

A Discrete Queue-Based Model for Capturing Memory and Soft-Decision Information in Correlated Fading Channels

Cecilio Pimentel, Fady Alajaji, *Senior Member, IEEE*, and Pedro Melo

Abstract—A discrete (binary-input 2^q -ary output) communication channel with memory is introduced to judiciously capture both the statistical memory and the soft-decision information of a time-correlated discrete fading channel (DFC) used with antipodal signaling and soft output quantization of resolution q . The discrete channel, which can be explicitly described via its binary input process and a 2^q -ary noise process, is shown to be symmetric, thus admitting a simple expression for its capacity when its noise is stationary ergodic. It is observed that considerable capacity gains can be achieved due to the channel's memory and the use of as few as 2 bits for soft-decision over interleaving the channel (to render it memoryless) and hard-decision demodulation ($q = 1$). The 2^q -ary noise process is next modeled via a queue-based (QB) ball-sampling mechanism to produce a mathematically tractable stationary ergodic Markovian noise source. The DFC is fitted by the QB noise model via an iterative procedure that minimizes the Kullback-Leibler divergence rate between the DFC and QB noise sources. Modeling results, measured in terms of channel noise correlation function and capacity reveal a good agreement between the two channels for a broad range of fading conditions.

Index Terms—Channel capacity, finite state Markov channels, queue-based Markovian noise, quantization, Rayleigh fading channels, soft-decision.

I. INTRODUCTION

WIRELESS communication channels are widely known to undergo time-varying multipath fading that is represented as a time-correlated random process. Due to the statistical dependence of successive fading samples, such channels exhibit memory. The development of iterative decoding schemes for codes (such as low-density parity-check and turbo codes) that achieve the capacity limit of memoryless channels, when they operate on binary (binary-input, binary-output) channels with memory, was considered in [1]–[4]. These works assume that the discrete channel (from the input of the

modulator to the output of the hard-quantized demodulator) is modeled as a binary finite state Markov channel (FSMC) and incorporate the FSMC structure in the decoding process in order to exploit the channel statistical memory. Significant performance gains were reported relative to traditional schemes that ignore the channel memory via perfect interleaving. In this latter case, the achievable rates are smaller than those of the original channel, since it is well known that memory increases capacity for a wide class of information stable channels [5], [6]. Furthermore, binary FSMC models have been shown to accurately approximate hard-decision demodulated correlated Rayleigh and Rician flat fading channels, e.g., see [7]–[13] and the references therein. On the other hand, recent studies have revealed that non-binary output information realized by softly quantizing the channel output can significantly increase capacity vis-a-vis hard output quantization for several classes of channels, including additive white Gaussian noise (AWGN) channels [14], [15], memoryless (fully interleaved) Rayleigh fading channels [16], additive colored Gaussian noise channels and intersymbol interference channels [14].

Motivated by the above results, we develop in this paper a simple non-binary-output FSMC model for time-correlated flat fading channels that effectively capture both their statistical memory and their soft-decision information. The new channel model may be used for designing novel coding/decoding schemes for soft-decision demodulated fading channels with memory that result in superior performance over systems that ignore the channel's memory (via interleaving) and/or soft-decision information (via hard demodulation). Specifically, we consider a discrete fading channel (DFC) which consists of a binary phase-shift keying (BPSK) modulator, a time-correlated flat Rayleigh fading channel and a q -bit soft-quantized coherent demodulator. We first show that for this channel there exists a 2^q -ary noise process (independent from the input process) such that its 2^q -ary output process can be written as an explicit function of the input and noise processes. We refer to this binary-input 2^q -ary output discrete channel as the non-binary noise discrete channel (NBND). We prove that the channel is symmetric in the sense that a uniformly distributed memoryless input process maximizes its block input-output mutual information. As a result, we derive a formula as well as simple (asymptotically tight) upper and lower bounds for the capacity of this channel when its noise is stationary and ergodic (which corresponds to a stationary ergodic underlying fading process). We also evaluate numerically the effect of

Paper approved by R. K. Mallik, the Editor for Diversity and Fading Channels of the IEEE Communications Society. Manuscript received April 28, 2011; revised October 13, 2011.

This work was supported in part by NSERC of Canada and CNPq of Brazil. Parts of this work were presented at the IEEE International Conference on Communications (ICC 2009), Dresden, Germany, June 2009, and at the IEEE Vehicular Technology Conference (VTC 2011 Spring), Budapest, Hungary, May 2011.

C. Pimentel and P. Melo are with the Department of Electronics and Systems, Federal University of Pernambuco, Recife, PE, 50711-970, Brazil (e-mail: cecilio@ufpe.br, pmelo@hotmail.com).

F. Alajaji is with the Department of Mathematics and Statistics, Queen's University, Kingston, ON K7L 3N6, Canada (e-mail: fady@mast.queensu.ca).
Digital Object Identifier 10.1109/TCOMM.2012.050912.110261

the quantizer parameters and channel correlation parameters on the capacity of the DFC. The capacity results indicate that exploiting both the channel's memory and soft-decision information is more worthwhile than ignoring either of them via channel interleaving or hard quantization.

We next model the noise process of the NBNDC via a 2^q -ary Markov process that generalizes the binary queue-based (QB) noise proposed in [6]. The resulting non-binary QB noise source is an M 'th order Markovian stationary ergodic process with $2^q + 2$ independent parameters. Closed-form expressions for several statistics and the entropy rate of the QB noise process are established. In related work, the modeling of non-binary channels via hidden Markov models has been considered in [17], [18]. However, unlike what is herein developed, these works do not provide explicit expressions for the channel statistics and capacity and are hence less amenable for tractable mathematical analysis. We proceed by investigating the appropriateness of the non-binary QB noise model for emulating the DFC. The $2^q + 2$ parameters of the QB noise are selected such that the Kullback-Leibler divergence rate between the DFC and the QB noise processes is minimized. The accuracy of the QB noise model is then measured in terms of the channel noise autocorrelation function and channel capacity. A good fit is obtained for a wide choice of fading conditions.

The rest of the paper is organized as follows. In Section II, we introduce the DFC and show that it can be represented via the NBNDC. We next investigate the capacity of the DFC in Section III from both analytical and numerical perspectives. We establish a capacity formula in terms of noise entropy rates as well as capacity upper and lower bounds (which are asymptotically exact with increasing blocklength). We also numerically illustrate the DFC capacity as a function of the channel's memory and soft-decision parameters. In Section IV, we present a mathematically tractable non-binary stationary ergodic Markov process with finite order to emulate the behavior of the NBNDC's infinite-memory non-binary noise process induced by the DFC. We show that this Markovian noise process, which is generated via a ball sampling mechanism involving a finite-queue, features closed-form formulas for its statistical and information-theoretic quantities. In Section V, we use this QB Markovian noise process to approximate the DFC noise process by choosing its parameters via an iterative algorithm so that the Kullback-Leibler rate distance between the two noise sources statistics is minimized. We then evaluate the goodness of the fit by comparing the channels' noise autocorrelation functions and capacities. Finally, we conclude the paper in Section VI.

II. THE NON-BINARY NOISE DISCRETE CHANNEL MODEL

A. Discrete Fading Channel with Soft-Decision Demodulation

We consider a DFC composed of a BPSK modulator, a time-correlated flat Rayleigh fading channel with AWGN, and a q -bit soft-quantized coherent demodulator. We define the input and output alphabets of the discrete channel by $\mathcal{X} = \{0, 1\}$, $\mathcal{Y} = \{0, 1, \dots, 2^q - 1\}$, respectively. Let $\{X_k\}$, $X_k \in \mathcal{X}$, $k = 1, 2, \dots$, be the input process to the discrete channel.

The sampled received symbol at the output of the matched filter at the k th signaling interval is written as

$$R_k = \sqrt{E_s} A_k S_k + N_k, \quad k = 1, 2, \dots$$

where $\{S_k\} = \{(2X_k - 1)\}$, E_s is the energy of the transmitted signal, $\{N_k\}$ is a sequence of independent and identically distributed zero-mean Gaussian random variables with variance $N_0/2$. Furthermore, $\{A_k\}$ is the channel's fading process with $A_k = |G_k|$, where $\{G_k\}$ is a time-correlated complex wide-sense stationary Gaussian process with zero-mean and autocorrelation function given by the Clarke's fading model [19] $R[k] = J_0(2\pi f_D T |k|)$, where $J_0(x)$ is the zero-order Bessel function of the first kind and $f_D T$ is the maximum Doppler frequency normalized by the signaling rate $1/T$. As a result, each fading random variable A_k is Rayleigh distributed with unit second moment. The processes $\{A_k\}$ and $\{N_k\}$ are independent of each other and of the input process. The random variable R_k is demodulated via a q -bit (uniform) scalar quantizer to yield the discrete channel output $Y_k \in \mathcal{Y}$ according to the following operation

$$Y_k = j, \quad \text{if } R_k \in (T'_{j-1}, T'_j)$$

for $j \in \mathcal{Y}$. The thresholds T'_j are uniformly spaced with step-size Δ , satisfying [16]

$$T'_j = \begin{cases} -\infty, & \text{if } j = -1 \\ (j + 1 - 2^{q-1})\Delta, & \text{if } j = 0, 1, \dots, 2^q - 2 \\ \infty, & \text{if } j = 2^q - 1. \end{cases}$$

Setting $\delta \triangleq \Delta/\sqrt{E_s}$ and $T_j \triangleq T'_j/\sqrt{E_s}$ as the normalized step-size and thresholds, respectively, we can write $T_j = (j + 1 - 2^{q-1})\delta$, for $j = 0, 1, \dots, 2^q - 2$. The conditional probability $q_{i,j}(a_k) \triangleq \Pr(Y_k = j \mid X_k = i, A_k = a_k)$, where $i \in \mathcal{X}$, $j \in \mathcal{Y}$ and $a_k \in [0, \infty)$, can be determined as follows

$$\begin{aligned} q_{i,j}(a_k) &= \Pr(T'_{j-1} < R_k < T'_j \mid X_k = i, A_k = a_k) \\ &= \Pr\left(T_{j-1} - (2i - 1)a_k < \frac{N_k}{\sqrt{E_s}} < T_j - (2i - 1)a_k\right) \\ &= Q(\sqrt{2\gamma}(T_{j-1} - (2i - 1)a_k)) - \\ &\quad Q(\sqrt{2\gamma}(T_j - (2i - 1)a_k)) \end{aligned} \quad (1)$$

where $\gamma = E_s/N_0$ is the signal-to-noise ratio (SNR) and $Q(x) = 1/\sqrt{2\pi} \int_x^\infty \exp\{-t^2/2\} dt$ is the Gaussian Q -function. Due to the symmetry of the BPSK constellation and the quantizer thresholds, we observe from (1) that $q_{i,j}(a_k) = q_{1-i, 2^q-1-j}(a_k)$. We may also write

$$q_{i,j}(a_k) = q_{0, \frac{j-(2^q-1)i}{(-1)^i}}(a_k)$$

for $i \in \mathcal{X}$, $j \in \mathcal{Y}$. For integer $n \geq 1$, let $\Pr(y^n \mid x^n, a^n) \triangleq \Pr(Y_1 = y_1, \dots, Y_n = y_n \mid X_1 = x_1, \dots, X_n = x_n, A_1 = a_1, \dots, A_n = a_n)$. Thus

$$\Pr(y^n \mid x^n, a^n) = \prod_{k=1}^n q_{x_k, y_k}(a_k) = \prod_{k=1}^n q_{0, \frac{y_k - (2^q-1)x_k}{(-1)^{x_k}}}(a_k).$$

The DFC is thus probabilistically specified in terms of the channel block (n -fold) conditional probability

$$\begin{aligned} P_{\text{DFC}}^{(n)}(y^n \mid x^n) &\triangleq \Pr(Y^n = y^n \mid X^n = x^n) \\ &= \mathbf{E}_{A_1 \dots A_n} \left[\prod_{k=1}^n q_{0, \frac{y_k - (2^q-1)x_k}{(-1)^{x_k}}}(A_k) \right] \end{aligned} \quad (2)$$

where $y^n = (y_1, \dots, y_n)$, $x^n = (x_1, \dots, x_n)$ and $\mathbf{E}_X[\cdot]$ denotes expectation with respect to the random variable X . For $n = 1$, a closed-form expression for $P_{\text{DFC}}^{(1)}(y|x)$, $y \in \mathcal{Y}$ and $x \in \mathcal{X}$, is given by [20]

$$P_{\text{DFC}}^{(1)}(j) = m(-T_{j-1}) - m(-T_j) \quad (3)$$

where $j = \frac{y - (2^q - 1)x}{(-1)^x} \in \mathcal{Y}$, and

$$m(T_j) = 1 - Q(T_j \sqrt{2\gamma}) - \frac{\left[1 - Q\left(\frac{T_j \sqrt{2\gamma}}{\sqrt{\frac{1}{\gamma} + 1}}\right) \right] e^{-\frac{T_j^2}{(\frac{1}{\gamma} + 1)}}}{\sqrt{\frac{1}{\gamma} + 1}}.$$

The expected value in (2) can be directly calculated for $n \leq 3$ since the joint probability density function of arbitrarily correlated Rayleigh and Rician random variables is only known in closed form for $n \leq 3$ (e.g., see [21], [22]); for $n > 3$, (2) can however be determined via simulations. Hence, it is important to provide an effective model for $P_{\text{DFC}}^{(n)}(\cdot | \cdot)$. We next provide an alternative representation for the DFC.

B. Alternative DFC Model: The Non-Binary Noise Channel

It is often convenient to express the channel output process $\{Y_k\}$ as an explicit function of the input process $\{X_k\}$ and a noise process $\{Z_k\}$, where $\{Z_k\}$ and $\{X_k\}$ are independent of each other (e.g., see [23, pp. 142-144]). In this case, one can model the noise process via a finite-state Markov model which is an important analytical tool for coding design and performance evaluation for channels with memory.

Consider a binary-input 2^q -ary output channel, which we refer to as the non-binary noise discrete channel (NBND), where the output $Y_k \in \mathcal{Y}$ is explicitly expressed in terms of the input $X_k \in \mathcal{X}$ and a noise $Z_k \in \mathcal{Z}$ via

$$Y_k = (2^q - 1)X_k + (-1)^{X_k} Z_k \quad (4)$$

for $k = 1, 2, \dots$, where the noise process $\{Z_k\}$ in (4) is independent of the input $\{X_k\}$ and is governed by the n -fold distribution $P_{\text{NBND}}^{(n)}(z^n) \triangleq P_{\text{NBND}}^{(n)}(Z_1 = z_1, \dots, Z_n = z_n)$, for $z_k \in \mathcal{Z}$, $k = 1, \dots, n$. It directly follows from (4) and the fact that the input and noise process are independent from each other that the NBND n -fold conditional probability is given for each $n \geq 1$ by

$$P_{\text{NBND}}^{(n)}(y^n | x^n) = P_{\text{NBND}}^{(n)}(z^n) \quad (5)$$

where

$$z_k = \frac{y_k - (2^q - 1)x_k}{(-1)^{x_k}}, \quad i = 1, \dots, n. \quad (6)$$

Now given $x^n \in \mathcal{X}^n$ and $y^n \in \mathcal{Y}^n$, whenever $P_{\text{NBND}}^{(n)}(z^n)$ is set to be equal to (2) for each $n \geq 1$ with each z_k as given by (6), we obtain that $P_{\text{DFC}}^{(n)}(y^n | x^n) = P_{\text{NBND}}^{(n)}(y^n | x^n)$ for each $n \geq 1$. Therefore, the NBND provides an alternative representation of the DFC. In Section V, we use this fact to fit the DFC given by (2) via an NBND whose noise process $\{Z_k\}$ is an M th order Markov source generated via a non-binary queue of length M .

Observation 1: In the special case when $q = 1$ (hard-decision demodulation), the NBND expression in (4) reduces

to the familiar expression $Y_k = X_k \oplus Z_k$, for $k = 1, 2, \dots$, (where \oplus denotes addition modulo 2) which is commonly adopted to model binary (binary-input binary-output) additive-noise discrete channels with memory (e.g., see [6], [8], [12]). Furthermore, if $\{Z_k\}$ is memoryless, then we obtain the memoryless binary symmetric channel (BSC) which represents the fully interleaved DFC.

The channel capacity of the NBND with stationary ergodic noise $\{Z_k\}$ is studied in the next section.¹

III. DFC CHANNEL CAPACITY

A. Capacity Formula and Bounds

Consider the NBND described by (4) where the noise process $\{Z_k\}$ is stationary ergodic. The channel capacity, in bits per channel use, for this information stable channel is given by [24], [25]

$$C = \lim_{n \rightarrow \infty} C^{(n)} = \sup_n C^{(n)} \quad (7)$$

where

$$C^{(n)} \triangleq \max_{p(x^n)} \frac{1}{n} I(X^n; Y^n)$$

where the maximum is taken with respect to all input distributions $p(x^n)$ and $I(X^n; Y^n)$ denotes the block mutual information between X^n and Y^n (e.g., cf. [24], [26]). $I(X^n; Y^n)$ can be written as follows

$$I(X^n; Y^n) = H(Y^n) - H(Y^n | X^n) = H(Y^n) - H(Z^n)$$

where $H(\cdot)$ denotes entropy and the last equality follows from (4) and the independence of X^n and Z^n . Thus

$$C^{(n)} = \frac{1}{n} \left(\max_{p(x^n)} [H(Y^n)] - H(Z^n) \right). \quad (8)$$

Note that $C^{(n)}$ is an increasing function of n [24]. The capacity achieving input distribution that maximizes $H(Y^n)$ is determined next.

Definition 1: Let $\mathcal{W} = \{0, 1, \dots, 2^q - 1\}$ and let $\{W_k\}$, $W_k \in \mathcal{W}$, be a process with n -fold probability distribution

$$\Pr(W^n = w^n) = \sum_{x^n \in \mathcal{X}^n} \Pr\left(Z^n = \frac{w^n - (2^q - 1)x^n}{(-1)^{x^n}}\right) \quad (9)$$

where $Z^n = (w^n - (2^q - 1)x^n)/(-1)^{x^n}$ denotes the tuple obtained from component-wise operations, i.e., $(Z_1 = (w_1 - (2^q - 1)x_1)/(-1)^{x_1}, \dots, Z_n = (w_n - (2^q - 1)x_n)/(-1)^{x_n})$. The probability assignment (9) is valid since

$$\begin{aligned} 1 &= \sum_{z^n \in \mathcal{Z}^n} \Pr(Z^n = z^n) \\ &= \sum_{w^n \in \mathcal{W}^n} \sum_{x^n \in \mathcal{X}^n} \Pr\left(Z^n = \frac{w^n - (2^q - 1)x^n}{(-1)^{x^n}}\right) \\ &= \sum_{w^n \in \mathcal{W}^n} \Pr(W^n = w^n). \end{aligned}$$

¹Note that the stationary noise $\{Z_k\}$ defined by (5) and (2) is ergodic when the covariance function of the fading process $\hat{G}(t)$ is asymptotically vanishing (such as the Clarke covariance function).

The process $\{W_k\}$ is stationary since $\{Z_k\}$ is stationary; indeed for any integer $m > 0$, $w^n \in \mathcal{W}^n$,

$$\begin{aligned} & \Pr(W_{1+m} = w_1, \dots, W_{n+m} = w_n) \\ &= \sum_{x^n \in \mathcal{X}^n} \Pr\left(Z_{1+m} = \frac{w_1 + (2^q - 1)x_1}{(-1)^{x_1}}, \dots, \right. \\ & \quad \left. Z_{n+m} = \frac{w_n - (2^q - 1)x_n}{(-1)^{x_n}}\right) \\ &= \sum_{x^n \in \mathcal{X}^n} \Pr\left(Z_1 = \frac{w_1 + (2^q - 1)x_1}{(-1)^{x_1}}, \dots, \right. \\ & \quad \left. Z_n = \frac{w_n - (2^q - 1)x_n}{(-1)^{x_n}}\right) \\ &= \Pr(W_1 = w_1, \dots, W_n = w_n). \end{aligned}$$

It follows from Definition 1 that the random variables Z_k and W_k are related as follows:

$$W_k = \min\{Z_k, 2^q - 1 - Z_k\}, \quad k = 1, 2, \dots \quad (10)$$

Example 1: Let $q = 2$ and $\alpha_j = \Pr(Z_k = j)$. The probability distribution of W_1 is given by

$$(\Pr(W_1 = 0), \Pr(W_1 = 1)) = (\alpha_0 + \alpha_3, \alpha_1 + \alpha_2).$$

Let $\alpha_{i,j} = \Pr(Z_k = i, Z_{k+1} = j)$. The probability distribution of W^2 is given by

$$\begin{aligned} & (\Pr(W^2 = 00), \Pr(W^2 = 01), \Pr(W^2 = 10), \Pr(W^2 = 11)) \\ &= (\alpha_{0,0} + \alpha_{0,3} + \alpha_{3,0} + \alpha_{3,3}, \alpha_{0,1} + \alpha_{0,2} + \alpha_{3,1} + \alpha_{3,2}, \\ & \quad \alpha_{1,0} + \alpha_{1,3} + \alpha_{2,0} + \alpha_{2,3}, \alpha_{1,1} + \alpha_{1,2} + \alpha_{2,1} + \alpha_{2,2}). \end{aligned}$$

Consider the $2^n \times 2^{qn}$ channel transition probability matrix $\mathbf{Q}^n = [P_{\text{NBND C}}^{(n)}(y^n | x^n)]$ corresponding to n channel uses, where each row (resp. column) of \mathbf{Q}^n is indexed by a sequence x^n (resp. y^n). We prove in Appendix A that \mathbf{Q}^n is symmetric since its columns can be partitioned into $2^{(q-1)n}$ submatrices, where each submatrix is of size $2^n \times 2^n$ with the property that its columns are permutations of each other and its rows are permutations of each other [24, p. 94]. This symmetry implies that the uniform input distribution over $\{0, 1\}^n$ maximizes the channel's block mutual information $I(X^n; Y^n)$ and hence achieves its capacity [24, p. 94].

Proposition 1: The value of $H(Y^n)$ under a uniform input distribution over $\mathcal{X}^n = \{0, 1\}^n$ is given by

$$\max_{p(x^n)} [H(Y^n)] = n + H(W^n). \quad (11)$$

Proof: We need to calculate

$$H(Y^n) = - \sum_{y^n \in \mathcal{Y}^n} \Pr(Y^n = y^n) \log_2 \Pr(Y^n = y^n) \quad (12)$$

for a uniform input distribution. In this case

$$\Pr(Y^n = y^n) = \frac{1}{2^n} \sum_{x^n \in \mathcal{X}^n} \Pr\left(Z^n = \frac{y^n - (2^q - 1)x^n}{(-1)^{x^n}}\right). \quad (13)$$

According to the symmetry of \mathbf{Q}^n , the probability (13) is the same for 2^n distinct values of y^n . Substituting (13) into (12) and using Definition 1, we have

$$\max_{p(x^n)} [H(Y^n)] = - \sum_{w^n \in \mathcal{W}^n} \Pr(W^n = w^n) \log_2 \left(\frac{\Pr(W^n = w^n)}{2^n} \right)$$

and hence (11) follows. \blacksquare

Substituting (11) into (8) yields

$$C^{(n)} = 1 + \frac{1}{n} [H(W^n) - H(Z^n)]. \quad (14)$$

Corollary 1: The channel capacity (7) is given by

$$\begin{aligned} C &= \lim_{n \rightarrow \infty} C^{(n)} \\ &= 1 + \lim_{n \rightarrow \infty} \frac{1}{n} [H(W^n) - H(Z^n)] \quad (15) \end{aligned}$$

$$= 1 + \mathcal{H}(W) - \mathcal{H}(Z) \quad (16)$$

in bits/channel use, where $\mathcal{H}(W) \triangleq \lim_{n \rightarrow \infty} (1/n)H(W^n)$ and $\mathcal{H}(Z) \triangleq \lim_{n \rightarrow \infty} (1/n)H(Z^n)$ denote the entropy rates of $\{W_n\}$ and $\{Z_n\}$, respectively.

Observation 2: In the case of hard-decision demodulation ($q = 1$), the noise process becomes binary and the entropy $H(W^n) = 0$ for all n , and the expression for channel capacity in (16) reduces to the well-known formula $C = 1 - \mathcal{H}(Z)$ (e.g., see [6]).

We can directly obtain upper and lower bounds on C using (15) and (7), respectively. Since $H(W^n)/n$ is decreasing in n for a stationary process $\{W_n\}$ (e.g., see [23], [26]), we have

$$C \leq \min \left\{ 1, 1 + \frac{1}{n} H(W^n) - \mathcal{H}(Z) \right\} \triangleq \bar{C}^{(n)}. \quad (17)$$

Combining (17) with the lower bound $C^{(n)} \leq C$ yields

$$C^{(n)} \leq C \leq \bar{C}^{(n)}$$

where as n increases, $C^{(n)}$ increases while $\bar{C}^{(n)}$ decreases; hence both bounds asymptotically coincide with C : $\lim_{n \rightarrow \infty} C^{(n)} = \lim_{n \rightarrow \infty} \bar{C}^{(n)} = C$. In the next subsection, we conduct a numerical capacity study of the DFC to determine the optimal values for the channel q -bit quantizer step-size δ and illustrate the potential gains in capacity due to the DFC's statistical memory and the use of soft-decision output quantization.

B. Capacity Numerical Study

We herein examine the behavior of the capacity of the DFC in terms of the quantizer parameters (q and δ), the SNR (γ) and the normalized Doppler frequency ($f_D T$). As shown in Section II, the DFC is an NBND C described by (4) with a stationary ergodic 2^q -ary noise $\{Z_k\}$ whose n -fold distribution is given by (2) for each n . As (2) cannot be determined for $n > 3$, we generate a realization of the noise process $\{Z_k\}$ via computer simulations for fixed DFC parameters (γ , $f_D T$, q , δ) and calculate $\{W_k\}$ using (10). The correlated Rayleigh fading samples are generated according to the method proposed in [27]. We then evaluate $\Pr(Z^n)$ and $\Pr(W^n)$ numerically for several values of n and compute the lower bound on the capacity $C^{(n)}$ using (14).

Fig. 1 presents $C^{(n)}$ versus the quantization step δ for several values of n for a DFC with parameters $\gamma = 10$ dB, $f_D T = 0.005$ and $q = 2$. One objective is to determine the optimal value of δ (in the sense of maximizing the channel capacity). We observe that for $n \geq 5$, $C^{(n)}$ is maximized for approximately $\delta = 0.2$. The curve $C^{(1)}$ corresponds to the

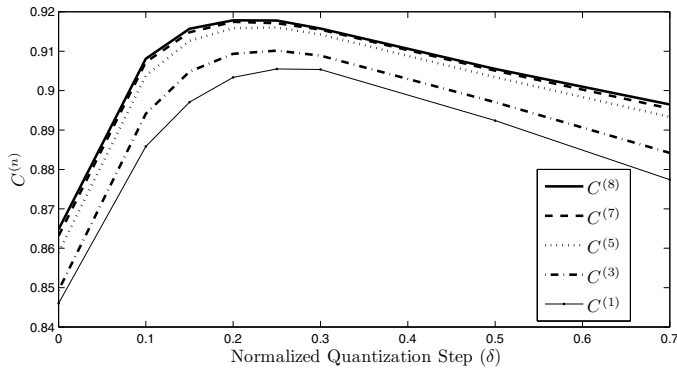


Fig. 1. $C^{(n)}$ versus the quantization step δ for different values of n ; DFC with $q = 2$, $f_D T = 0.005$, $\gamma = 10$ dB.

capacity of a memoryless DFC (the channel resulting when perfect interleaving is employed on the DFC). In this case, the optimal value of δ is 0.27. We may also obtain values of the capacity for channels with hard quantization ($q = 1$), as this channel is equivalent to a DFC with $\delta = 0$. For example, the capacity of this DFC with hard-quantized and perfectly interleaved (obtained from the curve $C^{(1)}$ with $\delta = 0$) is 0.846. We also remark that increasing n further than 7 does not improve accuracy. Indeed recall that as $n \uparrow \infty$, $C^{(n)} \uparrow C$ and yields the exact capacity. We denote the largest value of n for which the increase in $C^{(n)}$ is negligible (i.e., $< 10^{-2}$) by n^* . Table I summarizes the values of $C^{(n^*)}$ for $q = 1, 2$ (obtained for the optimal δ shown in the table) for selected values of n^* and γ . Note that the optimal values of δ provided in the table for $q > 1$ are different from those calculated in [14] for the memoryless DFC. We finally observe capacity gains due to the channel's memory and soft-decision quantization ($q > 1$) relative to hard-quantization ($q = 1$) and ideal channel interleaving (i.e., ignoring the channel's memory). For example, for $\gamma = 5$ dB, the capacity gain of $C^{(10)}$ ($q = 2$) over $C^{(10)}$ ($q = 1$) is 13 %, whereas it is 19 % when compared to the hard-quantized memoryless channel (with $C = 0.656$). For $\gamma = 2$ dB, the gains are 18.5% and 25.5%, respectively (in this case the capacity of the memoryless channel is $C = 0.51$). These capacity gains indicate that exploiting the noise memory and using soft output quantization (even with a resolution as low as $q = 2$ bits) are more worthwhile than ignoring noise memory via interleaving and using hard output quantization.

IV. NON-BINARY QUEUE-BASED MARKOVIAN NOISE

The binary queue-based channel was recently introduced in [6] to model a binary channel with a stationary ergodic binary M 'th-order additive Markov noise via a finite queue. We herein generalize the binary queue set up of [6] to produce a tractable non-binary Markovian noise model for the NBND. The non-binary queue-based (QB) noise process $\{Z_k\}_{k=1}^{\infty}$, $Z_k \in \mathcal{Y}$, is generated by slightly modifying the two-parcel procedure in [6]: given that we now operate on balls with $|\mathcal{Y}| = 2^q$ different colors (instead of only two colors). First, one of two parcels (an urn and a queue of size M) is selected with probability distribution $\{\varepsilon, 1 - \varepsilon\}$. The urn contains balls labeled with symbols in \mathcal{Y} satisfying the

TABLE I
LOWER BOUND ON CHANNEL CAPACITY, $C^{(n)}$, IN (BITS/CHANNEL USE)
FOR DFCs WITH $f_D T = 0.005$.

γ	n^*	$q = 2$		
		$C^{(n^*)}$	$C^{(n^*)}$	optimal δ
2 dB	11	0.54	0.64	0.5
5 dB	10	0.689	0.78	0.4
10 dB	7	0.86	0.915	0.2
15 dB	3	0.939	0.969	0.12

probability distribution $\rho = (\rho_0, \rho_1, \dots, \rho_{2^q-1})$. If the urn is selected, a noise symbol $Z_k = i$ is selected with probability ρ_i , $i \in \mathcal{Y}$. If the queue is selected, a noise symbol is selected with a probability distribution that depends on M and a bias parameter α ; see [6] for a detailed description of the procedure. The resulting QB noise process is a stationary ergodic M 'th order Markov source and has only $2^q + 2$ independent parameters (as opposed to a fully general Markovian process which would require in the order of 2^{qM} number of parameters): the size of the queue, M , the probability distribution of the balls in the urn, and correlation parameters ε and α , where $0 \leq \varepsilon < 1$, $\alpha \geq 0$. The state process $\{S_k\}_{k=-\infty}^{\infty}$ of the QB noise, defined by $S_k \triangleq (Z_k, Z_{k-1}, \dots, Z_{k-M+1})$, is a homogeneous first-order Markov process with an alphabet of size 2^{qM} .

Let $p_{\underline{i}\underline{j}}$ denote the conditional probability that $S_k = \underline{j}$ given that $S_{k-1} = \underline{i}$, where $\underline{i} = (i_0, \dots, i_{M-1})$, $\underline{j} = (j_0, \dots, j_{M-1})$, for $i_\ell, j_\ell \in \mathcal{Y}$, $\ell \in 0, \dots, M-1$. We let the rows and columns of the state transition probability matrix, denoted by \mathbf{P} , be indexed by the vectors \underline{i} and \underline{j} , respectively. It can be shown that the $(\underline{i}, \underline{j})$ th entry of the matrix $\mathbf{P} = [p_{\underline{i}\underline{j}}]$ is given as follows:

- For $M = 1$,

$$p_{\underline{i}\underline{j}} = p_{i_0 j_0} = \varepsilon \delta_{j_0, i_0} + (1 - \varepsilon) \rho_{j_0},$$

where

$$\delta_{i,j} = \begin{cases} 1, & \text{if } i = j \\ 0, & \text{if } i \neq j. \end{cases}$$

- For $M \geq 2$,

$$p_{\underline{i}\underline{j}} = \left(\sum_{\ell=0}^{M-2} \delta_{j_0, i_\ell} + \alpha \delta_{j_0, i_{M-1}} \right) \frac{\varepsilon}{M-1+\alpha} + (1-\varepsilon) \rho_{j_0}$$

if $j_{\ell+1} = i_\ell$, for $\ell = 0, \dots, M-2$, or otherwise $p_{\underline{i}\underline{j}} = 0$.

It can also be shown that the \underline{i} th component of the state stationary distribution column vector $\boldsymbol{\Pi} = [\pi_{\underline{i}}]$ is given by

$$\pi_{\underline{i}} = \frac{\prod_{\ell=0}^{2^q-1} \prod_{m=0}^{\xi_\ell-1} \left((1-\varepsilon) \rho_\ell + m \frac{\varepsilon}{M-1+\alpha} \right)}{\prod_{k=0}^{M-1} \left((1-\varepsilon) + k \frac{\varepsilon}{M-1+\alpha} \right)} \quad (18)$$

where $\prod_{k=0}^{-1} (\cdot) \triangleq 1$ and $\xi_\ell \triangleq \sum_{k=0}^{M-1} \delta_{i_k, \ell}$.

Example 2: For the case $q = 2$, $M = 1$, $\boldsymbol{\Pi} = [\rho_0, \rho_1, \rho_2, \rho_3]$ and \mathbf{P} is given by (19).

The QB noise block probability $P_{\text{QB}}^{(n)}(z^n) \triangleq \Pr(Z^n = z^n)$ is as follows:

$$\mathbf{P} = \begin{bmatrix} \varepsilon + (1 - \varepsilon) \rho_0 & (1 - \varepsilon) \rho_1 & (1 - \varepsilon) \rho_2 & (1 - \varepsilon) \rho_3 \\ (1 - \varepsilon) \rho_0 & \varepsilon + (1 - \varepsilon) \rho_1 & (1 - \varepsilon) \rho_2 & (1 - \varepsilon) \rho_3 \\ (1 - \varepsilon) \rho_0 & (1 - \varepsilon) \rho_1 & \varepsilon + (1 - \varepsilon) \rho_2 & (1 - \varepsilon) \rho_3 \\ (1 - \varepsilon) \rho_0 & (1 - \varepsilon) \rho_1 & (1 - \varepsilon) \rho_2 & \varepsilon + (1 - \varepsilon) \rho_3 \end{bmatrix}. \quad (19)$$

- For blocklength $n \leq M$

$$P_{\text{QB}}^{(n)}(z^n) = \frac{\prod_{\ell=0}^{2^q-1} \prod_{m=0}^{\xi'_\ell-1} \left((1 - \varepsilon) \rho_\ell + m \frac{\varepsilon}{M-1+\alpha} \right)}{\prod_{k=0}^{n-1} \left((1 - \varepsilon) + k \frac{\varepsilon}{M-1+\alpha} \right)} \quad (20)$$

where $\xi'_\ell = \sum_{k=1}^n \delta_{z_k, \ell}$.

- For blocklength $n \geq M + 1$

$$P_{\text{QB}}^{(n)}(z^n) = \prod_{i=M+1}^n \left[\left(\sum_{\ell=i-M+1}^{i-1} \delta_{z_i, z_\ell} + \alpha \delta_{z_i, z_{i-M}} \right) \times \frac{\varepsilon}{M-1+\alpha} + (1 - \varepsilon) \rho_{z_i} \right] \pi_{(z_1, \dots, z_M)} \quad (21)$$

where the expression for $\pi_{(z_1, \dots, z_M)}$ is given in (18).

Note that (20) yields the marginal noise distribution by setting $n = 1$: $P_{\text{QB}}^{(1)}(z_1) = \rho_{z_1}$, where $z_1 \in \{0, 1, \dots, 2^q - 1\}$. The correlation coefficient for the QB noise is a non-negative quantity given by

$$\begin{aligned} \text{Cor}_{\text{QB}} &= \frac{\mathbf{E}[Z_k Z_{k+1}] - \mathbf{E}[Z_k]^2}{\text{Var}(Z_k)} \\ &= \frac{\frac{\varepsilon}{M-1+\alpha}}{1 - (M-2+\alpha) \frac{\varepsilon}{M-1+\alpha}} \end{aligned} \quad (22)$$

where $\text{Var}(Z_k)$ denotes the variance of Z_k . We can prove that the QB noise autocorrelation function (ACF), defined as $R[m] = \mathbf{E}[Z_k Z_{k+m}]$, satisfies the formula (23), shown at the top of the next page. We conclude this section by establishing a closed-form expression for the entropy rate $\mathcal{H}_{\text{QB}}(Z)$ of the QB noise. Let w_0, \dots, w_{2^q-1} be non-negative integers such that $w_0 + \dots + w_{2^q-1} = M-1$ and let the probability distributions Γ_i , $i = 0, \dots, 2^q - 1$, be defined as

$$\Gamma_i = \left(\frac{w_0 \varepsilon}{M-1+\alpha} + (1 - \varepsilon) \rho_0, \dots, \frac{(w_i + \alpha) \varepsilon}{M-1+\alpha} + (1 - \varepsilon) \rho_i, \dots, \frac{w_{2^q-1} \varepsilon}{M-1+\alpha} + (1 - \varepsilon) \rho_{2^q-1} \right).$$

Then

$$\begin{aligned} \mathcal{H}_{\text{QB}}(Z) &= \sum_{w_0}^{M-1} \dots \sum_{w_{2^q-1}=0}^{M-1} \binom{M-1}{w_0, w_1, \dots, w_{2^q-1}} \times \\ &\quad \left[L_{(w_0+1, w_1, \dots, w_{2^q-1})} H(\Gamma_0) + L_{(w_0, w_1+1, \dots, w_{2^q-1})} H(\Gamma_1) \right. \\ &\quad \left. + \dots + L_{(w_0, w_1, \dots, w_{2^q-1}+1)} H(\Gamma_{2^q-1}) \right] \end{aligned} \quad (24)$$

where

$$L_{(w_0, w_1, \dots, w_{2^q-1})} = \frac{\prod_{\ell=0}^{2^q-1} \left(\prod_{m=0}^{w_\ell-1} (1 - \varepsilon) \rho_\ell + m \frac{\varepsilon}{M-1+\alpha} \right)}{\prod_{k=0}^{M-1} \left((1 - \varepsilon) + k \frac{\varepsilon}{M-1+\alpha} \right)}.$$

V. DFC MODELING VIA THE NBNDC WITH QB NOISE

In the following, we fit the DFC using the NBNDC with QB noise. For this purpose, given a DFC with fixed parameters, we estimate the $2^q + 2$ independent parameters of the QB noise process such that the QB noise block probability approximates well the DFC channel block probability of (2). Specifically, we select the parameters of the QB noise that minimize the Kullback-Leibler divergence rate (KLDL) between the QB and the DFC noise processes defined as [28]

$$D(P_{\text{DFC}} \| P_{\text{QB}}) \triangleq \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{z^n \in \mathcal{Y}^n} P_{\text{DFC}}^{(n)}(z^n) \log_2 \frac{P_{\text{DFC}}^{(n)}(z^n)}{P_{\text{QB}}^{(n)}(z^n)}$$

for identical one-dimensional probability distributions and noise correlation coefficients (we match the lower order statistics for both processes). Closed-form expressions for $P_{\text{QB}}^{(n)}(z^n)$ are given in (20) and (21), while $P_{\text{DFC}}(z^n)$ of (2) is calculated via computer simulations for $n > 1$. The minimization of the asymptotic KLDL quantity assures that both processes are statistically close for large blocklengths.²

Since the DFC noise process is stationary and the QB noise process is M 'th order Markovian, the minimization of the KLDL over the QB noise parameters reduces to minimizing [9], [28]

$$D_2^M(P_{\text{DFC}} \| P_{\text{QB}}) \triangleq - \sum_{z^{M+1}} P_{\text{DFC}}^{(M+1)}(z^{M+1}) \log_2 P_{\text{QB}}(z_{M+1} | z^M) \quad (25)$$

where $P_{\text{QB}}(z_{M+1} | z^M)$ is the QB conditional probability of the noise symbol z_{M+1} given the previous M symbols, which is evaluated using the QB noise block probability (21) and is given by

$$P_{\text{QB}}(z_{M+1} | z^M) = \left(\alpha \delta_{z_{M+1}, z_1} + \sum_{\ell=2}^M \delta_{z_{M+1}, z_\ell} \right) \times \frac{\varepsilon}{M-1+\alpha} + (1 - \varepsilon) \rho_{z_{M+1}}. \quad (26)$$

We match the one-dimensional probability distribution by setting $\rho_j = P_{\text{DFC}}^{(1)}(j)$, $j = 0, \dots, 2^q - 1$, where $P_{\text{DFC}}^{(1)}(j)$ is given by (3) in terms of the quantization parameters (δ and q) and γ . The remaining parameters (M , ε , α) are estimated as follows. We compute Cor_{DFC} and match the noise correlation coefficient $\text{Cor}_{\text{QB}} = \text{Cor}_{\text{DFC}}$. From (22) we can write the parameter α as

$$\alpha = \frac{\varepsilon + \text{Cor}_{\text{DFC}}(1 - M) + \text{Cor}_{\text{DFC}}(M - 2)\varepsilon}{\text{Cor}_{\text{DFC}}(1 - \varepsilon)}. \quad (27)$$

²The KLDL is an information measure that is widely employed to quantify the asymptotic statistical similarity between two sources. Thus the minimization of the KLDL between the noise sources of two channels attempts to match the two channels' statistical behavior for large blocklengths (note that large blocklengths are required to achieve reliable communication by the channel coding theorem).

$$R_{\text{QB}}[m] = \begin{cases} \mathbf{E}[Z_k^2], & \text{if } m = 0 \\ \frac{1}{1 - \frac{(M-2+\alpha)\varepsilon}{M-1+\alpha}} \left[\frac{\varepsilon}{M-1+\alpha} \mathbf{E}[Z_k^2] + (1-\varepsilon) \mathbf{E}[Z_k]^2 \right], & \text{if } 1 \leq m \leq M-1 \\ (1-\varepsilon) \mathbf{E}[Z_k]^2 + \frac{\varepsilon}{M-1+\alpha} \left[\sum_{i=1}^{M-1} R_{\text{QB}}[m-i] + \alpha R_{\text{QB}}[m-M] \right], & \text{if } m \geq M. \end{cases} \quad (23)$$

TABLE II
QB PARAMETERS FOR FITTING THE RAYLEIGH DFC WITH $q = 2$.

γ	$f_D T = 0.001$	$f_D T = 0.005$	$f_D T = 0.01$
2 dB ($\delta = 0.5$)	$M = 14$ $\varepsilon = 0.7978$ $\alpha = 0.7570$	$M = 11$ $\varepsilon = 0.7537$ $\alpha = 0.6362$	$M = 8$ $\varepsilon = 0.6846$ $\alpha = 0.5313$
5 dB ($\delta = 0.4$)	$M = 13$ $\varepsilon = 0.8388$ $\alpha = 0.7819$	$M = 10$ $\varepsilon = 0.7967$ $\alpha = 0.6318$	$M = 7$ $\varepsilon = 0.7260$ $\alpha = 0.5286$
10 dB ($\delta = 0.2$)	$M = 9$ $\varepsilon = 0.8060$ $\alpha = 0.7470$	$M = 7$ $\varepsilon = 0.7563$ $\alpha = 0.5932$	$M = 5$ $\varepsilon = 0.6765$ $\alpha = 0.4818$
15 dB ($\delta = 0.12$)	$M = 6$ $\varepsilon = 0.7580$ $\alpha = 0.7282$	$M = 5$ $\varepsilon = 0.7076$ $\alpha = 0.5511$	$M = 4$ $\varepsilon = 0.6371$ $\alpha = 0.399$

Minimization procedure: For fixed DFC parameters, we substitute (27) into (26) and the result into (25) and find the value of ε that minimizes (25) for each value of M . For that purpose, we apply the Newton-Raphson's method (e.g., see [29]) to the derivative of (25), resulting in the following iterative procedure for estimating ε . Given an iteration point ε_n , we show in Appendix B that the next iteration point is given by

$$\varepsilon_{n+1} = \varepsilon_n + \frac{\sum_{z^{M+1}} P_{\text{DFC}}(z^{M+1}) \frac{A_{z^{M+1}}}{A_{z^{M+1}} \varepsilon_n + B_{z^{M+1}}}}{\sum_{z^{M+1}} P_{\text{DFC}}(z^{M+1}) \frac{A_{z^{M+1}}^2}{(A_{z^{M+1}} \varepsilon_n + B_{z^{M+1}})^2}} \quad (28)$$

where

$$A_{z^{M+1}} = [1 + \text{Cor}_{\text{DFC}}(M-2)] \delta_{z_{M+1}, z_1} - \left(\sum_{\ell=2}^M \delta_{z_{M+1}, z_\ell} \right) \text{Cor}_{\text{DFC}} - (1 - \text{Cor}_{\text{DFC}}) \rho_{z_{M+1}} \quad (29)$$

and

$$B_{z^{M+1}} = \left(\delta_{z_{M+1}, z_1} (1-M) + \sum_{\ell=2}^M \delta_{z_{M+1}, z_\ell} \right) \text{Cor}_{\text{DFC}} + (1 - \text{Cor}_{\text{DFC}}) \rho_{z_{M+1}}. \quad (30)$$

From the constraint that $\alpha \geq 0$, we have from (27) that

$$\frac{\text{Cor}_{\text{DFC}}(M-1)}{1 + \text{Cor}_{\text{DFC}}(M-2)} \leq \varepsilon < 1.$$

In the algorithm, we used an empirically established estimate for the initial point ε_0 within this interval and observed convergence for all considered DFC parameters. We repeated this procedure for increasing values of M and chose a triplet (M , ε , α) to represent a specific DFC whenever $D_2^M(P_{\text{DFC}}||P_{\text{QB}})$ converges within a prescribed threshold (we use a threshold value κ varying between 10^{-3} and 5×10^{-3} for the results

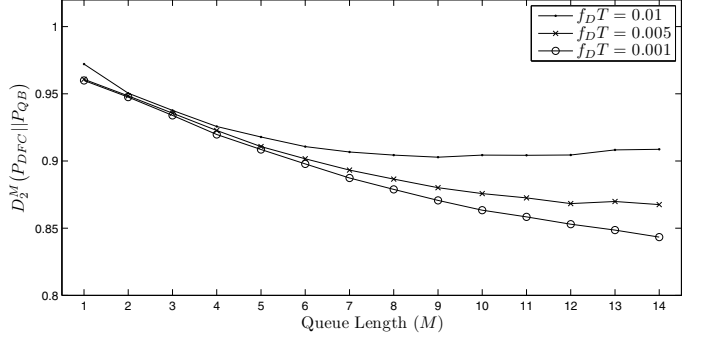


Fig. 2. $D_2^M(P_{\text{DFC}}||P_{\text{QB}})$ versus M . DFC with $q = 2$ and $\gamma = 5$ dB for $f_D T = 0.01$, $f_D T = 0.005$ and $f_D T = 0.001$.

of Table II³). Three typical curves for $D_2^M(P_{\text{DFC}}||P_{\text{QB}})$ as a function of M are shown in Fig. 2 for $q = 2$ and $\gamma = 5$ dB (with $f_D T = 0.01$, 0.005 and 0.001). The optimization procedure was carried out for DFCs with $q = 2$, four values of γ , and three values of $f_D T$ for each γ . We did not observe an important variation in the optimal value of δ for the considered $f_D T$ values, so we used the values provided in Table I. The vectors ρ calculated from (3) are $\rho = (0.6953, 0.1962, 0.085, 0.0235)$ for $\gamma = 2$ dB and $\delta = 0.5$, $\rho = (0.7822, 0.1536, 0.0538, 0.0104)$ for $\gamma = 5$ dB and $\delta = 0.4$, $\rho = (0.9239, 0.0528, 0.0188, 0.0045)$ for $\gamma = 10$ dB and $\delta = 0.2$, and $\rho = (0.9722, 0.0201, 0.0064, 0.0013)$ for $\gamma = 15$ dB and $\delta = 0.12$. Table II provides the remaining parameters of the QB noise process that fits a specific DFC.

We next use channel noise ACF and channel capacity as metrics for measuring the accuracy of the NBNDC with QB noise models of Table II in approximating the DFC. Figs. 3 and 4 compare the ACFs of the DFC and the QB noise processes for several values of M , for a DFC with $q = 2$ where $f_D T = 0.005$, $\gamma = 10$ dB and $\delta = 0.2$ in Fig. 3, while $f_D T = 0.001$, $\gamma = 15$ dB and $\delta = 0.12$ in Fig. 4. The figures show a generally good agreement⁴ between the ACF of the DFC and that of the QB noise model with the values of M described in Table II, although in the slower fading case (of $f_D T = 0.001$), QB models with larger values of M give a better ACF fit than the QB model of Table II. A similar behavior is also observed for all QB models listed in this table (curves not shown), thus indicating that QB models satisfactorily approximate the ACF of the DFC process. The expression of the ACF for the QB noise is given in (23), while

³Thus the values of M shown in Table II are chosen to be within κ from the M^* that locally minimizes $D_2^M(P_{\text{DFC}}||P_{\text{QB}})$ in order to avoid large values of M and hence keep a manageable total number of states.

⁴Note that for small values of m , there is a mismatch between between the ACFs of the DFC and QB models. This is due to the fact that, unlike the ACF of the DFC, the ACF of the QB model is constant for $0 \leq m \leq M-1$ (see the expression of $R_{\text{QB}}[m]$ in (23)).

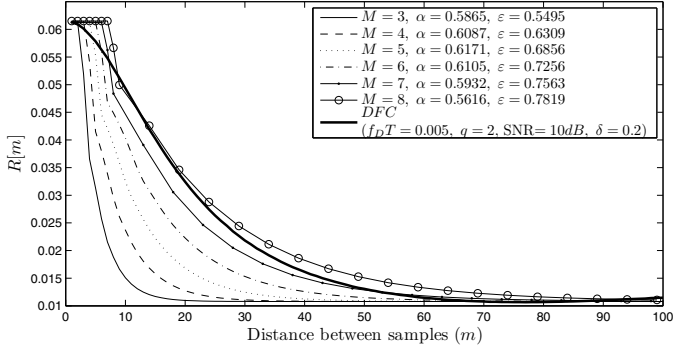


Fig. 3. Comparison of ACFs: DFC with $f_D T = 0.005$, $q = 2$, $\gamma = 10$ dB and $\delta = 0.2$ vs NBNDNC with QB noise.

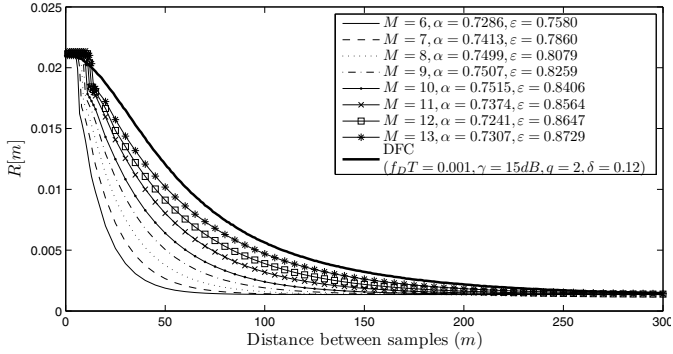


Fig. 4. Comparison of ACFs: DFC with $f_D T = 0.001$, $q = 2$, $\gamma = 15$ dB and $\delta = 0.12$ vs NBNDNC with QB noise.

the ACF of the DFC noise is calculated numerically from the $P_{\text{DFC}}^{(n)}(z^n)$ statistics of (2) obtained via simulations. In Fig. 5, we plot the upper and lower capacity bounds $C^{(n)}$ and $\bar{C}^{(n)}$ versus the blocklength n for the QB model fitting the DFC with $q = 2$, $\gamma = 10$ dB, $f_D T = 0.005$ and $\delta = 0.2$. We clearly see that as n increases, $C^{(n)}$ and $\bar{C}^{(n)}$ converge to yield a good estimate of capacity. Fig. 6 assesses the lower bound on the capacity of the DFC with the upper and lower bounds of the fitting QB noise model, where the DFC values of n^* and δ given in Table I for each SNR γ and the values of the QB noise are given in Table II. We observe a relatively close match in the capacity curves of both channels.

Finally, we have compared the QB model (of $2^q + 2$ parameters) with the fully general Markov model (of 2^{qM} parameters) to assess any significant performance discrepancy when both channels approximate the DFC. The full Markov model is matched to the DFC by equating its M -dimensional noise distribution to that of the DFC. The results indicate that the QB model performs close to the full Markov model for typical system conditions while requiring a substantially smaller number of parameters (unlike the full Markov model, the QB model's number of parameters does not grow exponentially with M , whose value can be considerable for slow fading and noisy conditions as shown in Table II; e.g., for typical values of $q = 2$ and $M = 7$, the QB model has only 6 parameters while the full Markov model requires $2^{14} = 16,384$ parameters). Thus the QB noise channel, in virtue of its closed-form statistical expressions, offers a practically effective model for the DFC.

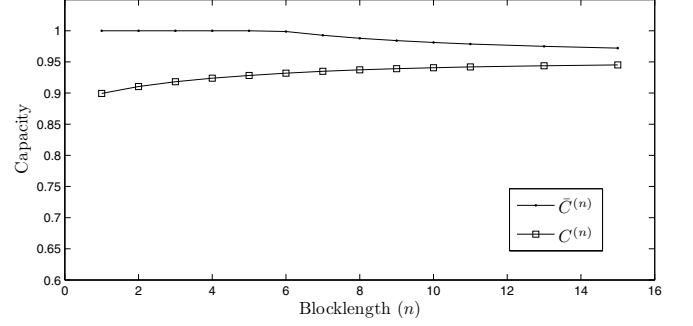


Fig. 5. Capacity of NBNDNC with QB noise: upper and lower bounds vs blocklength n . The NBNDNC models the DFC with $q = 2$, $\gamma = 10$ dB, $f_D T = 0.005$ and $\delta = 0.2$ and its QB noise parameters are $M = 7$, $\varepsilon = 0.7563$ and $\alpha = 0.5932$ (see Table II).

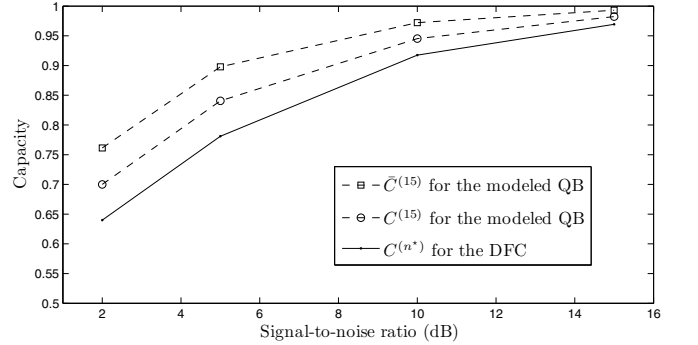


Fig. 6. Capacity comparison: DFC capacity lower bound (with $f_D T = 0.005$, $q = 2$ and n^* as given in Table I) vs NBNDNC with QB noise capacity lower and upper bounds (with $n = 15$).

VI. CONCLUSIONS

We introduced a binary-input 2^q -ary output discrete channel (denoted by NBNDNC) to properly represent both the statistical memory and the soft-decision information of BPSK-modulated time-correlated Rayleigh fading channels when they are coherently demodulated via a q -bit output quantizer. The NBNDNC's output is explicitly described in terms of its binary input and a 2^q -ary noise. This channel reduces to the familiar binary channel with modulo-2 additive noise when $q = 1$ (i.e., under hard-decision demodulation). We derived a simple formula (as well as asymptotically tight upper and lower bounds) for its capacity under stationary ergodic noise. We also generalized the binary queue-based (QB) noise process introduced in [6] to obtain a 2^q -ary QB stationary ergodic M th order Markovian noise model. We established closed-form analytical expressions for the statistics and the entropy rate of the 2^q -ary QB noise process in terms of its $2^q + 2$ parameters. Finally, we fitted the discrete soft-quantized correlated fading channel via the NBNDNC with QB noise and validated the modeling in terms of autocorrelation function and capacity.

In light of the capacity gains shown in Section III-B, an interesting direction for future work is to harness the structure of the NBNDNC model with QB noise for the construction of practical high-performing (in terms of error rate) channel codes that exploit both the channel's noise memory and soft-decision information for use over the correlated Rayleigh

fading channel. In a recent related work [30], it is shown that joint source-channel vector quantizers can be designed for the NBND C model with QB noise and can effectively be used over the underlying fading channel to achieve significantly improved performance in terms of mean-square error distortion. Extensions to modeling non-binary modulated (e.g., pulse amplitude modulation) correlated fading channels with soft-decision demodulation can also be studied in the future.

APPENDIX A

PROOF OF THE SYMMETRY OF THE TRANSITION MATRIX

We prove in this appendix that the matrix $\mathbf{Q}^n = [P_{\text{NBND C}}^{(n)}(y^n|x^n)]$ can be split into $2^{n(q-1)}$ submatrices of size $2^n \times 2^n$ in which the columns are permutations of each other and also the rows are permutations of each other, and thus \mathbf{Q}^n is symmetric [24].

Let $\mathcal{X} = \{0, 1\}$, $\mathcal{W} = \{0, 1, \dots, 2^{q-1} - 1\}$, and $\mathcal{Y} = \{0, 1, \dots, 2^q - 1\}$ be alphabets. We split the columns of \mathbf{Q}^n into $2^{n(q-1)}$ classes $C_{w^n}, w^n \in \mathcal{W}^n$, defined as

$$C_{w^n} = \left\{ y^n \in \mathcal{Y}^n \mid y^n = \frac{w^n - (2^q - 1)v^n}{(-1)^{v^n}}, v^n \in \mathcal{X}^n \right\}. \quad (31)$$

It is worth noting that each y^n can be associated to a single w^n and to a single v^n whose components are given by

$$w_k = \min\{y_k, 2^q - 1 - y_k\} \quad (32)$$

and

$$v_k = \begin{cases} 0, & \text{if } y_k \leq 2^{q-1} - 1 \\ 1, & \text{if } y_k > 2^{q-1} - 1. \end{cases} \quad (33)$$

Recall from (5) that the n -fold conditional probability satisfies

$$P_{\text{NBND C}}^{(n)}(y^n | x^n) = P_{\text{NBND C}}^{(n)}(z^n(x^n, y^n)) \quad (34)$$

where

$$z_k(x_k, y_k) = \frac{y_k - (2^q - 1)x_k}{(-1)^{x_k}}. \quad (35)$$

The next step is to prove that within each class C_{w^n} , its corresponding rows and columns in \mathbf{Q}^n are formed by permutations of $P_{\text{NBND C}}^{(n)}(y^n | x^n) = P_{\text{NBND C}}^{(n)}(z^n(x^n, y^n))$, for $y^n \in C_{w^n}$. For this purpose, it is sufficient to prove that $z^n(x^n, y^n)$ is permuted among the rows and columns of the submatrix formed by $y^n \in C_{w^n}$, i.e.

- 1) Within each class, there are only 2^n possibilities for $z^n(x^n, y^n)$.
- 2) If $x_1^n \neq x_2^n$, with y^n fixed, then $z^n(x_1^n, y^n) \neq z^n(x_2^n, y^n)$.
- 3) If $y_1^n \neq y_2^n$, with x^n fixed, then $z^n(x^n, y_1^n) \neq z^n(x^n, y_2^n)$.

1) Let us associate y^n to a class C_{w^n} using (32) and to a vector v^n using (33). We then substitute the components of $y^n \in C_{w^n}$ defined in (31) into (35), yielding

$$\begin{aligned} z_k(x_k, y_k) &= \frac{\frac{w_k - (2^q - 1)v_k}{(-1)^{v_k}} - (2^q - 1)x_k}{(-1)^{x_k}} \\ &= (-1)^{x_k + v_k} w_k - (2^q - 1) \frac{(-1)^{v_k} v_k + x_k}{(-1)^{x_k}}. \end{aligned} \quad (36)$$

Since $x_k, v_k \in \mathcal{X}$, within a class C_{w^n} , there are two possibilities for $z_k(x_k, y_k)$ in (36), for each k , namely

$$z_k(x_k, y_k) = \begin{cases} w_k, & \text{if } v_k = x_k \\ (2^q - 1) - w_k, & \text{if } v_k \neq x_k. \end{cases} \quad (37)$$

Hence, in a class C_{w^n} , there are only 2^n possibilities for $z^n(x^n, y^n)$. Thus, item 1 is satisfied.

2) Suppose that for $y^n \in \mathcal{Y}^n$ fixed and $x_1^n \neq x_2^n$, it may occur $z_k(x_{1,k}, y_k) = z_k(x_{2,k}, y_k)$ for every k . Hence

$$\frac{y_k - (2^q - 1)x_{1,k}}{(-1)^{x_{1,k}}} = \frac{y_k - (2^q - 1)x_{2,k}}{(-1)^{x_{2,k}}}. \quad (38)$$

So (38) implies that $x_1^n = x_2^n$, which is a contradiction. Hence, item 2 is satisfied.

3) Suppose now that for $x^n \in \mathcal{X}^n$ fixed and $y_1^n \neq y_2^n$, $z_k(x_k, y_{1,k}) = z_k(x_k, y_{2,k})$ for every k . Hence

$$\frac{y_{1,k} - (2^q - 1)x_k}{(-1)^{x_k}} = \frac{y_{2,k} - (2^q - 1)x_k}{(-1)^{x_k}}. \quad (39)$$

We conclude from (39) that $x_1^n = x_2^n$, which is a contradiction. Hence, item 3 is satisfied which completes the proof.

APPENDIX B

APPLICATION OF THE NEWTON-RAPHSON'S METHOD TO THE MINIMIZATION OF $D_2^M(P_{\text{DFC}}||P_{\text{QB}})$

In order to minimize the divergence between the QB noise distribution and the DFC distribution, we apply the recursion formula given by

$$\varepsilon_{n+1} = \varepsilon_n - \frac{\frac{d}{d\varepsilon} D_2^M(P_{\text{DFC}}||P_{\text{QB}})|_{\varepsilon=\varepsilon_n}}{\frac{d^2}{d\varepsilon^2} D_2^M(P_{\text{DFC}}||P_{\text{QB}})|_{\varepsilon=\varepsilon_n}} \quad (40)$$

where $D_2^M(P_{\text{DFC}}||P_{\text{QB}})$ is given in (25). Recall that, for M fixed, (25) is a function only of ε , if we use $P_{\text{QB}}(Z_{M+1}|Z^M)$ as in (26) and α as in (27). We then derive an expression for the derivative of $D_2^M(P_{\text{DFC}}||P_{\text{QB}})$ in relation to ε :

$$\begin{aligned} \frac{d}{d\varepsilon} D_2^M(P_{\text{DFC}}||P_{\text{QB}}) &= - \sum_{Z_{M+1}} P_{\text{DFC}}(Z^{M+1}) \times \\ &\quad \frac{d}{d\varepsilon} [\log_2 P_{\text{QB}}(Z_{M+1}|Z^M)]. \end{aligned} \quad (41)$$

Substituting (27) into (26) we obtain, after a few manipulations

$$P_{\text{QB}}(Z_{M+1}|Z^M) = \frac{1}{1 - \text{Cor}_{\text{DFC}}} (A_{Z_{M+1}} \varepsilon + B_{Z_{M+1}}) \quad (42)$$

where the expression for $A_{Z_{M+1}}$ and $B_{Z_{M+1}}$ are given in (29) and (30), respectively. We then have

$$\frac{d}{d\varepsilon} (-\log_2 P_{\text{QB}}(Z_{M+1}|Z^M)) = \frac{-A_{Z_{M+1}}}{A_{Z_{M+1}} \varepsilon + B_{Z_{M+1}}} \quad (43)$$

$$\frac{d^2}{d\varepsilon^2} (-\log_2 P_{\text{QB}}(Z_{M+1}|Z^M)) = \frac{-A_{Z_{M+1}}^2}{(A_{Z_{M+1}} \varepsilon + B_{Z_{M+1}})^2}. \quad (44)$$

We now substitute (43) into (41) and (44) into the derivative of (41). The resulting expressions are then substituted into (40), resulting in (28).

REFERENCES

- [1] J. Garcia-Frias and J. D. Villasenor, "Turbo decoding of Gilbert-Elliott channels," *IEEE Trans. Commun.*, vol. 50, pp. 357–363, Mar. 2002.
- [2] J. Garcia-Frias, "Decoding of low-density parity-check codes over finite-state binary Markov channels," *IEEE Trans. Commun.*, vol. 52, pp. 1840–1843, Nov. 2004.
- [3] A. W. Eckford, F. R. Kschischang, and S. Pasupathy, "Analysis of low-density parity check codes for the Gilbert-Elliott channels," *IEEE Trans. Inf. Theory*, vol. 51, pp. 3872–3889, Nov. 2005.
- [4] C. Nicola, F. Alajaji, and T. Linder, "Decoding LDPC codes over binary channels with additive Markov noise," in *Proc. 2005 Canadian Workshop Inform. Theory*, pp. 187–190.
- [5] R. L. Dobrushin and M. S. Pinsker, "Memory increases transmission capacity," *Probl. Pered. Inform.*, vol. 5, no. 1, pp. 94–95, 1969.
- [6] L. Zhong, F. Alajaji, and G. Takahara, "A binary communication channel with memory based on a finite queue," *IEEE Trans. Inf. Theory*, vol. 53, pp. 2815–2840, Aug. 2007.
- [7] W. Kumwilaisak, C.-C. J. Kuo, and D. Wu, "Fading channel modeling via variable-length Markov chain technique," *IEEE Trans. Veh. Technol.*, vol. 57, pp. 1338–1358, May 2008.
- [8] F. Babich, O. Kelly, and G. Lombardi, "Generalized Markov modeling for flat fading," *IEEE Trans. Commun.*, vol. 48, pp. 547–551, Apr. 2000.
- [9] L. Zhong, F. Alajaji, and G. Takahara, "A model for correlated Rician fading channels based on a finite queue," *IEEE Trans. Veh. Technol.*, vol. 57, pp. 79–89, Jan. 2008.
- [10] L. Wilhelmsson and L. B. Milstein, "On the effect of imperfect interleaving for the Gilbert-Elliott channel," *IEEE Trans. Commun.*, vol. 47, pp. 681–688, May 1999.
- [11] P. Sadeghi and P. Rapajic, "Capacity analysis for finite-state Markov mapping of Markov channels," *IEEE Trans. Commun.*, vol. 53, pp. 833–840, May 2005.
- [12] 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.
- [13] C. Pimentel, T. H. Falk, and L. Lisbôa, "Finite-state Markov modeling of correlated Rician-fading Channels," *IEEE Trans. Veh. Technol.*, vol. 53, pp. 1491–1501, Sep. 2004.
- [14] N. Phamdo and F. Alajaji, "Soft-decision demodulation design for COVQ over white, colored, and ISI Gaussian channels," *IEEE Trans. Commun.*, vol. 48, pp. 1499–1506, Sep. 2000.
- [15] J. Singh, O. Dabeer, and U. Madhow, "On the limits of communication with low-precision analog-to-digital conversion at the receiver," *IEEE Trans. Commun.*, vol. 57, pp. 3629–3639, Dec. 2009.
- [16] F. Alajaji and N. Phamdo, "Soft-decision COVQ for Rayleigh-fading channels," *IEEE Commun. Lett.*, vol. 2, pp. 162–164, June 1998.
- [17] F. Swarts and H. C. Ferreira, "Markov characterization of channels with soft decision outputs," *IEEE Trans. Commun.*, vol. 41, pp. 678–682, May 1993.
- [18] P. Kuczynski, A. Rigollé, W. H. Gerstacker, and J. B. Huber, "Hidden Markov modeling of error patterns and soft outputs for simulation of wideband CDMA transmission systems," *Int. J. Electron. Commun.*, vol. 58, pp. 256–267, 2004.
- [19] R. H. Clarke, "A statistical theory of mobile radio reception," *Bell Syst. Tech. J.*, vol. 47, pp. 957–1000, 1968.
- [20] G. Taricco, "On the capacity of the binary input Gaussian and Rayleigh fading channel," *Eur. Trans. Telecommun.*, vol. 7, pp. 201–208, Mar.-Apr. 1996.
- [21] Y. Chen and C. Tellambura, "Infinite series representations of the trivariate and quadrivariate Rayleigh distribution and their applications," *IEEE Trans. Commun.*, vol. 53, pp. 2092–2101, Dec. 2005.
- [22] P. Dharmawansa, N. Rajatheva, and C. Tellambura, "On the trivariate Rician distribution," *IEEE Trans. Commun.*, vol. 56, pp. 1993–1997, Dec. 2008.
- [23] R. W. Yeung, *Information Theory and Network Coding*. Springer-Verlag, 2008.
- [24] R. G. Gallager, *Information Theory and Reliable Communication*. Wiley, 1968.
- [25] S. Verdú and T. S. Han, "A general formula for channel capacity," *IEEE Trans. Inf. Theory*, vol. 40, pp. 1147–1157, July 1994.
- [26] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, 2nd edition. Wiley, 2006.
- [27] Y. R. Zheng and C. Xiao, "Improved models for the generation of multiple uncorrelated Rayleigh fading waveforms," *IEEE Commun. Lett.*, vol. 6, pp. 256–258, June 2002.
- [28] R. M. Gray, *Entropy and Information Theory*. Springer-Verlag, 1990.
- [29] F. B. Hildebrand, *Introduction to Numerical Analysis*, 2nd edition. McGraw Hill, 1974.
- [30] S. Shahidi, F. Alajaji, and T. Linder, "Exploiting memory and soft-decision information in channel optimized quantization for correlated fading channels," in *Proc. 2011 Canadian Workshop Inform. Theory*.



Cecilio Pimentel Cecilio Pimentel was born in Recife, Brazil, in 1966. He received the B.Sc. degree from the Federal University of Pernambuco, Recife, Brazil, in 1987; the M.Sc. degree from the Catholics University of Rio de Janeiro, Rio de Janeiro, Brazil, in 1990; and the Ph.D. degree from the University of Waterloo, Ontario, Canada, in 1996, all in electrical engineering. Since October 1996, he has been with the Department of Electronics and Systems at the Federal University of Pernambuco, where he is currently an Associate Professor. From 2007 to

2008, he was a Visiting Research Scholar at the Department of Mathematics and Statistics, Queen's University, Kingston, Canada. His research interests include digital communications, information theory, and error correcting coding.



Fady Alajaji (S'90–M'94–SM'00) received the B.E. degree with distinction from the American University of Beirut, Lebanon, and the M.Sc. and Ph.D. degrees from the University of Maryland, College Park, all in electrical engineering, in 1988, 1990 and 1994, respectively. He held a postdoctoral appointment in 1994 at the Institute for Systems Research, University of Maryland.

In 1995, he joined the Department of Mathematics and Statistics at Queen's University, Kingston, Ontario, where he is currently a Professor of Mathematics and Engineering. Since 1997, he has also been cross-appointed in the Department of Electrical and Computer Engineering at the same university. From 2003 to 2008, he served as chair of the Queen's Mathematics and Engineering program. His research interests include information theory, digital communications, error control coding, joint source-channel coding and data compression.

Dr. Alajaji currently serves as Area Editor and Editor for Source and Source-Channel Coding for the IEEE TRANSACTIONS ON COMMUNICATIONS. He served as co-chair of the 1999 Canadian Workshop on Information Theory, as co-chair of the Technical Program Committee (TPC) of the 2004 Biennial Symposium on Communications and as a TPC member of several international conferences and workshops. He received the Premier's Research Excellence Award from the Province of Ontario.



Pedro Melo Pedro Melo received the B.E. degree in electrical engineering from the Federal University of Pernambuco, Recife, Brazil, in 2010. He is now a M.Sc. student in the same university. As part of his M.Sc. research, he has been a visiting research student from September 2011 to December 2011 in Queen's University, Kingston, ON, Canada. His research interests include error correcting codes, mobile fading channels and digital signal processing.