1-\text{Cor}}}\right\}^{1-z_i} \\ &= L^{(M)}\prod_{i=M+1}^{n}\left\{\frac{d_{i-M}^{i-1}\frac{\text{Cor}}{1-\text{Cor}} + \text{BER}}{1+M\frac{\text{Cor}}{1-\text{Cor}}}\right\}^{z_i}\left\{\frac{\left(M-d_{i-M}^{i-1}\right)\frac{\text{Cor}}{1-\text{Cor}} + (1-\text{BER})}{1+M\frac{\text{Cor}}{1-\text{Cor}}}\right\}^{1-z_i} \end{aligned} \tag{3.43}$$

By comparing (3.41) with (3.38), and comparing (3.42) and (3.43) with (3.39), we conclude that the UQBC with memory order M and the QBC with memory order $M+1$ and $\alpha = 0$ have identical block transition probabilities for the same BER and Cor. Thus, the two channels have identical capacity under the above conditions. □

Theorem 3.2 *The capacity $C_{\mathrm{QBC}}^{(M)}$ of the QBC is strictly increasing in M and converges to a limit for fixed BER, Cor and $0 \leq \alpha \leq 1$.*

Proof For fixed BER and Cor, the capacity of the QBC is a function of memory order M and parameter α. Let $C_{\mathrm{QBC}}^{(M)}(\alpha)$ denote the capacity of the QBC. Thus, for $0 < \alpha < 1$, we have

$$\begin{aligned} C_{\mathrm{QBC}}^{(M)}(\alpha) &< C_{\mathrm{QBC}}^{(M)}(1) \quad \text{(by Theorem 3.1)} \\ &= C_{\mathrm{QBC}}^{(M+1)}(0) \quad \text{(by Lemma 3.5)} \\ &< C_{\mathrm{QBC}}^{(M+1)}(\alpha) \quad \text{(by Theorem 3.1).} \end{aligned} \tag{3.44}$$

When $\alpha = 0$, we have

$$C_{\mathrm{QBC}}^{(M)}(0) < C_{\mathrm{QBC}}^{(M)}(1) = C_{\mathrm{QBC}}^{(M+1)}(0). \tag{3.45}$$

When $\alpha = 1$, we have

$$C_{\mathrm{QBC}}^{(M)}(1) = C_{\mathrm{QBC}}^{(M+1)}(0) < C_{\mathrm{QBC}}^{(M+1)}(1). \tag{3.46}$$

Thus, the capacity $C_{\text{QBC}}^{(M)}$ of the QBC is strictly increasing in M for fixed BER, Cor and $0 \leq \alpha \leq 1$. Since $\{C_{\text{QBC}}^{(M)}\}$ is increasing and upper bounded (by 1), the limit of $C_{\text{QBC}}^{(M)}$ exists as $M \to \infty$. □

3.3 QBC Capacity versus Capacity of Other Channels with Memory

In this section, we compare in terms of capacity the QBC with the FMCC [2], the GEC [37] and a particular symmetric class of the Fritchman channel [22] under identical channel parameters.

3.3.1 Comparison with the Finite-Memory Contagion Channel

The noise process of the FMCC is stationary and hence identically distributed [2]; the channel's BER and Cor are as follows: $\text{BER} = \rho$ and $\text{Cor} = \delta/(1+\delta)$. By comparing the UQBC with the FMCC in terms of block transition probability, we obtain the following result.

Theorem 3.3 *The UQBC and the FMCC are identical; i.e., they have the same block transition probability for the same memory order M, BER and Cor. Therefore*

the two channels have identical capacity under the above conditions.

Proof In terms of M, BER and Cor, the FMCC block transition probability

$$\Pr\{Y^n = y^n \mid X^n = x^n\} = \Pr\{Z^n = z^n\},$$

where $z_i = y_i \oplus x_i$ for $i = 1, \cdots, n$, can expressed as follows [2]:

- For blocklength $n \leq M$,

$$\Pr\{Z^n = z^n\} = \frac{\prod_{j=0}^{d-1}\left(\mathrm{BER} + j\frac{\mathrm{Cor}}{1-\mathrm{Cor}}\right)\prod_{j=0}^{n-1-d}\left[(1-\mathrm{BER}) + j\frac{\mathrm{Cor}}{1-\mathrm{Cor}}\right]}{\prod_{j=1}^{n-1}\left(1 + j\frac{\mathrm{Cor}}{1-\mathrm{Cor}}\right)}, \quad (3.47)$$

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

- For blocklength $n \geq M + 1$,

$$\begin{aligned}&\Pr\{Z^n = z^n\}\\ &= L \prod_{i=M+1}^{n}\left(\frac{\mathrm{BER} + \lambda_{i-1}\frac{\mathrm{Cor}}{1-\mathrm{Cor}}}{1 + M\frac{\mathrm{Cor}}{1-\mathrm{Cor}}}\right)^{z_i}\left[\frac{(1-\mathrm{BER}) + (M-\lambda_{i-1})\frac{\mathrm{Cor}}{1-\mathrm{Cor}}}{1 + M\frac{\mathrm{Cor}}{1-\mathrm{Cor}}}\right]^{1-z_i}, \quad (3.48)\end{aligned}$$

  where

$$L = \frac{\prod_{j=0}^{\lambda_M - 1}\left(\mathrm{BER} + j\frac{\mathrm{Cor}}{1-\mathrm{Cor}}\right)\prod_{j=0}^{M-1-\lambda_M}\left[(1-\mathrm{BER}) + j\frac{\mathrm{Cor}}{1-\mathrm{Cor}}\right]}{\prod_{j=1}^{M-1}\left(1 + j\frac{\mathrm{Cor}}{1-\mathrm{Cor}}\right)},$$

  $\prod_{j=0}^{a}(\cdot) \stackrel{\triangle}{=} 1$ if $a < 0$, $z_i = x_i \oplus y_i$, and $\lambda_{i-1} = z_{i-1} + \cdots + z_{i-M}$ for $i = M+1, \cdots, n$.

Comparing (3.38) with (3.47) and (3.39) with (3.48), we observe that the FMCC and the UQBC have *identical* block transition probability for the same memory order

M, BER and Cor. Thus, the two channels have identical capacity under the above conditions. □

Observation: The above result appears at first somewhat surprising since the same Markov noise process seems to have been generated by two different experiments: the Polya contagion urn scheme and the finite queue scheme. However, upon further reflection, the equivalence of the two experiments becomes transparent when we equate the original $T = R + B$ balls in the FMCC urn scheme with the urn (with proportion p of red balls) in the UQBC scheme and the balls which are added and later removed in the FMCC urn scheme with the queue in the UQBC scheme. We can describe the FMCC urn scheme as a two-stage experiment. In the first stage we decide to pick from the original balls (with probability $T/(M\Delta + T)$) or from the "transient" balls (with probability $M\Delta/(M\Delta + T)$). This is equivalent to the first stage of the UQBC experiment where we choose either the urn (with probability $1-\varepsilon$) or the queue (with probability ε). Indeed, the FMCC parameter Δ no longer serves a purpose once the first stage is completed. In the second stage we choose a ball at random from the set we have chosen in the first stage. For the purposes of picking a color at random, we can view the transient balls (of which there are $M\Delta$) as just M balls, with each red ball representing Δ red balls and each black ball representing Δ black balls. Thus, if we have decided to choose from the transient balls in the first stage, we pick one of the M balls (each with "weight" Δ) at random (equivalent to picking a cell at

random in the UQBC scheme), then add a ball of that color and remove the ball that was added M draws ago (equivalent to pushing a ball of the chosen color into the front of the UQBC queue and pushing the last ball in the queue out).

We note that the correlation coefficient of the FMCC does not depend on memory order M ($\mathrm{Cor} = \delta/(1+\delta)$) while that of the UQBC depends on memory order M (see (3.12)). Therefore, the two channel models are different in their parameterizations although their block transition probabilities and capacities are identical under the same M, BER and Cor.

From Theorem 3.3 and [2], the following asymptotic expression for $C_{\mathrm{UQBC}}^{(M)}$ can be established as M approaches infinity while keeping BER and Cor fixed:

$$\lim_{M\to\infty} C_{\mathrm{UQBC}}^{(M)} = 1 - \int_0^1 h_b(z) f_Z(z) dz, \tag{3.49}$$

where $h_b(\cdot)$ is the binary entropy function and $f_Z(z)$ is the beta probability density function (pdf) $\beta_{u,v}(z)$ [45], [21, pp. 50] with parameters $u = \mathrm{BER}(1-\mathrm{Cor})/\mathrm{Cor}$ and $v = (1-\mathrm{BER})(1-\mathrm{Cor})/\mathrm{Cor}$; i.e.,

$$f_Z(z) = \beta_{u,v}(z) = \frac{\Gamma(u+v)}{\Gamma(u)\,\Gamma(v)} (1-z)^{(u-1)} z^{(v-1)} I_{[0,1]}(z),$$

where $I_{[0,1]}(z)$ denotes the indicator function on the set $[0,1]$.

From Theorem 3.2, we note that (3.49) is an upper bound of the capacity of the UQBC for a given M.

Corollary 3.1 For the same M, BER and Cor,

$$C_{\text{QBC}}^{(M)} < C_{\text{FMCC}}^{(M)} \quad (\text{when } 0 \leq \alpha < 1), \tag{3.50}$$

and

$$C_{\text{QBC}}^{(M)} > C_{\text{FMCC}}^{(M)} \quad (\text{when } \alpha > 1). \tag{3.51}$$

Proof (3.50) and (3.51) can be obtained directly from Theorems 3.1 and 3.3. □

3.3.2 Comparison with the Gilbert-Elliott Channel

The noise process of the GEC has infinite memory in the sense that it has an infinite dependency structure since it is a hidden Markov chain. Therefore, knowledge of the infinite past is necessary for the calculations of the GEC noise entropy rate and capacity. In practice, a finite past is used to approximate such calculations; a scheme for such calculation is proposed by Mushkin and Bar-David [37] based on increasingly tight upper and lower bounds (c.f., also the theoretical study in [27]). In particular, it is shown in [37] that the GEC capacity can be obtained by $C_{\text{GEC}} = \lim_{l\to\infty} \underline{C}_l$, where $\underline{C}_l$ is defined as $\underline{C}_l \stackrel{\triangle}{=} 1 - E[h_b(q_l)]$. The random variable $q_l(Z^{l-1})$, for $l \geq 2$, denotes the probability of a channel error at the lth use, conditioned on the previous noise samples, i.e., $q_l(Z^{l-1}) \stackrel{\triangle}{=} \Pr\{Z_l = 1 \mid Z^{l-1}\}$, where $Z^{l-1} = (Z_{l-1}, \ldots, Z_1)$. $\{\underline{C}_l\}_{l=1}^{\infty}$ is monotonically increasing with l since the binary entropy function $h_b(q)$ is concave and continuous over $[p_G, p_B]$. Thus, for $l \geq 1$, $\underline{C}_l$ can be computed to provide a lower

bound to the capacity C_{GEC}; i.e., $C_{\mathrm{GEC}} \geq \underline{C}_l$; as l increases, the bound becomes sharper. The following recursion holds for q_l [37]:

$$q_{l+1}(Z^l) = \nu(Z_l, q_l(Z^{l-1})), \tag{3.52}$$

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

$$\nu(0,q) \stackrel{\triangle}{=} \begin{cases} p_G + b(p_B - p_G) + \mu(q - p_G)[(1-p_B)/(1-q)], & p_B \neq 1, \\ (1-b)p_G + b, & p_B = 1,\ q \neq 1, \end{cases} \tag{3.53}$$

and

$$\nu(1,q) \stackrel{\triangle}{=} \begin{cases} p_G + b(p_B - p_G) + \mu(q - p_G)(p_B/q), & p_G \neq 0, \\ (1-g)p_B, & p_G = 0,\ q \neq 0, \end{cases} \tag{3.54}$$

for $p_G \leq q \leq p_B$, where $\mu \stackrel{\triangle}{=} 1 - g - b$. The initial value for the recursion is

$$q_1 \stackrel{\triangle}{=} \Pr\{Z_1 = 1\} = (\rho p_G + p_B)(\rho + 1)^{-1}, \tag{3.55}$$

where $\rho \stackrel{\triangle}{=} g/b$. The random sequence $\{q_l\}_{l=1}^{\infty}$ is a Markov process with initial value q_1 and transition probabilities

$$\Pr\{q_{l+1} = \alpha \mid q_l = \beta\} = \begin{cases} 1-\beta, & \alpha = \nu(0,\beta), \\ \beta, & \alpha = \nu(1,\beta). \end{cases} \tag{3.56}$$

By comparing $\underline{C}_2$ with the capacity of the UQBC with $M = 1$, $C_{\mathrm{UQBC}}^{(M=1)}$, we arrive at the following result.

Theorem 3.4 *For $M = 1$, and for the same BER and Cor, $C_{\mathrm{GEC}} \geq C_{\mathrm{UQBC}}^{(M=1)}$.*

Proof In terms of BER and Cor, the initial value for the random sequence q_l (3.55) is

$$q_1 = (\rho p_G + p_B)(\rho + 1)^{-1} = BER. \tag{3.57}$$

From the recursion (3.52), the function $\nu(\cdot, \cdot)$ in (3.53) and (3.54), and the transition probability (3.56), we obtain

$$\nu(1, q_1) = \text{BER} + (1 - \text{BER}) \cdot \text{Cor}, \quad \text{with} \quad \Pr[q_2 = \nu(1, q_1)] = \text{BER},$$

and

$$\nu(0, q_1) = \text{BER} \cdot (1 - \text{Cor}), \quad \text{with} \quad \Pr[q_2 = \nu(0, q_1)] = 1 - \text{BER},$$

where the GEC Cor is given by

$$\text{Cor} = \frac{\mu(\text{BER} - p_G)(p_B - \text{BER})}{\text{BER}(1 - \text{BER})}. \tag{3.58}$$

Thus,

$$\underline{C}_2 = 1 - E[h_b(q_2)]$$

$$= 1 - \{\text{BER} \cdot h_b[\text{BER} + (1 - \text{BER}) \cdot \text{Cor}] + (1 - \text{BER}) \cdot h_b[\text{BER} \cdot (1 - \text{Cor})]\}. \tag{3.59}$$

With memory order $M = 1$ in the UQBC, from (3.40) it follows that

$$C_{\text{UQBC}}^{(M=1)} = 1 - \sum_{\omega=0}^{1} L_{\omega}^{(1)} h_b[\omega \cdot \text{Cor} + (1 - \text{Cor}) \cdot \text{BER}]$$

$$= 1 - \{\text{BER} \cdot h_b[\text{Cor} + (1 - \text{Cor}) \cdot \text{BER}] + (1 - \text{BER}) \cdot h_b[\text{BER} \cdot (1 - \text{Cor})]\}. \tag{3.60}$$

Comparing (3.59) with (3.60), we observe that $\underline{C}_2 = C_{\text{UQBC}}^{(M=1)}$. Since $\underline{C}_2 \leq C_{\text{GEC}}$, we conclude that $C_{\text{GEC}} \geq C_{\text{UQBC}}^{(M=1)}$ for the same BER and Cor. □

Finally, it should be noted that when $M \geq 2$, $C_{\text{QBC}}^{(M)}$ can be either smaller or bigger than C_{GEC}, depending on the values of BER, Cor, M and α.

3.3.3 Comparison with the Symmetric Fritchman Channel (K, 1)-SFC

We define the symmetric Fritchman channel with K good states and one bad state ($(K, 1)$-SFC) by the following transition matrix on its states

$$\boldsymbol{P}_{(K,1)\text{-SFC}} = \begin{bmatrix} p_{00} & (1-p_{00})/K & \cdots & (1-p_{00})/K \\ (1-p_{00})/K & p_{00} & \cdots & (1-p_{00})/K \\ & \vdots & & \\ (1-p_{00})/K & \cdots & p_{00} & (1-p_{00})/K \\ (1-p_{11})/K & \cdots & (1-p_{11})/K & p_{11} \end{bmatrix}, \quad (3.61)$$

where p_{00} is the probability of staying in the current good state and p_{11} is the probability of staying in the bad state. By comparing the UQBC with memory 1 to the $(K, 1)$-SFC in terms of the probability of an arbitrary noise sequence, we obtain the following.

Lemma 3.6 For the same BER and Cor, and for any $K = 1, 2, \cdots$, the $(K,\ 1)$-SFC is statistically identical to the UQBC with $M = 1$. Hence $C_{(K,\ 1)\text{-SFC}} = C_{\text{UQBC}}^{(M=1)} \leq C_{\text{UQBC}}^{(M)} \leq C_{\text{QBC}}^{(M)},\ \forall\ M \geq 1$ and $\alpha \geq 1$.

Proof We observe that the good states all have the same stationary probability $(1-\text{BER})/K$ and they have the same transition pattern. Hence the good states can be combined into one large good state with stationary probability $(1-\text{BER})$; this makes the $(K,\ 1)$-SFC statistically the same as a 2-state Markov chain (the (1, 1)-FC or UQBC with memory 1). Hence $C_{(K,\ 1)\text{-SFC}} = C_{\text{UQBC}}^{(M=1)}$ and since memory increases capacity (Theorem 3.2) and α increases capacity (Theorem 3.1), we have

$$C_{(K,\ 1)\text{-SFC}} = C_{\text{UQBC}}^{(M=1)} \leq C_{\text{UQBC}}^{(M)} \leq C_{\text{QBC}}^{(M)}$$

$\forall\ M \geq 1$ and $\alpha \geq 1$. □

3.4 Capacity Numerical Results

We next compare numerically the capacities of the QBC, the GEC and the FC. For all these models, the capacities are calculated in terms of BER and Cor. Since the GEC is described by four parameters, we fix $p_G = 0.00002$ and $p_B = 0.92$ (since they allow for Cor to range from 0.1 to 0.9) and calculate the upper and lower bounds [37] for the capacity in terms of BER and Cor. Also, with $p_G = 0.00002$ and $p_B = 0.92$,

the upper and lower bounds to the noise entropy rate (and hence to C_{GEC}) converge quickly.

In [22], a general expression for the capacity of Fritchman channels with a single error-state and K good states ($(K,\,1)$-FC) is provided:

$$\begin{aligned} C_{(K,1)\text{-FC}} = 1 &+ \Pr\{Z_i = 1\} \sum_{z_{i+1}=0}^{1} \Pr\{Z_{i+1} = z_{i+1} | Z_i = 1\} \log_2 \Pr\{Z_{i+1} = z_{i+1} | Z_i = 1\} \\ &+ \sum_{m=1}^{\infty} \Big[\Pr\{Z_i = 1, Z_{i+1} = 0, \cdots, Z_{i+m} = 0\} \\ &\times \sum_{z_{i+m+1}=0}^{1} \Pr\{Z_{i+m+1} = z_{i+m+1} | Z_i = 1, Z_{i+1} = 0, \cdots, Z_{i+m} = 0\} \\ &\times \log_2 \Pr\{Z_{i+m+1} = z_{i+m+1} | Z_i = 1, Z_{i+1} = 0, \cdots, Z_{i+m} = 0\} \Big] \end{aligned} \tag{3.62}$$

where the probabilities and the conditional probabilities of error sequences can be calculated via (2.6). We employ the above expression to compute the capacity of the $(2,\,1)$-FC with the transition probability matrix

$$\boldsymbol{P}_{(2,\,1)\text{-FC}} = \begin{bmatrix} p_{00} & (1-p_{00})/2 & (1-p_{00})/2 \\ 0.1 & 0.5 & 0.4 \\ (1-p_{11})/2 & (1-p_{11})/2 & p_{11} \end{bmatrix}.$$

where p_{00} and p_{11} vary as BER and Cor vary.

As expected, we observe that memory increases the capacities of the UQBC for the same BER and Cor (see Figs. 3.2, 3.3, and 3.4). In those three figures, the curves for $C_{\text{UQBC}}^{(M\to\infty)}$ provide the upper bounds for the capacities of the UQBC. However, this pattern does not work for the QBC model generally, which depends on the value of α.

We illustrate the effect of the cell parameters M and α on the capacity of the QBC in Figs. 3.5 and 3.6. As expected from Theorems 3.1 and 3.2, the capacity increases with α for the same BER, Cor and memory order M, and the capacity increases with memory order M for the same BER and Cor when $0 \leq \alpha \leq 1$. However, this latter pattern does not generally hold for the QBC model as it depends on the condition that $\alpha < 1$ (c.f. Theorem 3.2). For example in Fig. 3.5, when $\alpha \geq 10$, the situation reverses and capacity decreases with the memory order. We also illustrate in Fig. 3.6 the adage "memory increases capacity" (for information stable channels [2]) by noting the large capacity gain of the QBC over the BSC for identical BER. Note that the BSC can indeed be looked at as the channel resulting when ideal interleaving is employed on the QBC; this indicates that effectively exploiting the noise memory in the system design is significantly better than ignoring it via interleaving.

The capacity of the QBC, the GEC and the (2, 1)-FC increase with Cor. As shown in Fig. 3.7, the QBC with $M = 2$ and $\alpha = 10$ has the largest capacity, whereas the UQBC with $M = 1$ (or (1, 1)-FC) has the smallest capacity. $C_{\text{UQBC}}^{(M=1)}$ is smaller than C_{GEC} as predicted by Theorem 3.4. When Cor=0.1, the GEC and the UQBC with $M = 1$ have nearly equal capacities, indicating that in this case we can replace the GEC with the less complex UQBC if our target is to achieve a capacity which is close to that of the GEC. For the same BER, the capacity of the QBC can be either smaller or bigger than that of the GEC and (2, 1)-FC, depending on the values of

Cor, M and α (see Fig. 3.7).

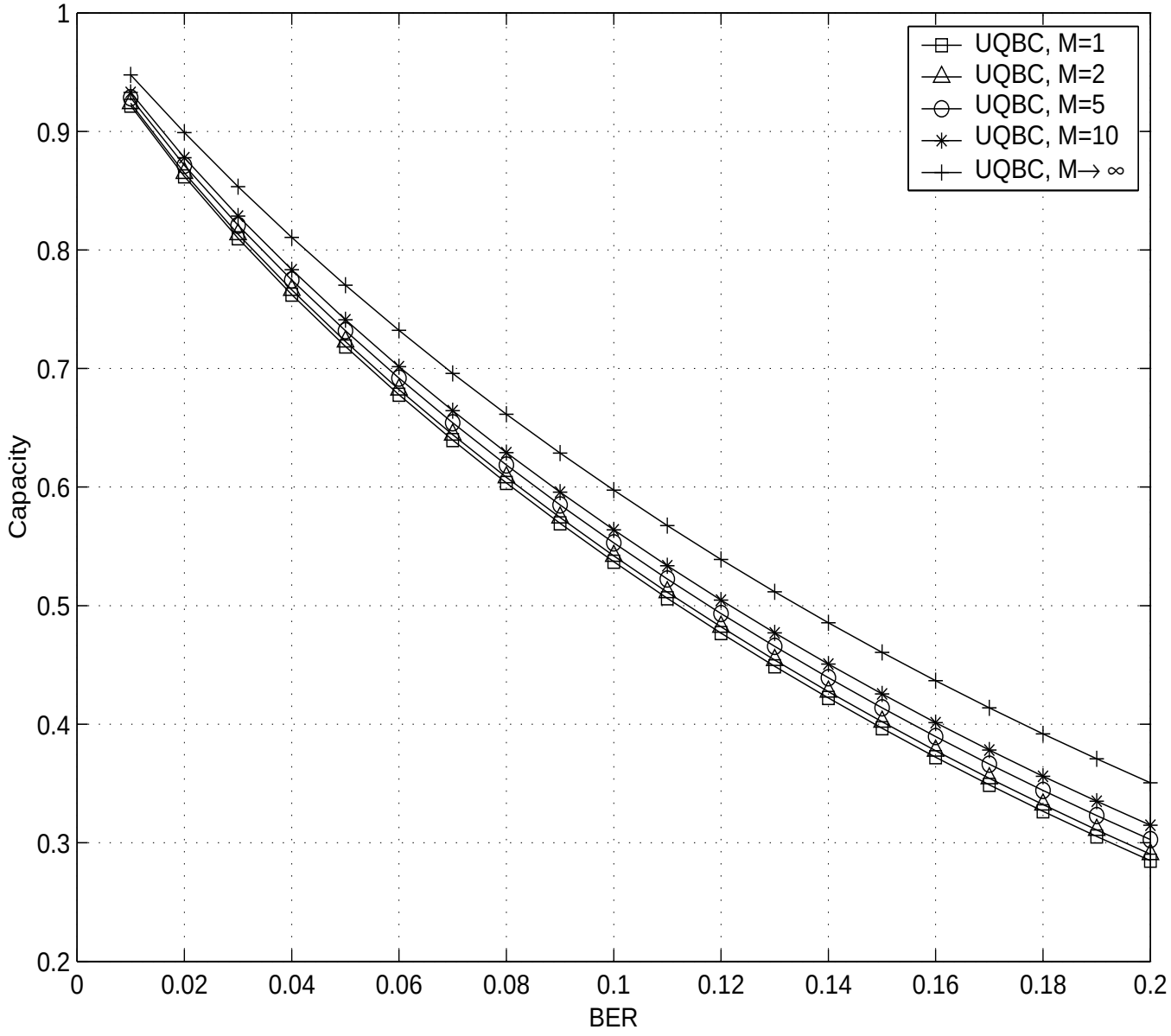


Figure 3.2: Capacity of the UQBC vs BER for Cor=0.1.

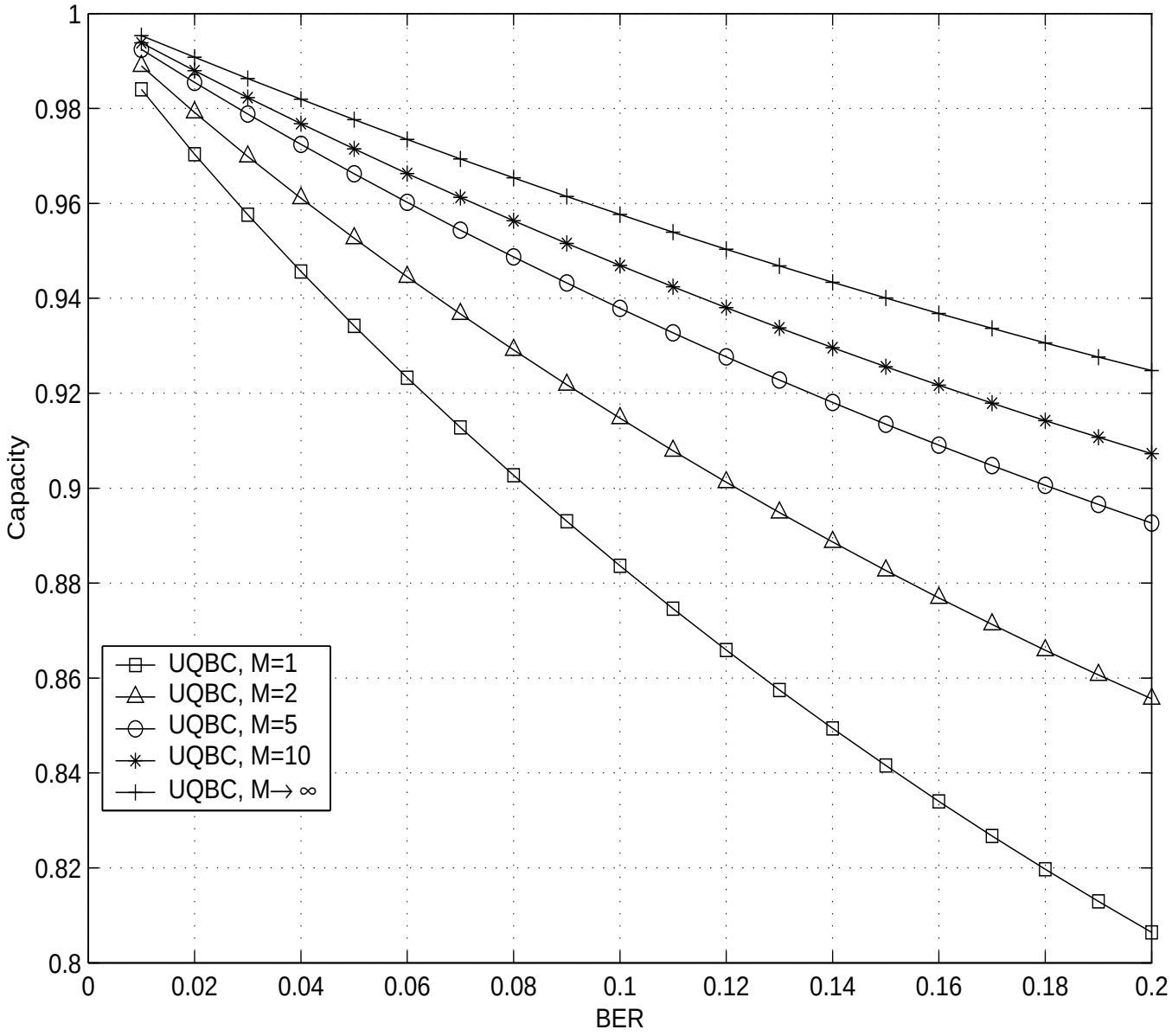


Figure 3.3: Capacity of the UQBC vs BER for Cor=0.9.

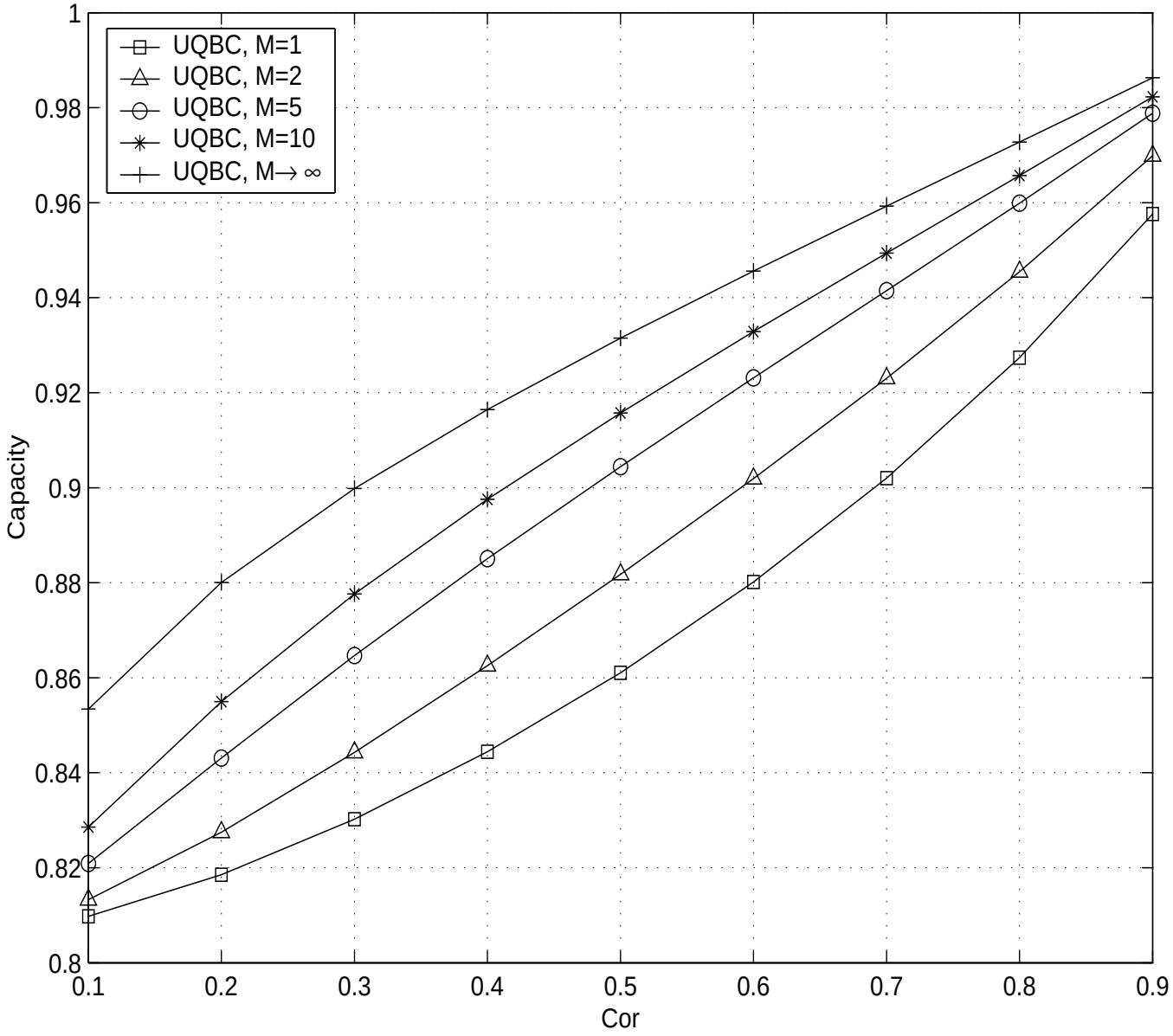


Figure 3.4: Capacity of the UQBC vs Cor for BER=0.03.

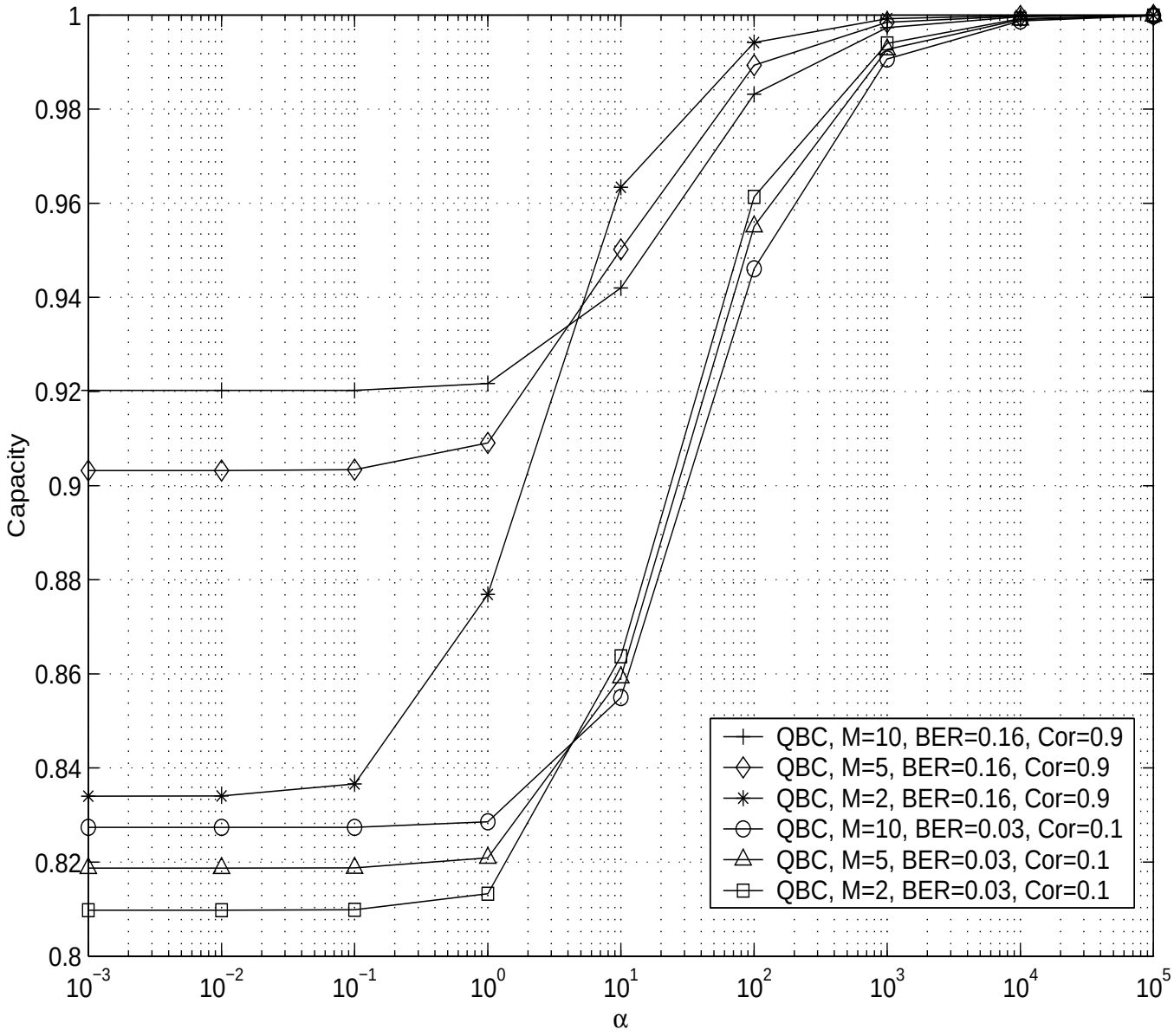


Figure 3.5: Capacity vs α for QBC.

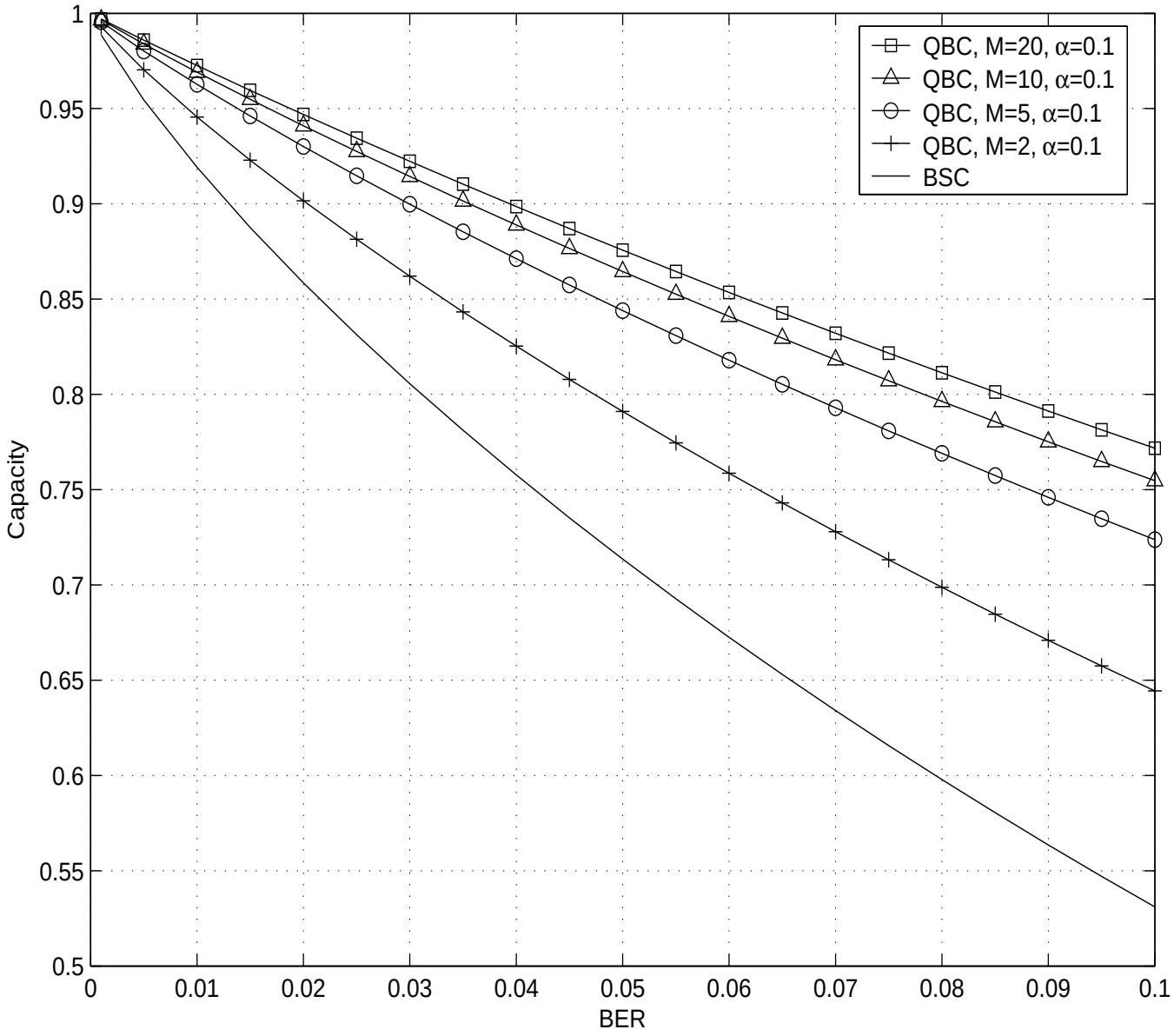


Figure 3.6: Capacity vs BER for QBC with Cor=0.5.

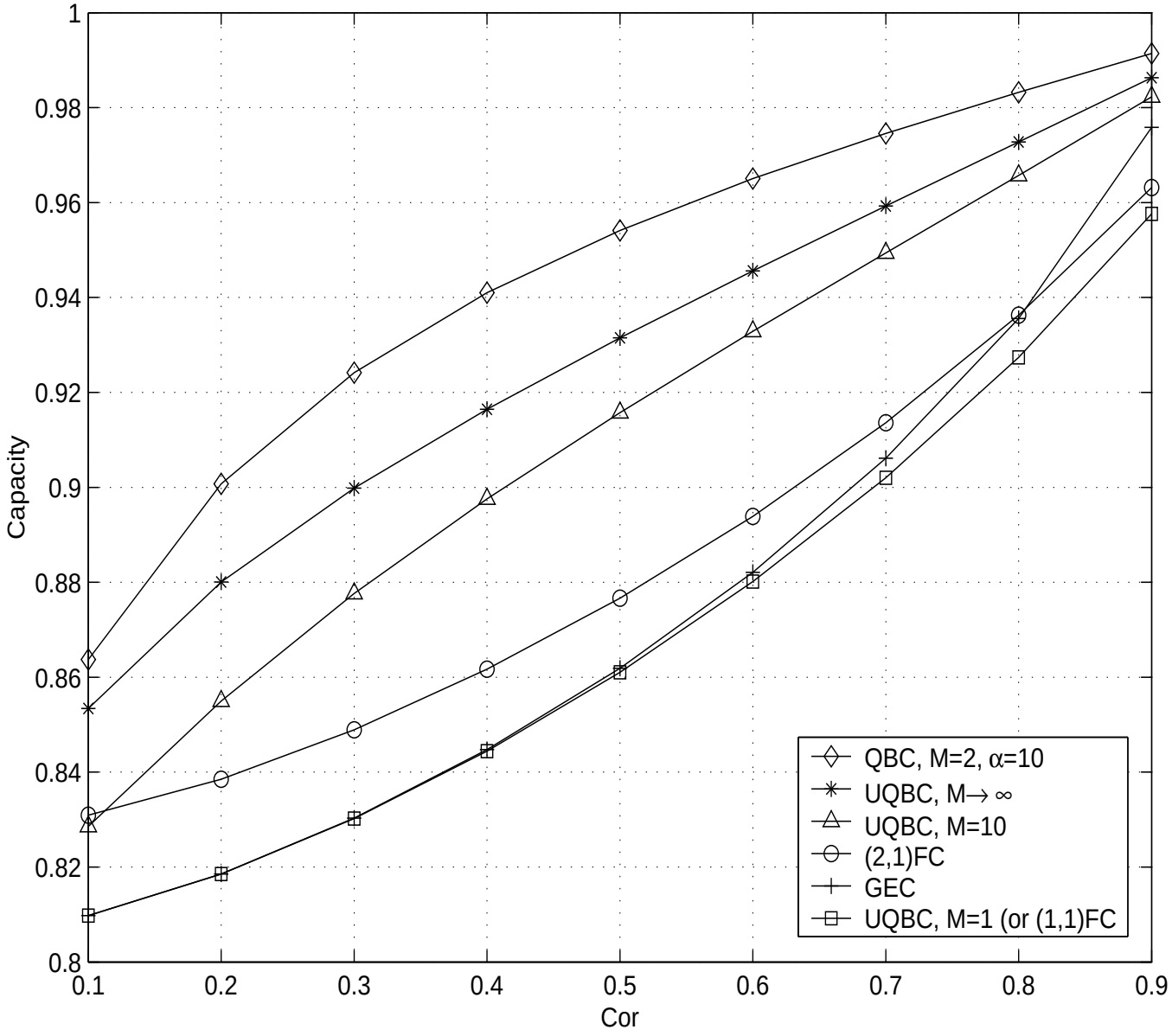


Figure 3.7: Capacity vs Cor for BER=0.03; $p_G = 0.00002$ and $p_B = 0.92$ (for GEC).

Chapter 4

Approximating the GEC via the QBC

We next consider the problem of fitting the GEC model via the QBC model. In other words for a given GEC, we construct a QBC that best approximates it.

4.1 Estimation of QBC Parameters

For a given GEC (the GC is a special case of the GEC if $p_G = 0$), we construct a QBC whose noise process is statistically "close" in the Kullback-Leibler sense to the noise process generated by the GEC. Specifically, given a GEC with fixed parameters b, g, p_B and p_G resulting in bit error rate $\mathrm{BER}_{\mathrm{GEC}}$ and correlation coefficient $\mathrm{Cor}_{\mathrm{GEC}}$,

we estimate the QBC parameters M, p, ε, and α that minimize the KLDR

$$\lim_{n\to\infty} \frac{1}{n} D_n(\mathrm{P}_{\mathrm{GEC}} \parallel \mathrm{P}_{\mathrm{QBC}}^{(M)}),$$

subject to the constraints

$$\mathrm{BER}_{\mathrm{QBC}} = \mathrm{BER}_{\mathrm{GEC}}$$

and

$$\mathrm{Cor}_{\mathrm{QBC}} = \mathrm{Cor}_{\mathrm{GEC}},$$

where $(1/n)D_n(\mathrm{P}_{\mathrm{GEC}} \parallel \mathrm{P}_{\mathrm{QBC}}^{(M)})$ is the normalized nth order Kullback-Leibler divergence (NKLD) between the n-fold GEC and QBC noise distributions, $\mathrm{P}_{\mathrm{GEC}}$ and $\mathrm{P}_{\mathrm{QBC}}^{(M)}$, respectively:

$$D_n(\mathrm{P}_{\mathrm{GEC}} \parallel \mathrm{P}_{\mathrm{QBC}}^{(M)}) = \sum_{z^n \in \{0,1\}^n} \mathrm{P}_{\mathrm{GEC}}(z^n) \log_2 \frac{\mathrm{P}_{\mathrm{GEC}}(z^n)}{\mathrm{P}_{\mathrm{QBC}}^{(M)}(z^n)},$$

where $\mathrm{P}_{\mathrm{GEC}}(z^n)$ is given by (2.2) and $\mathrm{P}_{\mathrm{QBC}}^{(M)}(z^n)$ is given by (3.7)-(3.8). It can be shown (e.g., see Proposition 2.4) that the KLDR between a stationary source and a Markov source does exist and can be expressed as

$$\begin{aligned}
& \lim_{n\to\infty} \frac{1}{n} D_n(\mathrm{P}_{\mathrm{GEC}} \parallel \mathrm{P}_{\mathrm{QBC}}^{(M)}) \\
= \; & -\lim_{n\to\infty} \frac{1}{n} H_{\mathrm{GEC}}(Z^n) \; - \sum_{z_1,\cdots,z_{M+1}} \mathrm{P}_{\mathrm{GEC}}(z_1,\cdots,z_{M+1}) \; \log_2 \mathrm{P}_{\mathrm{QBC}}^{(M)}(z_{M+1}|z_M,\cdots,z_1) \\
= \; & -\mathcal{H}_{\mathrm{GEC}}(Z) - E_{\mathrm{P}_{\mathrm{GEC}}}[\log_2 \mathrm{P}_{\mathrm{QBC}}^{(M)}(Z_{M+1}|Z^M)],
\end{aligned}$$

where $\mathcal{H}(\cdot)$ denotes the entropy rate and $\mathrm{P}_{\mathrm{QBC}}^{(M)}(z_{M+1}|z^M)$ is the QBC conditional error probability of symbol $M+1$ given the previous M symbols. Then the minimization

reduces to maximizing the second term

$$E_{\mathrm{P}_{\mathrm{GEC}}}[\log_2 \mathrm{P}^{(M)}_{\mathrm{QBC}}(Z_{M+1}|Z^M)]$$

(which is independent of n) over the QBC parameters. Note that in our approximation, we match the bit error rates and noise correlation coefficients of both channels to guarantee identical noise marginal distributions and identical probabilities of two consecutive errors (ones). Hence, given these constraints, the above optimization problem reduces to an optimization over only two QBC parameters: M and ε. This is achieved numerically by sequentially incrementing $M \geq 1$ and varying $0.0001 \leq \varepsilon \leq 0.9999$ for each given M.

4.2 Modeling Results

We evaluate how well the QBC model (obtained via the above KLDR minimization) fits or approximates the GEC according to three criteria: channel capacity, ACF,[1] and reliability function. The QBC ACF and capacity expressions are provided in Sections 3.1.3 and 3.1.4. Although the capacity of the GEC does not have an analytical expression, it can be determined accurately via the algorithm of [37] (see

[1]It is indeed observed in [60, 44] that the ACF is an effective tool to measure the agreement between models for channels with memory, including between finite-state channel models and discretized correlated fading channels.

Section 3.3.2). The ACF of the GEC can also be obtained directly from (2.2).

$$R[m] = \boldsymbol{\pi}^T \boldsymbol{P}(1) \left(\prod_{k=1}^{m-1} \boldsymbol{P} \right) \boldsymbol{P}(1)\mathbf{1}, \tag{4.1}$$

where $\boldsymbol{\pi}$, $\boldsymbol{P}(1)$, $\boldsymbol{P}$, $\boldsymbol{P}(1)$ and $\mathbf{1}$ are defined in Section 2.3.1.

We employ Propositions 3.1 and 3.2 to obtain the RCLB and SPUB on $E(R)$ for the QBC. However, Proposition 3.2 cannot be directly applied to hidden Markov channel models (such as the GEC) since the state is not a deterministic function of the previous state and previous letter. For the GC, if $z_i = 1$, the corresponding state is certainly B and thus the errors are completely defined by the conditional probability $p(k)$ of a run of $k-1$ zeros followed by a one subject to this sequence being preceded by a one, i.e., $p(k) = P\{Z_{i+1} = 0, \cdots, Z_{i+k-1} = 0, Z_{i+k} = 1 \mid Z_i = 1\}$. Therefore, we can use the following result [19] along with Proposition 3.1 to determine the RCLB and SPUB on $E(R)$ for the GC.

Proposition 4.1 For a channel with errors fully defined by $p(k)$ (such as the GC),

$$E_0^{(\infty)}(\rho) = \rho + (1+\rho)\gamma, \tag{4.2}$$

where γ is the solution of

$$\sum_{i=1}^{\tau} p(i)^{\frac{1}{1+\rho}} 2^{i\gamma} = 1, \tag{4.3}$$

where τ is a positive integer such that $p(k) = 0$ is assumed for $k > \tau$.

In our results for the GC error exponent, we set the value τ to be as large as 70,000. For the GEC, $E_0^{(\infty)}(\rho)$ does not admit (to our knowledge) a computable expression; hence we approximate it by calculating $E_0^{(n)}(\rho)$ for large n (we used $n = 21$ and noted that larger values of n result in only a minor change in $E_0^{(n)}(\rho)$), and we employ Proposition 3.1 to evaluate the GEC error exponent. Similarly, the GEC critical rate is estimated by computing $R_{cr}^{(n)}$ for large n, where $R_{cr}^{(n)} \stackrel{\triangle}{=} \left.\frac{\partial E_0^{(n)}(\rho)}{\partial \rho}\right|_{\rho=1}$.

A wide range of GEC/GC channel parameters is investigated with $0.01 \leq$ Cor ≤ 0.9 and $0.1\% \leq$ BER $\leq 10\%$. In Table 4.1 and Figs. 4.1-4.14, we present typical evaluation results for the approximation of the GEC/GC models via the QBC. In all, fitting results for eight values of Cor are shown (four for the GEC and four for the GC). To illustrate a realistic setting, the values of the pair (Cor,BER) for Cases **A** and **C** in Table 4.1 were chosen to match the conditions of the correlated Rayleigh fading channel studied in [44, Fig. 6.(b)] with normalized Doppler frequencies of 0.1 (fast fading) and 0.001 (slow fading), respectively.

We first observe in Figs. 4.1 and 4.2 that $(1/n)D_n(\mathrm{P}_{\mathrm{GEC}} \parallel \mathrm{P}_{\mathrm{QBC}}^{(M)})$ monotonically increases in n before asymptotically converging to the KLDR. Hence by choosing to minimize the KLDR to fit the GEC via the QBC, we are indeed addressing the worst-case scenario, as for finite n, the normalized nth order Kullback-Leibler distance is smaller than the KLDR and thus yields a closer statistical match between the two channels. We next notice that the values of the minimum KLDR in Figs. 4.1-4.2

are less than 0.01 for all cases, except for Case **G** (see (c) in Fig. 4.2). As a result, we observe a strong agreement in ACF, capacity and error exponents in Figs. 4.3-4.14 (this behavior was indeed observed for all computations). In particular, the value of the KLDR for Case **A** with Cor=0.0131 is less than 10^{-5} which indicates excellent matching of the two models, although in this case both channels behave like a BSC since they have a small correlation coefficient (see (a) in Fig. 4.1). In Figs. 4.3 and 4.4, we note that the ACF curves of the two channels are nearly identical, except for Cases **C** and **F** where the ACF curves for the GEC take a longer span of m before eventually converging.

Modeling results in terms of capacity are shown in Figs. 4.5-4.12, where the capacity of the GEC and its QBC approximation are shown for different BER values and for fixed Cor values. It is worth mentioning that in the capacity comparison figures, the QBC parameters are optimized for each value of BER. We clearly observe from the figures that the capacity curves of both channels match quite well and the capacity curves in Figs. 4.5-4.7 are almost identical. Note from the capacity figures that the largest Markovian memory M for the QBC model that best fits the GEC is 15 (see Fig. 4.7). Overall, we remark a strong match in capacity between the GEC and its QBC approximation - indeed, even in Fig. 4.12 where the fit is weakest for Cor = 0.9, the difference between the capacity curves is less than 1%.

We observe an excellent matching between the two channels in terms of error

exponents for low Cor values (see Fig. 4.13), while the agreement is weaker but still good for high values of Cor (see Fig. 4.14). Although we do not have exact RCLB and SPUB expressions for the GEC, the curves of $E_r(R)$ and $E_s(R)$ for $n = 21$ are close to the QBC exponents, particularly for Cases **A**, **C** and **D**. We also note a good fit between the critical rates of the channels.

Overall, our results indicate that the QBC provides a good fit for the GEC and GC channels in terms of ACF, capacity and reliability function for a wide range of channel conditions.

Cases	Cor	BER	GEC/GC parameters	QBC parameters
A	0.0131	0.00314	GEC: $p_G = 0.00259$, $p_B = 0.4523$	$M = 2$, $\varepsilon = 0.0145$, $p = 0.00314$, $\alpha = 0.1054$
B	0.2227	0.03	GC: $b = 0.02$, $g = 0.18$	$M = 4$, $\varepsilon = 0.4948$, $p = 0.03$, $\alpha = 0.4189$
C	0.248	0.012	GEC: $p_G = 0.00741$, $p_B = 0.6555$	$M = 4$, $\varepsilon = 0.5279$, $p = 0.012$, $\alpha = 0.3907$
D	0.341	0.1178	GEC: $p_G = 0.01$, $p_B = 0.5$	$M = 4$, $\varepsilon = 0.6397$, $p = 0.1178$, $\alpha = 0.4319$
E	0.5	0.03	GC: $b = 0.00367$, $g = 0.0636$	$M = 6$, $\varepsilon = 0.841$, $p = 0.03$, $\alpha = 0.2893$
F	0.7	0.03	GC: $b = 0.00228$, $g = 0.0547$	$M = 4$, $\varepsilon = 0.8864$, $p = 0.03$, $\alpha = 0.3441$
G	0.8	0.1	GEC: $p_G = 0.0101$, $p_B = 0.909$	$M = 4$, $\varepsilon = 0.9294$, $p = 0.1$, $\alpha = 0.2911$
H	0.9	0.03	GC: $b = 0.00161$, $g = 0.0495$	$M = 2$, $\varepsilon = 0.9251$, $p = 0.03$, $\alpha = 0.3723$

Table 4.1: GEC fitting via the QBC: GEC and QBC parameters. For the GEC, b and g are determined using (3.57) and (3.58); for the GC, $p_G = 0$ and p_B is determined using either (3.57) or (3.58).

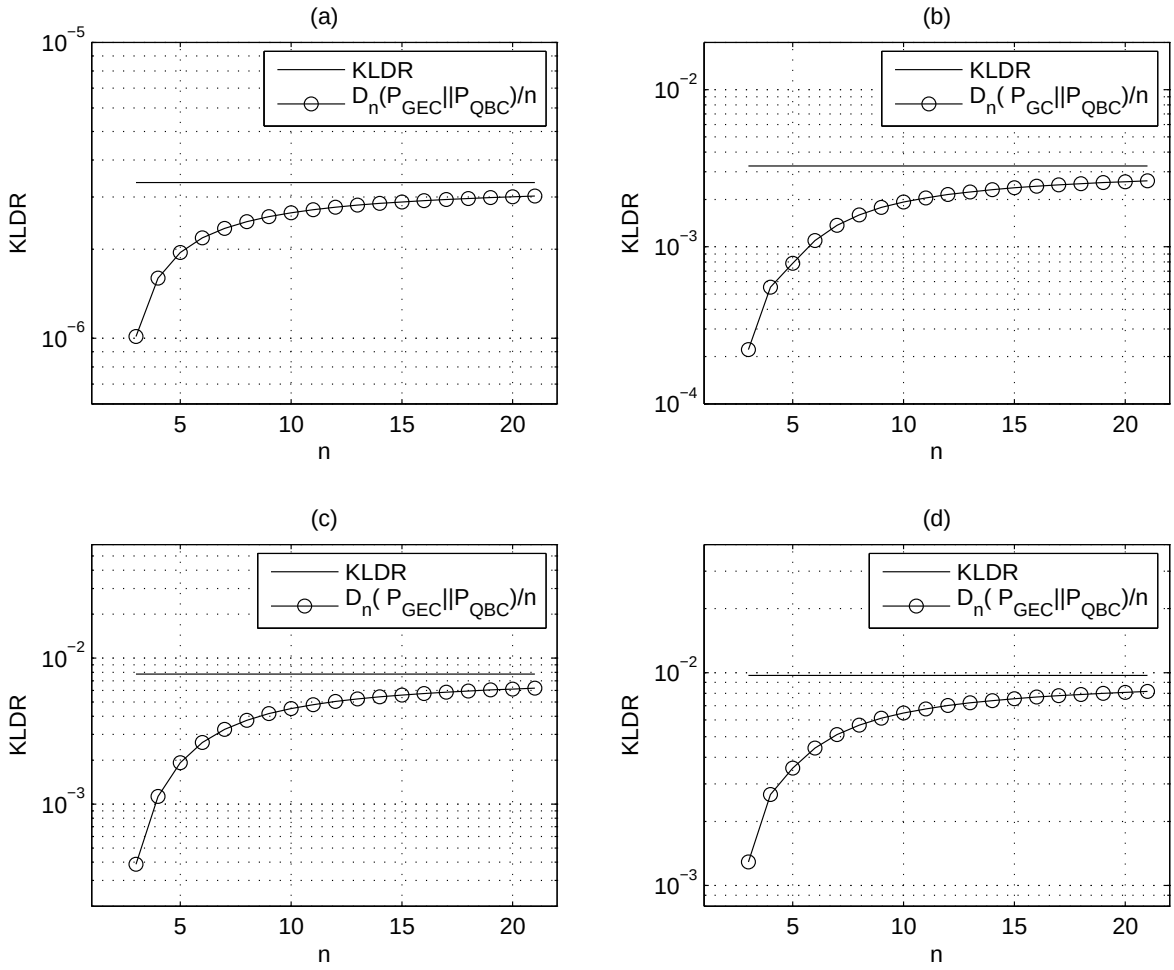


Figure 4.1: GEC fitting via the QBC: KLDR and $(1/n)D_n(\mathrm{P_{GEC}} \parallel \mathrm{P}^{(M)}_{\mathrm{QBC}})$ vs n. (a) Cor = 0.0131 and BER = 0.00314 (Case **A** in Table 4.1); (b) Cor = 0.2227 and BER = 0.03 (Case **B** in Table 4.1); (c) Cor = 0.248 and BER = 0.012 (Case **C** in Table 4.1); (d) Cor = 0.341 and BER = 0.1178 (Case **D** in Table 4.1).

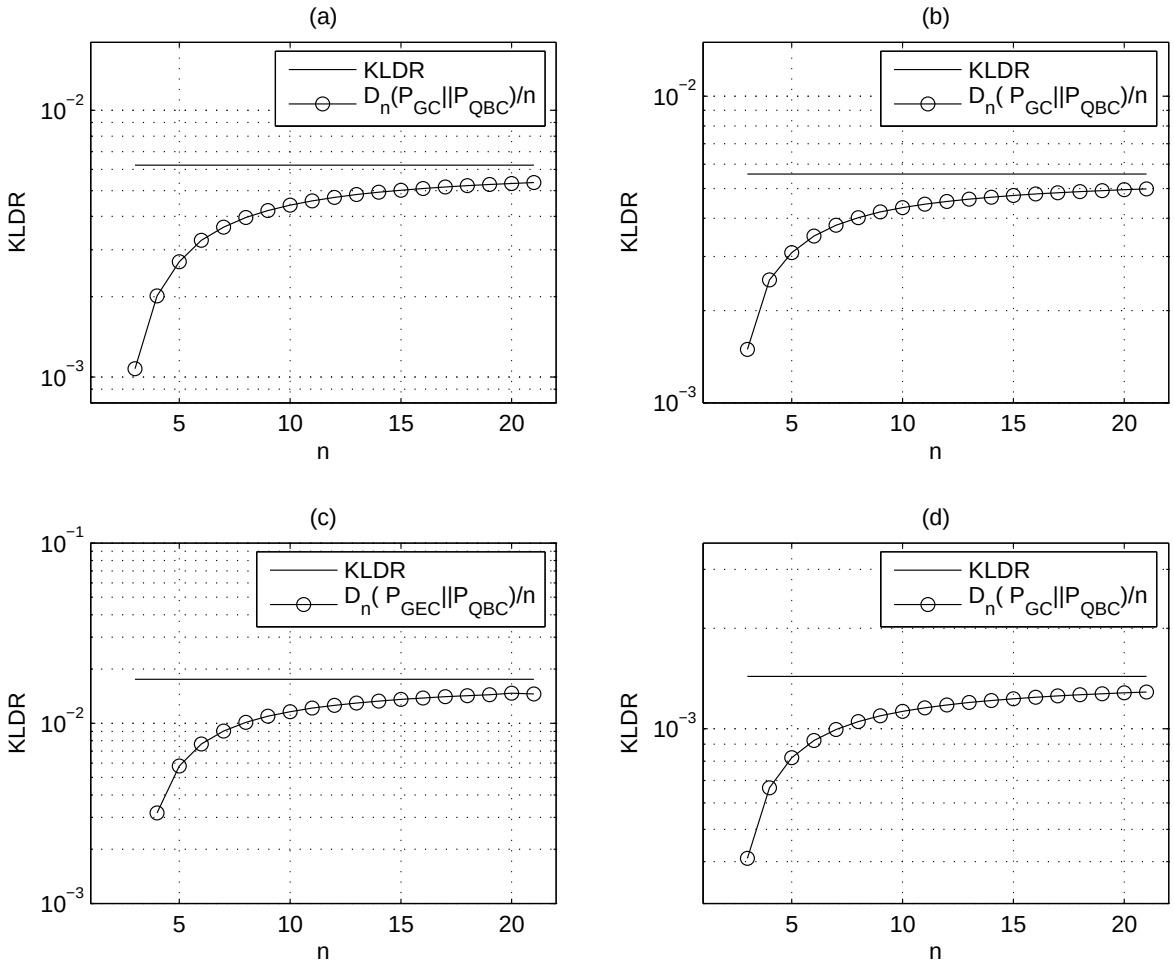


Figure 4.2: GEC fitting via the QBC: KLDR and $(1/n)D_n(\mathrm{P_{GEC}} \parallel \mathrm{P}^{(M)}_{\mathrm{QBC}})$ vs n. (a) Cor = 0.5 and BER = 0.03 (Case **E** in Table 4.1); (b) Cor = 0.7 and BER = 0.03 (Case **F** in Table 4.1); (c) Cor = 0.8 and BER = 0.1 (Case **G** in Table 4.1); (d) Cor = 0.9 and BER = 0.03 (Case **H** in Table 4.1).

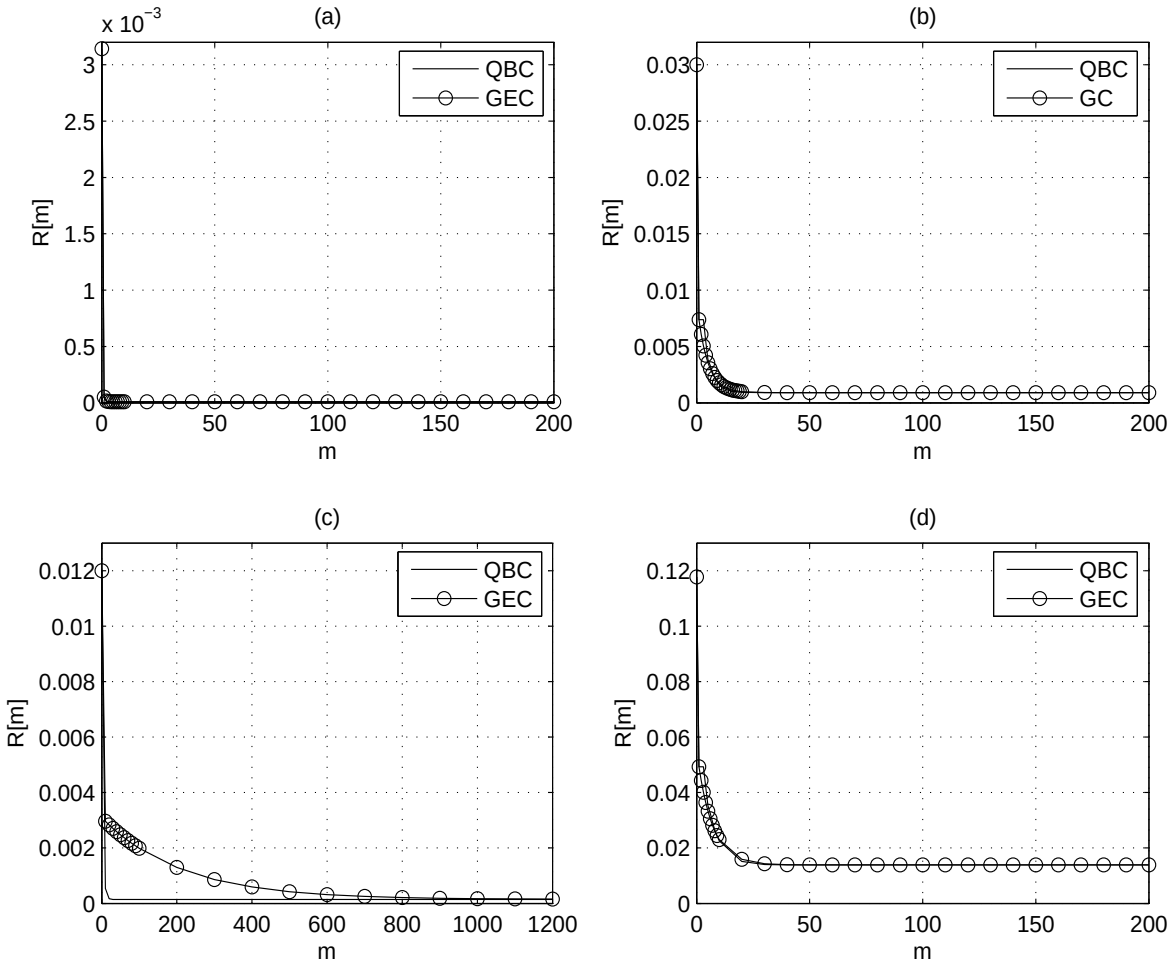


Figure 4.3: GEC fitting via the QBC: ACF vs m. (a) Cor = 0.0131 and BER = 0.00314 (Case **A** in Table 4.1); (b) Cor = 0.2227 and BER = 0.03 (Case **B** in Table 4.1); (c) Cor = 0.248 and BER = 0.012 (Case **C** in Table 4.1); (d) Cor = 0.341 and BER = 0.1178 (Case **D** in Table 4.1).

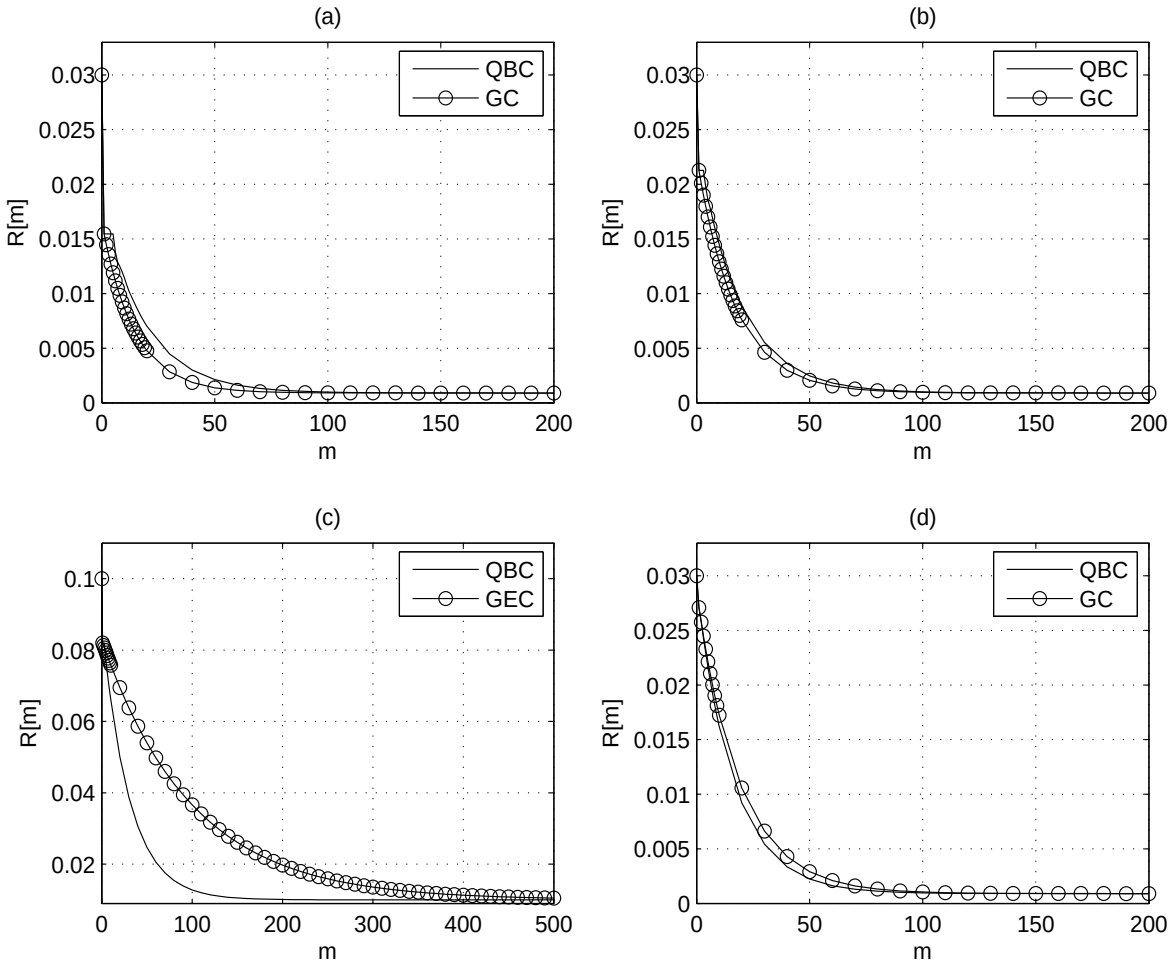


Figure 4.4: GEC fitting via the QBC: ACF vs m. (a) Cor = 0.5 and BER = 0.03 (Case **E** in Table 4.1); (b) Cor = 0.7 and BER = 0.03 (Case **F** in Table 4.1); (c) Cor = 0.8 and BER = 0.1 (Case **G** in Table 4.1); (d) Cor = 0.9 and BER = 0.03 (Case **H** in Table 4.1).

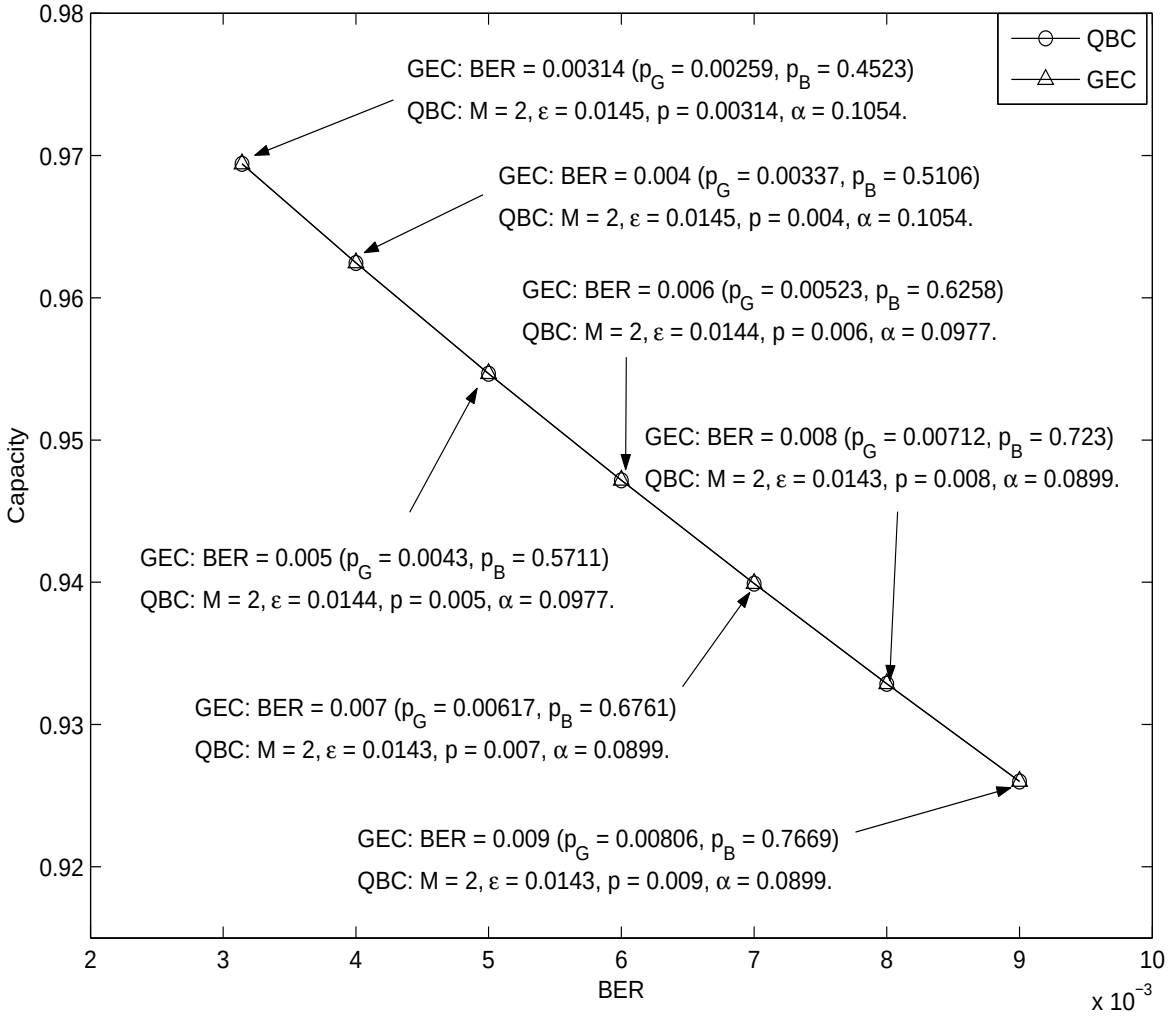


Figure 4.5: GEC fitting via the QBC: Capacity vs BER for Cor = 0.0131. For the GEC, b and g are determined using (3.57) and (3.58).

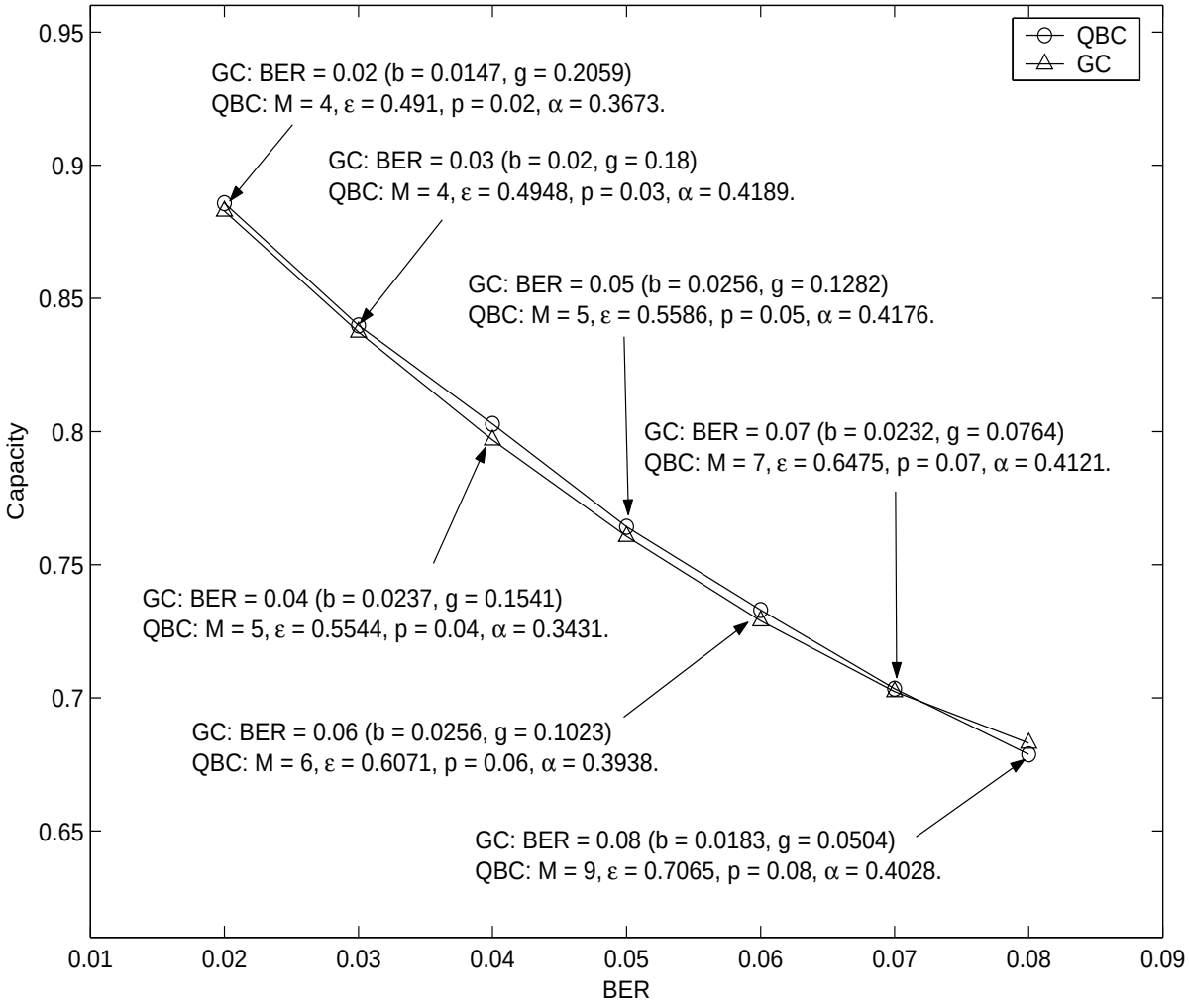


Figure 4.6: GC fitting via the QBC: Capacity vs BER for Cor $= 0.2227$. For the GC, $p_G = 0$ and p_B is determined using either (3.57) or (3.58).

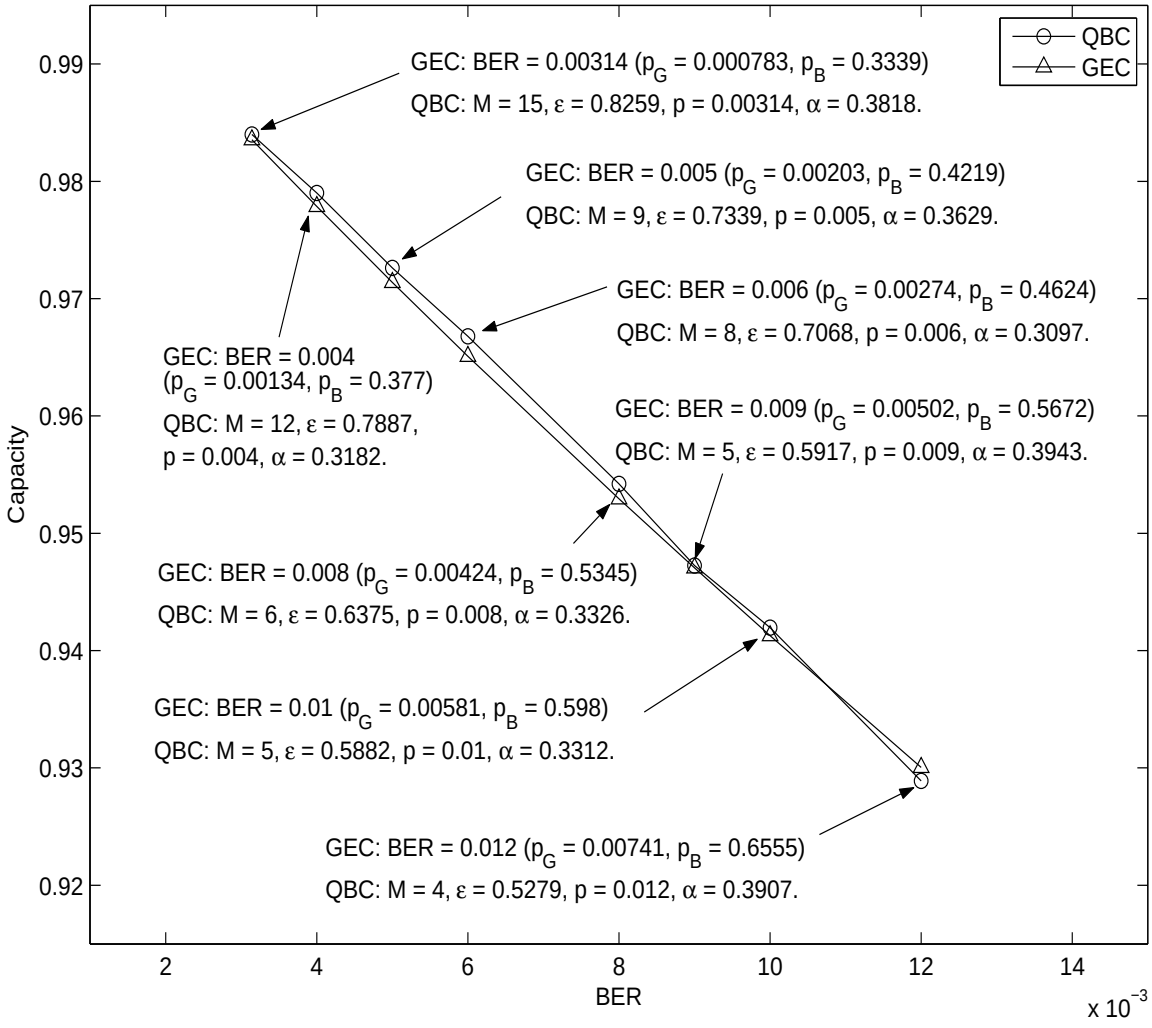


Figure 4.7: GEC fitting via the QBC: Capacity vs BER for Cor = 0.248. For the GEC, b and g are determined using (3.57) and (3.58).

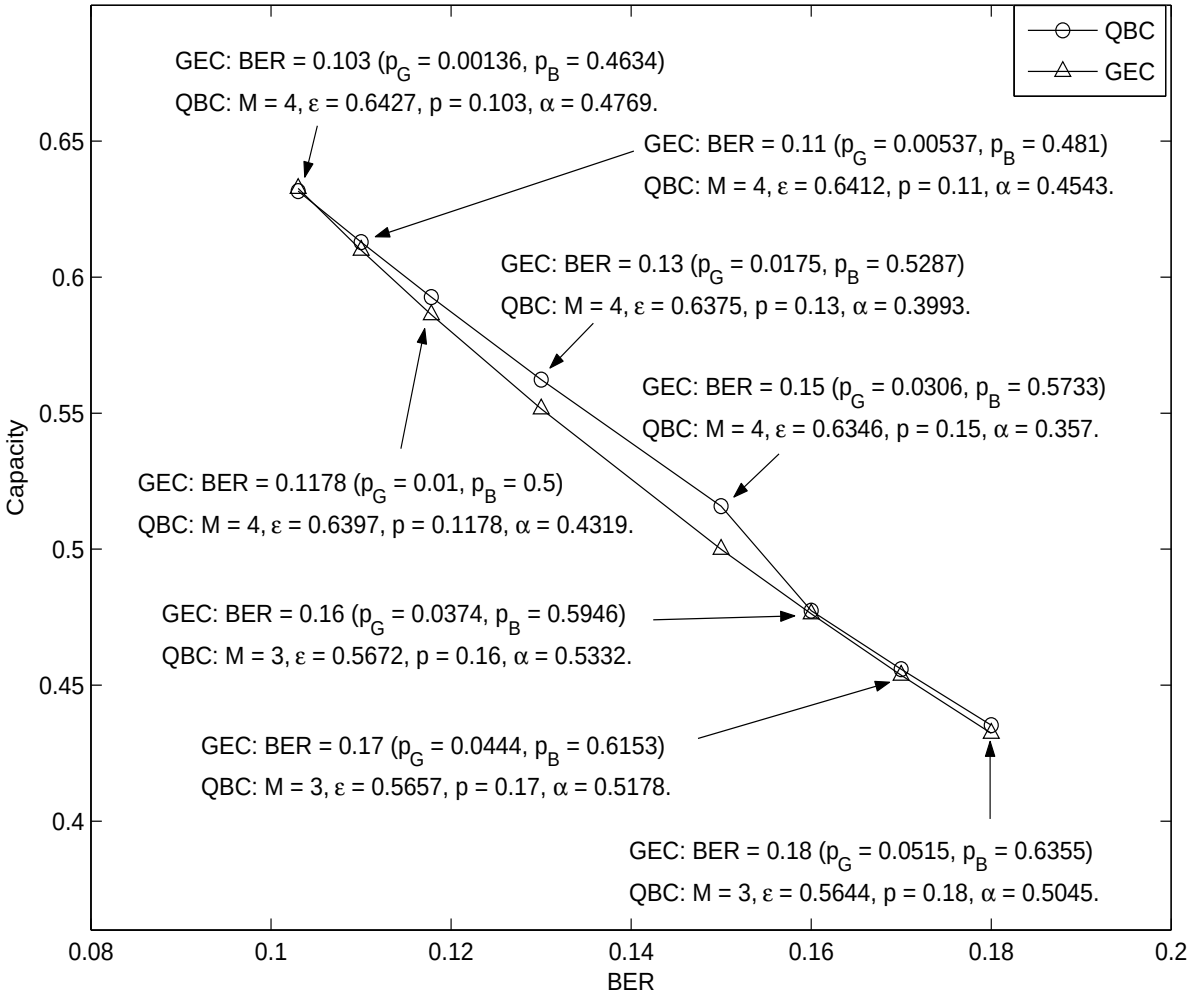


Figure 4.8: GEC fitting via the QBC: Capacity vs BER for Cor = 0.341. For the GEC, b and g are determined using (3.57) and (3.58).

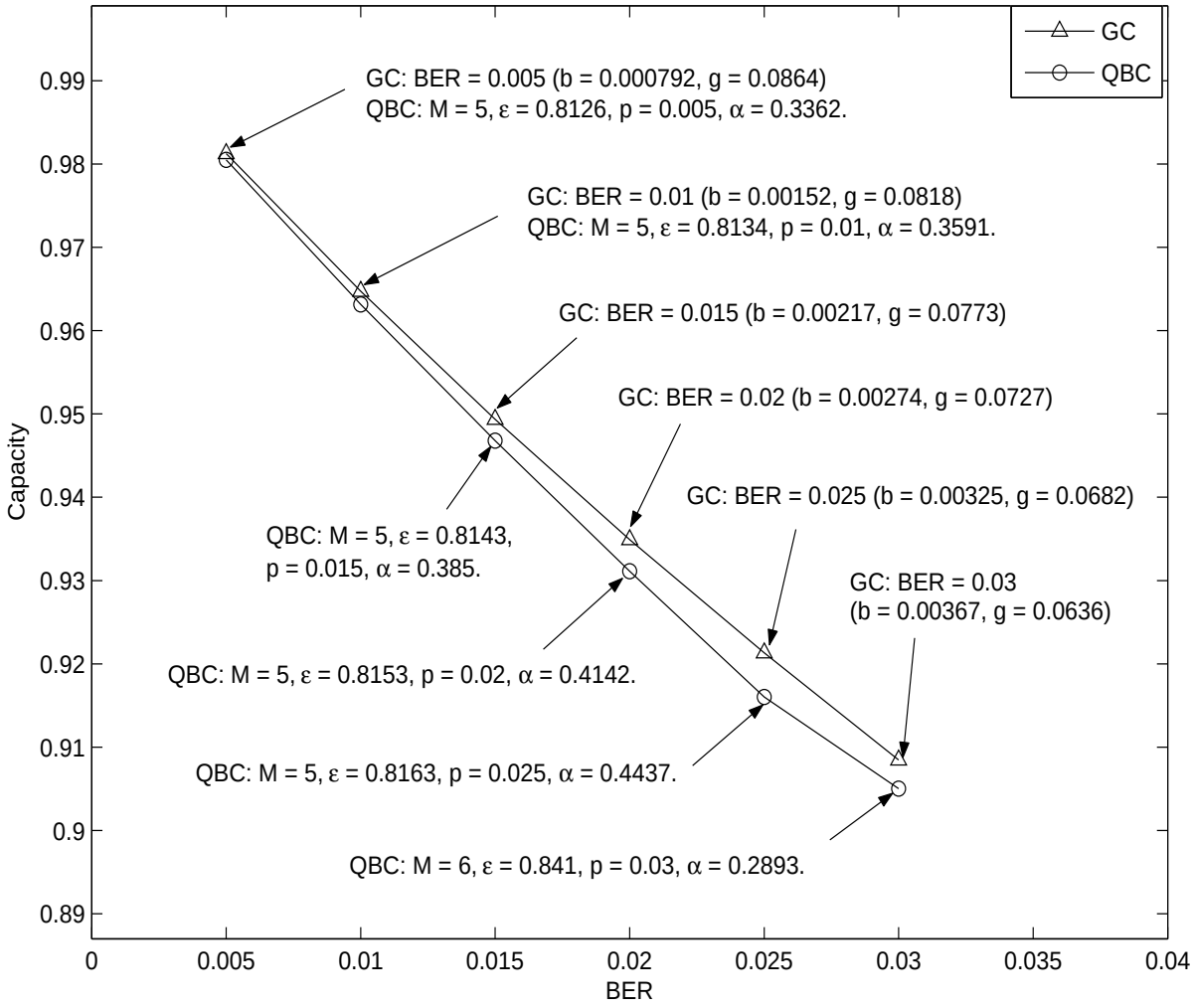


Figure 4.9: GC fitting via the QBC: Capacity vs BER for Cor = 0.5. For the GC, $p_G = 0$ and p_B is determined using either (3.57) or (3.58).

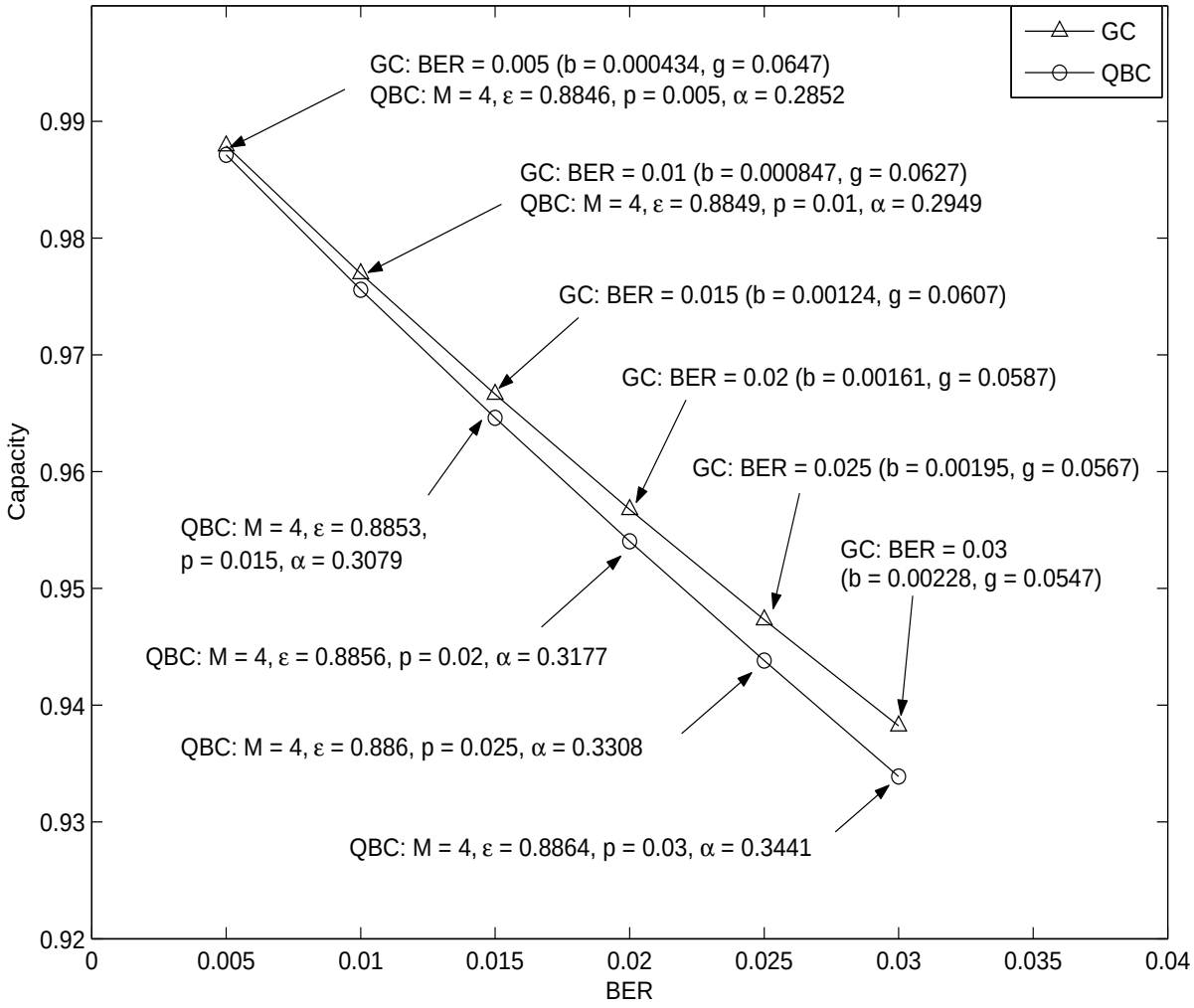


Figure 4.10: GC fitting via the QBC: Capacity vs BER for Cor = 0.7. For the GC, $p_G = 0$ and p_B is determined using either (3.57) or (3.58).

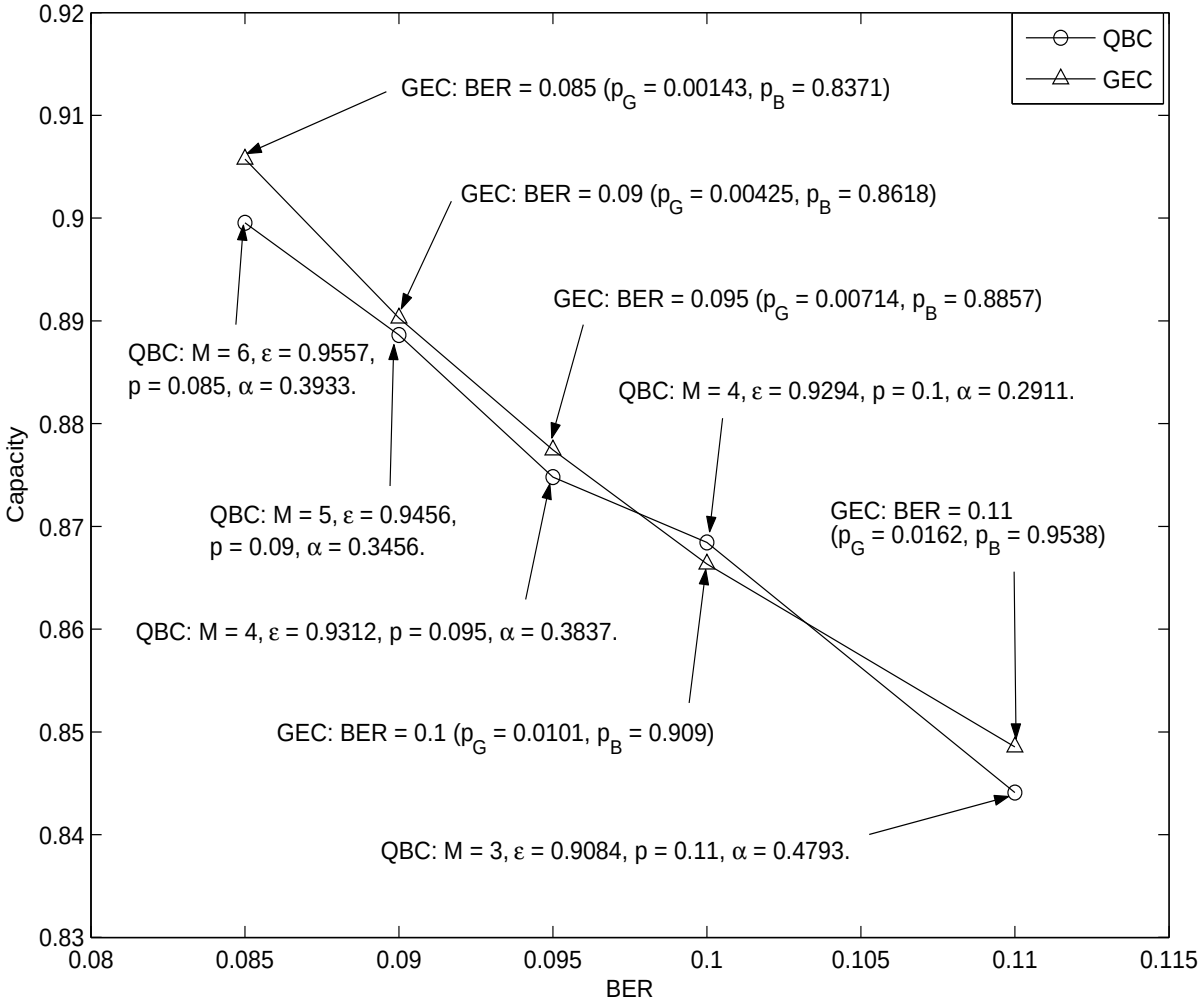


Figure 4.11: GEC fitting via the QBC: Capacity vs BER for Cor = 0.8. For the GEC, b and g are determined using (3.57) and (3.58).

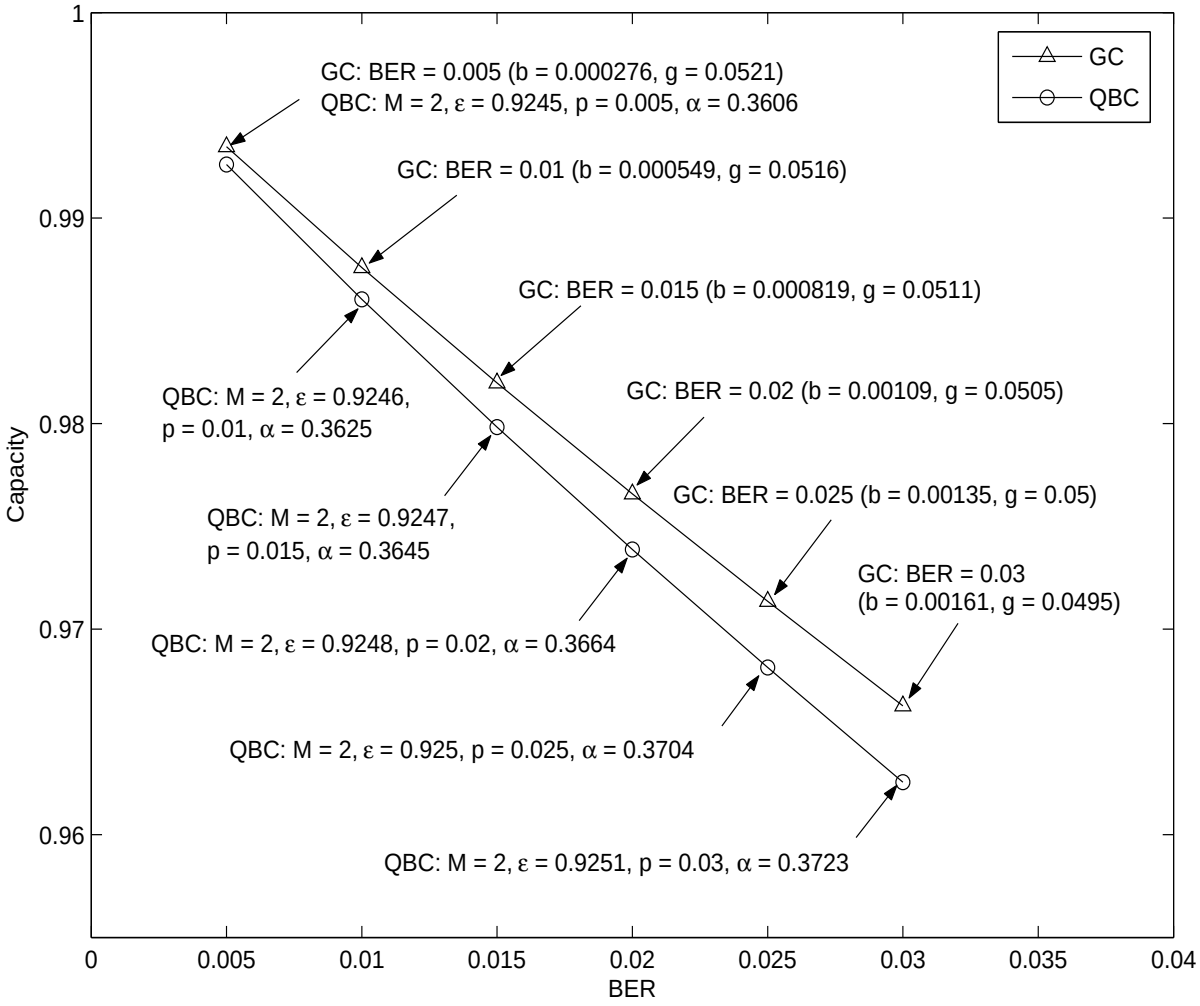


Figure 4.12: GC fitting via the QBC: Capacity vs BER for Cor = 0.9. For the GC, $p_G = 0$ and p_B is determined using either (3.57) or (3.58).

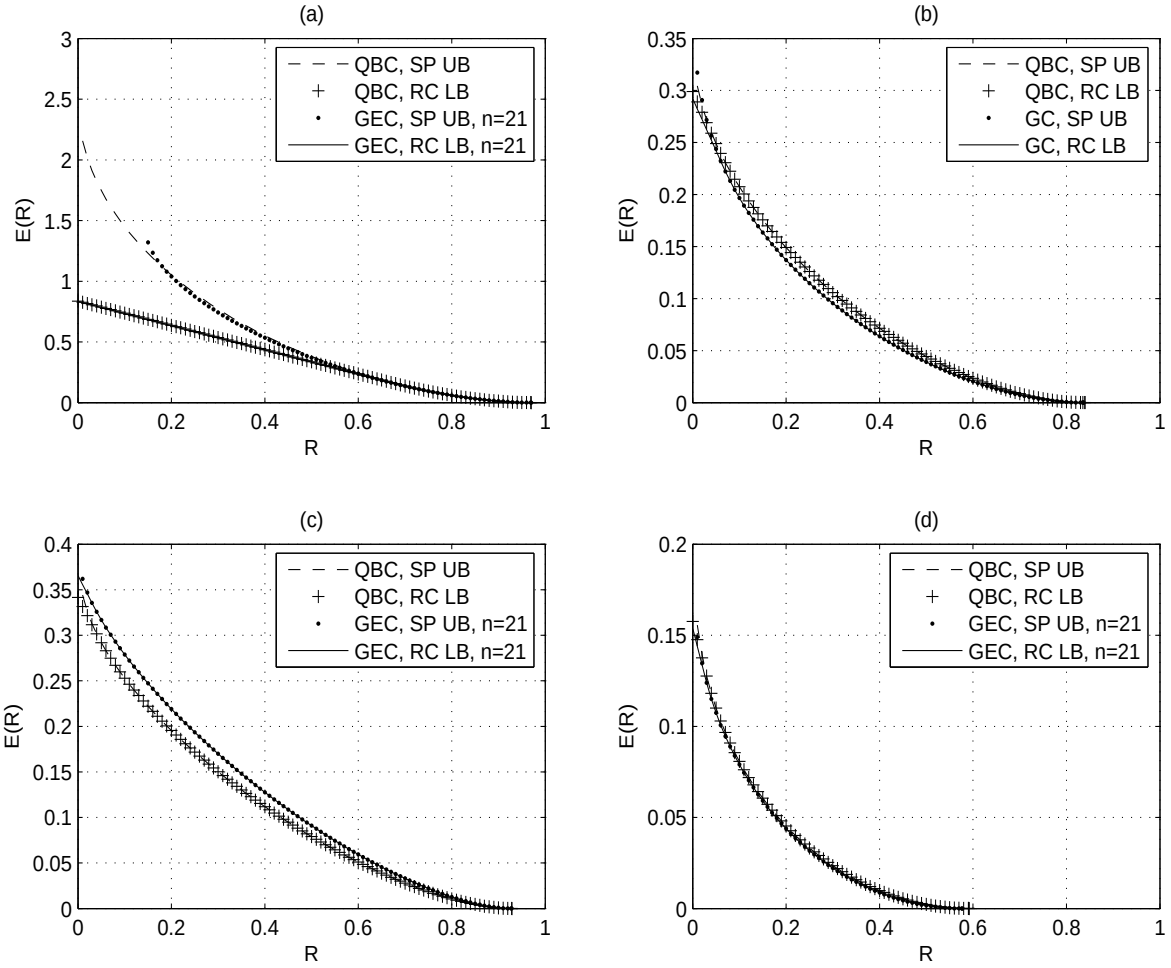


Figure 4.13: GEC fitting via the QBC: Bounds to $E(R)$ vs R. (a) Cor = 0.0131 and BER = 0.00314 (Case **A** in Table 4.1) with $R_{cr}^{(21)} = 0.66$ for the GEC and $R_{cr} = 0.67$ for the QBC; (b) Cor = 0.2227 and BER = 0.03 (Case **B** in Table 4.1) with $R_{cr} = 0.06$ for the GC and $R_{cr} = 0.05$ for the QBC; (c) Cor = 0.248 and BER = 0.012 (Case **C** in Table 4.1) with $R_{cr}^{(21)} = 0.03$ for the GEC and $R_{cr} = 0.04$ for the QBC; (d) Cor = 0.341 and BER = 0.1178 (Case **D** in Table 4.1) with $R_{cr}^{(21)} = 0.02$ for the GEC and $R_{cr} = 0.03$ for the QBC.

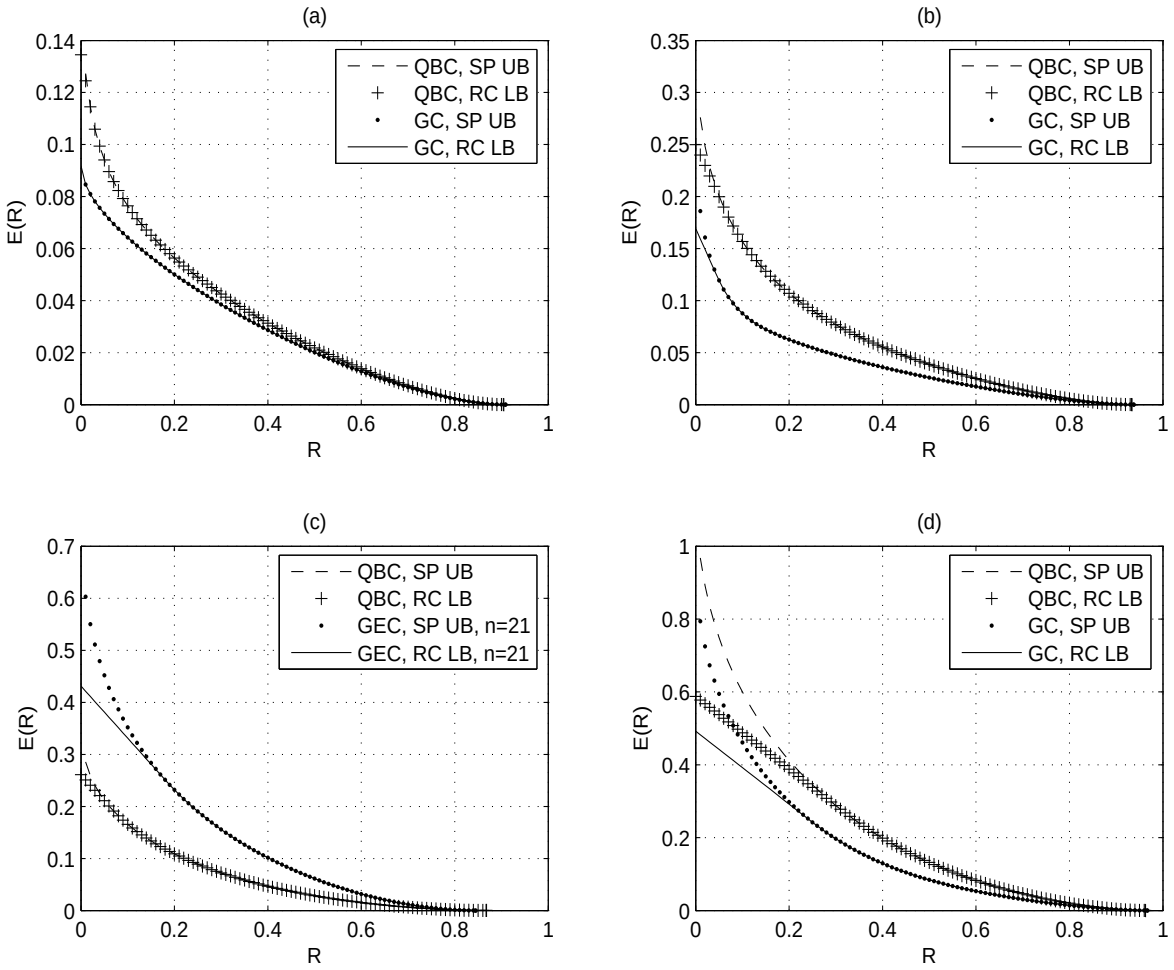


Figure 4.14: GEC fitting via the QBC: Bounds to $E(R)$ vs R. (a) Cor = 0.5 and BER = 0.03 (Case **E** in Table 4.1) with $R_{cr} = 0.0026$ for the GC and $R_{cr} = 0.02$ for the QBC; (b) Cor = 0.7 and BER = 0.03 (Case **F** in Table 4.1) with $R_{cr} = 0.04$ for the GC and $R_{cr} = 0.06$ for the QBC; (c) Cor = 0.8 and BER = 0.1 (Case **G** in Table 4.1) with $R_{cr}^{(21)} = 0.18$ for the GEC and $R_{cr} = 0.06$ for the QBC; (d) Cor = 0.9 and BER = 0.03 (Case **H** in Table 4.1) with $R_{cr} = 0.25$ for the GC and $R_{cr} = 0.31$ for the QBC.

Chapter 5

Fitting Rician Fading Channels via the QBC

We next consider the problem of fitting discretized Rayleigh and Rician fading channels via the QBC model using the KLDR-based approach introduced in the previous chapter. For the sake of comparison, we also model the fading channels via the GEC (which has the same number of parameters as the QBC) using the parameterization method of Pimentel *et. al.* in [44]. The accuracy of both methods is evaluated in terms of ACF and capacity.

5.1 Fading Channel Model

We consider a discrete (binary-input, binary-output) communication system, referred to as the discrete channel with Clarke's autocorrelation (DCCA) model, that employs binary frequency-shift keying modulation, a time-correlated Rician flat-fading channel, and a hard quantized noncoherent demodulation [44]. The complex envelope of the received signal at the input to the demodulator is corrupted by a multiplicative Rician fading and by an additive white Gaussian noise, i.e.,

$$\tilde{R}(t) = \sqrt{2E_s}\tilde{G}(t)\tilde{S}(t) + \tilde{N}(t),$$

where E_s is the symbol energy. $\tilde{S}(t)$ is the complex envelope of the symbol which can be expressed as

$$\tilde{S}(t) = \sum_{k=1}^{\infty} p_{a_k}(t - kT),$$

where the binary information bearing symbols a_k are embedded in the signals $p_i(t)$, $i = 0, 1$, which are equally probable orthogonal signals with unit energy. T is the symbol interval. $\tilde{N}(t)$ is the complex envelope of the white Gaussian noise with autocorrelation function given by [46] $\frac{1}{2}E[\tilde{N}(t+\tau)\tilde{N}^*(t)] = N_0\delta(\tau)$, where N_0 is the variance of $\tilde{N}(t)$. The complex envelope of the fading process $\tilde{G}(t) = \tilde{G}_I(t) + j\tilde{G}_Q(t)$ is a complex, wide sense stationary, Gaussian process with mean η, $j = \sqrt{-1}$, and the quadrature components $\tilde{G}_I(t)$ and $\tilde{G}_Q(t)$ are mutually independent Gaussian processes with the same covariance function $C(\tau)$ which, adopting Clarke's model

([14], [24]), can be expressed as

$$C(\tau) = \frac{1}{2}E[(\tilde{G}(t+\tau) - \eta)(\tilde{G}^*(t) - \eta)] = \sigma_g^2 J_0(2\pi f_D \tau),$$

where $J_0(x) = \sum_{k=0}^{\infty}(-1)^k(\frac{x^k}{2^k k!})^2$ is the zero-order Bessel function of the first kind, f_D is the maximum Doppler frequency experienced by the moving vehicle, and σ_g^2 is the variance of $\tilde{G}(t)$. At each signaling interval of length T, the demodulator forms two decision variables $\{0, 1\}$ and decides which signal was more likely to have been transmitted. A block diagram for visualizing the behavior of the overall system is given in Figure 5.1.

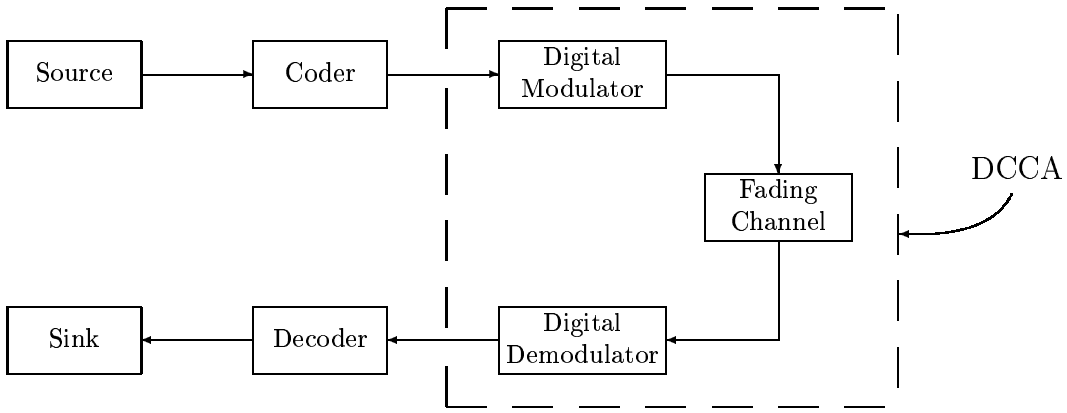


Figure 5.1: Overall system and the equivalent DCCA model.

The combination of digital modulator, fading channel, and digital demodulator yields the equivalent DCCA model. The study and analysis of the statistical behavior of the DCCA model is important since the design and construction of effective error control schemes for this simplified (binary-input, binary-output) model helps us better

exploit the system memory and achieve reliable communication over the underlying correlated fading channel.

The DCCA is represented as an additive noise channel with binary error process $\{Z_n\}_{n=1}^{\infty}$, where

$$Z_n = \begin{cases} 0 & \text{if the } n\text{th transmitted symbol is correctly received,} \\ 1 & \text{if the } n\text{th transmitted symbol is incorrectly received.} \end{cases}$$

The probability of an error sequence of length n, $z^n = (z_1, z_2, \cdots, z_n)$, can be obtained directly from [43] (with $\Omega = 1$).

$$\begin{aligned} &\mathrm{P}_{\mathrm{DCCA}}(z^n) \stackrel{\triangle}{=} \Pr\{Z^n = z^n\} \\ &= \sum_{l_1=z_1}^{1} \cdots \sum_{l_n=z_n}^{1} \left(\prod_{k=1}^{n} \frac{(-1)^{l_k+z_k}}{l_k+1} \right) \times \frac{\exp\{-\frac{E_s}{N_0} K_R \mathbf{1}^T \boldsymbol{F}((K_R+1)\boldsymbol{I} + \frac{E_s}{N_0}\bar{\boldsymbol{C}}\boldsymbol{F})^{-1}\mathbf{1}\}}{\det(\boldsymbol{I} + \frac{E_s}{N_0}(1+K_R)^{-1}\bar{\boldsymbol{C}}\boldsymbol{F})}, \end{aligned} \quad (5.1)$$

where $\boldsymbol{F}$ is a diagonal matrix defined as $\boldsymbol{F} = \mathrm{diag}(\frac{l_1}{l_1+1}, \cdots, \frac{l_n}{l_n+1})$, $\boldsymbol{I}$ is the identity matrix, $K_R = \eta/2\sigma_g^2$ is the Rician factor, and $\bar{\boldsymbol{C}}$ is the normalized covariance matrix with entries $\bar{C}_{ij} = J_0(2\pi f_D T|i-j|)$, $1 \le i, j \le n$.

The QBC is next used to model the equivalent binary error sequence of the DCCA, which represents the successes and failures that result from the transmission of symbols over the above fading channel.

5.2 Parameter Estimation

5.2.1 QBC Parameter Estimation

For a given DCCA, we construct a QBC whose noise process is statistically "close" in the Kullback-Leibler sense to the noise process generated by the DCCA. Specifically, given a DCCA with fixed average signal-to-noise ratio (SNR) E_s/N_0, normalized Doppler frequency $f_D T$ and Rician factor K_R resulting in bit error rate $\mathrm{BER}_{\mathrm{DCCA}}$ and correlation coefficient $\mathrm{Cor}_{\mathrm{DCCA}}$, we estimate the QBC parameters M, p, ε, and α that minimize the KLDR, $\lim_{n\to\infty}(1/n)D_n(\mathrm{P}_{\mathrm{DCCA}} \parallel \mathrm{P}_{\mathrm{QBC}}^{(M)})$, subject to the constraints

$$\mathrm{BER}_{\mathrm{QBC}} = \mathrm{BER}_{\mathrm{DCCA}} \text{ and } \mathrm{Cor}_{\mathrm{QBC}} = \mathrm{Cor}_{\mathrm{DCCA}},$$

where $(1/n)D_n(\mathrm{P}_{\mathrm{DCCA}} \parallel \mathrm{P}_{\mathrm{QBC}}^{(M)})$ is the NKLD between the n-fold DCCA and QBC noise distributions, where $\mathrm{P}_{\mathrm{QBC}}^{(M)}$ is given in closed form by (3.7) and (3.8) and $\mathrm{P}_{\mathrm{DCCA}}$ is given by (5.1).

As noted in the previous chapter, the KLDR between a stationary source and a Markov source does exist and can be expressed as

$$\lim_{n\to\infty}\frac{1}{n}D_n(\mathrm{P}_{\mathrm{DCCA}} \parallel \mathrm{P}_{\mathrm{QBC}}^{(M)}) = -\mathcal{H}_{\mathrm{DCCA}}(Z) - E_{\mathrm{P}_{\mathrm{DCCA}}}[\log_2 \mathrm{P}_{\mathrm{QBC}}^{(M)}(Z_{M+1}|Z^M)], \qquad (5.2)$$

where $\mathcal{H}(\cdot)$ denotes the entropy rate and $\mathrm{P}_{\mathrm{QBC}}^{(M)}(z_{M+1}|z^M)$ is the QBC conditional error probability of symbol $M+1$ given the previous M symbols. Then the minimization reduces to maximizing the second term in (5.2) (which is independent of n) over

the QBC parameters. Note that in our approximation, we match BER and Cor of both channels to guarantee identical noise marginal distributions and identical probabilities of two consecutive errors (ones). Hence, given these constraints, the above optimization problem reduces to an optimization over only two QBC parameters.

5.2.2 GEC Parameter Estimation

We next briefly describe the modeling method of the DCCA via the GEC introduced by Pimentel *et. al.* in [44]. For a given DCCA, the parameterization of the GEC is based on the following lemma.

Lemma 5.1 **[44]** *The probability of any observed sequence z^n generated by the* GEC *satisfies the following recurrence equation:*

$$\mathrm{P}_{\mathrm{GEC}}(z^n\varsigma\kappa) = c(\varsigma,\kappa)\mathrm{P}_{\mathrm{GEC}}(z^n\varsigma) + d(\varsigma,\kappa)\mathrm{P}_{\mathrm{GEC}}(z^n), \tag{5.3}$$

where $\mathrm{P}_{\mathrm{GEC}}(z^n\varsigma\kappa) \stackrel{\triangle}{=} \Pr\{Z_1 = z_1, \cdots, Z_n = z_n, Z_{n+1} = \varsigma, Z_{n+2} = \kappa\}$, *$\varsigma$ and κ are binary symbols,*

$$c(0,0) = (1-p_G)(1-b) + (1-p_B)(1-g), \;\; c(1,1) = p_G(1-b) + p_B(1-g), \tag{5.4}$$

$$d(0,0) = -(1-g-b)(1-p_G)(1-p_B), \;\; d(1,1) = -(1-g-b)p_G)p_B, \tag{5.5}$$

$c(1,0) = 1 - c(1,1)$, $c(0,1) = 1 - c(0,0)$, $d(0,1) = -d(0,0)$, and $d(1,0) = -d(1,1)$.

Lemma 5.1 shows that the parameters $c(\varsigma,\kappa)$ and $d(\varsigma,\kappa)$ can be calculated via a linear system of equations. For example, setting $z^n = \phi$, where ϕ is an empty sequence, $\mathrm{P}_{\mathrm{GEC}}(\phi) = 1$, and $z^n = \varsigma$ in (5.3), $c(\varsigma,\kappa)$ and $d(\varsigma,\kappa)$ can be determined by the probabilities of error sequences of length, at most 3:

$$c(\varsigma,\kappa) = \frac{\mathrm{P}_{\mathrm{GEC}}(\varsigma\varsigma\kappa) - \mathrm{P}_{\mathrm{GEC}}(\varsigma\kappa)\mathrm{P}_{\mathrm{GEC}}(\varsigma)}{\mathrm{P}_{\mathrm{GEC}}(\varsigma\varsigma) - \mathrm{P}^2_{\mathrm{GEC}}(\varsigma)}, \tag{5.6}$$

and

$$d(\varsigma,\kappa) = \frac{\mathrm{P}_{\mathrm{GEC}}(\varsigma\kappa)\mathrm{P}_{\mathrm{GEC}}(\varsigma\varsigma) - \mathrm{P}_{\mathrm{GEC}}(\varsigma\varsigma\kappa)\mathrm{P}_{\mathrm{GEC}}(\varsigma)}{\mathrm{P}_{\mathrm{GEC}}(\varsigma\varsigma) - \mathrm{P}^2_{\mathrm{GEC}}(\varsigma)}. \tag{5.7}$$

The GEC parameters follow by solving the nonlinear equations in (5.4)-(5.5) as follows.

Proposition 5.1 [**44**] *If* $\mathrm{P}_{\mathrm{GEC}}(01) \neq \mathrm{P}_{\mathrm{GEC}}(0)\mathrm{P}_{\mathrm{GEC}}(1)$, *the parameters of the* GEC *are uniquely determined by the four probabilities* $\mathrm{P}_{\mathrm{GEC}}(0)$, $\mathrm{P}_{\mathrm{GEC}}(00)$, $\mathrm{P}_{\mathrm{GEC}}(000)$ *and* $\mathrm{P}_{\mathrm{GEC}}(111)$. *The four parameters b, g,* p_G, *and* p_B *are given by the following.*

p_G *and* p_B *are the roots of the quadratic equation*

$$[-1 + c(1,1) + c(0,0)]x^2 + [1 - c(1,1) - c(0,0) + d(1,1) - d(0,0)]x - d(1,1) = 0,$$

and

$$b = \frac{c(0,0)p_B - c(1,1)(1-p_B) + (p_G - p_B)}{p_G - p_B},$$

$$g = \frac{c(0,0)p_G - c(1,1)(1-p_G) + (p_B - p_G)}{p_B - p_G},$$

Hence, if $\mathrm{P}_{\mathrm{DCCA}}(0)$, $\mathrm{P}_{\mathrm{DCCA}}(00)$, $\mathrm{P}_{\mathrm{DCCA}}(000)$ and $\mathrm{P}_{\mathrm{DCCA}}(111)$ are known, where $\mathrm{P}_{\mathrm{DCCA}}(z^n)$ is the probability of error sequences generated by the DCCA (see (5.1)), the parameters of the GEC can be obtained by (5.6), (5.7) and Proposition 5.1 by setting $\mathrm{P}_{\mathrm{GEC}}(z^n) = \mathrm{P}_{\mathrm{DCCA}}(z^n)$, $n = 1, 2, 3$.

5.3 Modeling Results and Discussions

We evaluate how well the QBC model fits or approximates the DCCA according to two criteria: channel capacity and ACF. The QBC ACF and capacity expressions are provided in Sections 3.1.3 and 3.1.4. The ACF of the DCCA can be obtained directly from (5.1):

$$R[m] = \frac{(1+K_R)^2}{\left(2+2K_R+\frac{E_s}{N_0}\right)^2 - \left(\frac{E_s}{N_0}\rho(m)\right)^2} \times \exp\left\{-\frac{2K_R\frac{E_s}{N_0}}{2+2K_R+\frac{E_s}{N_0}(\rho(m)+1)}\right\},$$

where $\rho(m) = J_0(2\pi m f_D T)$.

Although the capacity of the DCCA does not have a closed-form expression, it can be shown that the capacity of the DCCA has the following expression (see (2.1)).

$$C = \lim_{n\to\infty} \sup_{X^n} \frac{1}{n} I(X^n; Y^n) = 1 - \lim_{n\to\infty} \frac{1}{n} H(Z^n).$$

Defining $C_n \stackrel{\triangle}{=} 1 - (1/n)H(Z^n)$, we obtain a capacity lower bound which is computable for finite n.

For the sake of comparison, we also present modeling results via the GEC using the method of Pimentel *et. al.* in [44] (which we briefly described in Section 5.2.2).

Note that in [44], the authors also employ arbitrary Kth order Markov noise models to approximate the fading channels. However, unlike our QBC model which has only four parameters (as the GEC) and allows large values for its memory order M (since its noise is a special Mth order Markov process generated by our queue scheme), the Kth order Markov models of [44] suffer from the limitation of having a number of parameters that grows exponentially[1] with K. Therefore, we herein only compare our QBC-based modeling method with the GEC-based modeling method of [44] since both channels have identical number of parameters, hence identical degrees of freedom and complexity.

The capacity of the GEC is obtained via the algorithm in Section 3.3.2. The ACF of the GEC is provided in (4.1).

A wide range of DCCA channel parameters is investigated with SNR = 15 dB and 25 dB, $f_DT = 0.001$, 0.005, 0.01 and 0.1 for Rayleigh fading ($K_R = -\infty$ dB), and SNR = 15 dB and $f_DT = 0.001$, 0.005, 0.01 and 0.05 for Rician fading ($K_R = 5$ dB). The SNR, f_DT and K_R values (except for $f_DT = 0.005$) were chosen to match the conditions of the correlated Rician and Rayleigh fading channels studied in [44, Fig. 6,

[1]As a result, only models with memory order up to 6 are studied in [44]. Such models are shown to well approximate channels with fast and medium fading rates ($f_DT > 0.02$); but they are inadequate for slow fading rates. As we later show in this section, the QBC model can accommodate large values of the memory order; thus, it can provide an accurate approximation of channels with slow fading ($f_DT < 0.02$) in addition to medium and fast fading.

Fig. 7, Fig. 9 and Fig. 11]. The QBC and GEC parameters, obtained as explained in Sections 5.2.1 and 5.2.2, respectively, are provided in Tables 5.1-5.3.

Modeling results in terms of the ACF for the DCCA, its QBC approximation and its GEC approximation are shown in Figs. 5.2-5.7. We observe a strong agreement between the QBC and the DCCA in ACF in these figures (this behavior was indeed observed for all computations, especially for $f_D T = 0.1$, (b) in Figs. 5.3 and 5.5, where the ACF curves of the DCCA and its QBC approximation are identical). For slow and medium fading (Fig. 5.2, (a) of Fig. 5.4, Fig. 5.6 and (a) of Fig. 5.7), the ACF curve for the GEC takes a longer span of m before eventually converging, which indicates that the GEC model is infeasible for fitting very slow Rayleigh fading ($f_D T = 0.001$) and for fitting very slow to medium Rician fading ($f_D T = 0.001$, 0.005 and 0.01). We observe that the QBC has better performance than the Markov models in [44] (see (a) in Figs. 5.2), but with significantly smaller complexity since it is fully described by four parameters and allows us closed-form expressions for various fading characteristics. Compared with [44, (a) in Figs. 7], the QBC has similar performance as the Markov model of order 4 or 5, but with smaller complexity.

Note that since the QBC noise is a homogeneous Markov process, the KLDR between the DCCA and the QBC error processes exists and admits a simple expression. Hence, it is practical to minimize this KLDR by maximizing the expected value in (5.2) over the QBC parameters which is independent of n (see Section 5.2.1). How-

ever, this approach is not applicable to the GEC since the KLDR between the DCCA and the GEC noise processes does not admit a simple expression in general (as the GEC noise is hidden Markovian). The method of parameterization of the GEC of Section 5.2.2 is simple, but it only takes into account error sequences no longer than 3, which implies that this method is not appropriate for approximating slow fading.

Numerical results show that the largest Markovian memory M for the QBC model that best fits the DCCA is 20, while the largest Markovian memory K for the general Markov noise channel model considered in [44] is 6 because of its complexity. This explains why the QBC is more suitable for fitting slow fading with large memory than the Markov noise model considered in [44].

Modeling results in terms of capacity are shown in Figs. 5.8-5.9, where the lower bound for the capacity of the DCCA and the capacities of the QBC approximation and the GEC approximation are shown for different SNR values and f_DT values. We clearly observe from the figures that the capacity curves of the QBC and the lower bound curves for the capacity of the DCCA match quite well, and the capacities for $f_DT = 0.1$ (fast Rayleigh fading) are almost identical. The last observation can be explained by the fact that the DCCA has low memory at $f_DT = 0.1$ (fast fading); hence the lower bound for its capacity is tight (since $(1/n)H(Z^n) = H(Z_1)$ if Z^n is memoryless). Overall, we observe a strong match in capacity between the DCCA and its QBC approximation. In terms of capacity, the GEC has nearly as good a

performance as the QBC in fitting the DCCA.

Model	$f_D T$			
	0.001	0.005	0.01	0.1
QBC	$M = 20$ $\varepsilon = 0.8593$ $p = 0.0297$ $\alpha = 0.8959$	$M = 11$ $\varepsilon = 0.7602$ $p = 0.0297$ $\alpha = 0.3828$	$M = 7$ $\varepsilon = 0.6556$ $p = 0.0297$ $\alpha = 0.3387$	$M = 2$ $\varepsilon = 0.0893$ $p = 0.0297$ $\alpha = 0.131$
GEC	$b = 0.0000339$ $g = 0.000479$ $p_B = 0.3393$ $p_G = 0.00783$	$b = 0.000841$ $g = 0.0118$ $p_B = 0.3393$ $p_G = 0.00766$	$b = 0.00329$ $g = 0.045$ $p_B = 0.3395$ $p_G = 0.00713$	$b = 0.0324$ $g = 0.7466$ $p_B = 0.5199$ $p_G = 0.00849$

Table 5.1: QBC and GEC modeling parameters for $K_R = -\infty$ dB (Rayleigh) and SNR = 15 dB.

Model	f_DT			
	0.001	0.005	0.01	0.1
QBC	$M = 18$ $\varepsilon = 0.8506$ $p = 0.00314$ $\alpha = 0.2607$	$M = 6$ $\varepsilon = 0.6226$ $p = 0.00314$ $\alpha = 0.2525$	$M = 4$ $\varepsilon = 0.4666$ $p = 0.00314$ $\alpha = 0.2019$	$M = 2$ $\varepsilon = 0.0145$ $p = 0.00314$ $\alpha = 0.1054$
GEC	$b = 0.0000333$ $g = 0.00466$ $p_B = 0.3339$ $p_G = 0.000783$	$b = 0.000773$ $g = 0.1014$ $p_B = 0.334$ $p_G = 0.000622$	$b = 0.0025$ $g = 0.2887$ $p_B = 0.3343$ $p_G = 0.000279$	$b = 0.00103$ $g = 0.8338$ $p_B = 0.4523$ $p_G = 0.00259$

Table 5.2: QBC and GEC modeling parameters for $K_R = -\infty$ dB (Rayleigh) and SNR = 25 dB.

Model	f_DT			
	0.001	0.005	0.01	0.05
QBC	$M = 18$ $\varepsilon = 0.8195$ $p = 0.00853$ $\alpha = 0.9619$	$M = 17$ $\varepsilon = 0.8054$ $p = 0.00853$ $\alpha = 0.3971$	$M = 11$ $\varepsilon = 0.7219$ $p = 0.00853$ $\alpha = 0.3299$	$M = 3$ $\varepsilon = 0.3426$ $p = 0.00853$ $\alpha = 0.3726$
GEC	$b = 0.00000259$ $g = 0.000139$ $p_B = 0.3112$ $p_G = 0.00289$	$b = 0.0000646$ $g = 0.00347$ $p_B = 0.3113$ $p_G = 0.00288$	$b = 0.000257$ $g = 0.0137$ $p_B = 0.3115$ $p_G = 0.00284$	$b = 0.00542$ $g = 0.2533$ $p_B = 0.3185$ $p_G = 0.0019$

Table 5.3: QBC and GEC modeling parameters for $K_R = 5$ dB (Rician) and SNR = 15 dB.

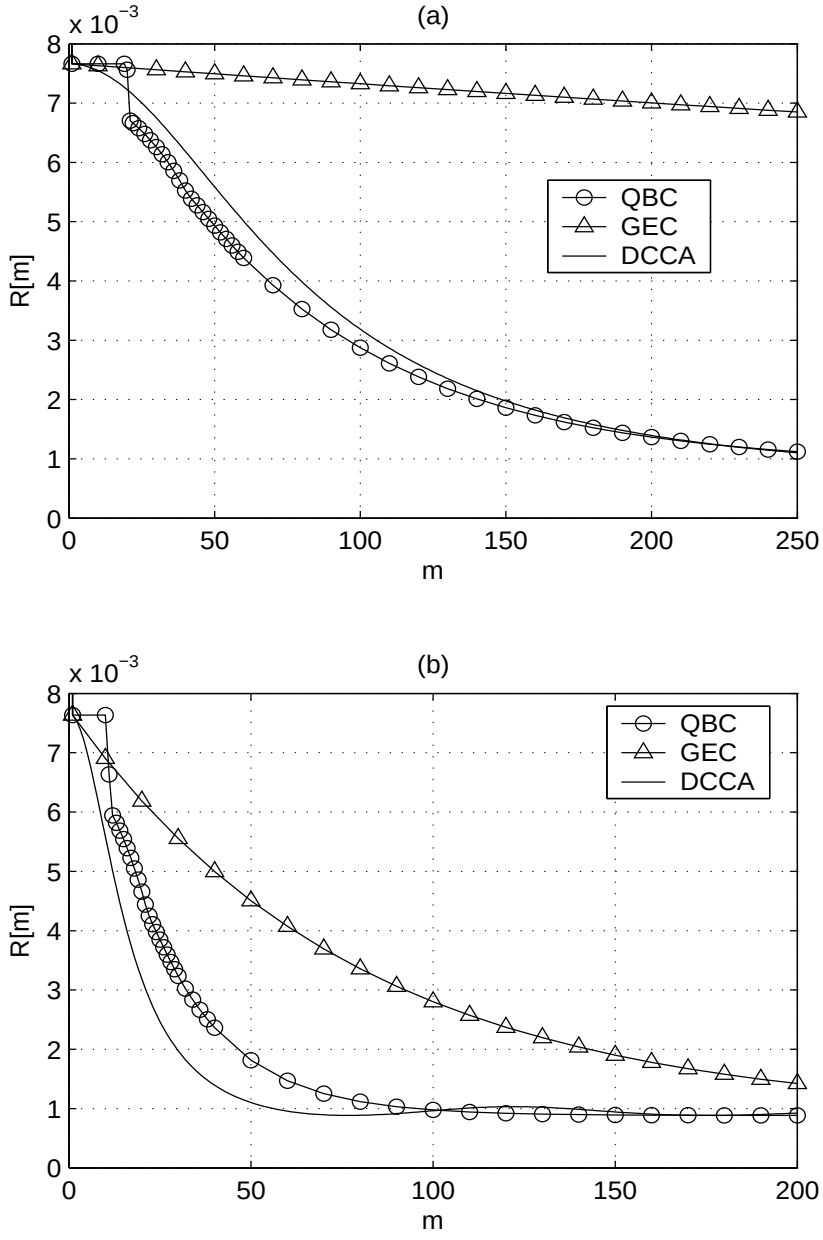


Figure 5.2: DCCA fitting via the QBC: ACF vs m for $K_R = -\infty$ dB (Rayleigh) and SNR = 15 dB. (a) $f_D T = 0.001$; (b) $f_D T = 0.005$.

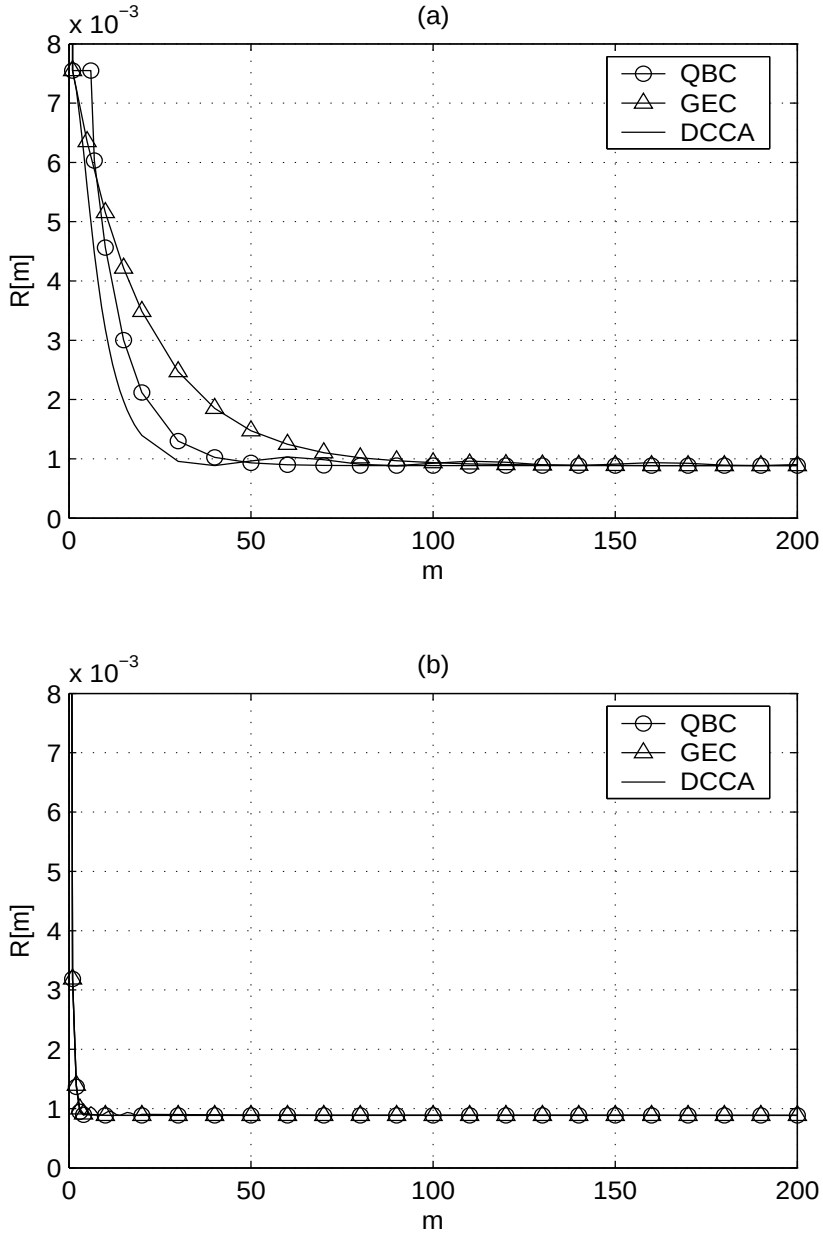


Figure 5.3: DCCA fitting via the QBC: ACF vs m for $K_R = -\infty$ dB (Rayleigh) and SNR = 15 dB. (a) $f_D T = 0.01$; (b) $f_D T = 0.1$.

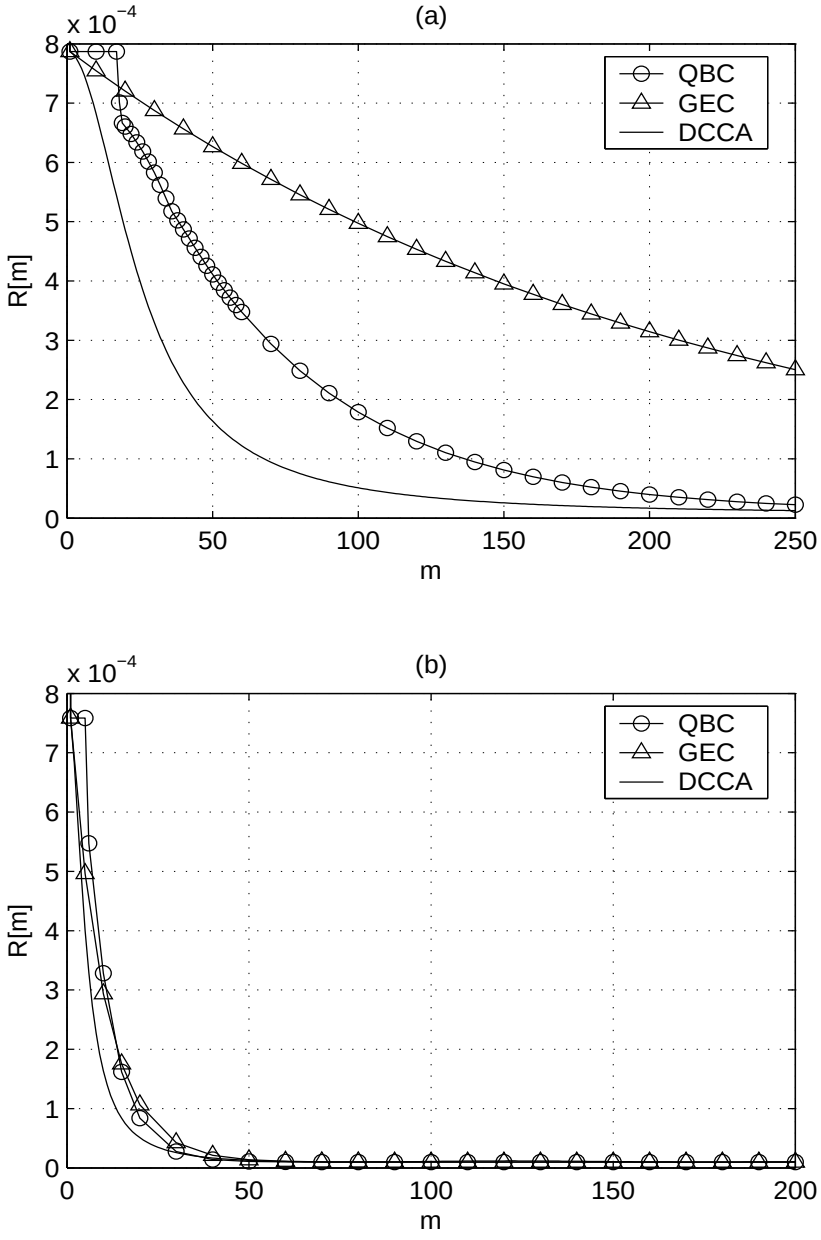


Figure 5.4: DCCA fitting via the QBC: ACF vs m for $K_R = -\infty$ dB (Rayleigh) and SNR = 25 dB. (a) $f_D T = 0.001$; (b) $f_D T = 0.005$.

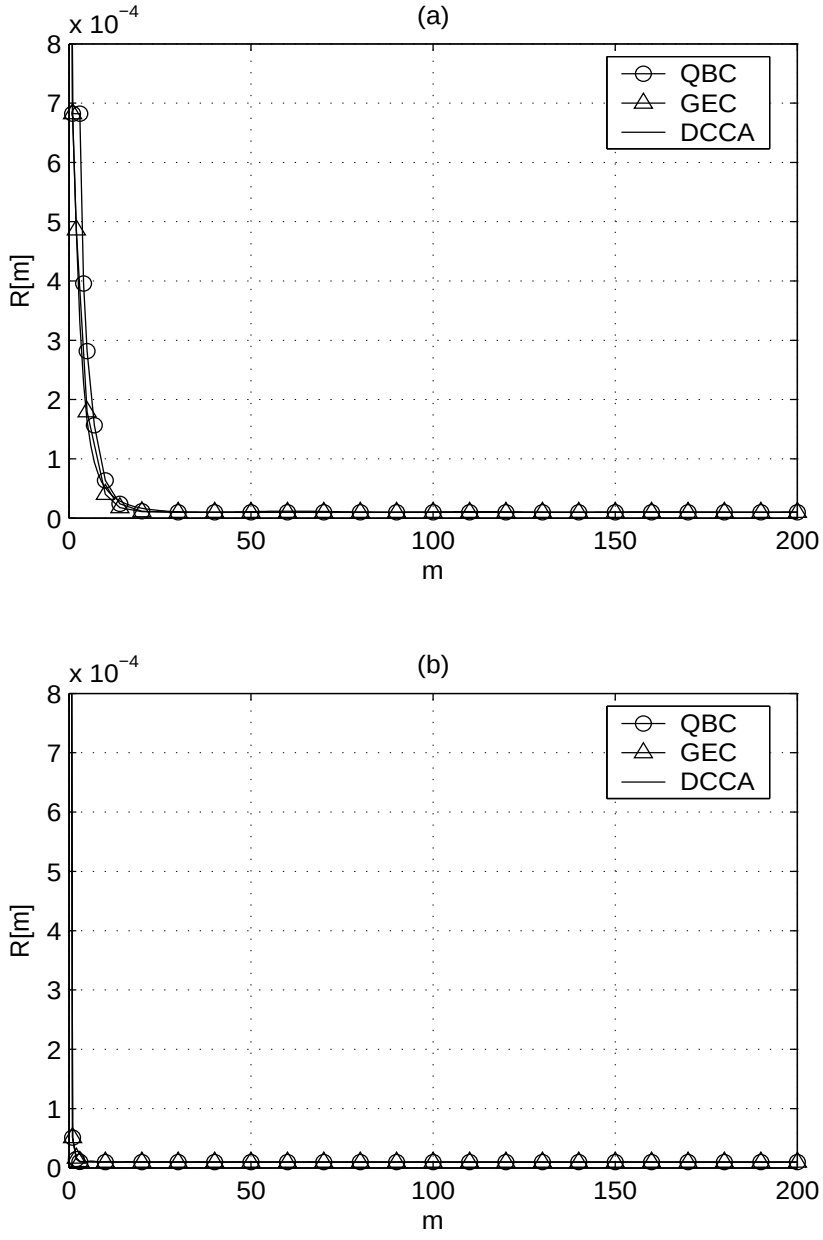


Figure 5.5: DCCA fitting via the QBC: ACF vs m for $K_R = -\infty$ dB (Rayleigh) and SNR = 25 dB. (a) $f_D T = 0.01$; (b) $f_D T = 0.1$.

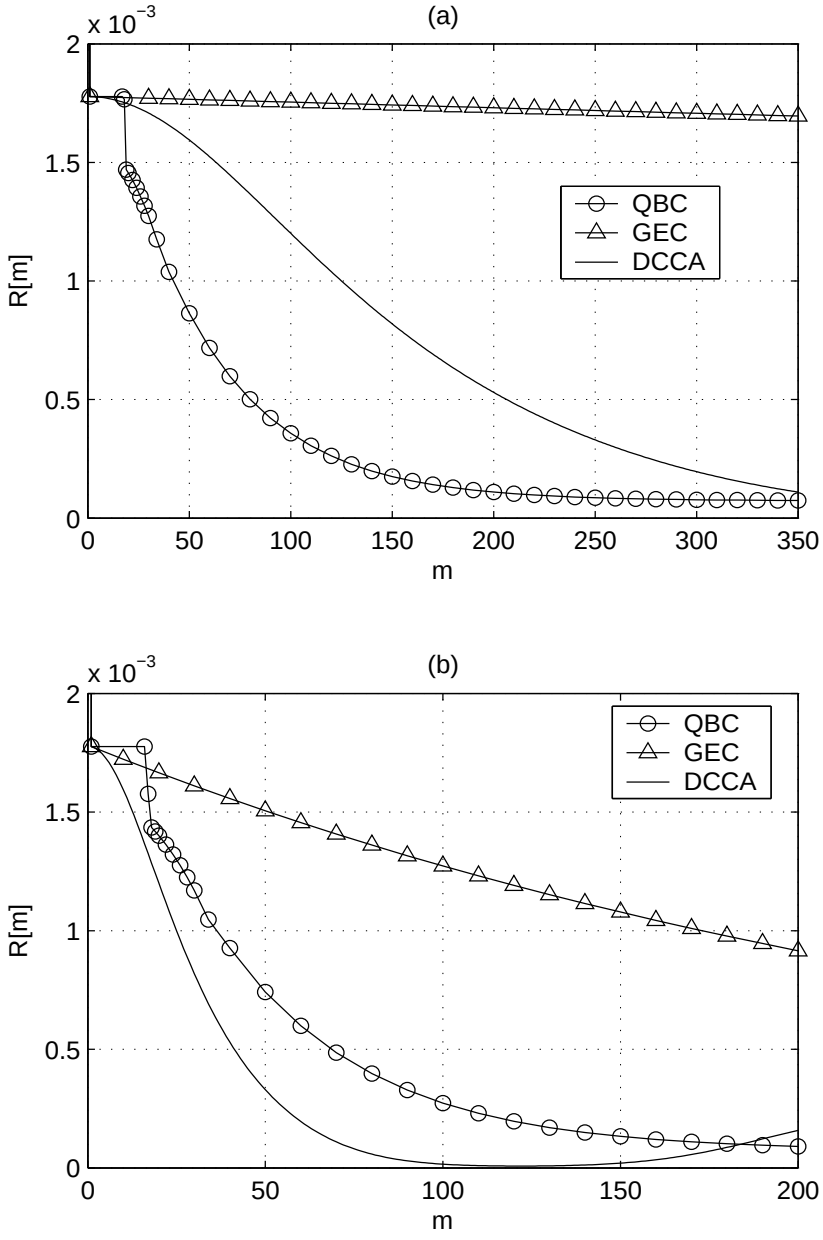


Figure 5.6: DCCA fitting via the QBC: ACF vs m for $K_R = 5$ dB (Rician) and SNR = 15 dB. (a) $f_D T = 0.001$; (b) $f_D T = 0.005$.

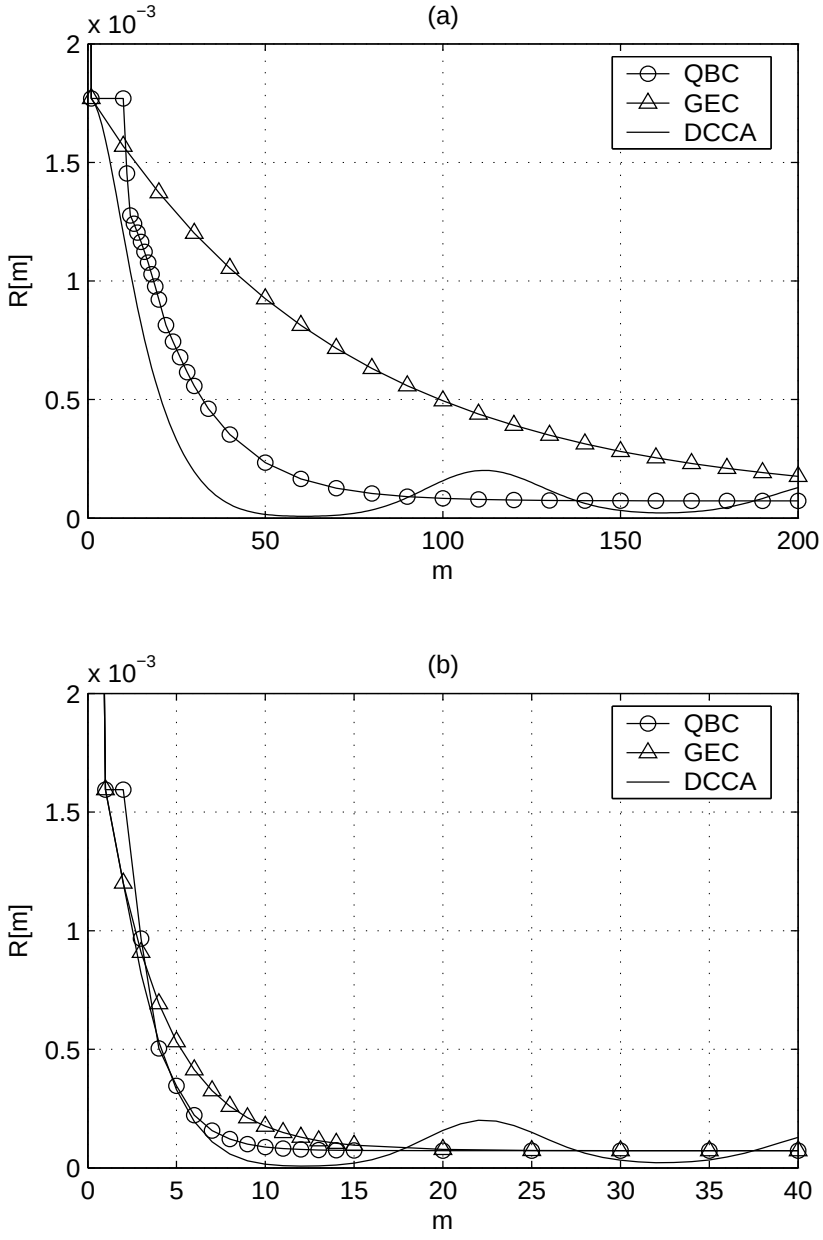


Figure 5.7: DCCA fitting via the QBC: ACF vs m for $K_R = 5$ dB (Rician) and SNR = 15 dB. (a) $f_D T = 0.01$; (b) $f_D T = 0.05$.

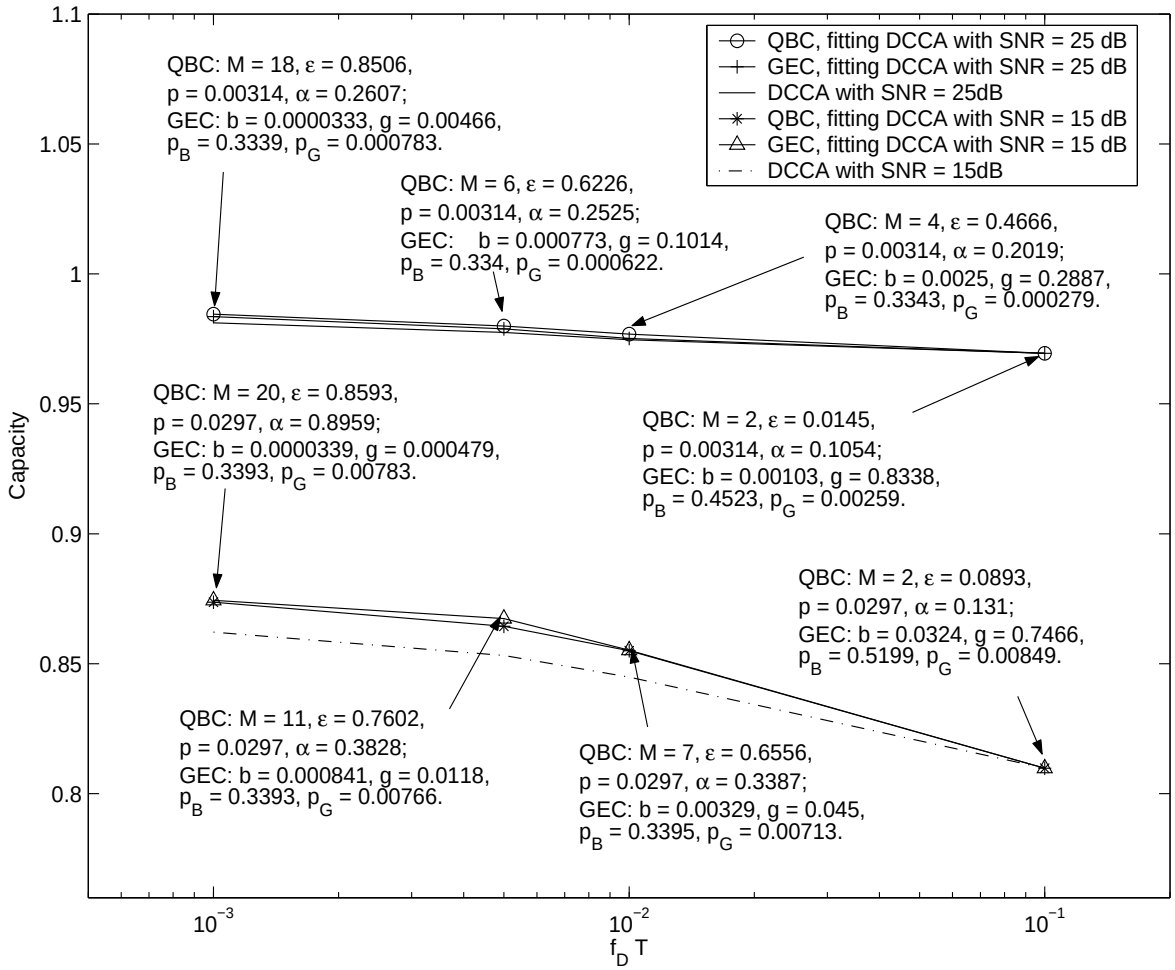


Figure 5.8: DCCA fitting via the QBC: Capacity vs normalized Doppler frequency $f_D T$ for Rayleigh fading channel.

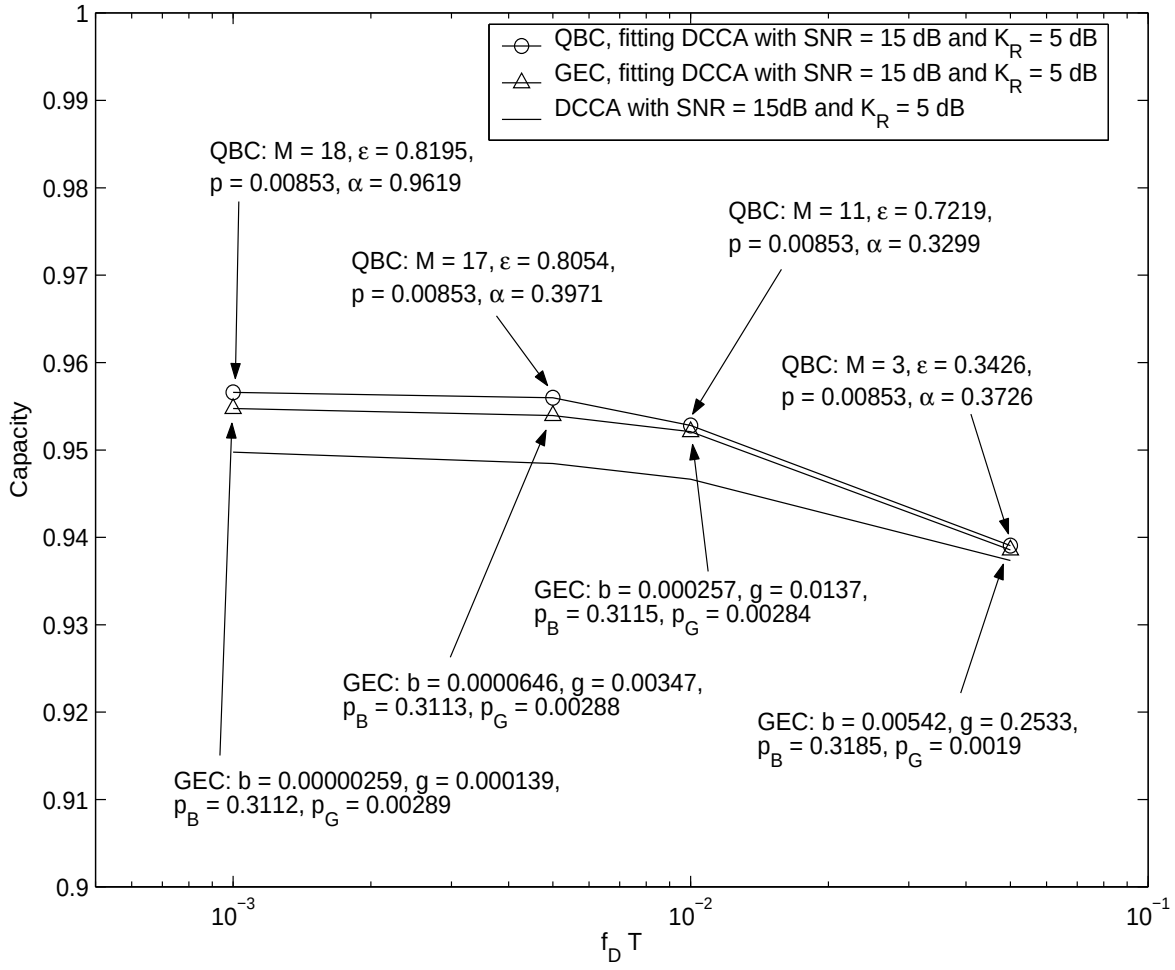


Figure 5.9: DCCA fitting via the QBC: Capacity vs normalized Doppler frequency $f_D T$ for Rician fading channel.

Chapter 6

Conclusions and Future Work

6.1 Summary

In this dissertation, we proposed a binary burst-noise channel based on a finite queue. The channel noise process is stationary and ergodic, and the state of the channel is a one-step Markov process with 2^M states.

We have derived the stationary noise distribution, the block transition probability and the capacity of the channel in terms of the parameters M, ε, p and α for the QBC. Its capacity is positive, increasing in α and asymptotically upper bounded by 1 for fixed M, BER and Cor. When $0 \leq \alpha \leq 1$, memory increases capacity for fixed α, BER and Cor. We also studied a particular case, the UQBC, by choosing $\alpha = 1$. The capacity of the UQBC is positive, increases with M and is asymptotically upper

bounded by $C_{\text{UQBC}}^{(M\to\infty)}$. We compared analytically the QBC with the FMCC, the GEC and the (K, 1)SFC. The UQBC and the FMCC were shown to have identical block transition probability for the same memory, BER and Cor; hence they have identical capacities under the above conditions. Therefore, $C_{\text{FMCC}} < C_{\text{QBC}}^{(M)}$ with $\alpha > 1$, and $C_{\text{FMCC}} > C_{\text{QBC}}^{(M)}$ with $0 \leq \alpha < 1$ for the same M, BER and Cor. Furthermore, when memory is 1, the capacity of the UQBC ($C_{\text{UQBC}}^{(M=1)}$) is smaller than that of the GEC (C_{GEC}) for the same BER and Cor, and the (K, 1)SFC is statistically identical to the UQBC with memory 1. Thus, the capacity of the (K, 1)SFC is not larger than that of the QBC with $\alpha \geq 1$ for the same BER, Cor and any M. It was observed via numerical computations that, for the set of considered parameters, the QBC with $M = 2$ and $\alpha = 100$ has the largest capacity, whereas the UQBC with $M = 1$ has the smallest capacity. The queue-based model can span a broad range of capacities by adjusting its parameters.

We also approximate the GEC and DCCA (the equivalent hard-decision demodulated Rician fading channel with memory) via the QBC model. The parameter estimation of the QBC was achieved by minimizing the KLDR between the probability of error sequences generated by the GEC/DCCA and the QBC, while maintaining identical BER and Cor. The results show that the curves of the capacity, the ACF, and the reliability function of the GEC and QBC match quite well and are almost identical for some cases with low memory order. A strong agreement is also shown

between the ACF and capacity curves of the QBC and the DCCA. This leads us to conclude that the QBC provides a very good approximation of the GEC under a variety of channel conditions; as well, the QBC matches the DCCA much better than the GEC and Markov models of [44] for slow and very slow fading conditions. This lead us to conclude that the QBC can characterize a large class of communication channels with memory, while remaining mathematically tractable.

6.2 Future Work

One possible direction for future work is the modeling and analysis of wired/wireless Internet traffic and channel coding, as an extension and application of this work. Sanneck and Carle [49] used an Mth order Markov chain process to describe the dependencies between packet losses. However, their models have a complexity (number of parameters) that grows exponentially with M, rendering it impractical for the modeling of packet loss processes with large memory. The QBC model, on the other hand, does not suffer from this limitation as it is fully described by only four parameters and allows single-letter analysis. The QBC can hence be employed to characterize the packet-loss patterns introduced by the Internet, especially to capture loss burstiness and distances between loss bursts. Another topic of future interest is the design, construction and analysis of channel codes for the QBC. One important objective in this problem is the judicious design of channel codes in order to fully exploit the channel

memory. Some results in this direction involving LDPC codes are reported in [40].

Bibliography

[1] L. Ahlin, "Coding methods for the mobile radio channel," in *Nordic Seminar Digital Land Mobile Commun.*, Espoo, Finland, Feb. 1985.

[2] F. Alajaji and T. Fuja, "A communication channel modeled on contagion," *IEEE Trans. Inform. Theory*, vol. 40, no. 6, pp. 2035–2041, Nov. 1994.

[3] F. Alajaji, N. Phamdo, N. Farvardin, and T. Fuja, "Detection of binary Markov sources over channels with additive Markov noise," *IEEE Trans. Inform. Theory,* vol. 42, pp. 230–239, Jan. 1996.

[4] F. Babich, O. E. Kelly and G. Lombardi, "A context-tree based model for quantized fading," *IEEE Commun. Lett.,* vol. 3, pp. 46–48, Feb. 1999.

[5] F. Babich and G. Lombardi, "A Markov model for the mobile propagation channel," *IEEE Trans. Veh. Technol.,* vol. 49, no. 1, pp. 63–73, Jan. 2000.

[6] F. Babich, O. E. Kelly and G. Lombardi, "Generalized Markov modeling for flat fading," *IEEE Trans. Commun.,* vol. 48, no. 4, pp. 547–551, Apr. 2000.

[7] H. Bischl and E. Lutz, "Packet error rate in the non-interleaved Rayleigh Channel," *IEEE Trans. Commun.,* vol. 43, no. 2/3/4, pp. 1375–1382, Feb./Mar./Apr. 1995.

[8] R. Blahut, *Principles and Practice of Information Theory*, Reading, MA: Addison Wesley, 1988.

[9] Q. Chen and K. P. Subbalakshmi, "Joint source-channel decoding for MPEG-4 coded video over wireless channels," *IEEE J. Selected Areas in Commun.*, vol. 21, no. 10, pp. 1780–1789, Dec. 2003.

[10] T. P.-C. Chen and T. Chen, "Error concealment aware rate shaping for wireless video transport," *EURASIP Signal Processing: Image Communication*, Special Issue on Recent Advances in Wireless Video, vol. 18, pp. 889–905, Nov. 2003.

[11] J.-Y. Chouinard, M. Lecours and G. Y. Delisle, "Estimation of Gilbert's and Fritchman's models parameters using the gradient method for digital mobile radio channels," *IEEE Trans. Veh. Technol.*, vol. 37, no. 3, pp. 158–166, Aug. 1988.

[12] M. J. Chu, and W. E. Stark, "Effect of Mobile velocity on communications in fading channels," *IEEE Trans. Veh. Technol.*, vol. 49, no. 1, pp. 202–210, Jan. 2000.

[13] V. Y. Y. Chu and P. Sweeney, "Characterizing error sequences of low earth orbit satellite channel and optimization with hybrid-ARQ schemes," *Proc. GLOBECOM'98*, Sydney, Australia, vol. 5, pp. 2930–2935, Nov. 1998.

[14] R. H. Clarke, "A statistical theory of mobile radio reception," *Bell Syst. Tech. J.*, vol. 47, pp. 957–1000, 1968.

[15] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, New York: Wiley, 1991.

[16] I. Csiszár and J. Körner, *Information Theory: Coding Theorems for Discrete Memoryless Systems*, New York: Academic Press, 1981.

[17] R. L. Dobrushin and M. S. Pinsker, "Memory increases transmission capacity", *Probl. Pered. Inform.*, vol. 5, no. 1, pp. 94–95, 1969.

[18] A. I. Drukarev and K. P. Yiu, "Performance of error-correcting codes on channels with memory," *IEEE Trans. Commun.*, vol. 34, no. 6, pp. 513–521, June 1986.

[19] M. Egarmin, "Upper and lower bounds for the probability of error in coding for discrete channels," *Probl. Pered. Inform.*, vol. 5, no. 1, pp. 23–39, 1969.

[20] E. O. Elliott, "Estimates of error rates for codes on burst-noise channel," *Bell Syst. Tech. J.*, vol. 42, pp. 1977–1997, Sept. 1963.

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

[22] B. D. Fritchman, "A binary channel characterization using partitioned Markov chains," *IEEE Trans. Inform. Theory*, vol. 13, no. 2, pp. 221–227, Apr. 1967.

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

[24] M. Gans, "A power spectral theory of propagation in the mobile radio environment," *IEEE Trans. Veh. Technol.*, vol. VT 21, pp. 27–38, Feb. 1972.

[25] E. N. Gilbert, "Capacity of a burst-noise channel," *Bell Syst. Tech. Jlow density parity check (.*, vol. 39, pp. 1253–1266, Sept. 1960.

[26] R. M. Gray, *Entropy and Information Theory*, New York: Springer–Verlag, 1990.

[27] B. M. Hochwald and P. R. Jelenković, "State learning and mixing in entropy of hidden Markov processes and the Gilbert-Elliott channel," *IEEE Trans. Inform. Theory*, vol. 45, no. 1, pp. 128–138, Jan. 1999.

[28] R. Iordache and I. Tabus, "Index assignment for transmitting vector quantized LSP parameters over binary Markov channels," *Proc. IEEE Int. Symp. Circuits and Systems*, Orlando, FL, USA, vol. 4, pp. 544–547, May 1999.

[29] R. Iordache, I. Tabus, and J. Astola, "Robust index-assignment using Hadamard transform for vector quantization transmission over finite-memory contagion channels," *Circuits, Systems, and Signal Processing*, vol. 21, no. 5, pp. 485–509, Sep.-Oct. 2002.

[30] C.-D. Iskander and P. T. Mathiopoulos, "Fast simulation of diversity Nakagami fading channels using finite-state Markov models," *IEEE Trans. on Broadcasting*, vol. 49, no. 3, pp. 269–277, Sept. 2003.

[31] L. N. Kanal and A. R. K. Sastry, "Models for channels with memory and their applications to error control," *Proc. of the IEEE* , vol. 66, no. 7, pp. 724–744, July 1978.

[32] G. K. Karagiannidis, D. A. Zogas and S. A. Kotsopoulos, "On the multivariate Nakagami-m distribution with exponential correlation," *IEEE Trans. Commun.*, vol. 51, pp. 1240–1244, Aug. 2003.

[33] —, "An efficient approach to multivariate Nakagami-m distribution using Green's matrix approximation," *IEEE Trans. Wireless Commun.*, vol. 2, pp. 883–889, Sept. 2003.

[34] A. Lapidoth and P. Narayan, "Reliable communication under channel uncertainty," *IEEE Trans. Inform. Theory*, vol. 44, no. 6, pp. 2148–2177, Oct. 1998.

[35] E. Lutz, D. Cygan, M. Dippold and W. Papke, "The land mobile satellite communication channel – recording, statistics, and channel model," *IEEE Trans. Veh. Technol.,* vol. 40, no. 2, pp. 375–386, May 1991.

[36] R. K. Mallik, "On multivariate Rayleigh and exponential distributions," *IEEE Trans. Inform. Theory*, vol. 48, pp. 1499–1515, June 2003.

[37] 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, Nov. 1989.

[38] V. Nagarajan and O. Milenkovic, "Performance analysis of structured LDPC codes in the Polya-urn channel with finite memory," *Proc. IEEE Canadian Conf. Elec. Comp. Eng.*, Niagara Falls, Canada, vol. 1, pp. 543–546, May 2004.

[39] M. Nakagami, "The m-distribution - a general formula for intensity distribution of rapid fading," in *Statistical Methods in Radio Wave Propagation*, W. G. Hoffman, Ed. Oxford, U.K.: Pergamon, pp. 3–36, 1960.

[40] C. Nicola, F. Alajaji, and T. Linder, "Decoding LDPC codes over binary channels with additive Markov noise," *Proc. of the 2005 Canadian Workshop on Information Theory*, Montreal, QC, Canada, pp. 187–190, June 2005.

[41] N. Phamdo, F. Alajaji and N. Farvardin, "Quantization of memoryless and Gauss-Markov sources over binary Markov channels," *IEEE Trans. Commun.*, vol. 45, pp. 668–675, June 1997.

[42] C. Pimentel and I. F. Blake, "Non-interleaved Reed-Solomon coding performance on finite state channels," *IEEE Int. Conf. Communication*, vol. 3, pp. 1493–1497, 1997.

[43] —, "Modeling burst channels using partitioned Fritchman's Markov models," *IEEE Trans. Veh. Technol.*, vol. 47, no. 3, pp. 885–899, Aug. 1998.

[44] C. Pimentel, T. H. Falk and L. Lisbôa, "Finite-state Markov modeling of correlated Rician fading channels," *IEEE Trans. Veh. Technol.,* vol. 53, no. 5, pp. 1491–1501, Sept. 2004.

[45] G. Polya, "Sur quelques points de la théorie des probabilités," *Ann. Inst. H. Poincarré*, vol. 1, pp. 117–161, 1931.

[46] J. G. Proakis, *Digital communications*. New York: McGraw-Hill, 2000.

[47] Z. Rached, F. Alajaji and L. L. Campbell, "Rényi's divergence and entropy rates for finite-alphabet Markov sources," *IEEE Trans. Inform. Theory*, vol. 47, pp. 1553–1561, May 2001.

[48] K. Sakakibara, "Performance analysis of the error-forecasting decoding for interleaved block codes on Gilbert-Elliott Channels," *IEEE Trans. Commun.,* vol. 48, no. 3, pp. 386–395, Mar. 2000.

[49] H. Sanneck and G. Carle, "A framework model for packet loss metrics based on loss runlengths," *Proc. of the SPIE/ACM SIGMM Multimedia Computing and Networking Conference 2000*, San Jose, CA, USA, pp. 177–187, Jan. 2000.

[50] C. E. Shannon, "A mathematical theory of communication," *Bell Sys. Tech. J.*, vol. 27, pp. 379–423, 623–656, 1948.

[51] G. Sharma, A. A. Hassan and A. Dholakia, "Performance evaluation of burst-error-correcting codes on a Gilbert-Elliott channel," *IEEE Trans. Commun.,* vol. 46, no. 7, pp. 846–849, July 1998.

[52] P. C. Shields, *The Ergodic Theory of Discrete Sample Paths*, The American Mathematical Society, 1996.

[53] M. K. Simon and M. -S. Alouini, "A simple single integral representation of the bivariate Rayleigh distribution," *IEEE Commun. Letters,* vol. 2, pp. 128–130, May 1998.

[54] —, "A unified performance analysis of digital communication with dual selective diversity over correlated Rayleigh and Nakagami-m fading channels," *IEEE Trans. Commun.,* vol. 47, pp. 33–43, Jan. 1999.

[55] —, *Digital Communication Over Fading Channels*. New York: Wiley, 2000.

[56] K. P. Subbalakshmi and J. Vaisey, "On the joint source-channel decoding of variable-length encoded sources: The additive-Markov case," *IEEE Trans. Commun.,* vol. 51, no. 9, pp. 1420–1425, Sept. 2003.

[57] F. Swarts and H. C. Ferreira, "Markov characterization of digital fading mobile VHF channels," *IEEE Trans. Veh. Technol.,* vol. 43, no. 4, pp. 977–985, Nov. 1994.

[58] C. C. Tan and N. C. Beaulieu, "Infinite series representations of the bivariate Rayleigh and Nakagami-m distribution," *IEEE Trans. Commun.,* vol. 45, pp. 1159–1161, Oct. 1997.

[59] —, "First-order Markov modeling for the Rayleigh fading channel," *Proc. IEEE GLOBECOM'98,* Sydney, Australia, vol. 6, pp. 3669–3674, Nov. 1998.

[60] —, "On first-order Markov modeling for the Rayleigh fading channel," *IEEE Trans. Commun.,* vol. 48, no. 12, pp. 2032–2040, Dec. 2000.

[61] S. Tsai, "Markov characterization of the HF channel," *IEEE Trans. Commun. Technol.*, vol. 17, no. 1, pp. 24–32, Feb. 1969.

[62] W. Turin, *Performance Analysis of Digital Transmission Systems.* New York: Computer Science, 1990.

[63] W. Turin and R. van Nobelen, "Hidden Markov modeling of flat fading channels," *IEEE J. Select. Areas Commun.*, vol. 16, pp. 1809–1817, Dec. 1998.

[64] S. Verdú and T. H. Han, "A general formula for channel capacity," *IEEE Trans. Inform. Theory*, vol. 40, no. 4, pp. 1147–1157, July 1994.

[65] A. J. Viterbi and J. K. Omura, *Principles of Digital Communication and Coding.* New York: McGraw-Hill, 1979.

[66] H. S. Wang and N. Moayeri, "Finite-state Markov channel – A useful model for radio communication channels," *IEEE Trans. Veh. Technol.,* vol. 44, no. 1, pp. 163–171, Feb. 1995.

[67] H. S. Wang and P.-C. Chang, "On verifying the first-order Markovian assumption for a Rayleigh fading channel," *IEEE Trans. Veh. Technol.,* vol. 45, no. 2, pp. 353–357, May 1996.

[68] L. Wilhelmsson and L. B. Milstein, "On the effect of imperfect interleaving for the Gilbert-Elliott channel," *IEEE Trans. Commun.,* vol. 47, no. 5, pp. 681–688, May 1999.

[69] Q. Zhang and S. Kassam, "Finite state Markov model for Rayleigh fading channels," *IEEE Trans. Commun.,* vol. 47, no. 11, pp. 1688–1692, Nov. 1999.

[70] L. Zhong, F. Alajaji, and G. Takahara, "A model for a binary burst-noise channel based on a finite queue," *Proc. of the 2001 Canadian Workshop on Information Theory*, Vancouver, BC, Canada, pp. 60–63, June 2001.

[71] —, "A comparative study of burst-noise communication channel models," *Proc. of the 2002 Biennial Symposium on Communications*, Kingston, ON, Canada, pp. 437–441, June 2002.

[72] —, "A queue-based model for binary communication channels," *Proc. of Allerton Conference on Commun., Contr., and Comp.*, Monticello, IL, USA, pp. 130–139, Oct. 2003.

[73] —, "An approximation of the Gilbert-Elliott channel via a queue-based channel model," *Proc. of the 2004 IEEE International Symposium on Information Theory*, Chicago, IL, USA, pp. 63, June 2004.

[74] —, "A queue-based model for wireless Rayleigh fading channels with memory," to appear in *Proc. of IEEE 62nd Semiannual Vehicular Technology Conference*, Dallas, TX, USA, Sept. 2005.

[75] —, "A binary communication channel with memory based on a finite queue," submitted to *IEEE Transaction on Information Theory*, 49 pages.

[76] —, "A binary queue-based model for wireless Rician fading channels with memory," under preparation.

[77] M. Zorzi, R. R. Rao and L. B. Milstein, "ARQ error control on fading mobile radio channels," *IEEE Trans. Veh. Technol.,* vol. 46, pp. 445–455, May 1997.

[78] M. Zorzi and R. R. Rao, "On the statistics of block errors in bursty channels," *IEEE Trans. Commun.,* vol. 45, pp. 660–667, June 1997.