On the Queueing Behavior of Random Codes over a Gilbert-Elliot Erasure Channel

Parimal Parag, Jean-Francois Chamberland, Henry D. Pfister, Krishna R. Narayanan
Department of Electrical and Computer Engineering, Texas A&M University

***Abstract*—This paper considers the queueing performance of a system that transmits coded data over a time-varying erasure channel. In our model, the queue length and channel state together form a Markov chain that depends on the system parameters. This gives a framework that allows a rigorous analysis of the queue as a function of the code rate. Most prior work in this area either ignores block-length (e.g., fluid models) or assumes error-free communication using finite codes. This work enables one to determine when such assumptions provide good, or bad, approximations of true behavior. Moreover, it offers a new approach to optimize parameters and evaluate performance. This can be valuable for delay-sensitive systems that employ short block lengths.**

I. Introduction

Forward error-correcting codes have played an instrumental role in the many successes of digital communications over the past decades [1]. The fact that it is possible to transmit digital information reliably at a positive rate over an unknown noisy channel is now universally acknowledged [2]. The main cost of improving reliability is the use of increasingly long codewords [3]. One situation where the valuable lessons of classical coding theory may not apply directly is the general area of delay-constrained communications. If system specifications dictate that almost all information bits should be made available at the destination shortly after they arrived at the transmitter, it may not be possible to aggregate a large number of them before encoding and transmission. In some cases, stringent delay requirements will force a system designer to resort to short block codes or short constraint-length convolutional codes.

From a coding perspective, using short codewords on channels with memory creates two impediments. First, decoders are designed to correct the most-likely error patterns and the probability of seeing atypical error patterns cannot be neglected for short block lengths. Second, if the coherence time of the channel is longer than a codeword transmission interval, then optimal code rate may depend heavily on the channel state, which is unknown to the transmitter. Together, these factors impair the rapid transmission of information.

Coding performance as a function of block-length and code-rate has been assessed in the information theory literature using the reliability function [3]. This criterion focuses on the exponential rate at which the error probability decays with block length, known as the error exponent, as a function of information rate. The concept of a reliability function can also be extended to variable-length codes in the presence of feedback [4]. More recently, consideration has been given to the reliability function for bits with fixed delay, as opposed to coded blocks, in the presence of feedback [5].

While remarkable, these results remain asymptotic in nature and do not necessarily capture overall system behavior adequately. For delay-sensitive applications and short codewords, three interrelated effects come into play. The probability of decoding failure for every codeword is not negligible. Packet retransmissions lead to queue buildups at the source and, thereby, induce longer latencies. Channel correlation over time introduces dependencies among successive decoding attempts, which further perturb queueing behavior and end-to-end delay. This is especially true when decoding failures are likely to occur in sequence [6]. Thus, a queueing analysis is necessary when considering the behavior of communication systems subject to very stringent delay requirements.

For delay-sensitive systems with short codewords, the natural tradeoff between code-rate and probability of decoding failure is hard to characterize [7]. In a non-asymptotic regime where information is queued at the source, transmitting data at a rate slightly below Shannon capacity may lead to poor performance. Recent results in the literature hint at the fact that, for delay-constrained communication, optimal code-rate selection depends heavily on block-length and channel correlation [8], [9]. These findings are especially important for real-time traffic and live interactive sessions, as these applications are sensitive to latency and require the use of short codewords.

Guidelines for code-rate selection in the context of delay-sensitive traffic were previously obtained for an erasure channel with memory [10]. The approach favored therein, which permits a complete characterization of queueing behavior, consists in building a Markov model for the evolution of the system. Crucial assumptions that facilitate analysis can be summarized as follows: the packet arrival process at the source is Bernoulli, the packet lengths are i.i.d. geometric, the error protection uses random codes, and the channel evolution is governed by a Markov chain.

In this article, we adopt a similar formulation and extend results that were obtained for the correlated erasure case to a more encompassing Gilbert-Elliot framework. This latter class of erasure channels is common to the literature on channels with memory, and subsumes earlier work based on similar

This material is based upon work supported, in part, by the National Science Foundation (NSF) under Grant No. 0830696, by the Texas Norman Hackerman Advanced Research Program under Grant No. 000512-0168-2007, and by Qatar Telecom (Qtel). Any opinions, findings, conclusions, and recommendations expressed in this material are those of the authors and do not necessarily reflect NSF's, THECB's or Qtel's views.

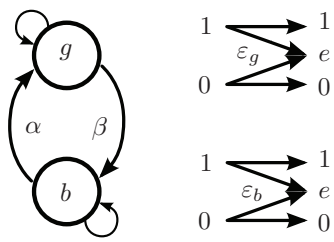

Fig. 1. A Gilbert-Elliot bit erasure channel is employed to model the operation of a communication link with memory. This model captures both the uncertainty associated with transmitting bits over a noisy channel and correlation over time typical of several communication channels.

concepts. We also present an in-depth analysis of system performance using different criteria that reflect the needs of various contemporary applications. This research is significant because it offers a new perspective on the selection of code-rate and block-length for delay-sensitive systems and provides a rigorous investigation into the effects of time-correlation on the queued performance of real-time wireless connections.

II. Channel Abstraction and Coding

Throughout, we assume that coded bits are sent from the transmitter to the destination over a Gilbert-Elliot erasure channel. This channel can be in one of two states: a *good* state g in which every bit is erased with probability ε_g and a *bad* state b in which every bit is erased with probability ε_b, independently of other bits. Our naming scheme implies $\varepsilon_b \geq \varepsilon_g$. Transitions between channel states occur according to a Markov process. The probability of transitioning to state g given that the Markov chain is currently in state b is denoted by α. The likelihood of the reverse transition from g to b is symbolized by β. Under alphabetical state ordering, the parameters of this Markov chain can be expressed in the form of a probability transition matrix,

$$\mathbf{P} = \begin{bmatrix} 1-\alpha & \alpha \\ \beta & 1-\beta \end{bmatrix}. \tag{1}$$

A graphical interpretation of the communication channel under consideration appears in Fig. 1.

The state of the channel at time n is a random variable, which we denote by C_n. Moreover, the succession of states over time, $\{C_n : n \in \mathbb{N}\}$, forms a Markov process. Finding the conditional probability $\Pr(C_{n+1} = d|C_n = c)$ amounts to selecting an entry in $\mathbf{P}$. Likewise, $\Pr(C_{n+N} = d|C_n = c)$ can be obtained by locating the corresponding entry in $\mathbf{P}^N$, the Nth power of $\mathbf{P}$. We note that this Markov chain converges to its stationary distribution at an exponential rate that depends on the second eigenvalue of $\mathbf{P}$ (i.e. $1-\alpha-\beta$).

In our analysis, a packet of length L is sectioned into M data segments each containing K information bits. Packing loss is treated implicitly since the last data segment of each packet is zero padded to K bits. Every segment is encoded separately into a codeword of length N, which is subsequently stored in the queue for eventual transmission over the Gilbert-Elliot erasure channel. Decoding failures are handled through immediate retransmission of the missing data.

A. Distribution of Erasures

A quantity that is of fundamental importance in our analysis is the conditional probability of decoding failure at the destination. An intermediary step in identifying this probability is to derive an expression for E, the number of erasures within a codeword of length N. This, in turn, depends on the number of visits to each state within N consecutive realizations of the channel. More specifically, we are interested in conditional probabilities of the form

$$\Pr(E = e, C_{N+1} = d|C_1 = c), \tag{2}$$

where $e \in \mathbb{N}_0$ and $c, d \in \{b, g\}$. The generating function for these conditional probabilities is based on generalizing the entries of $\mathbf{P}$ to the vector space of real polynomials in x with

$$\mathbf{P}_x = \begin{bmatrix} (1-\alpha)(1-\varepsilon_b+\varepsilon_b x) & \alpha(1-\varepsilon_b+\varepsilon_b x) \\ \beta(1-\varepsilon_g+\varepsilon_g x) & (1-\beta)(1-\varepsilon_g+\varepsilon_g x) \end{bmatrix}.$$

Let $[\![x^j]\!]$ be the operator which maps a polynomial in x to the coefficient of x^j. Then, the conditional probability (2) is given, in terms of the Nth power of $\mathbf{P}_x$, by

$$\Pr(E = e, C_{N+1} = d|C_1 = c) = [\![x^e]\!] \left[\mathbf{P}_x^N\right]_{c,d}.$$

It is worth mentioning that one can employ this method or alternative combinatorial means to obtain closed-form expressions for the desired conditional probabilities [11], [10].

B. Probability of Decoding Failure

During every transmission, a segment of K information bits is encoded using a code defined by a random parity-check matrix $\mathbf{H}$ of size $(N-K) \times K$, where each matrix entry is selected independently and uniformly from $\{0, 1\}$. Maximum likelihood decoding is used at the destination.

Random coding has the benefit that the probability of decoding failure depends only on the number of erasures and not on the locations of the erasures. Consequently, the decoding failure probability is a function of the number of erasures E in the block. Once the value of E is known, we can derive the desired probability as follows. Conditioned on $E = e$, decoding at the destination will succeed if and only if the submatrix of $\mathbf{H}$ formed by choosing the e erased columns has rank e [12]. Furthermore, the probability that a random $e \times p$ matrix over $\mathbb{F}_2$, where $p = N-K$ stands for the number of parity bits, has rank e is equal to $\prod_{i=0}^{e-1}\left(1-2^{i-p}\right)$. Thus, given e erasures within a codeword of length N, the probability of decoding failure can be written as

$$P_{\mathrm{f}}(N-K, e) \triangleq 1 - \prod_{i=0}^{e-1}\left(1-2^{i-(N-K)}\right).$$

The average probability of decoding failure at the destination is therefore equal to $P_{\mathrm{f}}(N-K) \triangleq \mathbb{E}\left[P_{\mathrm{f}}(N-K, E)\right]$, where the expectation over E depends implicitly on all possible channel realizations within a block. While the average probability of decoding failure offers a good measure of performance, it alone does not capture the queueing behavior of the system. Indeed, correlation among decoding-failure events may also alter the behavior of the queue at the transmitter.

III. Arrival and Departure Processes

Having introduced a precise model for the physical layer, we turn to the description of the arrival and departure processes at the queue. In our framework, the block-length, which we denote by N, remains fixed throughout and every codeword transmission requires N consecutive uses of the channel. Each data packet is broken into length-K data segments that are separately encoded into blocks. In terms of system characterization, N is fundamental in that it determines the sampling period of our Markov chain.

We assume that the packet arrival process is i.i.d. Bernoulli with parameter γ. This implies that, during each codeword transmission interval, a new packet arrives at the source with probability γ. The number of bits in each data packet is assumed to be an i.i.d. random process whose marginal distribution is geometric with parameter ρ. Therefore, the probability that a packet contains exactly ℓ bits becomes

$$\Pr(L = \ell) = (1-\rho)^{\ell-1}\rho \qquad \ell = 1, 2, \ldots$$

where $\rho \in (0, 1)$. These assumptions on the structure of the arrival process and the packet-length distribution are crucial for the construction of a tractable Markov model for our communication system. They enable a rigorous analysis of the queue and lead to meaningful guidelines for system design and optimization.

Departures from the queue are governed by the underlying Gilbert-Elliot channel and the design-rate $r = K/N$ of our random linear code. The number of information bits contained in every codeword is therefore $K = rN$. A low-rate code will, in general, have a smaller probability of decoding failure than the same system with a higher rate code. Still, the successful decoding of a codeword associated with a high-rate code leads to the transmission of a larger amount of data bits. These competing considerations create a natural tradeoff between information content and probability of decoding failure. Accordingly, the code-rate r, or equivalently the number of information bits K, is a parameter that should be optimized.

Once a code rate is selected, the number of successfully decoded codewords needed to complete the transmission of a given packet is $M = \lceil L/rN \rceil$. Since L is geometric, we find that M also has a geometric distribution, albeit with parameter

$$\rho_r = \sum_{\ell=1}^{rN} (1-\rho)^{\ell-1}\rho = 1 - (1-\rho)^{rN}.$$

The probability that a data packet requires the successful transmission of m data segments of size rN is equal to

$$\Pr(M = m) = (1-\rho_r)^{m-1}\rho_r \qquad m = 1, 2, \ldots$$

For a head packet to depart from the queue, the destination must successfully decode the most recent codeword it received, and this codeword must carry the final segment of information corresponding to this packet. Implicit to our system model is the ability of the destination to acknowledge the reception of a codeword through instantaneous feedback. Based on this side information, the transmitter is able to remove data segments and packets from the queue after successful transmission.

IV. Queueing Behavior

The number of data packets in the queue at the onset of block s is denoted by Q_s. The state of the Gilbert-Elliot channel at this same instant is represented by C_{sN+1}. Together, these two quantities form the state of our Markov process, $U_s = (C_{sN+1}, Q_s)$. We emphasize that the cardinality of this state space is countable, with U_s belonging to $\{b, g\} \times \mathbb{N}_0$. Furthermore, the Markov chain underlying the evolution of our system possesses a special structure; it forms an instance of a discrete-time *quasi-birth-death process*. Fortunately, there are many established techniques to study such mathematical objects. We present one possible approach in Section IV-A.

The transition probability from U_s to U_{s+1} is given by

$$\begin{aligned}&\Pr(U_{s+1} = (d, q_{s+1}) | U_s = (c, q_s)) \\ &\quad = \sum_{e \in \mathbb{N}_0} \Pr\left(Q_{s+1} = q_{s+1} | E = e, Q_s = q_s\right) \times \\ &\qquad \Pr\left(E = e, C_{(s+1)N+1} = d | C_{sN+1} = c\right).\end{aligned} \tag{3}$$

Recall that a methodology was introduced in Section II-A to derive the distribution of $\left(E, C_{(s+1)N+1}\right)$ conditioned on the value of C_{sN+1}. Obtaining expressions for probabilities of the type $\Pr\left(Q_{s+1} = q_{s+1} | E = e, Q_s = q_s\right)$ remains.

We first consider conditional events $\{Q_s = q_s\}$ for which $q_s > 0$; admissible values for Q_{s+1} are then limited to values in $\{q_s - 1, q_s, q_s + 1\}$. Two factors can affect the length of the queue, the arrival of a new data packet and the completion of a packet transmission. The latter occurrence will only take place if a codeword is successfully decoded at the destination and the head packet has no additional data segment left at the source. Keeping these facts in mind, we get

$$\begin{aligned}&\Pr\left(Q_{s+1} = q_s + 1 | E = e, Q_s = q_s\right) \\ &\quad = \gamma\big(P_{\mathrm{f}}(N-K, e) + (1 - P_{\mathrm{f}}(N-K, e))(1-\rho_r)\big) \\ &\Pr\left(Q_{s+1} = q_s | E = e, Q_s = q_s\right) = \gamma\left(1 - P_{\mathrm{f}}(N-K, e)\right)\rho_r \\ &\quad + (1-\gamma)\big(P_{\mathrm{f}}(N-K, e) + (1 - P_{\mathrm{f}}(N-K, e))(1-\rho_r)\big) \\ &\Pr(Q_{s+1} = q_s - 1 | E = e, Q_s = q_s) \\ &\quad = (1-\gamma)\left(1 - P_{\mathrm{f}}(N-K, e)\right)\rho_r.\end{aligned}$$

When the queue is empty, $\{Q_s = 0\}$, only two possibilities can occur,

$$\begin{aligned}\Pr(Q_{s+1} = 1 | E = e, Q_s = 0) &= \gamma \\ \Pr(Q_{s+1} = 0 | E = e, Q_s = 0) &= 1 - \gamma.\end{aligned}$$

Collecting these findings and using (3), we get the probability transition matrix of the Markov process $\{U_s\}$. A graphical rendition of the state transitions appears in Fig. 2.

To proceed with the analysis of our queued system, a compact representation of the conditional probabilities defined in (3) is apropos. For $q \in \mathbb{N}$ and $c, d \in \{b, g\}$, we introduce the following mathematical notation,

$$\begin{aligned}\mu_{cd} &= \Pr(U_{s+1} = (d, q-1) | U_s = (c, q)) \\ \kappa_{cd} &= \Pr(U_{s+1} = (d, q) | U_s = (c, q)) \\ \lambda_{cd} &= \Pr(U_{s+1} = (d, q+1) | U_s = (c, q)).\end{aligned}$$

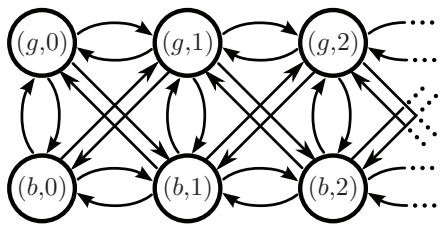

Fig. 2. State space and transition diagram for the aggregate queued process $\{U_s\}$; self-transitions are intentionally omitted.

Similarly, when the queue is empty, we use $\kappa^0_{cd} = \Pr(U_{s+1} = (d,0)|U_s = (c,0))$ and $\lambda^0_{cd} = \Pr(U_{s+1} = (d,1)|U_s = (c,0))$. Collectively, these labels define the 12 transition probabilities associated with a non-empty queue, and the 8 transition probabilities subject to the non-negativity constraint at zero.

We are ready to derive the equilibrium distribution of our system. We note that, if the channel state is ergodic and the queue is stable, then the Markov chain $\{U_s\}$ is positive recurrent and possesses a unique stationary distribution [13]. Let $U = (C, Q)$ be a random vector with the following probability distribution,

$$\Pr(U = (c,q)) = \lim_{s\to\infty} \Pr(U_s = (c,q)).$$

We employ the semi-infinite vector π as a convenient notation for the equilibrium distribution of our system, with

$$\pi(2q+i) = \begin{cases} \Pr(C = b, Q = q) & \text{if } i = 1 \\ \Pr(C = g, Q = q) & \text{if } i = 2, \end{cases}$$

for $i \in \{1,2\}$ and $q \in \mathbb{N}_0$. The states $\{(b,q),(g,q)\}$ are known as the qth level of the Markov chain and $\pi_q \triangleq [\pi(2q+1)\ \pi(2q+2)]$ is the stationary distribution associated with the qth level.

Using this compact notation, we can write the Chapman-Kolmogorov equations as $\pi\mathbf{T} = \pi$, where $\mathbf{T}$ is the probability transition matrix associated with $\{U_s\}$. One possible approach to solve for the stationary distribution of our Markov model is to employ spectral representation and ordinary generating functions [10]. In this article, we adopt an alternate means and apply the matrix geometric method [14], [15].

A. Matrix Geometric Method

We can represent the probability transition matrix $\mathbf{T}$ as a semi-infinite matrix of the form

$$\mathbf{T} = \begin{pmatrix} \mathbf{C}_1 & \mathbf{C}_0 & \mathbf{0} & \mathbf{0} & \cdots \\ \mathbf{A}_2 & \mathbf{A}_1 & \mathbf{A}_0 & \mathbf{0} & \cdots \\ \mathbf{0} & \mathbf{A}_2 & \mathbf{A}_1 & \mathbf{A}_0 & \cdots \\ \mathbf{0} & \mathbf{0} & \mathbf{A}_2 & \mathbf{A}_1 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix} \tag{4}$$

where the submatrices $\mathbf{C}_1$, $\mathbf{C}_0$, $\mathbf{A}_2$, $\mathbf{A}_1$, and $\mathbf{A}_0$ are 2×2 real matrices. More specifically, we have

$$\mathbf{A}_0 = \begin{bmatrix} \lambda_{bb} & \lambda_{bg} \\ \lambda_{gb} & \lambda_{gg} \end{bmatrix} \qquad \mathbf{A}_1 = \begin{bmatrix} \kappa_{bb} & \kappa_{bg} \\ \kappa_{gb} & \kappa_{gg} \end{bmatrix}$$
$$\mathbf{A}_2 = \begin{bmatrix} \mu_{bb} & \mu_{bg} \\ \mu_{gb} & \mu_{gg} \end{bmatrix}.$$

When the queue is empty, the relevant submatrices become

$$\mathbf{C}_0 = \begin{bmatrix} \lambda^0_{bb} & \lambda^0_{bg} \\ \lambda^0_{gb} & \lambda^0_{gg} \end{bmatrix} \qquad \mathbf{C}_1 = \begin{bmatrix} \kappa^0_{bb} & \kappa^0_{bg} \\ \kappa^0_{gb} & \kappa^0_{gg} \end{bmatrix}.$$

Note that the Markov chain associated with (4) belongs to the class of processes with repetitive structure. The following theorem characterizes its stationary distribution.

Theorem 4.1: Consider a positive recurrent Markov chain on a countable state space with transition matrix $\mathbf{T}$ given by (4). Let the positive matrix $\mathbf{R}$ be defined as the limit, starting from $\mathbf{R}_0 = \mathbf{0}$, of the matrix recursion

$$\mathbf{R}_{j+1} = (\mathbf{A}_0 + \mathbf{R}_j^2\mathbf{A}_2)(\mathbf{I} - \mathbf{A}_1)^{-1}.$$

Then, the qth-level stationary distribution π_q satisfies $\pi_{q+1} = \pi_q\mathbf{R}$ for $q \geq 1$ with $\pi_1 = \pi_0\mathbf{Z}$ and

$$\mathbf{Z} = (\mathbf{I} - \mathbf{C}_1)\mathbf{A}_2^{-1}$$
$$\pi_0 = \begin{bmatrix} \frac{\beta}{\alpha+\beta} & \frac{\alpha}{\alpha+\beta} \end{bmatrix} \left(\mathbf{I} + \mathbf{Z}(\mathbf{I} - \mathbf{R})^{-1}\right)^{-1}.$$

Corollary 4.2: The decay rate of the complementary cumulative distribution function of the queue satisfies

$$\lim_{\tau\to\infty} \tau^{-1} \log \Pr(Q \geq \tau) = \log \varrho(\mathbf{R}),$$

where $\varrho(\mathbf{R})$ is the spectral radius of $\mathbf{R}$.

V. Performance Evaluation

This mathematical characterization makes it possible to compute a wide range of advanced performance criteria for the system under consideration, including average packet error rate and outage capacity. Herein, we focus on two measures that are most relevant to delay-sensitive communications. First, we look at the probability that the queue exceeds a threshold, $\Pr(Q > \tau)$, where τ is relatively small. Second, we examine the decay rate of the complementary cumulative distribution function, as discussed in Corollary 4.2. Again, we emphasize that the tail decay in buffer occupancy is given by the dominant eigenvalue of $\mathbf{R}$.

For illustrative purposes, we select the following parameters. The Gilbert-Elliot erasure channel is defined by $\alpha = 0.02$, $\beta = 0.005$, $\varepsilon_b = 0.49$, and $\varepsilon_g = 0.0025$. This generates an average erasure probability of 0.1. The channel memory decays at an exponential rate of $(1 - \alpha - \beta) = 0.975$. The blocklength is fixed at $N = 114$ and the arrival process is defined by the arrival probability $\gamma = 0.25$ and average packet length $\rho^{-1} = 195$. If codewords are transmitted every 4.615 ms, then this corresponds to an arrival rate of roughly 10.6 Kbits/sec and an ergodic channel capacity of roughly 22.2 Kbits/sec. These parameters are selected to loosely match the operation of a wireless GSM relay link.

System performance as a function of the number of information bits per codeword, K, is shown in Fig. 3. Each curve represents the complementary cumulative distribution function evaluated at a different threshold value, $\Pr(Q > \tau)$.

As expected, the probability of the queue exceeding a prescribed threshold decreases as τ increases. More interestingly, it is instructive to notice that $K = 83$ appears uniformly

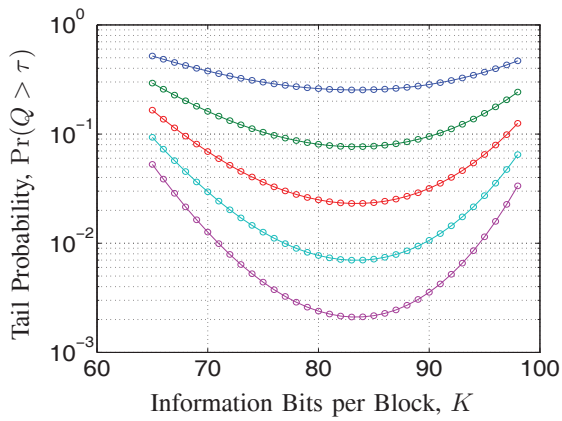

Fig. 3. This figure shows tail probabilities in the equilibrium packet distribution of the queue, $\Pr(Q > \tau)$, for threshold values $\tau \in \{5, 10, 15, 20, 25\}$. The minimums occur uniformly at $rN = 83$ for all threshold values.

optimal for all values of τ. Further supporting evidence for this observation is offered by looking at the asymptotic decay rate in tail occupancy, displayed in Fig. 4. When the arrival rate $\gamma\rho_r^{-1}$ is between 47.5 and 60, one finds that $K = 83$ is also optimal in terms of tail decay. This robustness property is very encouraging, as it simplifies system design.

An important observation that does not appear on these two figures is the fact that, for short block lengths, the optimal value of K depends heavily on the channel parameters α, β, ε_g and ε_b. A naive conjecture would place $K = rN$ close to the Shannon limit $0.9 \times 114 = 102.6$, but this is much larger than the optimal value of $K = 83$. A more sophisticated approach is to maximize the throughput of a system with an infinite-backlog. After some calculation, one finds that this leads to $K = 87$, which is much closer to the true optimum. But, as the channel memory parameter $(1-\alpha-\beta)$ varies, the optimal value of K changes substantially. In fact, as $(1-\alpha-\beta) \to 1$, K approaches N.

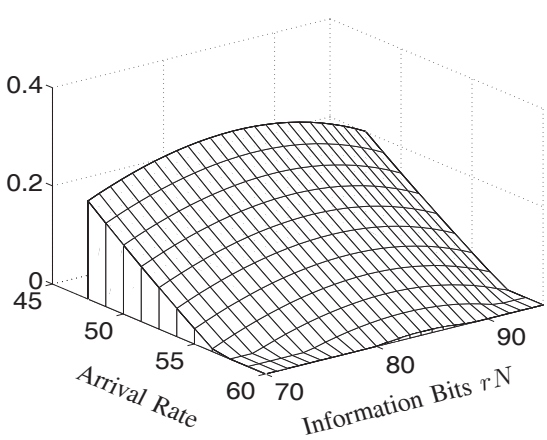

Fig. 4. This figure shows tail decay rate, $-\lim_{\tau\to\infty} \tau^{-1} \log \Pr(Q \geq \tau)$, as a function of the number of information bits rN and the average arrival rate γ/ρ_r in bits per cycle.

VI. Conclusions

This work provides a unified approach that links queueing performance with the operation of a communication system at the physical layer. The methodology and results are developed for the Gilbert-Elliot erasure channel, but can be generalized to more intricate finite-state channels with memory. For example, the simple performance characterization of random codes over erasure channels extends naturally to hard-decision decoding of BCH codes over Gilbert-Elliot error channels. For fixed parameters, the optimal code rate appears relatively insensitive to target threshold τ in the queue. Still, channel memory and cross-over probabilities can affect this optimal operating point. More generally, the optimal code rate seems to be linked to ratio between the codeword time and the coherence time of the channel.

References

[1] J. Costello, D.J., J. Hagenauer, H. Imai, and S. Wicker, "Applications of error-control coding," *IEEE Trans. Inf. Theory*, vol. 44, no. 6, pp. 2531–2560, October 1998.

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

[3] R. G. Gallager, *Information Theory and Reliable Communication*. Wiley, 1968.

[4] M. V. Burnashev, "Data transmission over a discrete channel with feedback," *Problems of Information Transmission*, vol. 12, no. 4, pp. 250–265, 1976.

[5] A. Sahai, "Why block length and delay behave differently for channel coding with feedback," *IEEE Trans. Inf. Theory*, vol. 54, no. 5, pp. 1860–1886, May 2008.

[6] L. Liu, P. Parag, J. Tang, W.-Y. Chen, and J.-F. Chamberland, "Resource allocation and quality of service evaluation for wireless communication systems using fluid models," *IEEE Trans. Inf. Theory*, vol. 53, no. 5, pp. 1767–1777, May 2007.

[7] A. Ephremides and B. Hajek, "Information theory and communication networks: an unconsummatedunion," *IEEE Trans. Inf. Theory*, vol. 44, no. 6, pp. 2416–2434, October 1998.

[8] D. Wu and R. Negi, "Effective capacity: a wireless link model for support of quality of service," *IEEE Trans. Wireless Commun.*, vol. 2, no. 4, pp. 630–643, July 2003.

[9] L. Ying, S. Yang, and R. Srikant, "Optimal delay-throughput tradeoffs in mobile ad hoc networks," *IEEE Trans. Inf. Theory*, vol. 54, no. 9, pp. 4119–4143, September 2008.

[10] P. Parag, J.-F. Chamberland, H. D. Pfister, and K. R. Narayanan, "Code rate, queueing behavior and the correlated erasure channel," in *IEEE Information Theory Workshop on Information Theory*, Cairo, Egypt, January 2009.

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

[12] T. Richardson and R. Urbanke, *Modern Coding Theory*. Cambridge University Press, 2008.

[13] J. R. Norris, *Markov Chains*, ser. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, 1998.

[14] M. F. Neuts, *Matrix-Geometric Solutions in Stochastic Models: An Algorithmic Approach*. Dover Publications, 1995.

[15] G. Latouche and V. Ramaswami, *Introduction to Matrix Analytic Methods in Stochastic Modeling*, ser. ASA-SIAM Series on Statistics and Applied Probability. Society for Industrial Mathematics, 1987.