A Binary Communication Channel with Memory Based on a Finite Queue*

Libo Zhong, Fady Alajaji and Glen Takahara

Department of Mathematics and Statistics

Queen's University, Kingston, Ontario, Canada, K7L 3N6

Abstract

A model for a binary additive noise communication channel with memory is introduced. 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 error exponent) 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, such as the well-known Gilbert-Elliott channel (GEC), the Fritchman channel and the finite-memory contagion channel. We also 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 between the probability of noise sequences generated by the GEC and the QBC, while maintaining identical bit error rates and correlation coefficients. The accuracy of fitting the GEC via the QBC is evaluated in terms of ACF, channel capacity and error exponent. Numerical results indicate that the QBC provides a good approximation of the GEC for various channel conditions; it thus offers an interesting alternative to the GEC while remaining mathematically tractable.

Index Terms: Modeling of communication channels, binary channels with memory, finite-state channels, additive bursty noise, Markov noise, Gilbert-Elliott channel, fading, error statistics, correlation, capacity, error exponent, Kullback-Leibler divergence rate.

*Submitted to the *IEEE Transactions on Information Theory*, January 2005; Revised March 2006. This work was supported in part by the Natural Sciences and Engineering Research Council (NSERC) of Canada and the Premier Research Excellence Award (PREA) of Ontario. Parts of this work were presented at the Seventh Canadian Workshop on Information Theory, Vancouver, BC, Canada, June 2001, at the Forty First Annual Allerton Conference on Communication, Control, and Computing, Allerton Park, IL, USA, September 2003, and at the 2004 IEEE International Symposium on Information Theory, Chicago, IL, USA, June 2004.

1 Introduction

It is well known that most real-world communication channels have memory and experience noise and fading distortions in a bursty fashion [24]. In order to design effective communication systems for such channels, it is critical to fully understand their behavior. 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, and simple enough to allow mathematically tractable system analysis.

During the past several decades, a variety of channel models with memory have been proposed and studied for the modeling of wireless channels. The most commonly used models for representing the "discretized" version (under hard-decision demodulation) of binary-input fading channels with memory [29, 30] are the Gilbert-Elliott channel (GEC) [19, 14] and the Fritchman channel (FC) [16]. They were for example employed to model high-frequency channels [37], mobile radio channels [6, 35], low earth orbit satellite channels [7] and magnetic tape recorders [11]. These models, which have been partly adopted for historical reasons (since being introduced in the 1960s, they were the first available models for channels with memory), are described by binary additive noise sources generated via finite-state hidden Markov models (HMMs).[1] 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. This may partly explain why, to date, few (to the best of our knowledge) 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 (e.g, see [12, 18] where iterative decoding of low-density parity-check codes (LDPC) over the GEC is studied and it is shown that exploiting the channel memory at the decoder leads to significant performance gains). 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 [25]. 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 [10, 1]). It is therefore vital to construct channel models which can well represent the behaviour of real-world channels while remaining analytically tractable for design purposes.

[1] A description of other lesser known, but related, finite or infinite state HMM based channel models is provided in [24].

[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 [25]).

In [1, 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 [31]. In such a channel, every error (or "infection", if we use the contagion interpretation) effectively increases the probability of future errors ([31], [15, 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. The FMCC model was adopted in several joint source-channel coding studies (e.g., [2, 4, 5, 22, 23, 28, 34] where the channel statistics are incorporated into the system design in order to exploit the noise memory. The performance of structured LDPC codes for the FMCC was also recently investigated in [26]. In light of the simple structure of the FMCC, the authors make a detailed density evolution analysis and provide a complete characterization of the convergence region for iterative decoding within the channel parameter space.

We introduce a new binary additive noise channel based on a finite queue of length M. The proposed queue-based channel (QBC) model also features an Mth order Markov noise source that is fully characterized by four parameters, making it more sophisticated than the FMCC for channel modeling (as it has an additional degree of freedom) while remaining mathematically tractable. It is also important to point out that Pimentel *et. al.* recently showed in a numerical study [30] that the class of binary channel models with additive Mth order Markov noise (to which both the QBC and FMCC models belong) is a good approximation, in terms of the error autocorrelation function (ACF) and variational distance, of 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 Mth order Markov noise channels considered in [30] 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 [30, Fig. 11]). The QBC model, on the other hand, does not suffer from this limitation as it is fully described by only four parameters and it can accommodate very large values of the memory M. Indeed, it was recently shown in [40, 41] that the QBC provides a significantly better approximation of the correlated Rayleigh and Rician fading channels (in terms of fitting the error ACF) than the GEC and the Markov models of [30]; this is observed for a wide range of channel conditions, including very slow fading. The QBC 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.

We also 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 error exponent (or reliability function). 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.

The rest of this paper is organized as follows. Preliminaries on previous channel models are presented in Section 2. In Section 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 selected with equal probability; 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 FMCC, the GEC and a particular class of the Fritchman channel. In Section 4, the problem of approximating the GEC via the QBC is considered. Finally, concluding remarks are given in Section 5.

Hereafter, a discrete-time binary additive noise communication channel 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's 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 paper, a given noise process $\{Z_n\}_{n=1}^{\infty}$ will be generated according to one of the following models: the GEC, the FC, the FMCC or the QBC.

2 Preliminaries: Previous Binary Channel Models

2.1 Gilbert-Elliott Channel

The GEC model belongs to the family of finite-state channels, which is thoroughly studied in [17, 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 (see Fig. 1). 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 (with $p_G < p_B$) are the crossover probabilities in the good and bad states, respectively (the Gilbert Channel (GC) [19] is obtained when $p_G = 0$, i.e., it behaves like a noiseless BSC in the good state). After every symbol transmission, the Markov chain makes a state transition according to the transition probability matrix

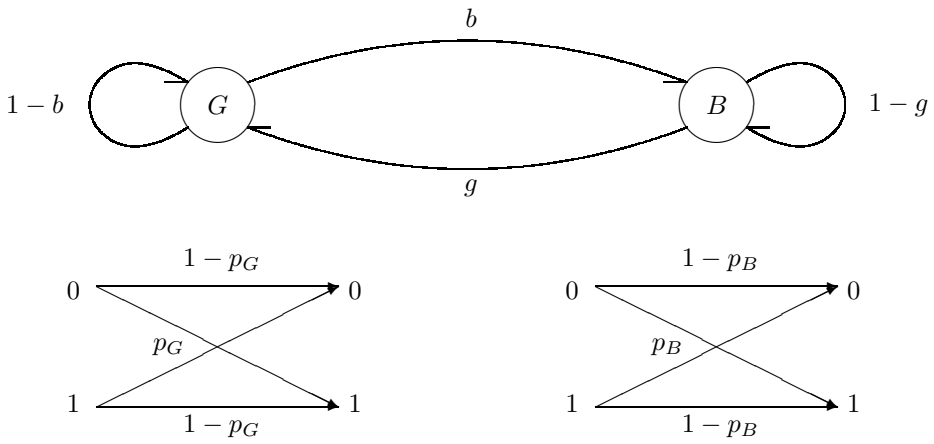

Figure 1: The Gilbert-Elliott channel model.

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

where $0 < b < 1$ and $0 < g < 1$. A useful approach for calculating the probability of an error or noise sequence for the Gilbert-Elliott Channel is discussed in [29]. 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

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

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 $s_k = j$. The vector $\boldsymbol{\pi} = [\pi_0 \;\; \pi_1]^T$ indicates the stationary distribution vector of the GEC, and $\mathbf{1}$ is the 2-dimensional vector with all ones.

2.2 Fritchman Channel

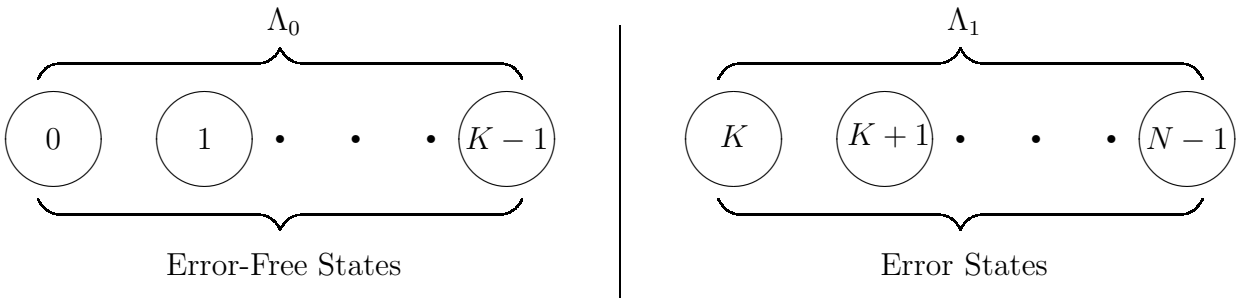

Figure 2: Partitioning of the state space of the Fritchman channel model.

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) [16]. Corresponding to this partition, the state transition probability matrix $\boldsymbol{P}$ and the stationary vector $\boldsymbol{\pi}$ can be written 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 [29], Pimentel and Blake represent the probability of a noise sequence of length n by

$$\Pr(z_1, z_2, \cdots, z_n) = \boldsymbol{\pi}_{z_1}^T \left(\prod_{l=1}^{n-1} \boldsymbol{P}_{z_l z_{l+1}} \right) \mathbf{1}. \tag{2}$$

2.3 Finite-Memory Contagion Channel

The noise process of the FMCC [1] 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

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

It can be shown that the noise process $\{Z_i\}_{i=M+1}^{\infty}$ is a stationary ergodic Markov source of order M [1].

3 Queue-Based Channel with Memory

We first present the queue-based binary channel with memory whose additive noise process is generated according to a ball sampling involving the following two parcels.

- **Parcel 1** is a queue of length M as shown in Fig. 3, that contains initially M balls, either red or black.

Figure 3: 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:

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

Intuitively, the above procedure, which generates each noise sample Z_n, consists of randomly switching between two sub-experiments: one sub-experiment, chosen with probability ε, involving a queue of length M (Parcel 1); and one sub-experiment, chosen with probability $1 - \varepsilon$, involving a large urn (Parcel 2). The noise sample Z_n is then determined by the outcome of the selected sub-experiment. Also, as indicated above, this outcome alters the composition of the queue (by placing a new ball of the same color as the outcome in the first cell of the queue, thus moving all the remaining balls by one cell to the right and flushing out the ball in the last cell).

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 $n \geq M+1$,

$$\begin{aligned}
&\Pr^{(M)}(Z_n = 1 \mid Z_{n-1} = z_{n-1}, \cdots, Z_1 = z_1) \\
&= \varepsilon \left(z_{n-1} \cdot \frac{1}{M-1+\alpha} + \cdots + z_{n-M+1} \cdot \frac{1}{M-1+\alpha} + z_{n-M} \cdot \frac{\alpha}{M-1+\alpha} \right) + (1-\varepsilon)p \\
&= \Pr(Z_n = 1 \mid Z_{n-1} = z_{n-1}, \cdots, Z_{n-M} = z_{n-M}),
\end{aligned} \tag{3}$$

where $z_l \in \{0,1\}$, $l = 1, \cdots, n-1$. Hence $\{Z_n\}_{n=1}^{\infty}$ is a homogeneous (or time-invariant [15]) 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. The channel state process $\{\underline{S}_n\}_{n=1}^{\infty}$ is a homogeneous Markov process with stationary distribution

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

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

$$p_{ij}^{(M)} = \begin{cases}
\left(M - \omega_i^{(M)} - 1 + \alpha\right) \frac{\varepsilon}{M-1+\alpha} + (1-\varepsilon)(1-p), & \text{if } j = \frac{i}{2}, \text{ and } i \text{ is even}, \\
\left(M - \omega_i^{(M)}\right) \frac{\varepsilon}{M-1+\alpha} + (1-\varepsilon)(1-p), & \text{if } j = \lfloor \frac{i}{2} \rfloor, \text{ and } i \text{ is odd}, \\
\omega_i^{(M)} \frac{\varepsilon}{M-1+\alpha} + (1-\varepsilon)p, & \text{if } j = \frac{i+2^M}{2}, \text{ and } i \text{ is even}, \\
\left(\omega_i^{(M)} - 1 + \alpha\right) \frac{\varepsilon}{M-1+\alpha} + (1-\varepsilon)p, & \text{if } j = \lfloor \frac{i+2^M}{2} \rfloor, \text{ and } i \text{ is odd}, \\
0, & \text{otherwise};
\end{cases} \tag{4}$$

where $\omega_i^{(M)}$ is the number of "ones" in the M-bit binary representation of the decimal integer i.

We note 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 [1]). It can be shown (see Appendix A) that the stationary

distribution $\boldsymbol{\pi}^{(M)}$ of the process is given by

$$\pi_i^{(M)} = \frac{\prod_{j=0}^{M-1-\omega_i^{(M)}} \left[j\frac{\varepsilon}{M-1+\alpha} + (1-\varepsilon)(1-p) \right] \prod_{j=0}^{\omega_i^{(M)}-1} \left[j\frac{\varepsilon}{M-1+\alpha} + (1-\varepsilon)p \right]}{\prod_{j=0}^{M-1} \left[1 - (\alpha+j)\frac{\varepsilon}{M-1+\alpha} \right]}, \tag{5}$$

for $i = 0, 1, 2, \cdots, 2^M - 1$, where $\prod_{j=0}^{a}(\cdot) \stackrel{\triangle}{=} 1$ if $a < 0$. We note that the stationary distribution (5) is identical to the stationary distribution of the channel state for the FMCC ([1], 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 (4) 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 $n \leq M$,

$$\Pr^{(M)}(Z^n = z^n) = \frac{\prod_{j=0}^{n-d_1^n-1} \left[j\frac{\varepsilon}{M-1+\alpha} + (1-\varepsilon)(1-p) \right] \prod_{j=0}^{d_1^n-1} \left[j\frac{\varepsilon}{M-1+\alpha} + (1-\varepsilon)p \right]}{\prod_{j=M-n}^{M-1} \left[1 - (\alpha+j)\frac{\varepsilon}{M-1+\alpha} \right]}, \tag{6}$$

  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 $n \geq M+1$,

$$\begin{aligned} \Pr^{(M)}(Z^n = z^n) &= L^{(M)} \prod_{i=M+1}^{n} \left[\left(d_{i-M+1}^{i-1} + \alpha z_{i-M} \right) \frac{\varepsilon}{M-1+\alpha} + (1-\varepsilon)p \right]^{z_i} \\ &\quad \left\{ \left[\left(M - 1 - d_{i-M+1}^{i-1} \right) + \alpha(1 - z_{i-M}) \right] \frac{\varepsilon}{M-1+\alpha} + (1-\varepsilon)(1-p) \right\}^{1-z_i}, \end{aligned} \tag{7}$$

  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.

$$\text{BER} = \Pr(Z_i = 1) = \Pr(Z_1 = 1) = p, \tag{8}$$

and

$$\text{Cor} = \frac{Cov(Z_i, Z_{i+1})}{\sqrt{Var(Z_i)Var(Z_{i+1})}} = \frac{Cov(Z_2, Z_1)}{Var(Z_1)}, \tag{9}$$

where $Cov(Z_2, Z_1) = E[Z_2 Z_1] - E[Z_2]E[Z_1] = \Pr(Z_2 = 1, Z_1 = 1) - p^2$ and $Var(Z_1) = E[Z_1^2] - E[Z_1]^2 = \Pr[Z_1 = 1] - p^2 = p - p^2$. To obtain $Cov(Z_2, Z_1)$, we use (7) 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 (6) (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

$$Cov(Z_2, Z_1) = p\frac{\frac{\varepsilon}{M-1+\alpha} + (1-\varepsilon)p}{1 - (M-2+\alpha)\frac{\varepsilon}{M-1+\alpha}} - p^2 = p(1-p)\frac{\frac{\varepsilon}{M-1+\alpha}}{1 - (M-2+\alpha)\frac{\varepsilon}{M-1+\alpha}}. \tag{10}$$

Note that, since $\alpha = 1$ when $M = 1$, (10) also holds for $M = 1$. Thus, for $M \geq 1$ we have

$$\text{Cor} = \frac{Cov(Z_2, Z_1)}{Var(Z_1)} = \frac{\frac{\varepsilon}{M-1+\alpha}}{1 - (M-2+\alpha)\frac{\varepsilon}{M-1+\alpha}}. \tag{11}$$

From (8) and (11), the parameters p and ε can be expressed in terms of BER, Cor, M and α as follows:

$$p = \text{BER}, \tag{12}$$

and

$$\varepsilon = \frac{(M-1+\alpha)\text{Cor}}{1 + (M-2+\alpha)\text{Cor}}. \tag{13}$$

Similarly, since $\frac{\varepsilon}{1-\varepsilon} = \frac{(M-1+\alpha)\text{Cor}}{1-\text{Cor}}$, (6) and (7) can be written in terms of BER, Cor, M and α as follows:

- For blocklength $n \leq M$,

$$\Pr^{(M)}(Z^n = z^n) = \frac{\prod_{j=0}^{n-d_1^n-1}\left[j\frac{\text{Cor}}{1-\text{Cor}} + (1-\text{BER})\right]\prod_{j=0}^{d_1^n-1}\left(j\frac{\text{Cor}}{1-\text{Cor}} + \text{BER}\right)}{\prod_{j=0}^{n-1}\left(1 + j\frac{\text{Cor}}{1-\text{Cor}}\right)}. \tag{14}$$

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

$$\begin{aligned}\Pr^{(M)}(Z^n = z^n) &= L^{(M)} \prod_{i=M+1}^{n} \left[\frac{\left(d_{i-M+1}^{i-1} + \alpha z_{i-M}\right) \frac{\text{Cor}}{1-\text{Cor}} + \text{BER}}{1 + (M-1+\alpha)\frac{\text{Cor}}{1-\text{Cor}}} \right]^{z_i} \\ &\quad \left\{ \frac{\left[M-1-d_{i-M+1}^{i-1} + \alpha(1-z_{i-M})\right] \frac{\text{Cor}}{1-\text{Cor}} + (1-\text{BER})}{1 + (M-1+\alpha)\frac{\text{Cor}}{1-\text{Cor}}} \right\}^{1-z_i}, \end{aligned} \tag{15}$$

where

$$L^{(M)} = \frac{\prod_{j=0}^{M-1-d_1^M} \left[j\frac{\text{Cor}}{1-\text{Cor}} + (1-\text{BER}) \right] \prod_{j=0}^{d_1^M-1} \left(j\frac{\text{Cor}}{1-\text{Cor}} + \text{BER} \right)}{\prod_{j=0}^{M-1} \left(1 + j\frac{\text{Cor}}{1-\text{Cor}} \right)}. \tag{16}$$

The stationary distribution (5) can be expressed in terms of M, BER and Cor as follows.

$$\pi_i^{(M)} = \frac{\prod_{j=0}^{M-\omega_i^{(M)}-1} \left[j\frac{\text{Cor}}{1-\text{Cor}} + (1-\text{BER}) \right] \prod_{j=0}^{\omega_i^{(M)}-1} \left(j\frac{\text{Cor}}{1-\text{Cor}} + \text{BER} \right)}{\prod_{j=0}^{M-1} \left(1 + j\frac{\text{Cor}}{1-\text{Cor}} \right)}, \tag{17}$$

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

In [23], 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 1 For fixed BER and Cor, the stationary distribution $\pi_i^{(M)}$ obeys the following recursion:

$$\pi_i^{(M)} = \pi_{2i}^{(M+1)} + \pi_{2i+1}^{(M+1)}, \text{ for } i = 0, 1, 2, \cdots, 2^M - 1. \tag{18}$$

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 (17),

$$\pi_{2i}^{(M+1)} = \frac{\prod_{j=0}^{M-\omega_i^{(M)}} \left[j\frac{\text{Cor}}{1-\text{Cor}} + (1-\text{BER}) \right] \prod_{j=0}^{\omega_i^{(M)}-1} \left(j\frac{\text{Cor}}{1-\text{Cor}} + \text{BER} \right)}{\prod_{j=0}^{M} \left(1 + j\frac{\text{Cor}}{1-\text{Cor}} \right)}, \tag{19}$$

and

$$\pi_{2i+1}^{(M+1)} = \frac{\prod_{j=0}^{M-\omega_i^{(M)}-1} \left[j\frac{\text{Cor}}{1-\text{Cor}} + (1-\text{BER}) \right] \prod_{j=0}^{\omega_i^{(M)}} \left(j\frac{\text{Cor}}{1-\text{Cor}} + \text{BER} \right)}{\prod_{j=0}^{M} \left(1 + j\frac{\text{Cor}}{1-\text{Cor}} \right)}. \tag{20}$$

Summing (19) and (20) we get

$$\begin{aligned}\pi_{2i}^{(M+1)} + \pi_{2i+1}^{(M+1)} &= \frac{\prod_{j=0}^{M-\omega_i^{(M)}-1}\left[j\frac{\text{Cor}}{1-\text{Cor}} + (1-\text{BER})\right]\prod_{j=0}^{\omega_i^{(M)}-1}\left(j\frac{\text{Cor}}{1-\text{Cor}} + \text{BER}\right)}{\prod_{j=0}^{M}\left(1 + j\frac{\text{Cor}}{1-\text{Cor}}\right)} \\ &\quad \times \left[\left(M - \omega_i^{(M)}\right)\frac{\text{Cor}}{1-\text{Cor}} + (1-\text{BER}) + \omega_i^{(M)}\frac{\text{Cor}}{1-\text{Cor}} + \text{BER}\right] \\ &= \frac{\prod_{j=0}^{M-\omega_i^{(M)}-1}\left[j\frac{\text{Cor}}{1-\text{Cor}} + (1-\text{BER})\right]\prod_{j=0}^{\omega_i^{(M)}-1}\left(j\frac{\text{Cor}}{1-\text{Cor}} + \text{BER}\right)}{\prod_{j=0}^{M-1}\left(1 + j\frac{\text{Cor}}{1-\text{Cor}}\right)}. \end{aligned} \tag{21}$$

By comparing (21) and (17), we get

$$\pi_i^{(M)} = \pi_{2i}^{(M+1)} + \pi_{2i+1}^{(M+1)}. \tag{22}$$

□

3.1.3 Autocorrelation Function

The autocorrelation function (ACF) of a binary stationary process $\{Z_n\}_{n=1}^{\infty}$ is defined by $R[m] = E\{Z_i Z_{i+m}\} = \Pr(Z_i = 1, Z_{i+m} = 1)$. It can be shown that the ACF of the QBC satisfies the following.

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

It can also be shown that $R[m]$ is strictly decreasing for $m > M$ and that $\lim_{m\to\infty} R[m] = p^2$; thus, Z_i and Z_{i+m} are asymptotically independent.

3.1.4 Channel Capacity

The information capacity for channels with memory is defined by Shannon's familiar expression [20, p. 287], [38]

$$C = \lim_{n\to\infty} \sup_{X^n} \frac{1}{n} I(X^n; Y^n), \tag{23}$$

where $I(X^n; Y^n)$ denotes the block mutual information [8, p. 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. For the wide class of channels with memory that are information stable [38] (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 [33], 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 (23). 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 (23). Thus,

$$C_{\text{QBC}}^{(M)} \;=\; 1-\mathcal{H}^{(M)}(Z), \tag{24}$$

where $\mathcal{H}^{(M)}(Z)$ is the entropy rate of the Mth order Markov noise process. It can be obtained using (4) and (5), as follows.

$$\begin{aligned}
\mathcal{H}^{(M)}(Z) &\stackrel{\triangle}{=} \lim_{n\to\infty}\frac{1}{n}H^{(M)}(Z_1,\cdots,Z_n)\\
&= H^{(M)}(Z_{M+1}\mid Z_M,Z_{M-1},\cdots,Z_1)\\
&= H^{(M)}(\underline{S}_2\mid\underline{S}_1)\\
&= -\sum_{i,j=0}^{2^M-1}\pi_i^{(M)}p_{ij}^{(M)}\log_2 p_{ij}^{(M)}\\
&= \sum_{\omega=0}^{M-1}\binom{M-1}{\omega}L_\omega^{(M)}h_b\left[\omega\frac{\varepsilon}{M-1+\alpha}+(1-\varepsilon)p\right]\\
&\quad+\sum_{\omega=1}^{M}\binom{M-1}{\omega-1}L_\omega^{(M)}h_b\left[(\omega-1+\alpha)\frac{\varepsilon}{M-1+\alpha}+(1-\varepsilon)p\right],
\end{aligned}\tag{25}$$

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 $a<0$, 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 $C_{\text{QBC}}^{(M)}$ is positive since $\mathcal{H}^{(M)}(Z)<1$ for fixed M, ε, p and α. Using (24) and (25), we obtain the following expression for $C_{\text{QBC}}^{(M)}$.

$$\begin{aligned}
C_{\text{QBC}}^{(M)} &= 1-\sum_{\omega=0}^{M-1}\binom{M-1}{\omega}L_\omega^{(M)}h_b\left[\omega\frac{\varepsilon}{M-1+\alpha}+(1-\varepsilon)p\right]\\
&\quad-\sum_{\omega=1}^{M}\binom{M-1}{\omega-1}L_\omega^{(M)}h_b\left[(\omega+\alpha-1)\frac{\varepsilon}{M-1+\alpha}+(1-\varepsilon)p\right].
\end{aligned}\tag{26}$$

In terms of the channel parameters M, BER, Cor and α, the capacity in (26) can be written as

$$\begin{aligned}
C_{\text{QBC}}^{(M)} &= 1-\sum_{\omega=0}^{M-1}\binom{M-1}{\omega}L_\omega^{(M)}h_b\left[\frac{\omega\frac{\text{Cor}}{1-\text{Cor}}+\text{BER}}{1+(M-1+\alpha)\frac{\text{Cor}}{1-\text{Cor}}}\right]\\
&\quad-\sum_{\omega=1}^{M}\binom{M-1}{\omega-1}L_\omega^{(M)}h_b\left[\frac{(\omega-1+\alpha)\frac{\text{Cor}}{1-\text{Cor}}+\text{BER}}{1+(M-1+\alpha)\frac{\text{Cor}}{1-\text{Cor}}}\right],
\end{aligned}\tag{27}$$

where

$$L_{\omega}^{(M)} = \frac{\prod_{j=0}^{M-1-\omega}\left[j\frac{\mathrm{Cor}}{1-\mathrm{Cor}}+(1-\mathrm{BER})\right]\prod_{j=0}^{\omega-1}\left(j\frac{\mathrm{Cor}}{1-\mathrm{Cor}}+\mathrm{BER}\right)}{\prod_{j=0}^{M-1}\left(1+j\frac{\mathrm{Cor}}{1-\mathrm{Cor}}\right)}, \tag{28}$$

which is not a function of α .

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

$$C_{\mathrm{QBC}}^{(M)} = 1 - h_b(p) = 1 - h_b(\mathrm{BER}). \tag{29}$$

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

Proof Notice that for various α, ε has to change to keep Cor fixed from (11). Rewriting (27), we get

$$C_{\mathrm{QBC}}^{(M)} = 1 - \sum_{\omega=0}^{M-1}\binom{M-1}{\omega}\left(L_{\omega}^{(M)} h_b\left[f(\alpha)\right] + L_{\omega+1}^{(M)} h_b\left[g(\alpha)\right]\right), \tag{30}$$

where $L_{\omega}^{(M)}$ is expressed in (28), $h_b(\cdot)$ 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

$$\frac{d\,C_{\mathrm{QBC}}^{(M)}}{d\,\alpha} = -\sum_{\omega=0}^{M-1}\binom{M-1}{\omega}\left[L_{\omega}^{(M)}\frac{d\,h_b\,(f)}{d\,f}\frac{d\,f}{d\,\alpha} + L_{\omega+1}^{(M)}\frac{d\,h_b\,(g)}{d\,g}\frac{d\,g}{d\,\alpha}\right]. \tag{31}$$

Since

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

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

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

$$\frac{d\,C_{\mathrm{QBC}}^{(M)}}{d\,\alpha} = -V^{(M)}\sum_{\omega=0}^{M-1}\left\{\binom{M-1}{\omega}L_{\omega}^{(M)}\left[-\left(\omega\frac{\mathrm{Cor}}{1-\mathrm{Cor}}+\mathrm{BER}\right)\right]\right.$$

$$\times \log_2 \left[\frac{(M-1+\alpha-\omega)\frac{\text{Cor}}{1-\text{Cor}} + (1-\text{BER})}{\omega \frac{\text{Cor}}{1-\text{Cor}} + \text{BER}} \right]$$
$$+ L_{\omega+1}^{(M)} \left[(M-\omega-1)\frac{\text{Cor}}{1-\text{Cor}} + (1-\text{BER}) \right]$$
$$\times \log_2 \left[\frac{(M-\omega-1)\frac{\text{Cor}}{1-\text{Cor}} + (1-\text{BER})}{(\omega+\alpha)\frac{\text{Cor}}{1-\text{Cor}} + \text{BER}} \right] \Bigg\} \tag{32}$$

where

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

Notice, from (28), that

$$W_\omega \stackrel{\triangle}{=} L_\omega^{(M)} \times \left(\omega \frac{\text{Cor}}{1-\text{Cor}} + \text{BER} \right) = L_{\omega+1}^{(M)} \times \left[(M-\omega-1)\frac{\text{Cor}}{1-\text{Cor}} + (1-\text{BER}) \right]. \tag{33}$$

Hence, (32) can be rewritten as

$$\begin{aligned} \frac{d\, C_{\text{QBC}}^{(M)}}{d\, \alpha} &= -V^{(M)} \sum_{\omega=0}^{M-1} \Bigg\{ \binom{M-1}{\omega} W_\omega \\ &\quad \log_2 \left[\frac{\omega \frac{\text{Cor}}{1-\text{Cor}} + \text{BER}}{(\omega+\alpha)\frac{\text{Cor}}{1-\text{Cor}} + \text{BER}} \times \frac{(M-\omega-1)\frac{\text{Cor}}{1-\text{Cor}} + (1-\text{BER})}{(M-1+\alpha-\omega)\frac{\text{Cor}}{1-\text{Cor}} + (1-\text{BER})} \right] \Bigg\}. \end{aligned} \tag{34}$$

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 (27) 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 (13) for a fixed Cor. 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 Error Exponent

The channel error exponent or reliability function $E(R)$ 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 [3, 17]). In other words, it is 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 the established lower bounds are claimed for $\liminf_{n\to\infty} -\frac{1}{n}\log P_e^*(R,n)$ and the upper bounds are claimed for $\limsup_{n\to\infty} -\frac{1}{n}\log P_e^*(R,n)$. $E(R)$ is a non-negative and non-increasing function of R for all $0 \le R < C$, and $E(R) = 0$ for $R \ge C$, 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).[3]

For our binary channels with additive noise, we can apply the following random coding lower bound (RCLB) [17] for $E(R)$.[4]

Proposition 1 *(Random coding lower bound (RCLB) [17])* For a binary channel with additive noise described by $\{p^{(n)}(z^n)\}$,

$$E(R) \ge E_r(R) \tag{35}$$

where

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

and

$$E_0^{(n)}(\rho) \;=\; \rho - \frac{1+\rho}{n}\log_2\left[\sum_{z^n\in\{0,1\}^n} p^{(n)}(z^n)^{\frac{1}{1+\rho}}\right]. \tag{36}$$

If the additive noise is a stationary ergodic Markov source, then a sphere-packing upper bound (SPUB) for $E(R)$ also holds (e.g., see [13] or [42]). Furthermore, $E_0^{(\infty)}(\rho) \stackrel{\triangle}{=} \lim_{n\to\infty} E_0^{(n)}(\rho)$ exists and admits a computable expression [17, 32]. This is summarized in the following proposition.

Proposition 2 *(Sphere-packing upper bound (SPUB) for Markov noise channels)* For a binary channel with stationary ergodic additive Markov noise with transition matrix $[p_{ij}]$ and n-fold distribution $p^{(n)}(z^n)$,

$$E(R) \le E_s(R),$$

[3] Recall that even for discrete memoryless channels, $E(R)$ is not exactly known at low rates; so only upper and lower bounds to $E(R)$ can be examined (see for example [3, 9, 17, 39]).

[4] This bound (with the appropriate modifications) actually holds for more general (non-additive) discrete channels with memory (see [17, Theorem 5.6.1]).

where

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

and

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

where $\lim_{n\to\infty}(1/n)H_\alpha(p^{(n)})$ is the *Rényi entropy rate* [32] of the Markov noise with parameter α ($\alpha > 0$ and $\alpha \neq 1$),

$$H_\alpha(p^{(n)}) \stackrel{\triangle}{=} \frac{1}{1-\alpha} \log_2 \left[\sum_{z^n} \left(p^{(n)}(z^n) \right)^\alpha \right],$$

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

It directly follows that for binary channels with additive stationary ergodic Markov noise, the above RCLB and SPUB for $E(R)$ are tight when $R \geq R_{cr}$, where R_{cr} is the critical rate given by

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

The above results also directly hold for Markov noise sources of order M (c.f., [32]). 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 1 and 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 (14) and (15) as follows.

- For blocklength $n \leq M$,

$$\Pr^{(M)}(Z^n = z^n) = \frac{\prod_{j=0}^{n-d_1^n-1} \left(j \frac{\text{Cor}}{1-\text{Cor}} + (1-\text{BER}) \right) \prod_{j=0}^{d_1^n-1} \left(j \frac{\text{Cor}}{1-\text{Cor}} + \text{BER} \right)}{\prod_{j=0}^{n-1} \left(1 + j \frac{\text{Cor}}{1-\text{Cor}} \right)}. \tag{37}$$

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

$$\Pr^{(M)}(Z^n = z^n) = L^{(M)} \prod_{i=M+1}^{n} \left(\frac{d_{i-M}^{i-1} \frac{\text{Cor}}{1-\text{Cor}} + \text{BER}}{1 + M \frac{\text{Cor}}{1-\text{Cor}}} \right)^{z_i} \left[\frac{\left(M - d_{i-M}^{i-1} \right) \frac{\text{Cor}}{1-\text{Cor}} + (1-\text{BER})}{1 + M \frac{\text{Cor}}{1-\text{Cor}}} \right]^{1-z_i} \tag{38}$$

  where $L^{(M)}$ is given by (16).

Similarly, from (27), the capacity of the UQBC in terms of M, BER, and Cor is given by

$$C_{\mathrm{UQBC}}^{(M)} = 1 - \sum_{\omega=0}^{M} \binom{M}{\omega} L_{\omega}^{(M)} h_b \left(\frac{\omega \frac{\mathrm{Cor}}{1-\mathrm{Cor}} + \mathrm{BER}}{1 + M \frac{\mathrm{Cor}}{1-\mathrm{Cor}}} \right), \tag{39}$$

where $L_{\omega}^{(M)}$ is given by (28).

Lemma 2 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 (14) and (15),

- For blocklength $n \leq M$,

$$\Pr^{(M+1)}(Z^n = z^n) = \frac{\prod_{j=0}^{n-d_1^n-1} \left(j \frac{\mathrm{Cor}}{1-\mathrm{Cor}} + (1 - \mathrm{BER}) \right) \prod_{j=0}^{d_1^n-1} \left(j \frac{\mathrm{Cor}}{1-\mathrm{Cor}} + \mathrm{BER} \right)}{\prod_{j=0}^{n-1} \left(1 + j \frac{\mathrm{Cor}}{1-\mathrm{Cor}} \right)}. \tag{40}$$

- For blocklength $n = M+1$,

$$\begin{aligned} & \Pr^{(M+1)}(Z^n = z^n) \stackrel{\triangle}{=} L^{(M+1)} \\ = & \frac{\prod_{j=0}^{M+1-d_1^{M+1}-1} \left[j \frac{\mathrm{Cor}}{1-\mathrm{Cor}} + (1 - \mathrm{BER}) \right] \prod_{j=0}^{d_1^{M+1}-1} \left(j \frac{\mathrm{Cor}}{1-\mathrm{Cor}} + \mathrm{BER} \right)}{\prod_{j=0}^{M} \left(1 + j \frac{\mathrm{Cor}}{1-\mathrm{Cor}} \right)} \\ = & L^{(M)} \left(\frac{d_1^M \frac{\mathrm{Cor}}{1-\mathrm{Cor}} + \mathrm{BER}}{1 + M \frac{\mathrm{Cor}}{1-\mathrm{Cor}}} \right)^{z_{M+1}} \left[\frac{\left(M - d_1^M \right) \frac{\mathrm{Cor}}{1-\mathrm{Cor}} + (1 - \mathrm{BER})}{1 + M \frac{\mathrm{Cor}}{1-\mathrm{Cor}}} \right]^{1-z_{M+1}}. \end{aligned} \tag{41}$$

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

$$\begin{aligned} & \Pr^{(M+1)}(Z^n = z^n) \\ = & L^{(M+1)} \prod_{i=M+2}^{n} \left(\frac{d_{i-M}^{i-1} \frac{\mathrm{Cor}}{1-\mathrm{Cor}} + \mathrm{BER}}{1 + M \frac{\mathrm{Cor}}{1-\mathrm{Cor}}} \right)^{z_i} \left[\frac{\left(M - d_{i-M}^{i-1} \right) \frac{\mathrm{Cor}}{1-\mathrm{Cor}} + (1 - \mathrm{BER})}{1 + M \frac{\mathrm{Cor}}{1-\mathrm{Cor}}} \right]^{1-z_i} \\ = & L^{(M)} \prod_{i=M+1}^{n} \left(\frac{d_{i-M}^{i-1} \frac{\mathrm{Cor}}{1-\mathrm{Cor}} + \mathrm{BER}}{1 + M \frac{\mathrm{Cor}}{1-\mathrm{Cor}}} \right)^{z_i} \left[\frac{\left(M - d_{i-M}^{i-1} \right) \frac{\mathrm{Cor}}{1-\mathrm{Cor}} + (1 - \mathrm{BER})}{1 + M \frac{\mathrm{Cor}}{1-\mathrm{Cor}}} \right]^{1-z_i}. \end{aligned} \tag{42}$$

By comparing (40) with (37), and comparing (41) and (42) with (38), 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 2 The capacity $C_{\mathrm{QBC}}^{(M)}$ of the QBC is strictly increasing in M 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 $C_{\text{QBC}}^{(M)}(\alpha)$ denote the capacity of the QBC. Thus, for $0 < \alpha < 1$, we have

$$\begin{aligned} C_{\text{QBC}}^{(M)}(\alpha) &< C_{\text{QBC}}^{(M)}(1) \ \text{(by Theorem 1)} \\ &= C_{\text{QBC}}^{(M+1)}(0) \ \text{(by Lemma 2)} \\ &< C_{\text{QBC}}^{(M+1)}(\alpha) \ \text{(by Theorem 1).} \end{aligned} \tag{43}$$

When $\alpha = 0$, we have

$$C_{\text{QBC}}^{(M)}(0) < C_{\text{QBC}}^{(M)}(1) = C_{\text{QBC}}^{(M+1)}(0). \tag{44}$$

When $\alpha = 1$, we have

$$C_{\text{QBC}}^{(M)}(1) = C_{\text{QBC}}^{(M+1)}(0) < C_{\text{QBC}}^{(M+1)}(1). \tag{45}$$

Thus, the capacity $C_{\text{QBC}}^{(M)}$ of the QBC is strictly increasing in M for fixed BER, Cor and $0 \leq \alpha \leq 1$. □

Observation: Note that for $\alpha > 1$, capacity may *not increase* with M (see Fig. 4). An interesting open problem is the question of how does increasing memory affect the capacity of a channel model in a given class of models with memory (e.g., the class of unrestricted Mth order Markov channels) and for providing sufficient conditions so that increasing memory increases capacity, when comparing two models within the class. Theorem 2 does address this problem by providing such sufficient conditions for the class of QBC models.

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 [1], the GEC [25] and a particular symmetric class of the Fritchman channel [16] 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 [1]; the channel's BER and Cor are as follows: BER$= \rho$ and Cor$= \delta/(1+\delta)$. By comparing the UQBC with the FMCC in terms of block transition probability, we obtain the following result.

Theorem 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 [1]:

- For blocklength $n \leq M$,

$$\Pr(Z^n = z^n) \quad = \quad \frac{\prod_{j=0}^{d-1} \left(\mathrm{BER} + j\frac{\mathrm{Cor}}{1-\mathrm{Cor}}\right) \prod_{j=0}^{n-1-d} \left[(1-\mathrm{BER}) + j\frac{\mathrm{Cor}}{1-\mathrm{Cor}}\right]}{\prod_{j=1}^{n-1} \left(1 + j\frac{\mathrm{Cor}}{1-\mathrm{Cor}}\right)}, \tag{46}$$

  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 $n \geq M + 1$,

$$\Pr(Z^n = z^n) \quad = \quad L \prod_{i=M+1}^{n} \left(\frac{\mathrm{BER} + \lambda_{i-1}\frac{\mathrm{Cor}}{1-\mathrm{Cor}}}{1 + M\frac{\mathrm{Cor}}{1-\mathrm{Cor}}}\right)^{z_i} \left[\frac{(1-\mathrm{BER}) + (M-\lambda_{i-1})\frac{\mathrm{Cor}}{1-\mathrm{Cor}}}{1 + M\frac{\mathrm{Cor}}{1-\mathrm{Cor}}}\right]^{1-z_i}, \tag{47}$$

  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 (37) with (46) and (38) with (47), 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 (11)). 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 and [1], the following asymptotic expression for $C_{\mathrm{UQBC}}^{(M)}$ can be established as M approaches infinity while keeping BER and Cor fixed:

$$\lim_{M\to\infty} C_{\mathrm{UQBC}}^{(M)} = 1 - \int_0^1 h_b(z) f_Z(z) dz, \tag{48}$$

where $h_b(\cdot)$ is the binary entropy function and $f_Z(z)$ is the beta probability density function (pdf) $\beta_{u,v}(z)$ [31], [15, p. 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 2, we note that (48) is an upper bound of the capacity of the UQBC for a given M.

Corollary 1 For the same M, BER and Cor,

$$C_{\mathrm{QBC}}^{(M)} < C_{\mathrm{FMCC}}^{(M)} \quad (\text{when } 0 \le \alpha < 1), \tag{49}$$

and

$$C_{\mathrm{QBC}}^{(M)} > C_{\mathrm{FMCC}}^{(M)} \quad (\text{when } \alpha > 1). \tag{50}$$

Proof (49) and (50) can be obtained directly from Theorems 1 and 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 [25] based on increasingly tight upper and lower bounds (c.f., also the theoretical study in [21]). In particular, it is shown in [25] that the GEC capacity can be obtained by $C_{\mathrm{GEC}} =$

$\lim_{l\to\infty} \underline{C}_l$, where $\underline{C}_l$ is defined as $\underline{C}_l \triangleq 1 - E[h_b(q_l)]$. 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., $q_l(Z^{l-1}) \triangleq \Pr[Z_l = 1 \mid Z^{l-1}]$, 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., $C_{\text{GEC}} \geq \underline{C}_l$; as l increases, the bound becomes sharper. The following recursion holds for q_l [25]:

$$q_{l+1}(Z^l) = \nu(Z_l, q_l(Z^{l-1})), \tag{51}$$

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

$$\nu(0,q) \triangleq \begin{cases} p_G + b(p_B - p_G) + \mu(q - p_G)[(1-p_B)/(1-q)], & p_B \neq 1, \\ (1-b)p_G + b, & p_B = 1,\ q \neq 1, \end{cases} \tag{52}$$

and

$$\nu(1,q) \triangleq \begin{cases} p_G + b(p_B - p_G) + \mu(q - p_G)(p_B/q), & p_G \neq 0, \\ (1-g)p_B, & p_G = 0,\ q \neq 0, \end{cases} \tag{53}$$

for $p_G \leq q \leq p_B$, where $\mu \triangleq 1 - g - b$. The initial value for the recursion is

$$q_1 \triangleq \Pr[Z_1 = 1] = (\rho p_G + p_B)(\rho + 1)^{-1}, \tag{54}$$

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

$$\Pr[q_{l+1} = \alpha \mid q_l = \beta] = \begin{cases} 1-\beta, & \alpha = \nu(0,\beta), \\ \beta, & \alpha = \nu(1,\beta). \end{cases} \tag{55}$$

By comparing $\underline{C}_2$ with the capacity of the UQBC with $M = 1$, $C_{\text{UQBC}}^{(M=1)}$, we arrive at the following result.[5]

Theorem 4 For $M = 1$, and for the same BER and Cor, $C_{\text{GEC}} \geq C_{\text{UQBC}}^{(M=1)}$.

Proof In terms of BER and Cor, the initial value for the random sequence q_l (54) is

$$q_1 = (\rho p_G + p_B)(\rho + 1)^{-1} = BER. \tag{56}$$

From the recursion (51), the function $\nu(\cdot,\cdot)$ in (52) and (53), and the transition probability (55), we obtain

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

and

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

[5]Note that the UQBC with $M = 1$ is indeed a special case of the GEC, realized by setting $p_G = 0$ and $p_B = 1$.

where the GEC Cor is given by

$$\text{Cor} = \frac{\mu(\text{BER} - p_G)(p_B - \text{BER})}{\text{BER}(1 - \text{BER})}. \tag{57}$$

Thus,

$$\begin{aligned} \underline{C}_2 &= 1 - E[h_b(q_2)] \\ &= 1 - (\text{BER} \cdot h_b[\text{BER} + (1 - \text{BER}) \cdot \text{Cor}] + (1 - \text{BER}) \cdot h_b[\text{BER} \cdot (1 - \text{Cor})]) . \end{aligned} \tag{58}$$

With memory order $M = 1$ in the UQBC, from (39) it follows that

$$\begin{aligned} C_{\text{UQBC}}^{(M=1)} &= 1 - \sum_{\omega=0}^{1} L_{\omega}^{(1)} h_b[\omega \cdot \text{Cor} + (1 - \text{Cor}) \cdot \text{BER}] \\ &= 1 - (\text{BER} \cdot h_b[\text{Cor} + (1 - \text{Cor}) \cdot \text{BER}] + (1 - \text{BER}) \cdot h_b[\text{BER} \cdot (1 - \text{Cor})]) . \end{aligned} \tag{59}$$

Comparing (58) with (59), 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

$$\boldsymbol{P}_{(K,1)\text{-SFC}} = \begin{bmatrix} p_{00} & (1-p_{00})/K & \cdots & (1-p_{00})/K \\ (1-p_{00})/K & p_{00} & \cdots & (1-p_{00})/K \\ & \vdots & & \\ (1-p_{00})/K & \cdots & p_{00} & (1-p_{00})/K \\ (1-p_{11})/K & \cdots & (1-p_{11})/K & p_{11} \end{bmatrix}, \tag{60}$$

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 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 big good state with stationary probability

(1−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 2) and α increases capacity (Theorem 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 [25] 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 [16], a general expression for the capacity of Fritchman channels with a single error-state and K good states ($(K,\,1)$-FC) is provided:

$$\begin{aligned} C_{(K,1)\text{-FC}} \;=\; & 1 - \Bigg[- \Pr(1) \sum_{z_n=0}^{1} \Pr(z_n \mid 1) \log_2 \Pr(z_n \mid 1) \\ & - \sum_{m=1}^{\infty} \left[\Pr(10^m) \sum_{z_n=0}^{1} \Pr(z_n \mid 10^m) \log_2 \Pr(z_n \mid 10^m) \right] \Bigg], \end{aligned} \tag{61}$$

where 10^m denotes the binary sequence with a single 1 and m 0s, and the probabilities and the conditional probabilities of error sequences can be calculated via (2). 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.

We illustrate the effect of the cell parameters M and α on the capacity of the QBC in Figs. 4 and 5. As expected from Theorems 1 and 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 2). For example in Fig. 4, when $\alpha \geq 10$, the situation reverses and capacity decreases with the memory order. We also illustrate in Fig. 5 the adage "memory increases capacity" (for information stable channels [1]) 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. 6, the QBC with $M = 2$ and $\alpha = 10$ has the biggest 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 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. 6).

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 BER_{GEC} and correlation coefficient Cor_{GEC}, we estimate the QBC parameters M, p, ε, and α that minimize the KLDR

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

subject to the constraints

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

and

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

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

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

where $\mathrm{P}_{\mathrm{GEC}}(z^n)$ is given by (1) and $\mathrm{P}_{\mathrm{QBC}}^{(M)}(z^n)$ is given by (6)-(7). It can be shown (e.g., see [20]) that the Kullback-Leibler divergence rate 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

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

(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.

4.2 Modeling Results

We evaluate how well the QBC model (obtained via the above KDLR minimization) fits or approximates the GEC according to three criteria: channel capacity, ACF,[6] and the RCLB bound for the error exponent. The QBC capacity and ACF expressions are provided in Section 3. Although the capacity of the GEC does not have an analytical expression, it can be determined accurately via the algorithm of [25]. The ACF of the GEC can also be computed [30].

We employ Propositions 1 and 2 to obtain both the RCLB and SPUB on $E(R)$ for the QBC. For the GEC, we compute the RCLB bound using Proposition 1. However, with the exception of the special GC case, $E_0^{(\infty)}(\rho)$ is not known to exist (to our knowledge) for the general GEC since the channel state is not a deterministic function of the previous state and noise 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., $p(k) = P(0^{k-1}1 \mid 1)$. Therefore, we can use the following result (e.g., [13]) along with Proposition 1 to determine the RCLB on $E(R)$ for the GC.

Proposition 3 For a channel with errors fully defined by $p(k)$ (such as the GC), $E_0^{(\infty)}(\rho)$ exists and is given by

$$E_0^{(\infty)}(\rho) = \rho + (1+\rho)\gamma, \tag{62}$$

[6] It is indeed observed in [36, 30] 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.

where γ is the solution of

$$\sum_{i=1}^{\tau} p(i)^{\frac{1}{1+\rho}} 2^{i\gamma} = 1, \tag{63}$$

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

In our results for the RCLB on the error exponent of the GC, we set the value τ to be as large as 70,000. Since for the more general GEC, the lim inf of $E_0^{(n)}(\rho)$ does not admit a computable expression, 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 1 to evaluate the GEC RCLB. Similarly, the GEC critical rate is estimated by computing $R_{cr}^{(n)}$ for large n, where $R_{cr}^{(n)} \triangleq \frac{\partial E_0^{(n)}(\rho)}{\partial \rho}\Big|_{\rho=1}$.

A wide range of GEC/GC channel parameters is investigated with $0.01 \leq \text{Cor} \leq 0.9$ and $0.3\% \leq \text{BER} \leq 10\%$. In Table 1 and Figs. 7-20, 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 1 were chosen to match the conditions of the correlated Rayleigh fading channel studied in [30, Fig. 6.(b)] with normalized Doppler frequencies of 0.1 (fast fading) and 0.001 (slow fading), respectively.

We first observe in Figs. 7 and 8 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. 7-8 are less than 0.01 for all cases, except for Case **G** (see (c) in Fig. 8). As a result, we observe a strong agreement in ACF, capacity and error exponents in Figs. 9-20 (this behaviour 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. 7). In Figs. 9 and 10, 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. 11-18, 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. 11-13 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. 13). Overall, we remark a strong match in capacity between the GEC and its QBC approximation - indeed, even in Fig. 18 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. 19), while the agreement is weaker but still good for high values of Cor (see Fig. 20). Although we do not have an exact RCLB expression for the GEC, the curves of $E_r(R)$ 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 for a wide range of channel conditions.

5 Conclusion

In this paper, we introduced a binary burst-noise channel based on a finite queue of length M. The resulting channel noise process is a stationary ergodic Markov source of order M. We derived the stationary noise distribution, the block transition probability and the capacity of the channel in terms of the QBC parameters M, ε, p and α. The capacity is positive, and strictly increasing in α for fixed M, BER and Cor. When $0 \leq \alpha \leq 1$, the memory order M strictly increases capacity for fixed α, BER and Cor. We also examined a particular case, the UQBC, by choosing $\alpha = 1$. The capacity of the UQBC is positive, increases with M and so is 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 probabilities for the same memory, BER and Cor; hence they have identical capacity 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 $M = 1$, the UQBC capacity ($C_{\text{UQBC}}^{(M=1)}$) 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 order 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 $\alpha = 10$ has the biggest 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 studied the problem of fitting the GEC via the QBC. Numerical results lead us to conclude that, for various channel conditions, the QBC provides a good approximation of the GEC in terms of KLDR, channel capacity, ACF, and error exponent. It hence offers an attractive alternative to the GEC by virtue of its tractability for mathematical analysis as it admits a closed form expression for its capacity and a transparent formula for its n-fold statistics.

An important problem for future work is the judicious construction of practical capacity-achieving channel codes for the QBC that directly operate on the channel (without the use of channel interleaving) and fully exploit the noise memory structure. Initial progress in this direction is achieved in [27], where a belief propagation decoder is designed for the decoding of (regular) LDPC codes over the QBC. It is noted that considerable performance improvement is achievable over the standard decoding strategy for the equivalent memoryless channel (resulting from ideal channel interleaving) which does not exploit the channel memory. The decoder studied in [27] is an extension of the sum-product algorithm (originally designed for memoryless channels) and is similar to the decoding algorithm developed for the GEC (e.g., see [12, 18]).

The next step for future work is to further close the gap vis-à-vis the Shannon limit by constructing LDPC codes optimized for the QBC and analyzing their performance using density evolution. We expect that such study will be substantially facilitated by the QBC's simple statistical structure and information-theoretic properties – such as the closed form expression for channel capacity and its monotonicity with respect to two channel parameters (cf. Theorems 1 and 2). Finally, in light of the demonstrated capability for the QBC to accurately model correlated fading channels [40, 41], once capacity-achieving LDPC codes are successfully designed for the QBC, they can also directly be applied to the fading channels (as fitted by the QBC), thus achieving important performance gains over the current coding methods for fading channels.

A Derivation of the Stationary Distribution of the QBC

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 (5) satisfies the following:

$$\pi_{\frac{i}{2}}^{(M)} = \pi_i^{(M)} \left[\frac{\left(M - \omega_i^{(M)} - 1 + \alpha\right)\varepsilon}{M - 1 + \alpha} + (1-\varepsilon)(1-p) \right] + \pi_{i+1}^{(M)} \left[\frac{\left(M - \omega_i^{(M)} - 1\right)\varepsilon}{M - 1 + \alpha} + (1-\varepsilon)(1-p) \right], \tag{64}$$

and

$$\pi_{\frac{i+2M}{2}}^{(M)} = \pi_i^{(M)} \left[\frac{\omega_i^{(M)}\varepsilon}{M - 1 + \alpha} + (1-\varepsilon)p \right] + \pi_{i+1}^{(M)} \left[\frac{\left(\omega_i^{(M)} + \alpha\right)\varepsilon}{M - 1 + \alpha} + (1-\varepsilon)p \right], \tag{65}$$

which are directly obtained from the transition probabilities (4). First we show that (64) 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,

$$\pi_{i+1}^{(M)} = \pi_i^{(M)} \frac{\omega_i^{(M)} \frac{\varepsilon}{M-1+\alpha} + (1-\varepsilon)p}{(M - \omega_i^{(M)} - 1)\frac{\varepsilon}{M-1+\alpha} + (1-\varepsilon)(1-p)}. \tag{66}$$

Therefore, (64) 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 (65) 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 (66), we see that (65) 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. It is evident that $\sum_{i=0}^{2^M-1} \pi_i^{(M)} = 1$ is satisfied, as (5) has the same form as the stationary distribution of the FMCC channel state given in [1]. Thus, it is verified that (5) is the stationary distribution of the QBC.

Acknowledgment

The original idea of using a (uniformly operated) queue to model channel noise with finite memory was suggested by Rajiv Laroia to Fady Alajaji in 1992.

References

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

[2] 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.

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

[4] 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.

[5] 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.

[6] 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.

[7] 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, vol. 5, pp. 2930–2935, Nov. 1998.

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

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

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

[11] 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.

[12] A. W. Eckford, F. R. Kschischang and S. Pasupathy, "Analysis of low-density parity-check codes for the Gilbert-Elliott channel," *IEEE Trans. Inform. Theory*, vol. 51, no. 11, pp. 3872-3889, Nov. 2005.

[13] V. 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.

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

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

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

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

[18] J. Garcia-Frias, "Decoding of low-density parity-check codes over finite-state binary Markov channels," *IEEE Trans. Commun.*, vol. 52, no. 11, pp. 1840-1843, Nov. 2004.

[19] E. N. Gilbert, "Capacity of a burst-noise channel," *Bell Syst. Tech. J.*, vol. 39, pp. 1253–1266, Sept. 1960.

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

[21] 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.

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

[23] 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.

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

[25] 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.

[26] 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, May 2004.

[27] C. Nicola, F. Alajaji and T. Linder, "Decoding LDPC codes over binary channels with additive Markov noise," *Proc. Canadian Workshop Inform. Theory*, pp. 187-190, Montreal, Canada, June 2005.

[28] 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.

[29] C. Pimentel and I. F. Blake, "Modeling burst channels using partitioned Fritchman's Markov models," *IEEE Trans. Veh. Technol.*, vol. 47, no. 3, pp. 885–899, Aug. 1998.

[30] 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.

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

[32] 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.

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

[34] 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.

[35] 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.

[36] C. Tan and N. C. Beaulieu, "On first-order Markov modeling for the Rayleigh fading channel," *IEEE Trans. Commun.,* vol. 48, no. 12, pp. 2032–2040, Dec. 2000.

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

[38] 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.

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

[40] L. Zhong, *A Binary Burst-Noise Communication Channel Modeled by a Finite Queue: Information Theoretic Properties and Applications to Wireless Communications*, Ph.D. Dissertation, Queen's University, Kingston, Ontario, Canada, June 2005.

[41] L. Zhong, F. Alajaji and G. Takahara, "A queue-based model for wireless Rayleigh fading channels with memory," *Proc. IEEE 62nd Semiannual Vehicular Technology Conf.*, pp. 1362-1366, Dallas, Texas, Sept. 2005.

[42] Y. Zhong, F. Alajaji and L. L. Campbell, "On the joint source-channel coding error exponent for systems with memory," *Proc. IEEE Trans. Inform. Theory,* pp. 1386-1390, Adelaide, Australia, Sept. 2005.

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 1: GEC fitting via the QBC: GEC and QBC parameters. For the GEC, b and g are determined using (56) and (57); for the GC, $p_G = 0$ and p_B is determined using either (56) or (57).

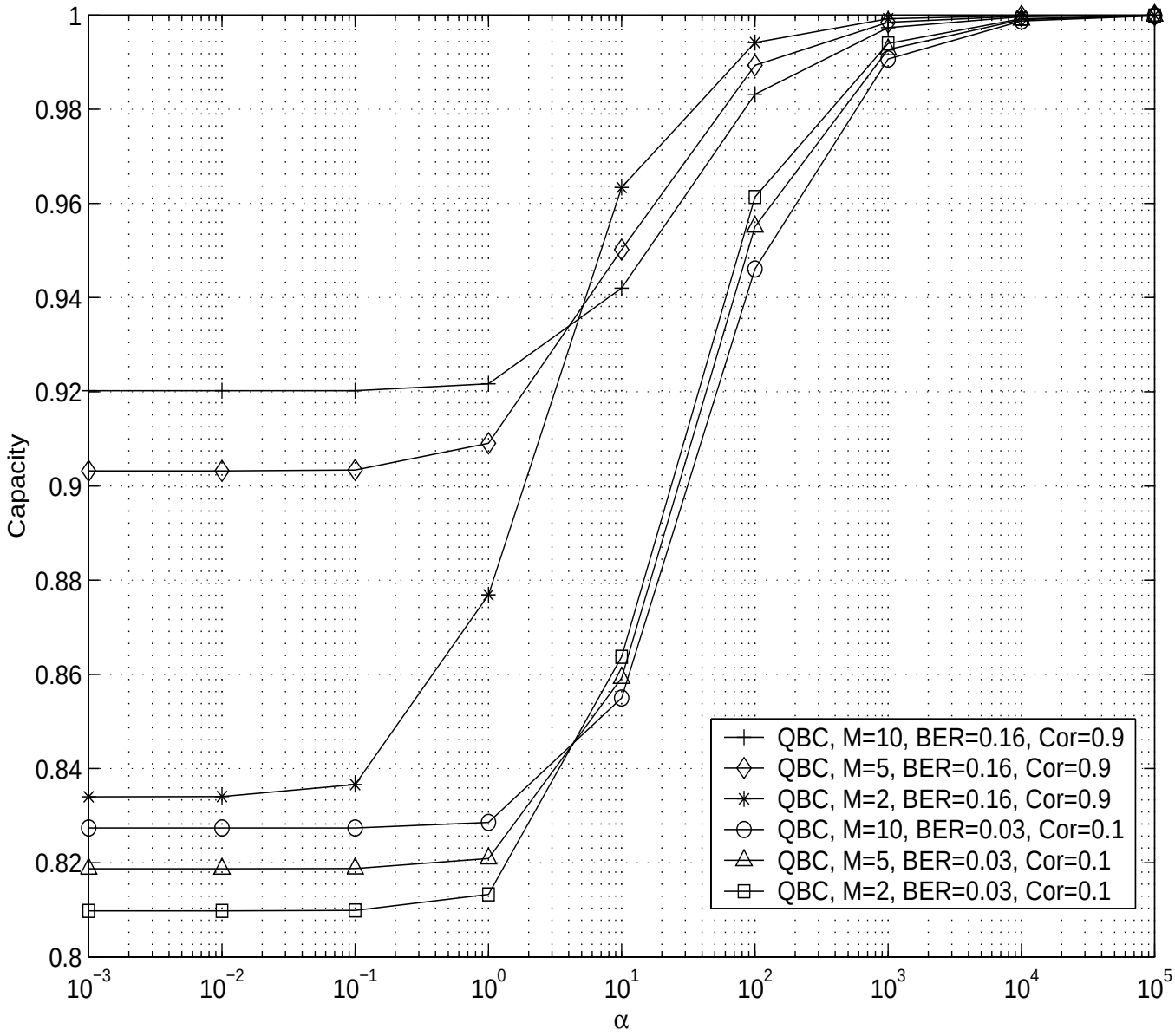

Figure 4: Capacity vs α for QBC.

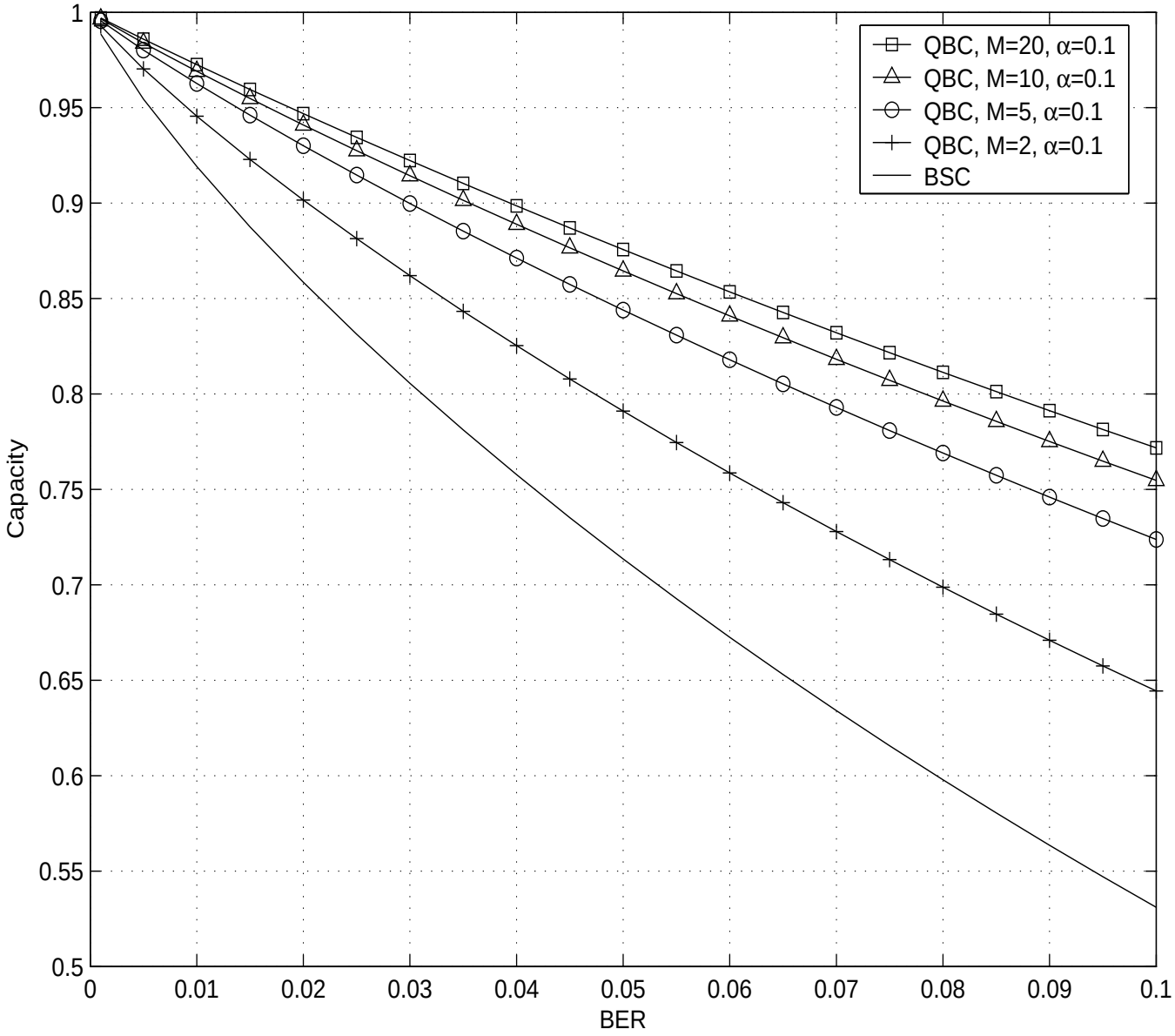

Figure 5: Capacity vs BER for QBC with Cor=0.5.

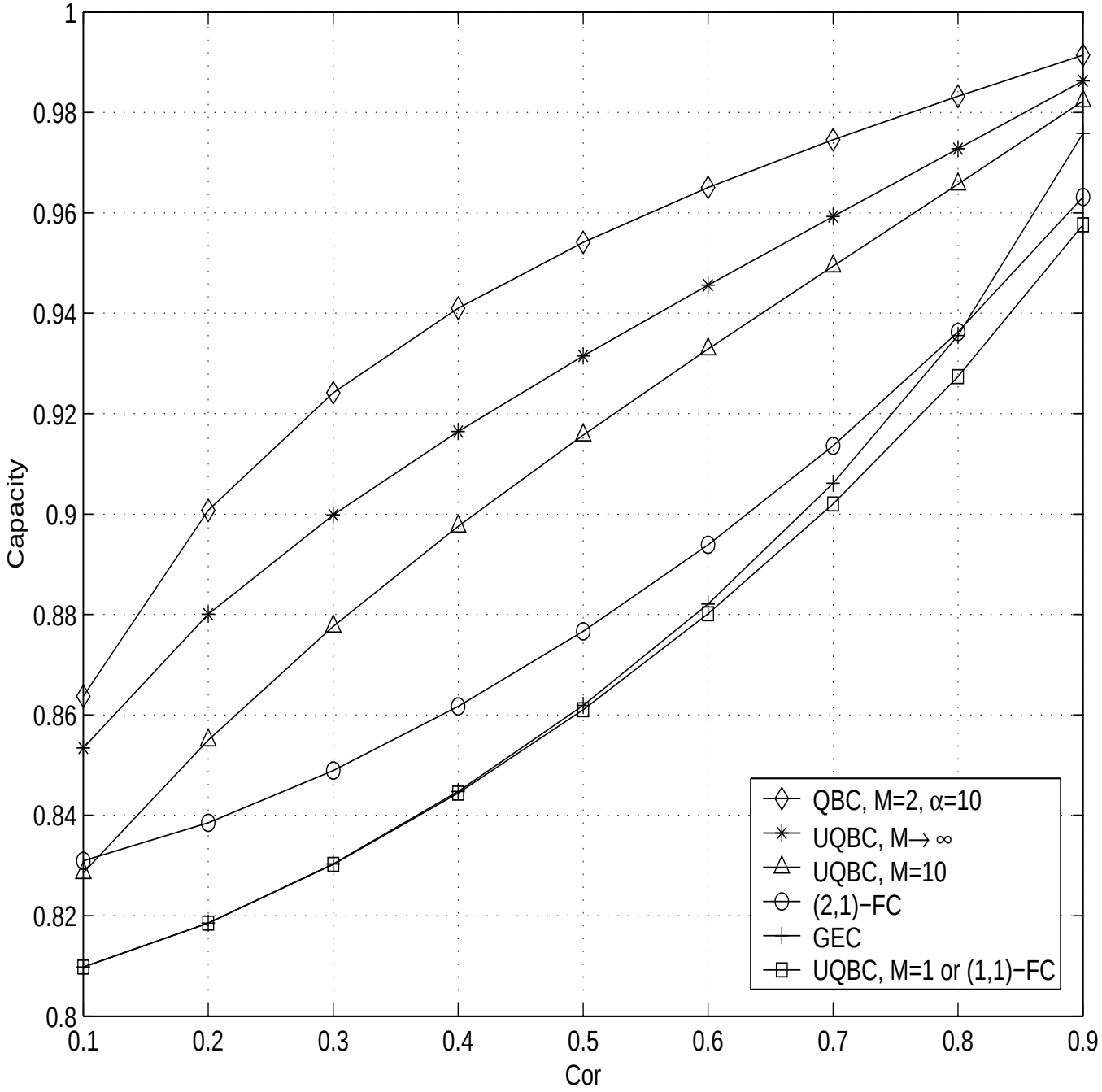

Figure 6: Capacity vs Cor for BER=0.03; $p_G = 0.00002$ and $p_B = 0.92$ (for GEC).

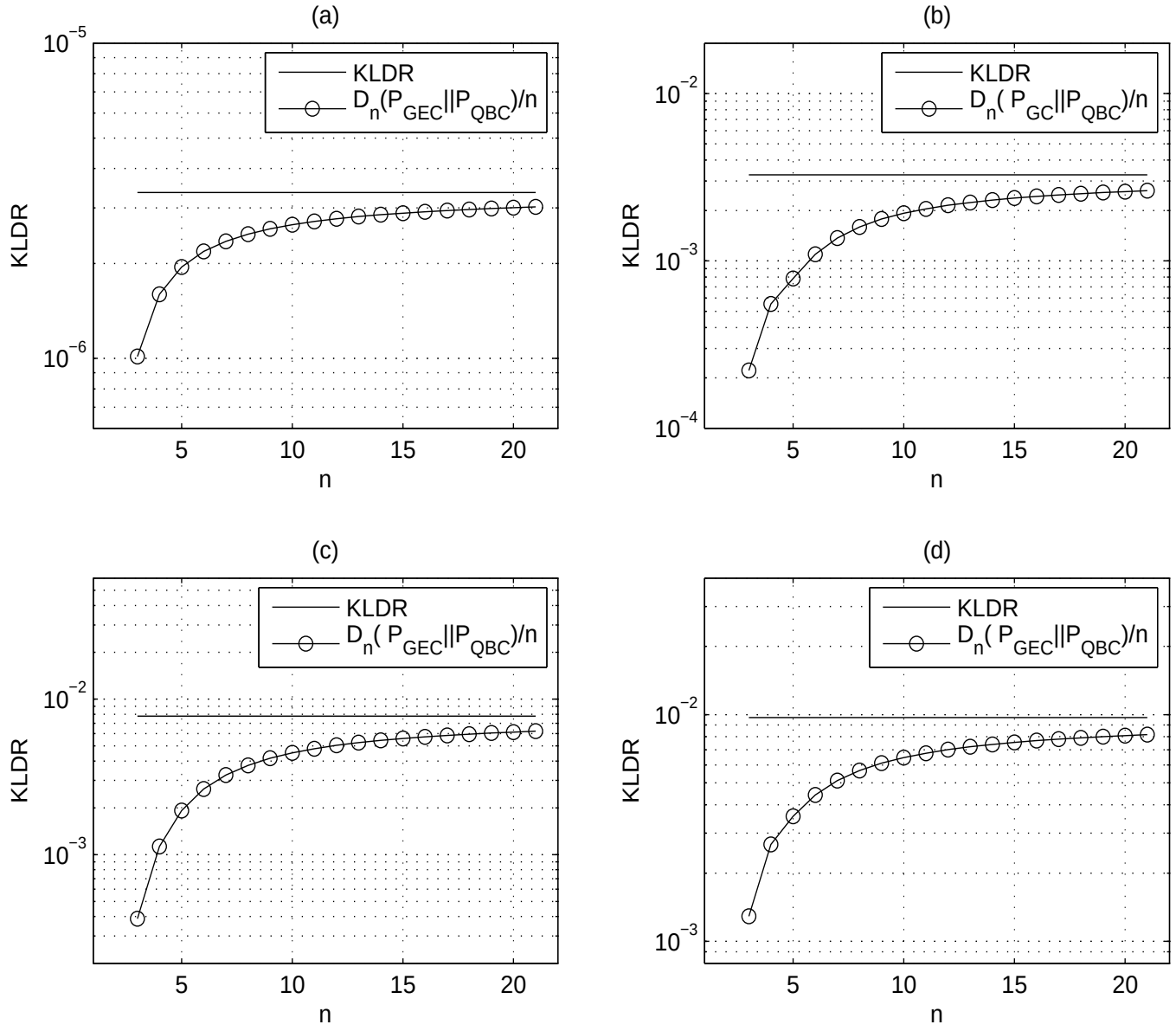

Figure 7: GEC fitting via the QBC: KLDR and $(1/n)D_n(\mathrm{P}_{\mathrm{GEC}} \parallel \mathrm{P}_{\mathrm{QBC}}^{(M)})$ vs n. (a) Cor = 0.0131 and BER = 0.00314 (Case **A** in Table 1); (b) Cor = 0.2227 and BER = 0.03 (Case **B** in Table 1); (c) Cor = 0.248 and BER = 0.012 (Case **C** in Table 1); (d) Cor = 0.341 and BER = 0.1178 (Case **D** in Table 1).

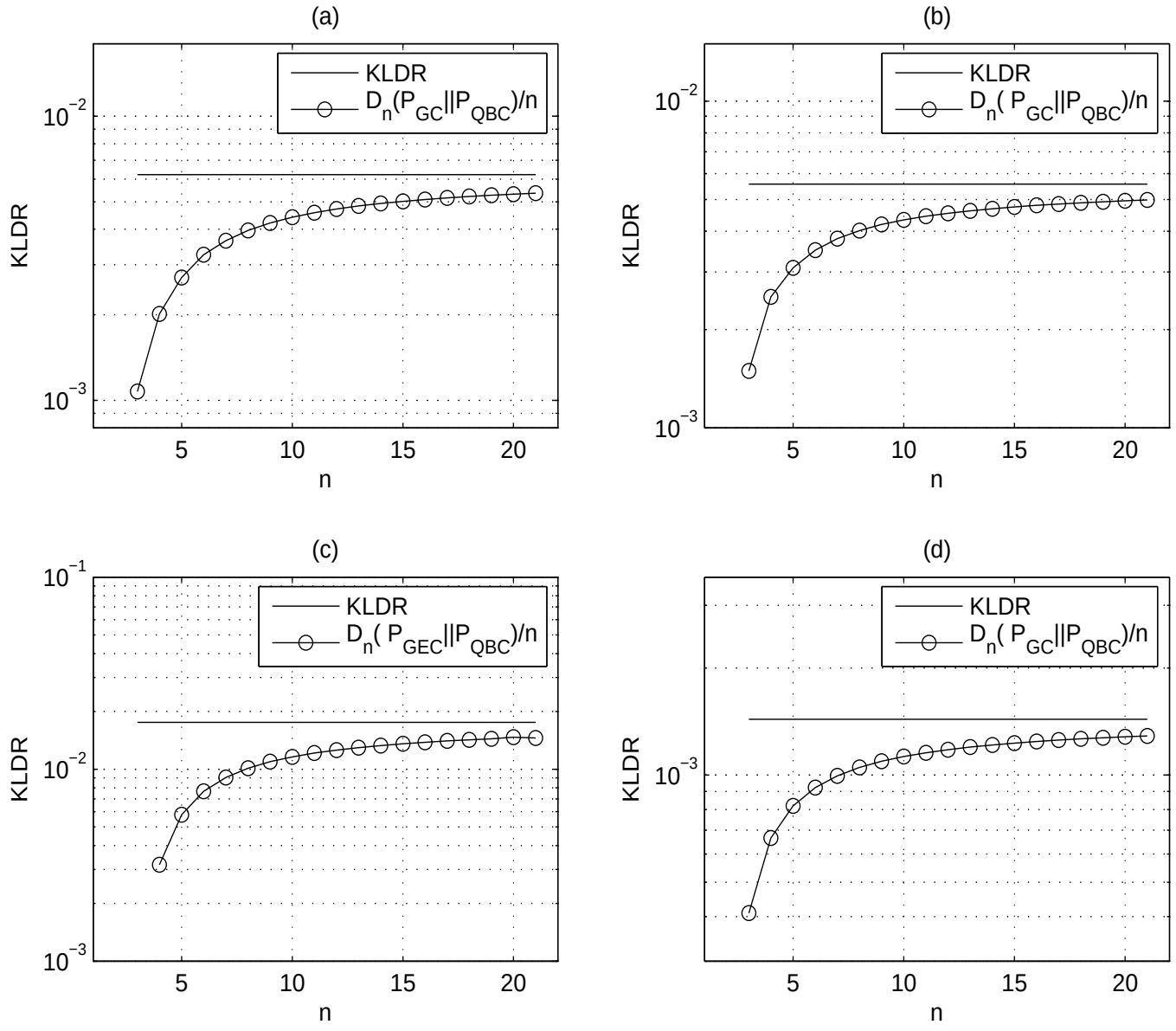

Figure 8: GEC fitting via the QBC: KLDR and $(1/n)D_n(\mathrm{P}_{\mathrm{GEC}} \parallel \mathrm{P}_{\mathrm{QBC}}^{(M)})$ vs n. (a) Cor = 0.5 and BER = 0.03 (Case **E** in Table 1); (b) Cor = 0.7 and BER = 0.03 (Case **F** in Table 1); (c) Cor = 0.8 and BER = 0.1 (Case **G** in Table 1); (d) Cor = 0.9 and BER = 0.03 (Case **H** in Table 1).

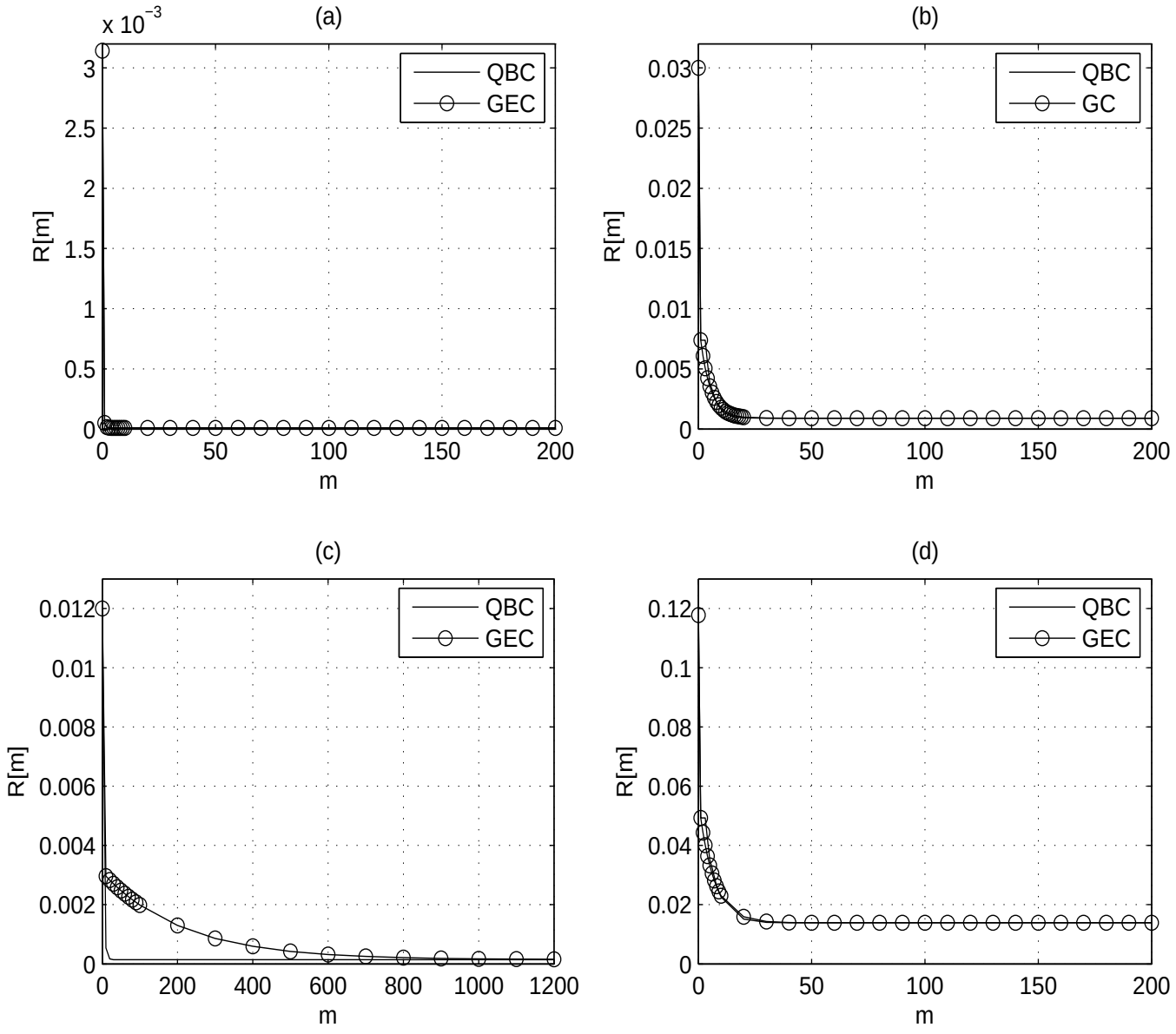

Figure 9: GEC fitting via the QBC: ACF vs m. (a) Cor = 0.0131 and BER = 0.00314 (Case **A** in Table 1); (b) Cor = 0.2227 and BER = 0.03 (Case **B** in Table 1); (c) Cor = 0.248 and BER = 0.012 (Case **C** in Table 1); (d) Cor = 0.341 and BER = 0.1178 (Case **D** in Table 1).

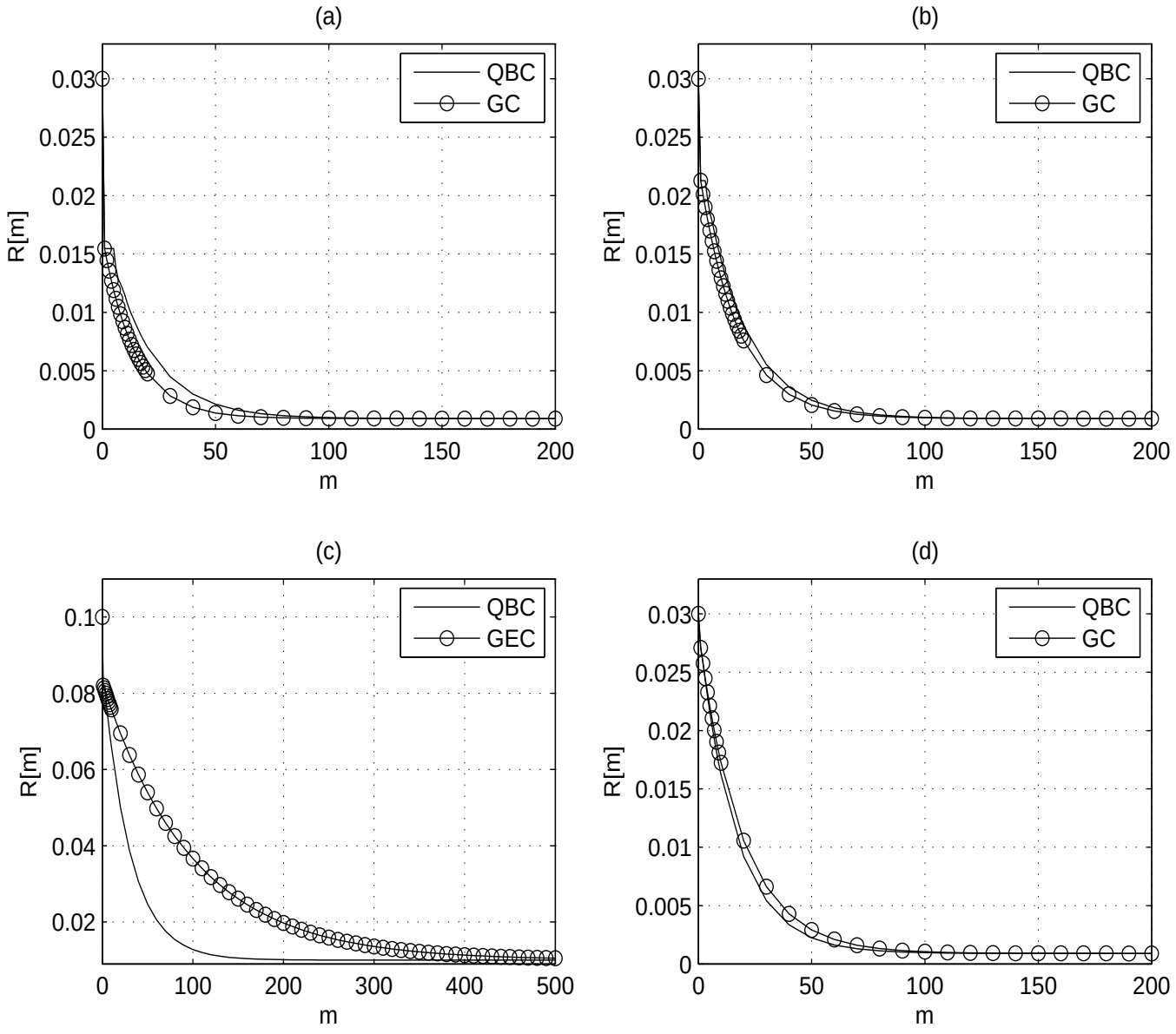

Figure 10: GEC fitting via the QBC: ACF vs m. (a) Cor = 0.5 and BER = 0.03 (Case **E** in Table 1); (b) Cor = 0.7 and BER = 0.03 (Case **F** in Table 1); (c) Cor = 0.8 and BER = 0.1 (Case **G** in Table 1); (d) Cor = 0.9 and BER = 0.03 (Case **H** in Table 1).

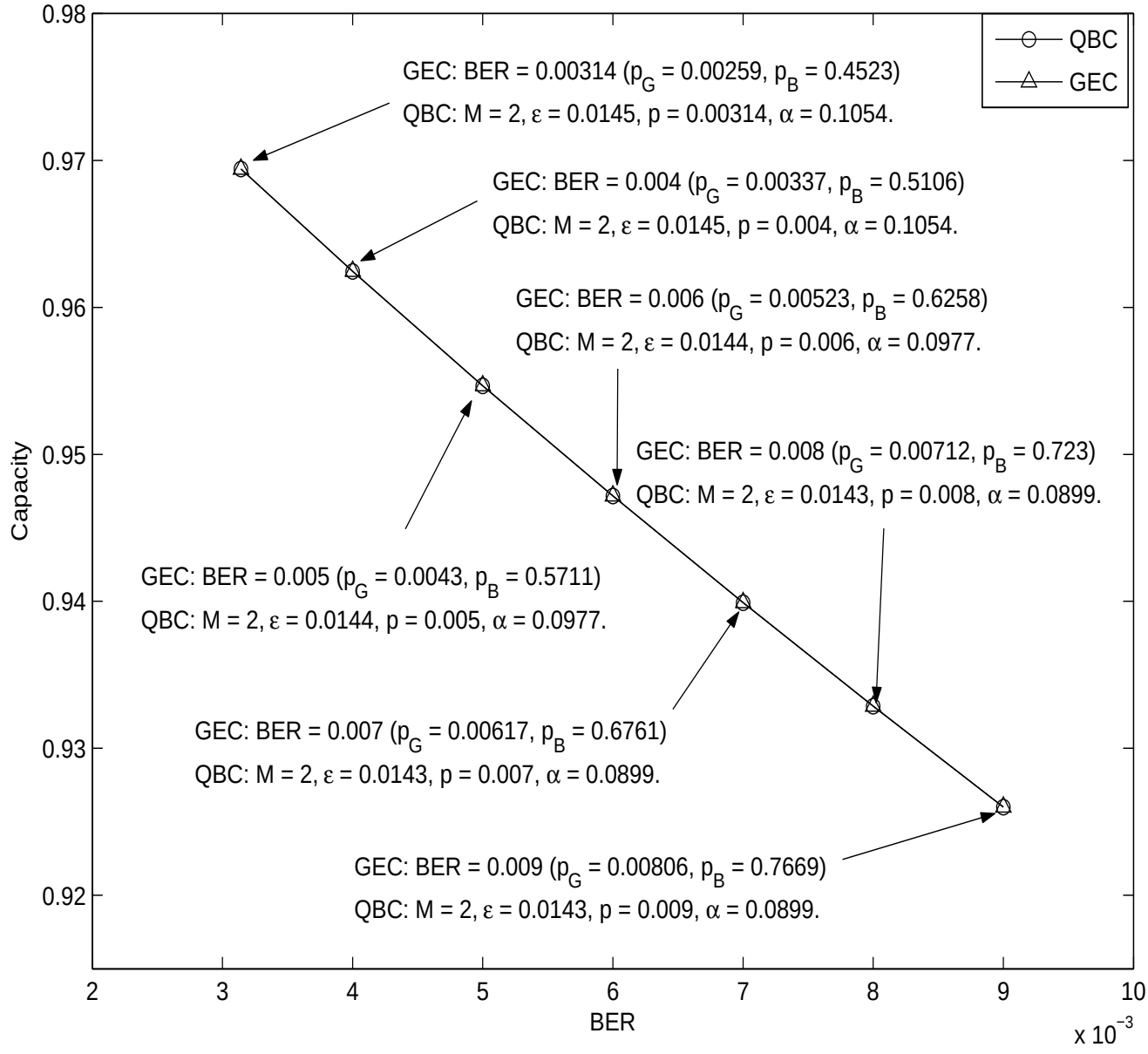

Figure 11: GEC fitting via the QBC: Capacity vs BER for Cor = 0.0131. For the GEC, b and g are determined using (56) and (57).

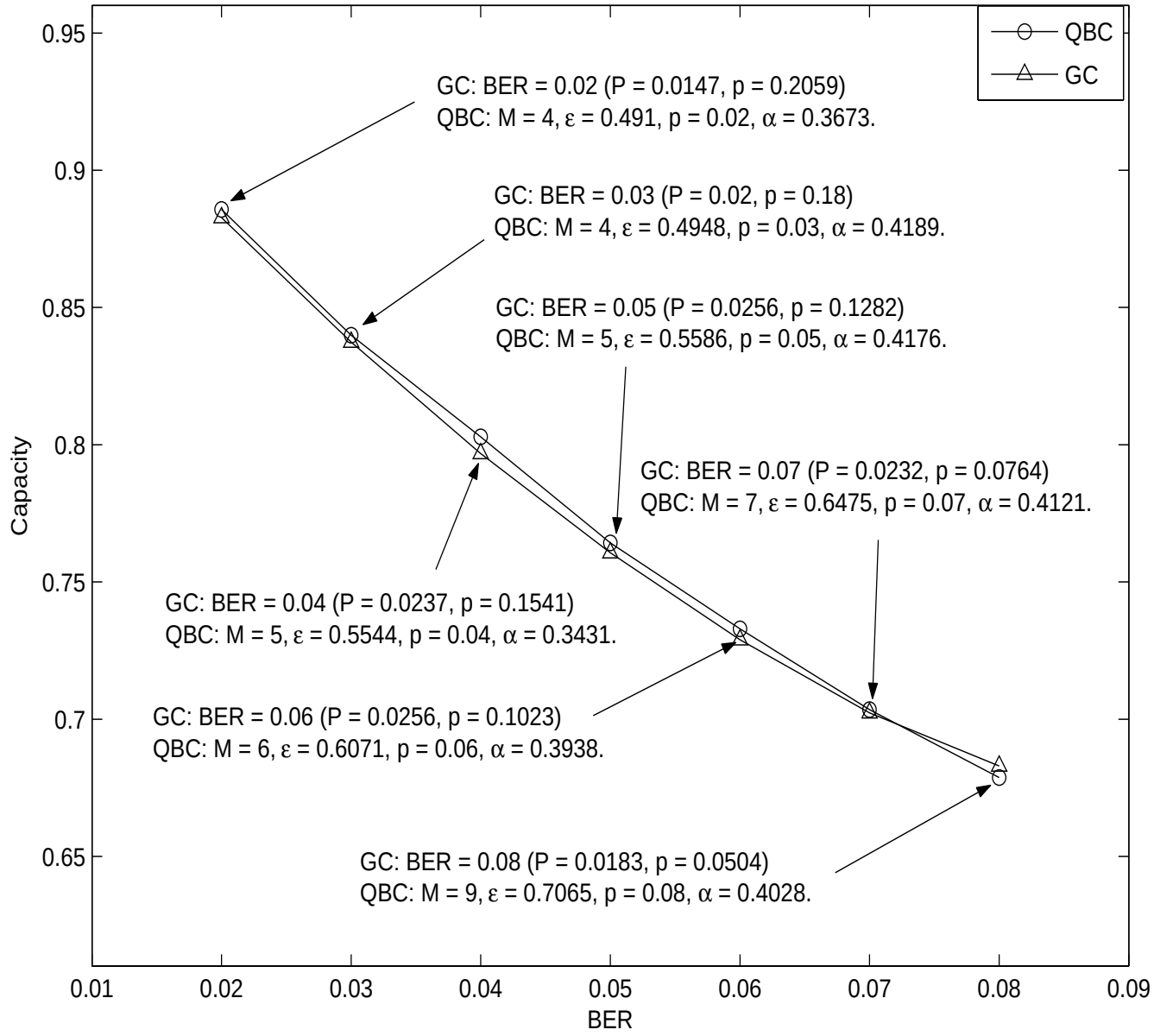

Figure 12: GC fitting via the QBC: Capacity vs BER for Cor = 0.2227. For the GC, $p_G = 0$ and p_B is determined using either (56) or (57).

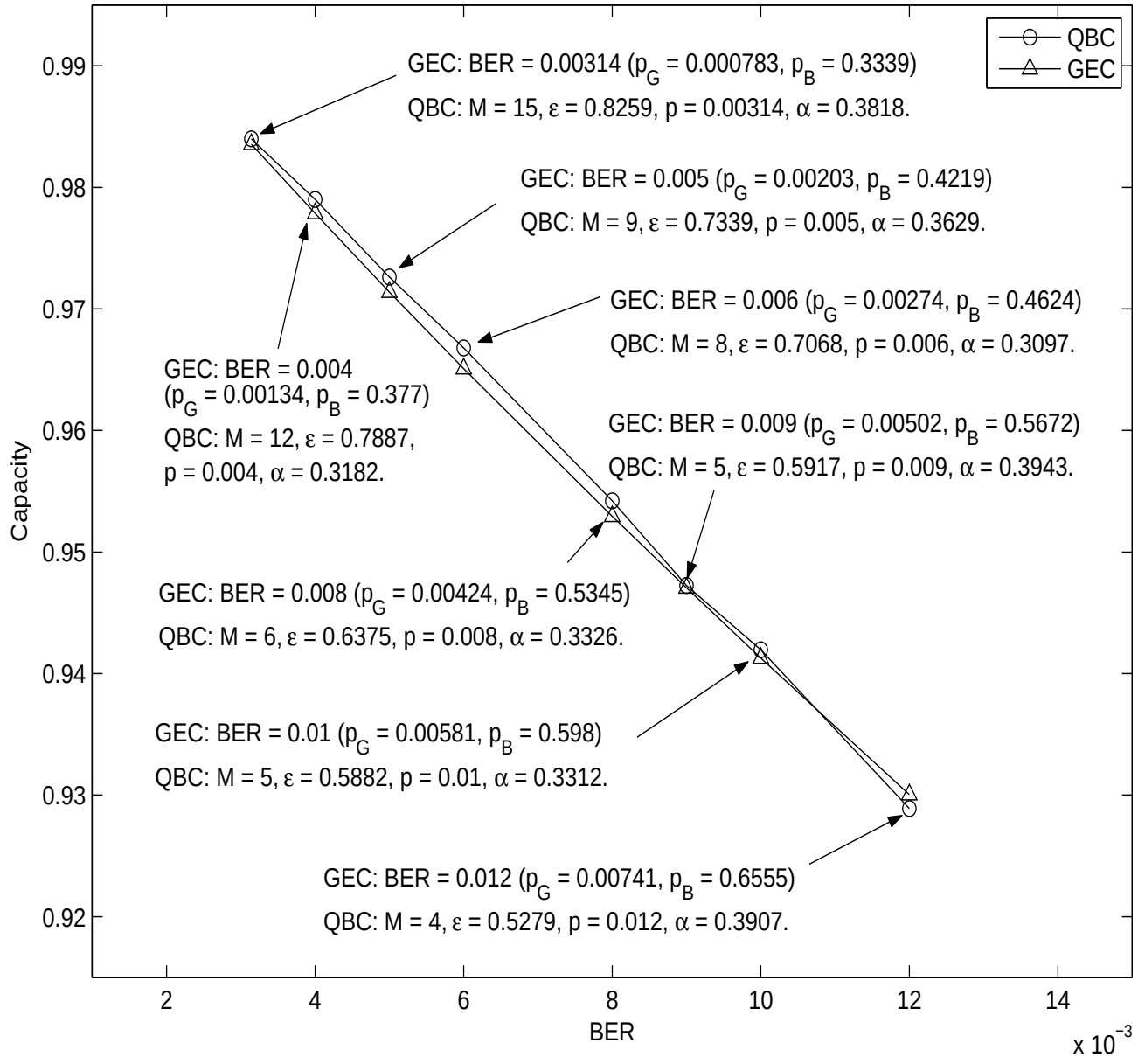

Figure 13: GEC fitting via the QBC: Capacity vs BER for Cor = 0.248. For the GEC, b and g are determined using (56) and (57).

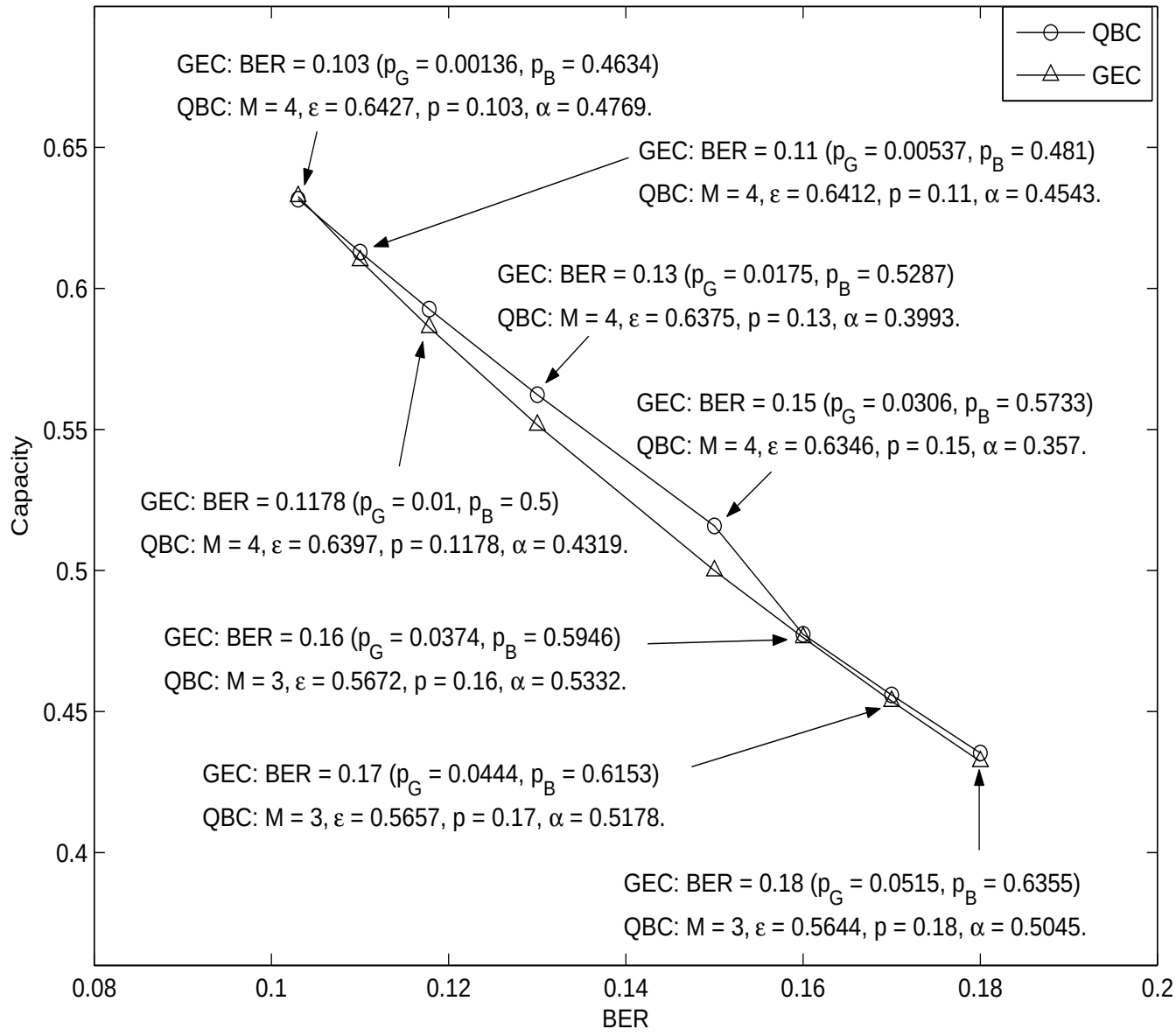

Figure 14: GEC fitting via the QBC: Capacity vs BER for Cor = 0.341. For the GEC, b and g are determined using (56) and (57).

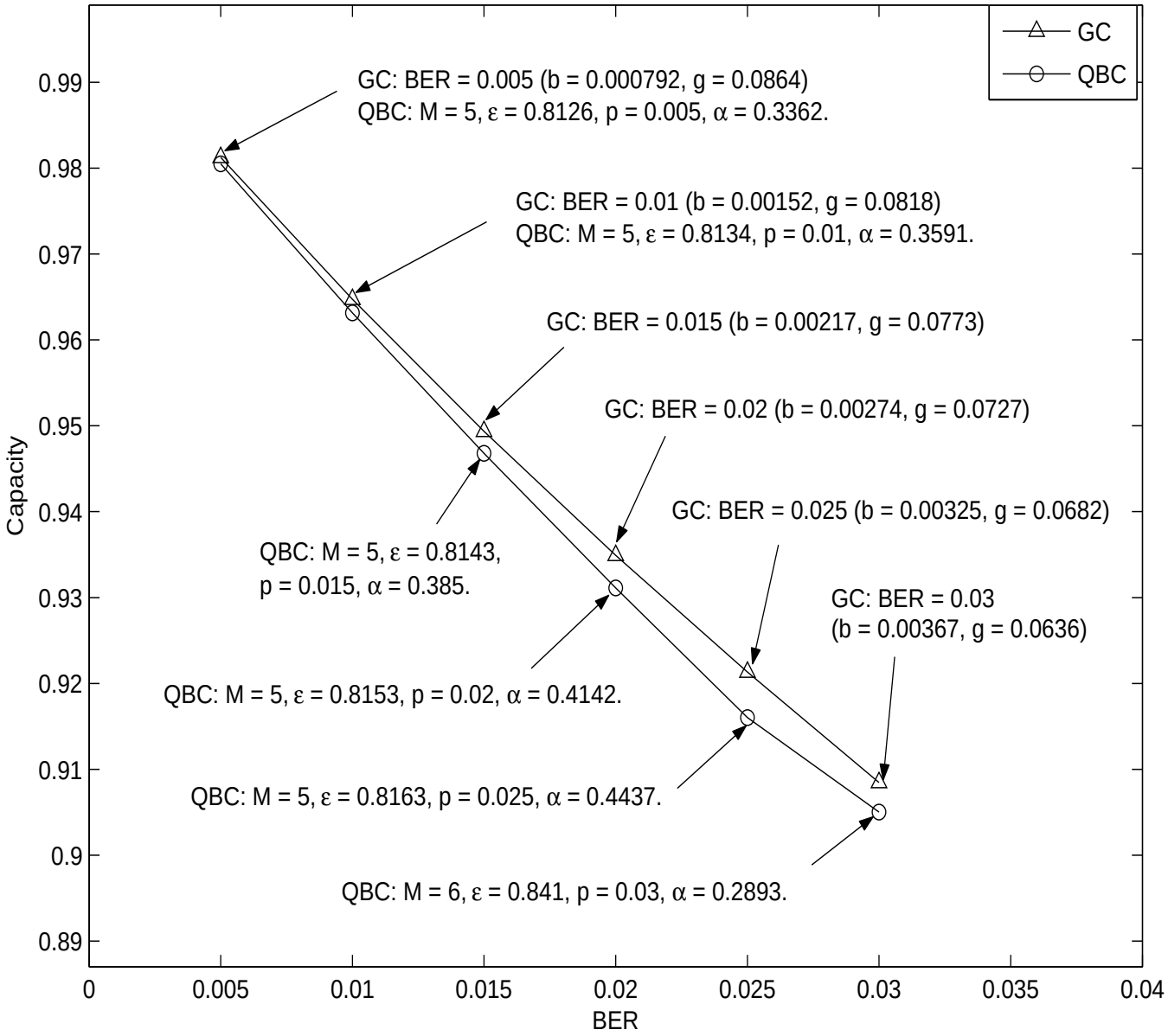

Figure 15: GC fitting via the QBC: Capacity vs BER for Cor = 0.5. For the GC, $p_G = 0$ and p_B is determined using either (56) or (57).

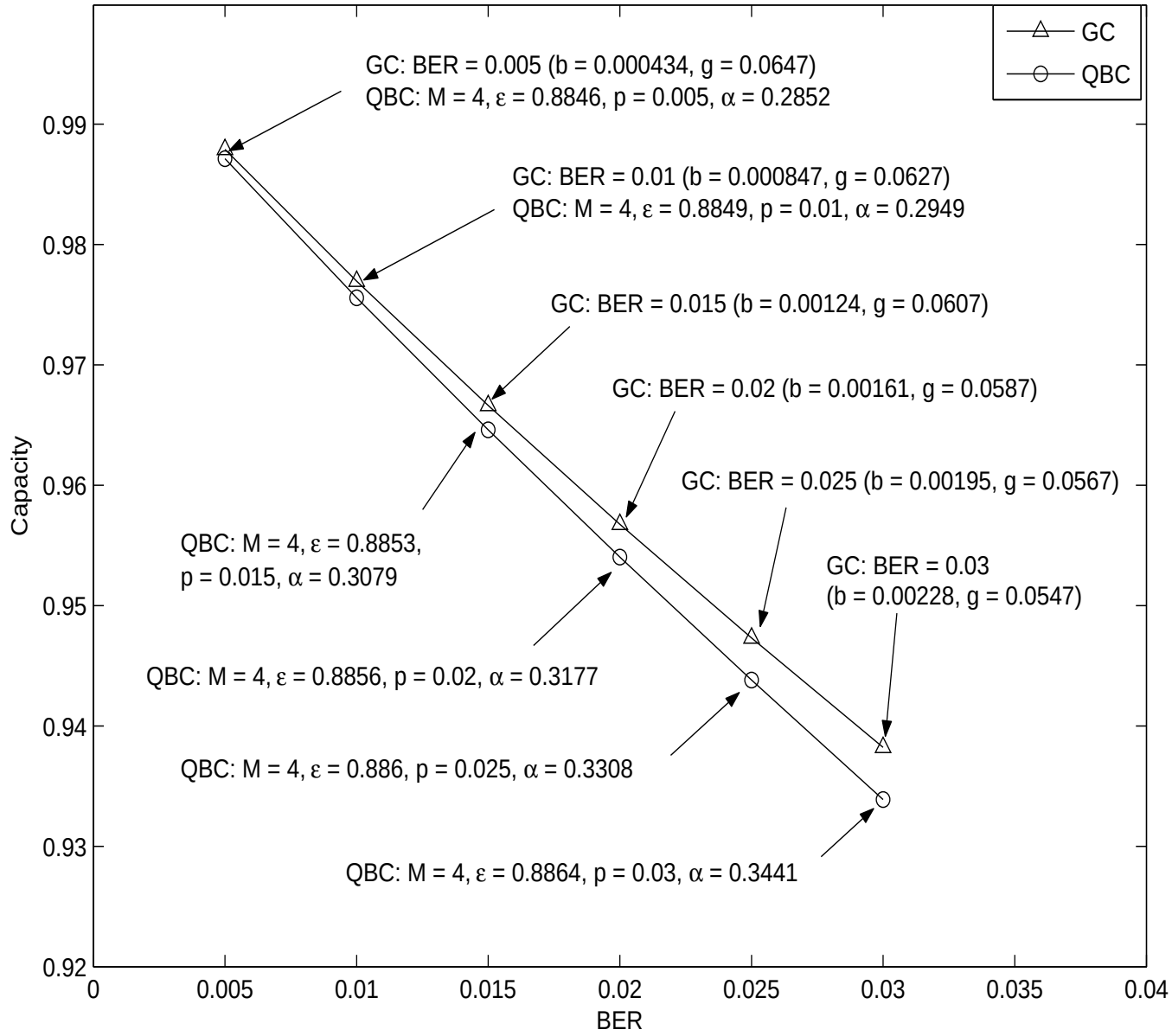

Figure 16: GC fitting via the QBC: Capacity vs BER for Cor = 0.7. For the GC, $p_G = 0$ and p_B is determined using either (56) or (57).

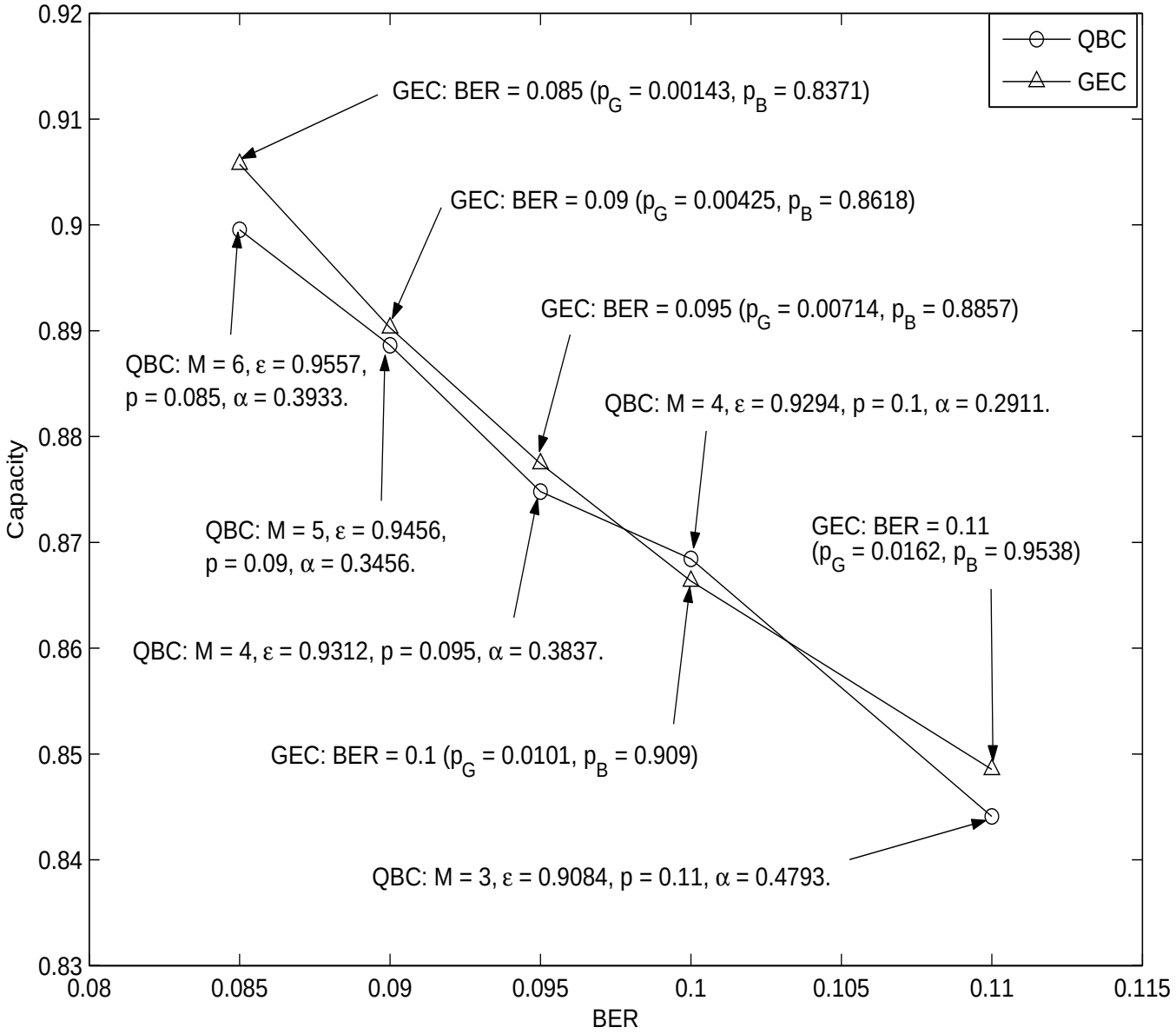

Figure 17: GEC fitting via the QBC: Capacity vs BER for Cor = 0.8. For the GEC, b and g are determined using (56) and (57).

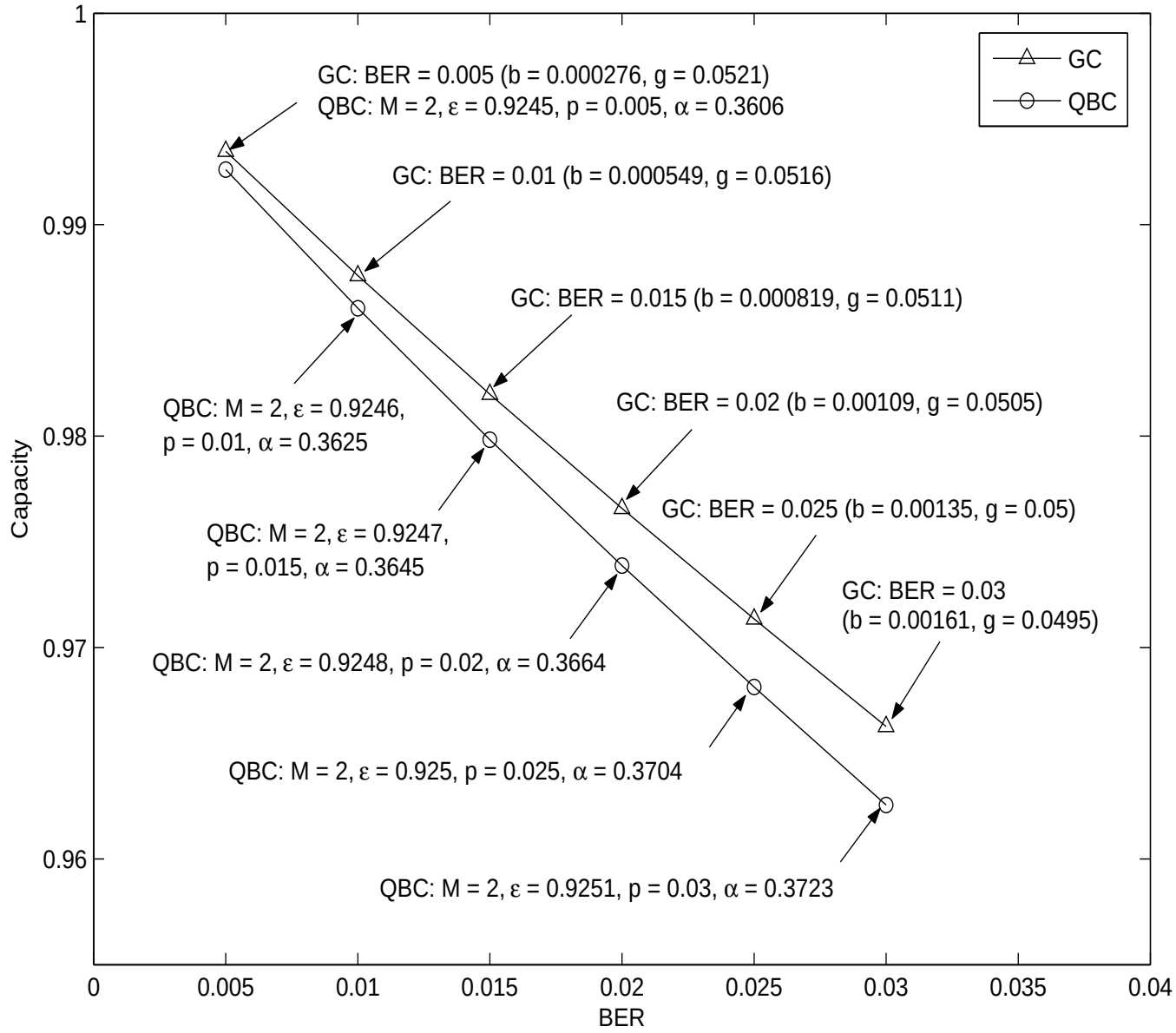

Figure 18: GC fitting via the QBC: Capacity vs BER for Cor = 0.9. For the GC, $p_G = 0$ and p_B is determined using either (56) or (57).

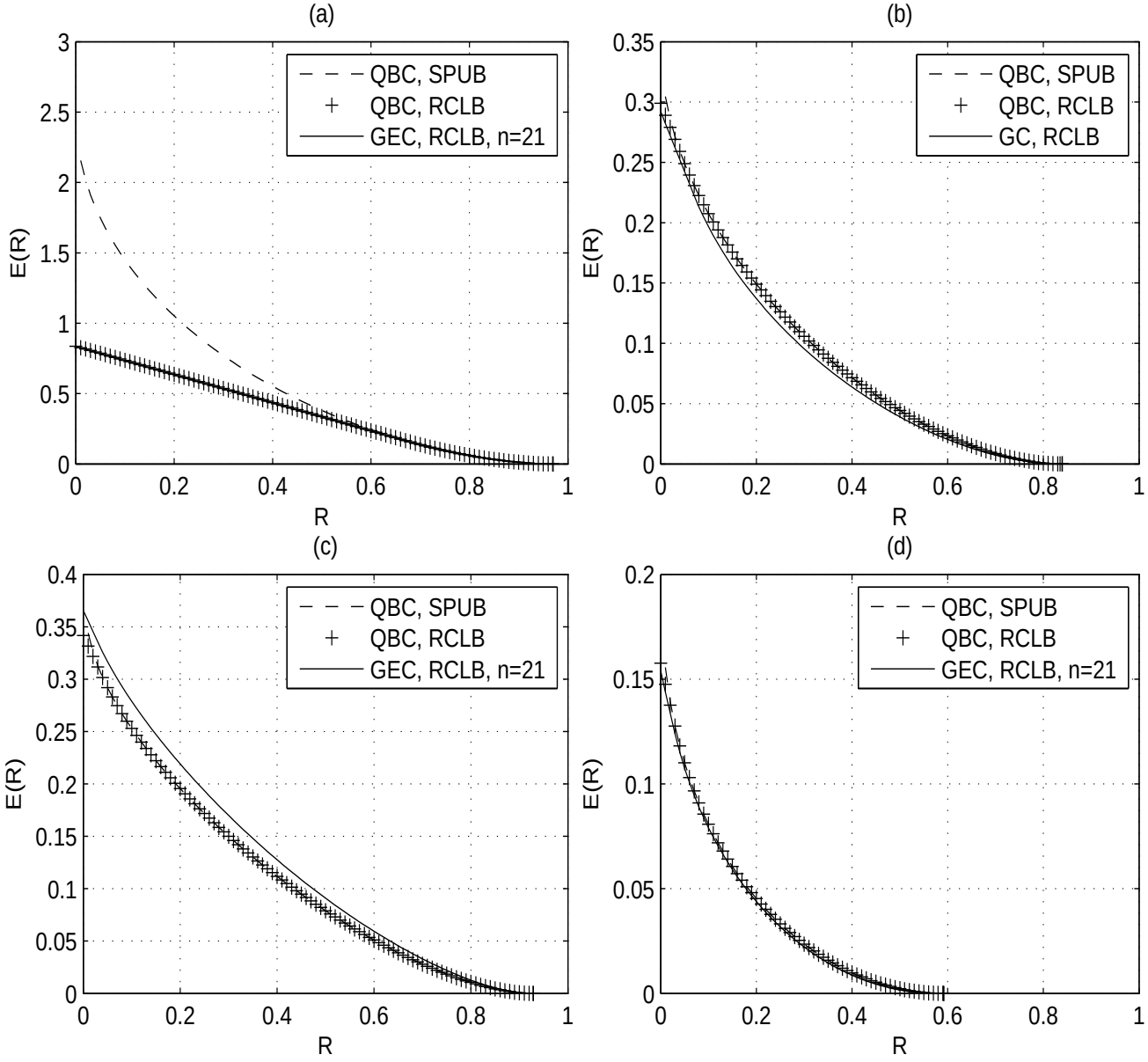

Figure 19: GEC fitting via the QBC: Bounds to $E(R)$ vs R. (a) Cor = 0.0131 and BER = 0.00314 (Case **A** in Table 1) with $R_{cr}^{(21)} = 0.66$ for the GEC and $R_{cr} = 0.67$ for the QBC; (b) Cor = 0.2227 and BER = 0.03 (Case **B** in Table 1) with $R_{cr} = 0.06$ for the GC and $R_{cr} = 0.05$ for the QBC; (c) Cor = 0.248 and BER = 0.012 (Case **C** in Table 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 1) with $R_{cr}^{(21)} = 0.02$ for the GEC and $R_{cr} = 0.03$ for the QBC.

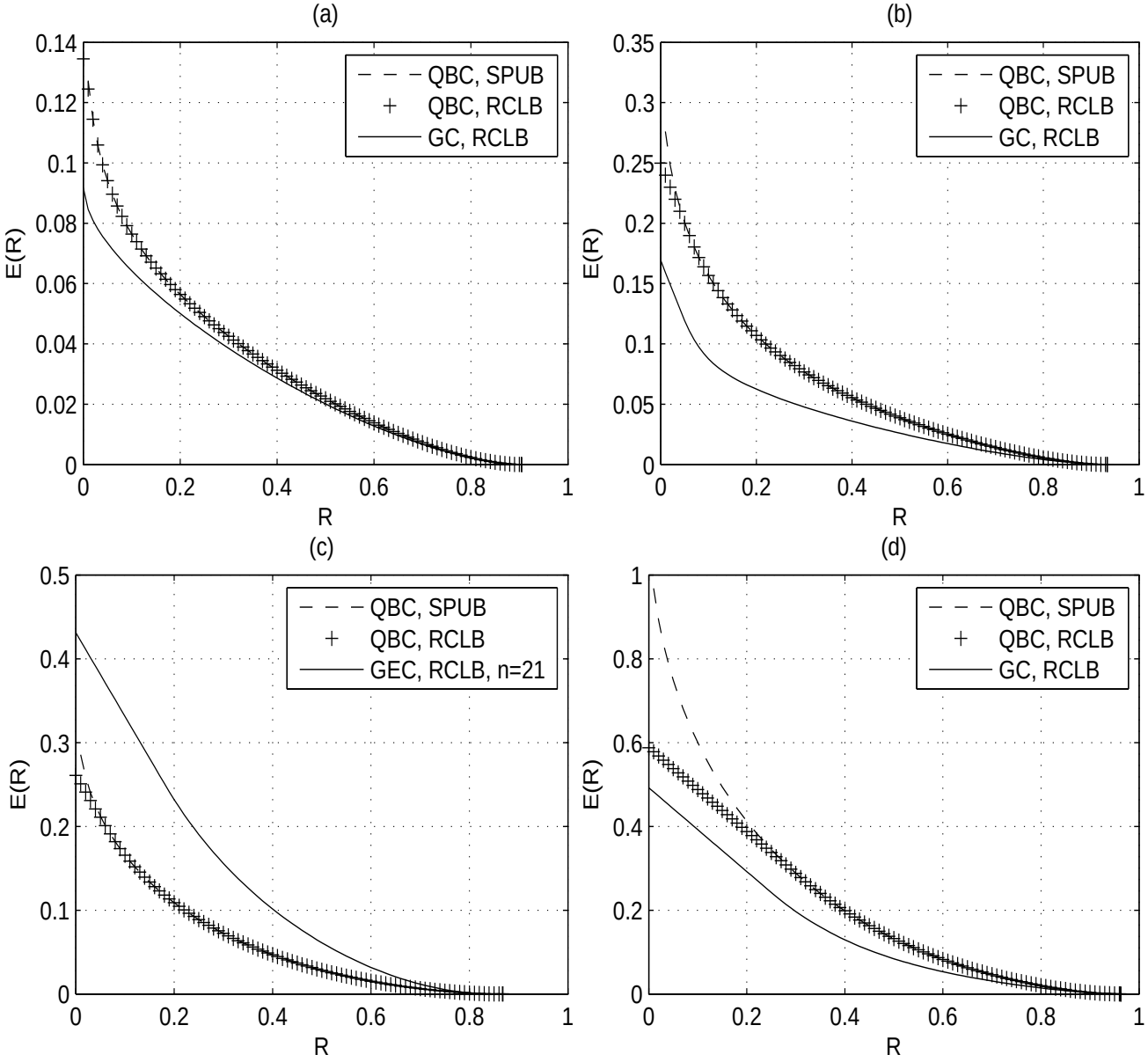

Figure 20: GEC fitting via the QBC: Bounds to $E(R)$ vs R. (a) Cor = 0.5 and BER = 0.03 (Case **E** in Table 1) with $R_{cr} = 0.0026$ for the GC and $R_{cr} = 0.02$ for the QBC; (b) Cor = 0.7 and BER = 0.03 (Case **F** in Table 1) with $R_{cr} = 0.04$ for the GC and $R_{cr} = 0.06$ for the QBC; (c) Cor = 0.8 and BER = 0.1 (Case **G** in Table 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 1) with $R_{cr} = 0.25$ for the GC and $R_{cr} = 0.31$ for the QBC.