CCIT Report #468 February 2004

On the Reliability Exponent of the Exponential Telephone Signaling Channel

Yakir Ovadia * Shraga I. Bross †

February 4, 2004

Abstract

Abstract—A lower bound on the reliability exponent of the memoryless exponential server timing channel with noiseless feedback is provided. The lower bound depends on whether fixed or random transmission-time, as well as on whether fixed block-length or variable block-length, codes are considered (with block-length denoting the number of recorded departures). On the other hand we show that Arikan's one-way sphere-packing bound for fixed transmission-time codes [8] applies as well to fixed transmission-time codes for the case at hand.

Index Terms – Point process channel, reliability exponent, noiseless feedback, timing channel, telephone signaling channel.

I. Introduction

The exponential-server timing channel is a model for a single-server queue with a first-in first-out service discipline wherein the server's service times are independent and exponentially distributed with mean $1/\mu$ (see [1]-[7] and references therein). Transmission begins at time 0 with the queue containing a possibly non-zero amount u_0 of unfinished work. The model allows u_0 to be a random variable except that it must be independent of the message transmitted. The message to be transmitted is encoded via a codeword $\tilde{\boldsymbol{x}} = (\tilde{x}_1, \tilde{x}_2, \ldots, \tilde{x}_n)$ which consists of n nonnegative components that determine the interarrival times of packets to the queue. The receiver observes the interdeparture times $\boldsymbol{y} = (y_1, y_2, \ldots, y_n)$ of packets from the queue and makes its decision based on this.

* All Optical, Yoqneam 20692, Israel email:yakir@alloptical.co.il

† Department of Electrical Engineering Technion, Haifa 32000, Israel email:shraga@ee.technion.ac.il

Given the interarrival vector $\tilde{\boldsymbol{x}}$, and the initial unfinished work u_0, the conditional probability density of observing y_0 and thereafter the interdeparture vector $\boldsymbol{y}$ is

$$\begin{aligned} p(\boldsymbol{y}, y_0 | \tilde{\boldsymbol{x}}, u_0) &= \delta(y_0 - u_0) \prod_{k=1}^{n} e_\mu(y_k - w_k) \\ w_k &= \max\left\{0\ ,\ \sum_{\ell=1}^{k} \tilde{x}_\ell - \sum_{\ell=0}^{k-1} y_\ell\right\} \end{aligned} \tag{1}$$

Here $e_\mu(t) = \mu e^{-\mu t}, t \geq 0$ denotes the exponential density with mean $1/\mu$, w_k is the amount of time for which the server is idle between the $(k-1)$st departure and the kth arrival, and $\delta(\cdot)$ is the delta function.

Two definitions of a block code for the timing channel have been considered in [2], depending, respectively, on whether it is a random or a fixed transmission-time scheme.

Definition 1: An (n, M, T, P_e)-*code* consists of a codebook of M codewords and a decoder. Each codeword is a vector of n nonnegative interarrival times $\tilde{\boldsymbol{x}}$. The decoder, after observing the n departures, selects the correct codeword with probability greater than $1-P_e$, under equiprobable codewords and $p(\boldsymbol{y}|\tilde{\boldsymbol{x}})$. The nth departure occurs on the average (under equiprobable codewords and $p(\boldsymbol{y}|\tilde{\boldsymbol{x}})$) no later than T.

Definition 2: An (n, M, T, P_e)-*window-code* consists of a codebook of M codewords and a decoder. Each codeword is a vector of n nonnegative interarrival times $\tilde{\boldsymbol{x}}$. The nth arrival of every codeword occurs before time T. The decoder, after observing the departures in $[0, T]$, selects the correct codeword with probability greater than $1 - P_e$, under equiprobable codewords and $p(\boldsymbol{y}|\tilde{\boldsymbol{x}})$.

Thus the term "window-code" refers to a code whose decoder observes only those departures that occur in the window $[0, T]$, and the cost of transmission for the window-code is T seconds.

We shall use the average probability of error criterion assuming that the message D is drawn according to a uniform distribution over $\{1, \ldots, M\}$. Consequently, the error probability for the code is defined as

$$P_e = \Pr\{D \neq \psi(\boldsymbol{Y})\} = \frac{1}{M} \sum_{m=1}^{M} \Pr\{\psi(\boldsymbol{Y}) \neq m | D = m\}\ ,$$

where ψ is the decoding function.

The rate R is achievable if, for every $\gamma > 0$, there exists a sequence of $(n, M^{(n)}, T^{(n)}, P_e^{(n)})$-codes/window-codes that satisfies $(\ln M^{(n)})/T^{(n)} > R - \gamma$ for all sufficiently large n, and $\lim_{n\to\infty} P_e^{(n)} = 0$.

By [1, Thm. 4, Thm. 6] and [2, Thm. 2] any rate not larger than

$$C(\mu,\lambda) \quad = \quad \lambda \ln \frac{\mu}{\lambda} \quad , \quad 0 \le \lambda \le \mu \ , \tag{2}$$

is achievable with departure-rate λ codes/window-codes on the continuous-time exponential-server timing channel. Thus the rate $R = \mu/e$ nats per second is achievable with departure-rate $\lambda = \mu/e$ packets per second.

The coding theorem of Anantharam-Verdú shows that an arbitrarily small average error probability is possible for any rate smaller than $C(\lambda, \mu)$, but it does not tell us how large the coding duration T must be in order to achieve a specified error probability. A partial answer to this question is provided by examining the error exponent of the channel.

Definition 3: A number $E \ge 0$ is called an *achievable error exponent* at rate $R > 0$ for the exponential timing channel if, for every $\delta > 0$ and sufficiently large T, there exists a code with coding duration T such that $M \ge \exp\{T(R-\delta)\}$ and $P_e \le \exp\{-T(E-\delta)\}$.

The fixed (resp. random) transmission-time *reliability exponent* at rate $R > 0$, referred as $E(R)$, is the largest achievable error exponent at that rate, and formally it is defined as follows.

$$E(R) = \limsup_{T\to\infty} -\frac{\ln P_e(R,T)}{T} \ ,$$

where $P_e(R,T)$ is the minimum value of the Maximum-Likelihood (ML) decoding error P_e over all (n, M, T) fixed (resp. random) transmission-time codes for the given channel such that $M = \lceil e^{RT} \rceil$ and $n \ge 1$ is arbitrary. The zero-rate reliability exponent is defined as

$$E(0) = \sup_{R>0} E(R) \ .$$

With regard to the performance of (n, M, T, P_e) window-codes for the exponential timing channel, Arikan established in [8] the following result.

Proposition 1 [8]: Let $E_{sp}(R)$ for $0 < R < C = \mu/e$ be defined parametrically by

$$\begin{aligned} E_{sp}(R) \quad &= \quad \frac{\mu}{(1+\rho)^{(1+\rho)/\rho}}[\rho - \log(1+\rho)] \\ R \quad &= \quad \frac{\mu}{(1+\rho)^{(1+\rho)/\rho}} \log(1+\rho)^{1/\rho} \end{aligned} \tag{3}$$

as ρ ranges over $(0, \infty)$, and let $E_{sp}(0) = \mu$.

For $R_c \triangleq (\mu/4)\log 2 < R < C$, let $E_r(R) = E_{sp}(R)$, while for $0 \le R \le R_c$, let $E_r(R) = \mu/4 - R$. Then the reliability exponent of the exponential-server timing channel with mean service time $1/\mu$ satisfies

$$E_r(R) \le E(R) \le E_{sp}(R)$$

for all $0 \leq R < C$.

Thus, for rates below the critical rate R_c, the one-way fixed transmission-time reliability exponent is yet unknown, and in particular this result establishes the bound $\mu/4 \leq E(0) \leq \mu$. Recently Wagner and Anantharam [9] improved this estimate by showing that $E(0) \geq \mu/2$.

The Telephone Signaling Channel (TSC) is a memoryless variant of the timing channel that incorporates noiseless feedback and is formally defined as follows. The queue starts with $u_0 = 0$. Upon placing the kth customer ($k \geq 1$), the transmitter waits (for s_k seconds) until the corresponding kth departure, and then after a period of time x_{k+1} which depends on the message being sent (and possibly on previous departure times), places the $(k+1)$st customer. The times $s_1, s_2, \ldots$ are independent and identically distributed. While the capacity of the exponential TSC (i.e. an exponential timing channel with noiseless feedback) has been shown in [1]-[2] to be the same as that of the one-way exponential timing channel, the reliability exponent for this model is yet unknown. It should be mentioned in this connection that, for the ideal Poisson channel with noiseless feedback — a model that is equivalent to the TSC with customer recall [1] — Lapidoth presented in [11] a code construction that attains the reliability exponent, in the case of a capacity reducing average-power constraint.

In this work we investigate the reliability exponent of the exponential TSC under various coding strategies. The main contribution of this paper is a pair of achievable exponents the former with fixed transmission-time schemes and the latter with random transmission-time schemes. All results establish tighter lower bounds than $E_r(R)$ over some rate intervals. This demonstrates that, akin to the ideal Poisson channel [10, 11], whereas all variants of the timing channel exhibit the same capacity [1, 2, 4] the reliability exponent of these channels is not necessarily the same.

The paper is organized as follows. Section II presents definitions of the families of codes considered herein and the corresponding achievable exponents with each of these ensembles. The section concludes with a discussion of the coding implications of some of the main results. Section III.A is devoted to the performance analysis of some fixed block-length code ensembles either in a fixed transmission-time or a random transmission-time scheme, while Section III.B considers, respectively, the performance of some variable-length code ensembles. In section III.C we argue the applicability of Arikan's one-way sphere-packing bound to fixed transmisson-time codes for our channel.

II. Definitions and Main Results

Notation

As feedback allows for encoding schemes which constitute of either transmission with variable block-length or transmission with fixed block-length, we shall henceforth adopt the following notation. Random variables will be denoted by capital letters, while their realizations will be denoted by the respective lower case letters. Whenever the dimension of a random vector is clear from the context the random vector will be denoted by a bold face letter, that is, $\boldsymbol{X}$ denotes the random vector $(X_1, X_2, \dots, X_n)$, and $\boldsymbol{x} = (x_1, x_2, \dots, x_n)$ will designate a specific sample value of $\boldsymbol{X}$. However, when we consider transmission schemes wherein the block-length is a random variable and therefore we'd find it important to emphasize explicitly the dimension of a random vector — $X_1^{n_1}$ shall denote the random vector $(X_{1,1}, X_{1,2}, \dots, X_{1,n_1})$, and $x_1^{n_1} = (x_{1,1}, x_{1,2}, \dots, x_{1,n_1})$ will designate a specific sample value of $X_1^{n_1}$.

A. Fixed Transmission-Time

In a TSC (as observed in [1]-[2]) the encoder, who knows precisely when a customer departs from the server, can make use of this information in order to avoid queuing altogether. In that case the channel model boils down to the memoryless additive noise channel

$$Y_j = X_j + S_j \ , \quad j = 1, 2, \dots$$

This observation suggests that an (M, T, P_e)-*feedback window-code* code may be defined as follows:

1. A codebook of M encoding functions

$$f_j : \{1, \dots, M\} \times \mathbb{R}_+^{j-1} \to \mathbb{R}_+, \quad j = 1, 2, \dots, k$$

   where $f_j(m, y^{j-1})$ is the waiting time chosen by the encoder before loading the jth customer into the queue when attempting to convey message m to the receiver after obtaining the previous $j-1$ customer interdepartures $y^{j-1} = (y_1, \dots, y_{j-1})$. Here each transmitted codeword is a variable length sequence of nonnegative real numbers x^k (where k is a random variable), and transmission is terminated by the time T.

2. A decoding function $\psi : \mathbb{R}_+^{\tilde{k}} \to \{1, \dots, M\}$ such that given the channel output $\boldsymbol{y}$ during $[0, T]$ it produces the correct message with probability greater than $1 - P_e$. Here $\tilde{k}$ is the number of the recorded departures during $[0, T]$.

In this definition of a code feedback is used not only to avoid queuing as rather it allows the encoder to choose the jth waiting time as a function of the message m as well as the past $j - 1$ interdeparture times.

First we argue that the attainable exponent with the feedback code defined above is not smaller than the attainable exponent with the one-way window-code over the channel (1). To see that, notice that if given a codeword $\tilde{\boldsymbol{x}}(m)$, intended for one-way transmission, and any corresponding realization $\boldsymbol{y}$ of the departure process for the channel (1) a "corresponding codeword" $\boldsymbol{x}(m)$ for the TSC is obtained by setting $x_j(m) = w_j$. Here a corresponding codeword is in the sense that had the vector of service times been identical in both cases the same interarrival and departure vectors would have been observed for both channels. Therefore, any lower bound on the achievable exponent for the one-way channel (1) is also achievable on the feedback channel (e.g. $E_r(R)$, $E(0) \geq \mu/2$).

The codes we use to calculate the various lower bounds for the reliability function of the memoryless feedback channel are somewhat "degenerate" in the sense that they don't exploit all degrees of freedom provided by feedback. The encoder makes use of feedback to avoid queuing and to control completely the idling times of the server. Feedback, however, is not used to choose the waiting times. Nevertheless we show, the non-obvious fact, that Arikan's lower bound is still achievable with this family of codes and during a fixed transmission interval. On the other hand we show that Arikan's one-way sphere-packing bound for fixed transmission-time codes applies as well for the TSC even if feedback is used to choose the waiting times.

Formally, the family of fixed transmission-time variable block-length codes considered henceforth is defined as follows.

Definition 4: An (M, T, P_e)-*feedback window-code* consists of a codebook of M codewords and a decoder. Each transmitted codeword is a variable length sequence of nonnegative real numbers x^k (where k is a random variable). Transmitting the codeword $x^k(m)$ means that the first arrival occurs at $t = x_1(m)$. The jth component, $x_j(m)$, is the amount of time the encoder will wait after the $(j-1)$st departure, before loading the jth customer into the queue, $j = 2, \ldots, k$. The transmission is terminated by the time T. The decoder, after observing departures in $[0, T]$, selects the correct codeword with probability greater than $1 - P_e$, under equiprobable codewords and $p(\boldsymbol{y}|\boldsymbol{x})$.

Thus, the code in Definition 4 allows for the exchange of a message based on the transmission of a variable block-length codeword (i.e. a codeword consisting of a random number of packets) during a fixed transmission-time T. The number $\tilde{k}$ of recorded departures during $[0, T]$ will admit either the value $\tilde{k} = k$ or $\tilde{k} = k - 1$, depending on the kth service time s_k.

An explicit decoder the performance of which will be considered herein is similar to the one suggested in [2, Section II.C]— it chooses the codeword that explains the received sequence of departures with the minimum sum of service times. For a candidate codeword, $\boldsymbol{x}$, let

$$d(\boldsymbol{x},\boldsymbol{y}) \triangleq \begin{cases} \max\{0,\ T - x_1\}, & \tilde{k}=0 \\ \sum_{j=1}^{\tilde{k}} d'(y_j - x_j) + \max\left\{0,\ T - \sum_{j=1}^{\tilde{k}} y_j - x_{\tilde{k}+1}\right\}, & \tilde{k}>0 \end{cases} \tag{4}$$

where

$$d'(r) \triangleq \begin{cases} r, & \text{if } r \geq 0 \\ +\infty, & \text{if } r < 0 \end{cases}$$

Given the codebook $\{\boldsymbol{x}(1), \boldsymbol{x}(2), \ldots, \boldsymbol{x}(M)\}$, the decoder ψ_d maps $\boldsymbol{y}$ to

$$\psi_d(\boldsymbol{y}) \triangleq \begin{cases} i, & \text{if } d(\boldsymbol{x}(i),\boldsymbol{y}) < \min_{j\neq i} d(\boldsymbol{x}(j),\boldsymbol{y}) \\ 0, & \text{if no such } i \text{ exists} \end{cases} \tag{5}$$

where an output 0 is interpreted as a decoding error.

Proposition 2: The fixed transmission-time variable block-length reliability exponent $\hat{E}_f(R)$ of the exponential TSC with service rate μ is lower bounded, for every $0 \leq R < C$, as follows

$$\hat{E}_f(R) \geq \max\left\{\hat{E}_r^{(f)}(R)\ ,\ \hat{E}_{ex}^{(f)}(R)\right\}\ . \tag{6}$$

Here the first component on the r.h.s. of (6), $\hat{E}_r^{(f)}(R) \equiv E_r(R)$, is identical Arikan's exponent for the one-way exponential timing channel as given by (3). The second component on the r.h.s. of (6) is defined by

$$\hat{E}_{ex}^{(f)}(R) \ = \ \sup_{\rho\geq 1}\{\mu a(\rho)/2 - \rho R\}\ , \tag{7}$$

where $a(\rho)$ is the solution of

$$2a^{\rho}(2-a)^{\rho-1} \ = \ 1\ . \tag{8}$$

In particular $\hat{E}_{ex}^{(f)}(0) = \mu/2$ (see Appendix I).

A sketch of the functions $\hat{E}_r^{(f)}(R)$ and $\hat{E}_{ex}^{(f)}(R)$ is given in Fig. 1.

Thus, as far as fixed transmission-time is considered, variable block-length feedback codes attain at least the corresponding fixed-time one-way exponent even if queuing is avoided.

Next, we argue that the sphere-packing bound, obtained by Arikan in [8, Section III] for the one-way exponential timing channel, is valid for fixed transmission-time codes over the TSC.

Proposition 3: The reliability exponent of the fixed transmission-time exponential TSC with service rate μ is bounded from above, for every $0 \leq R \leq C = \mu/e$, by

$$\hat{E}_f(R) \leq E_{sp}(R)\ .$$

A sketch of the function $E_{sp}(R)$ is given in Fig. 1.

Finally, we consider fixed block-length fixed transmission-time codes, that are defined as follows.

Definition 5: An (n, M, T, P_e)-*feedback window-code* consists of a codebook of M codewords and a decoder. Each codeword is a vector of n nonnegative real numbers $\boldsymbol{x}$. Transmitting the codeword $\boldsymbol{x}(m)$ means that the first arrival occurs at $t = x_1(m)$. The kth component, $x_k(m)$, is the amount of time the encoder will wait after the $(k-1)$st departure, before loading the kth customer into the queue, $k = 2, \ldots, n$. The arrival of the nth customer occurs before time T. The decoder, after observing departures in $[0, T]$, selects the correct codeword with probability greater than $1 - P_e$, under equiprobable codewords and $p(\boldsymbol{y}|\boldsymbol{x})$.

The number $\tilde{n}$ of recorded departures during $[0, T]$ will admit either the value $\tilde{n} = n$ or $\tilde{n} = n - 1$. Nevertheless, without loss of optimality, we restrict the code to have $\tilde{n} = n$, thus we assume that the *departure* of the nth customer occurs before time T. Consequently, the decoding rule is similar to (4)-(5) with the change

$$d(\boldsymbol{x}, \boldsymbol{y}) \triangleq \sum_{j=1}^{n} d'(y_j - x_j)\,. \tag{9}$$

It is conceivable that the restriction to n-length codewords that match into a fixed time interval would penalize the error exponent. The next result provides an achievable exponent with this ensemble.

Proposition 4: The fixed block-length fixed transmission-time reliability exponent of the exponential TSC with service rate μ is lower bounded, for every $0 \le R < C$, as follows

$$\tilde{E}_f(R) \ge \max\left\{\tilde{E}_r^{(f)}(R)\ ,\ \tilde{E}_{ex}^{(f)}(R)\right\}\ , \tag{10}$$

where the first component on the r.h.s. of (10) is defined by

$$\tilde{E}_r^{(f)}(R) = \sup_{\substack{0<r\lesssim 1\\ 0<a\lesssim 1\\ 0<\rho\lesssim 1}}\ \sup_{0\le\theta_\mu<\frac{a}{(1-a)(1+\rho)}}\ \min\left\{\tilde{E}_r^{(0)}(R)\ ,\ \tilde{E}_r^{(1)}\right\} \tag{11}$$

with

$$\tilde{E}_r^{(0)}(R) =$$
$$-\mu\frac{ra\left[\theta_\mu(1+\rho)+1\right]}{a\left[\theta_\mu(1+\rho)+1\right]+(1-a)(1+\rho)}\ln\left\{\frac{a^{(1+\rho)}e^{-\frac{\theta_\mu(1-a)(1+\rho)}{a(\theta_\mu+1/(1+\rho))}}}{a-\theta_\mu(1-a)(1+\rho)}\right\}-\rho R$$

$$\tilde{E}_r^{(1)} = \mu\left\{1 - r - \frac{ra\left[\theta_\mu(1+\rho)+1\right]}{a\left[\theta_\mu(1+\rho)+1\right]+(1-a)(1+\rho)} \cdot \ln\left[\frac{a\left[\theta_\mu(1+\rho)+1\right]+(1-a)(1-r)(1+\rho)}{ra\left[\theta_\mu(1+\rho)+1\right]}\right]\right\} \tag{12}$$

and the second component on the r.h.s. of (10) is defined by

$$\tilde{E}_{ex}^{(f)}(R) = \sup_{\substack{0<r\leq 1\\ 0\leq a\leq 1\\ 0<B_\mu}} \min\left\{\sup_{\substack{1\leq\rho\\ 0\leq\theta_B<1}} \tilde{E}_{ex}^{(0)}(R)\ ,\ \tilde{E}_{ex}^{(1)}\right\} \tag{13}$$

with

$$\tilde{E}_{ex}^{(0)}(R) = -\mu\frac{rB_\mu\rho}{B_\mu+(1-a)}\ln\left\{\frac{e^{-2\theta_B(1-a)}\left[\frac{a^2}{2\rho}+B_\mu\frac{(1-a\theta_B)^2}{1-\theta_B}\right]}{B_\mu(1-\theta_B)+\frac{1}{2\rho}}\right\} - \rho R$$

$$\tilde{E}_{ex}^{(1)} = \mu\left\{1 - r - \frac{rB_\mu}{B_\mu+(1-a)}\ln\left[\frac{B_\mu+(1-a)(1-r)}{rB_\mu}\right]\right\}\ . \tag{14}$$

In particular $\tilde{E}_{ex}^{(f)}(0) = \mu/2$ *(see Appendix I).*

A sketch of the functions $\tilde{E}_r^{(f)}(R)$ and $\tilde{E}_{ex}^{(f)}(R)$ is given in Fig. 2.

Our results indicate on the structure of good low-rate codes for the exponential-server timing channel in the presence of feedback. It is our hope that this might reveal some insight into the structure of good low-rate codes for the one-way model as well.

B. Random Transmission-Time

If unbounded random transmission-time codes with mean coding duration T are allowed, a better error exponent can be achieved.

Definition 6: An (n, M, T, P_e) *-feedback-code* consists of a codebook of M codewords and a decoder. Each codeword is a vector of n nonnegative real numbers $\boldsymbol{x}$. Transmitting the codeword $\boldsymbol{x}(m)$ means that the first arrival occurs at $t = x_1(m)$. The kth component, $x_k(m)$, is the amount of time the encoder will wait after the $(k-1)$st departure, before loading the kth customer into the queue, $k = 2, \ldots, n$. The last customer exits on the average (under equiprobable codewords and $p(\boldsymbol{y}|\boldsymbol{x})$) no later than T. The decoder, after observing the n departures, selects the correct codeword with probability greater than $1 - P_e$, under equiprobable codewords and $p(\boldsymbol{y}|\boldsymbol{x})$.

The code in Definition 6 has a decoder that observes all n departures; the cost of transmission in that case is the expected time of the nth departure. The decoding rule is similar to (4)-(5) with the change made in (9).

Proposition 5: The fixed block-length random transmission-time reliability exponent of the exponential TSC with service rate μ is lower bounded, for every $0 \leq R < C$, as follows

$$E_f(R) \geq \max\left\{E_r^{(f)}(R)\ ,\ E_{ex}^{(f)}(R)\right\}\ , \tag{15}$$

where

$$\begin{aligned} E_r^{(f)}(R) &= \sup_{\substack{0<a\leq 1\\ 0<\rho\leq 1}} \ \sup_{0\leq\theta_\mu<\frac{a}{(1-a)(1+\rho)}} E_{r'}(R) \\ E_{ex}^{(f)}(R) &= \sup_{\substack{0\leq a\leq 1\\ 0<B_\mu}} \ \sup_{\substack{1\leq\rho\\ 0\leq\theta_B<1}} E_{ex'}(R)\ , \end{aligned} \tag{16}$$

and

$$\begin{aligned} E_{r'}(R) &= -\mu\frac{a\left[\theta_\mu(1+\rho)+1\right]}{a\left[\theta_\mu(1+\rho)+1\right]+(1-a)(1+\rho)}\ln\left\{\frac{a^{(1+\rho)}e^{-\frac{\theta_\mu(1-a)(1+\rho)}{a(\theta_\mu+1/(1+\rho))}}}{a-\theta_\mu(1-a)(1+\rho)}\right\}-\rho R \\ E_{ex'}(R) &= -\mu\frac{B_\mu\rho}{B_\mu+(1-a)}\ln\left\{\frac{e^{-2\theta_B(1-a)}\left[\frac{a^2}{2\rho}+B_\mu\frac{(1-a\theta_B)^2}{1-\theta_B}\right]}{B_\mu(1-\theta_B)+\frac{1}{2\rho}}\right\}-\rho R\ . \end{aligned} \tag{17}$$

In particular $E_{ex}^{(f)}(0) = \mu$ (see Appendix I).

A sketch of the functions $E_r^{(f)}(R)$ and $E_{ex}^{(f)}(R)$ is given in Fig. 3.

Next, we consider variable block-length random transmission-time list-size L codes, that are defined as follows.

Definition 7: An (M, T, P_e)-*list-size L feedback code* consists of a codebook of M codewords and a decoder. Each transmitted codeword is a variable length sequence of nonnegative real numbers x^k (where k is a random variable). Transmitting the codeword $x^k(m)$ means that the first arrival occurs at $t = x_1(m)$. The jth component, $x_j(m)$, is the amount of time the encoder will wait after the $(j-1)$st departure, before loading the jth customer into the queue, $j = 2, \ldots, k$. The transmission is terminated once the decoder can form a list of no more than $L\,(L \geq 1)$ most probable codewords such that the correct codeword is within this list with probability greater than $1 - P_e$ under equiprobable codewords and $p(\boldsymbol{y}|\boldsymbol{x})$. The kth customer exits on the average no later than T.

Thus, the code in Definition 7 allows for the exchange of a message based on the transmission of a variable length codeword during a random transmission-time with expected duration of T. In the special case of $L = 1$ and $P_e = 0$, the decoder selects the correct codeword a.s. and the corresponding rate establishes a lower bound for $C_{ze}^{(r)}$ — the random transmission-time *zero error capacity.*

Proposition 6: The random transmission-time zero error capacity of the exponential TSC is lower bounded by

$$C_{ze}^{(r)} \geq \sup_{f_X(x)} -\frac{\mathsf{E}\left[\ln\left(F_X(Y)\right)\right]}{\mathsf{E}[Y]} \ , \tag{18}$$

where the supremum is taken over all laws $f_X(x)$ having nonnegative support, and the mean is taken w.r.t. the law $f_Y(y)$ defined by $f_Y(y) = \int_0^y f_X(x) e_\mu(y-x)dx$. In particular

$$C_{ze}^{(r)} \geq 0.1829\mu \ . \tag{19}$$

It is not surprising that the random transmission-time exponent is superior to the fixed transmission-time one, as Lapidoth already demonstrated in [11] that, if the expected transmission time is the relevant quantity, zero error transmission is feasible for the ideal Poisson channel with noiseless feedback whereas the fixed transmission-time exponent is finite.

C. Discussion of Results

We analyze the performance of various coding strategies for the memoryless exponential-server timing channel with noiseless feedback under the restriction that feedback is used but to aviod queuing.

For fixed transmission-time codes, the largest random coding exponent is obtained with a variable block-length scheme. In that case the random coding exponent equals Arikan's one-way result, while at low rates a tighter estimate using expurgation is provided. It is an open problem to see if the same expurgated exponent can be achieved with one-way codes; as yet the Wagner-Anantharam result [9] and our result yield the same lower bound of $\mu/2$ at $R = 0$.

Since the one-way sphere-packing bound holds for the memoryless feedback channel, even if feedback is exploited to define the transmitted codeword, it follows that the aforementioned code ensemble attains the reliability exponent of the feedback channel at rates $R_c \leq R < C$.

Let $\lambda_c(R)$ and $\mu_c(R)$ be defined parametrically, for $0 < R < C$, by

$$\begin{aligned} R &= \frac{\mu}{(1+\rho)^{(1+\rho)/\rho}} \ln(1+\rho)^{1/\rho} \\ \lambda_c &= \frac{\mu}{(1+\rho)^{(1+\rho)/\rho}} \\ \mu_c &= \frac{\mu}{1+\rho} \ , \end{aligned}$$

as ρ ranges over $(0, \infty)$. The function λ_c determines the departure rate for the critical channel that achieves the sphere packing exponent when μ_c is the server's service rate. For

$0 < \rho \leq 1$, $\lambda_c(R)$ is also the departure rate that attains the one-way random coding exponent [8].

Similarly, for the memoryless feedback channel, when considering the subcode that achieves the sphere packing exponent, λ_c determines the departure rate for the critical channel when μ_c is the server's service rate. On the other hand, the same subcode has a departure rate denoted as $\lambda^{(\mu)}$ when μ is the service rate. Recalling that the codewords' components in our feedback code determine precisely the waiting times, we now choose a codebook with codewords that their components have an average waiting time of β. Consequently, $\beta + \frac{1}{\mu} = \frac{1}{\lambda^{(\mu)}}$. Since the same codebook is used in both service rates, we also know that $\beta + \frac{1}{\mu_c} = \frac{1}{\lambda_c}$. Eliminating β we conclude that

$$\lambda^{(\mu)} = \frac{\mu}{(1+\rho)^{(1+\rho)/\rho} - \rho} .$$

On the other hand, when feedback is used just to avoid queuing, our results show that a random coding exponent, that coincides with the sphere-packing exponent at rates $R_c \leq R < C$, is achieved with a departure rate that is defined via the same parameter $0 < \rho \leq 1$ as follows

$$\lambda_{vl} = \frac{\mu}{(1+\rho)^{(1+\rho)/\rho} - \rho} .$$

These results imply that for $R_c \leq R < C$, λ_{vl} is the optimal departure rate. Any code with different departure rate may exhibit an inferior error exponent. Thus, in this feedback scheme the server is busy for a larger fraction of time except for rates approaching capacity in which case $\lambda_{vl}(C) = \lambda_c(C)$. It should be emphasized however that even when $\lambda_{vl}(R) > \lambda_c(R)$ the waiting times are drawn according to the law (22) which is a mixture of a point mass at zero and an exponential distribution with mean $1/\lambda_c(R)$.

The variable block-length scheme can be implemented as follows. Both the encoder and decoder generate *a priori* a codebook the block-length of which equals some fixed integer n. Then the loading of customers into the queue is terminated either when the coding interval $[0, T]$ is completely consumed by $k \leq n$ customer arrivals or when the full block-length $k = n$ is exhausted before the coding interval ends, while in both cases the decoder observes departures in $[0, T]$. Our analysis in sections III.B.2-III.B.3 shows that if the a priori block-length is chosen properly as $n \geq n_{min} = \alpha(R)T$ (where $\alpha(R)$ is defined subsequently) the achievability of the claimed error exponent is guaranteed. Thus, the suggested variable block-length scheme doesn't require the availability of a common source of randomness at the transmitter and receiver, as a *deterministic* fixed block-length codebook is sufficient for this purpose.

For rates $R_c \leq R < C$, $\alpha(R)$ is defined parametrically via the above parameter $0 < \rho \leq 1$ as $\alpha(\rho) = \frac{\rho\mu}{(1+\rho)^{(1+\rho)/\rho}\ln(1+\rho)}$. For rates $R_l = \frac{1}{4}\ln\frac{4}{3} < R < R_c$, $\alpha(R) = \alpha(R_c)$ and also $\lambda_{vl}(R) = \lambda_{vl}(R_c)$, while for rates $0 < R \leq R_l$, $\alpha(R)$ and $\lambda_{vl}(R)$ are defined by the following set of equations

$$\begin{aligned} R &= -\frac{\mu a(2-a)\left[\ln a + \ln(2-a)\right]}{2[a+2\rho(1-a)]} \\ \lambda_{vl} &= \frac{\mu a}{a+2\rho(1-a)} \\ \alpha &= -\frac{\mu a}{2\rho\ln a} , \end{aligned}$$

as ρ ranges over $(1,\infty)$, and a ranges over $\left(\frac{1}{2},1\right)$, and both satiafy the relationship

$$1 = 2a^{\rho}(2-a)^{\rho-1}.$$

A sketch of the functions $\alpha(R)$, $\lambda_{vl}(R)$ and $\lambda_c(R)$ is given in Fig. 4. As expected, $\alpha(R) \geq \lambda_{vl}(R) \geq \lambda_c(R)$.

In section III.A.2, a specific decoding scheme is suggested, analyzing it in the context of fixed transmission-time fixed block-length codes, gives rise to some upper bounds on the probability of error which coincide with the various upper bounds obtained via Gallager's random coding and expurgation expressions. This decoder can be analyzed in the context of variable block-length codes as well, again resulting with the same exponents obtained via Gallager's approaches.

Next we wish to discuss our results as well as those of [8, 9] with regard to the analogy, first mentioned in [1], between the exponential timing channel (ETC) and the ideal Poisson channel (IPC) (also known as the direct detection photon channel, point process channel, etc.). The input to the IPC is a waveform $\lambda(t)$ that determines the rate of a corresponding Poisson process which the receiver observes. We consider henceforth the special case of an IPC and an input peak constraint $0 \leq \lambda(t) \leq \mu$. As observed in [8], the capacity of both the ETC and the IPC is given by μ/e. The capacity of the IPC is achieved by using waveforms with average power μ/e while the capacity-achieving departure rate for the ETC is $\lambda = \mu/e$. Furthermore, this analogy extends further to the random coding exponent $E_r(R)$ that coincides for both channels as well as to the sphere-packing exponent $E_{sp}(R)$ which is the same for both models.

By showing that the zero-rate error exponent of the ETC is at least $\mu/2$ Wagner and Anantharam conclude in [9] that the ETC is strictly more reliable than the IPC. The reason for this can be explained as follows. In the IPC the receiver observes a doubly-stochastic point process in which case, by activating $\lambda(t) = \mu$ for a period of Δ, the encoder controls just the expected number of arrivals during that period of time, namely $\mathsf{E}[N_\Delta] = \mu\Delta$, while it is incapable of generating a precise predetermined number of arrivals as the timing channel encoder does. On the other hand, by setting $\lambda(t) = 0$, the Poisson channel encoder is capable of terminating the arrival process at any given time, while the timing channel encoder once placing customers into the queue it cannot recall them any further. Since at low rates optimal timing channel codes, in the sense of the reliability exponent, exhibit vanishing departure rates the latter disadvantage of the timing channel encoder becomes insignificant while the ability to place a single customer that generates just a single departure (the departure time of which is drawn Exponential(μ) following the arrival time) becomes crucial. Therefore, the ETC is more reliable than the IPC at low rates.

The situation is completely different, however, in the presence of noiseless feedback. Here, as demonstrated by Lapidoth in [11], the IPC encoder is capable of synthesizing single arrivals the arrival time of which is drawn Exponential(μ) once the encoder sets $\lambda(t) = \mu$, while it also preserves the ability of recalling such an ongoing arrival process at any time. Consequently, any exponent that is achievable for the exponential telephone signaling channel (ETSC) is achievable for the IPC with noiseless feedback. In fact, the last argument establishes that the sphere-packing exponent $E_{sp}(R)$ obtained by Lapidoth for the IPC with noiseless feedback (when the average-power is constrained to be $\leq \mu/e$) is also the sphere-packing exponent for the ETSC, although our proof for Proposition 3 (in section III.C) follows different arguments.

Thus, as stated in [1], the IPC and the ETC are fundamentally different and caution must be excercised when drawing conclusions regarding the respective relaibility of variants of these two models.

Finally, for random transmission-time codes, our lower bound with fixed length codes is better than all lower bounds we obtained with the aforementioned fixed transmission-time codes, and in particular $E_f(0) \geq \mu$. More importantly, with variable block-length codes, a lower bound on the random transmission-time zero error capacity is established. The determination of the exact value of the random transmission-time zero error capacity is an open problem.

III. Proofs

A. Fixed Block-Length Schemes

This section considers the performance of a family of n-length block codes, as per Definitions 5 and 6, on the TSC. Specifically, we consider first Gallager's ML decoding random coding exponent for these codes. Then we evaluate an attainable exponent with the decoding rule (4)-(5) and the change in (9), and finally we compute the expurgated exponent for these codes. In all cases a distinction is made depending on whether it is a fixed or a random transmission-time code.

To this end consider an ensemble of codes, wherein a code in the ensemble with codewords $\boldsymbol{x}(1), \ldots, \boldsymbol{x}(M)$ is assigned the probability $\prod_{m=1}^{M} q(\boldsymbol{x}(m))$, where

$$q(\boldsymbol{x}) = \alpha^{-1}\phi(\boldsymbol{x}) \prod_{k=1}^{n} f_X(x_k) \tag{20}$$

with

$$\phi(\boldsymbol{x}) = \begin{cases} 1, & \beta n - \sqrt{n} < \sum_{k=1}^{n} x_k < \beta n + \sqrt{n} \\ 0, & \text{otherwise} \end{cases} \tag{21}$$

The density $f_X(x_k)$ is chosen to be a mixture of a mass point at the origin and an exponential density with parameter B, that is

$$f_X(x) = a\delta(x) + (1-a)Be^{-Bx} \ , \ 0 \le a \le 1 \ , \quad x \ge 0 \ , \tag{22}$$

and β in (21) is chosen as $\beta = (1-a)/B$.

The term α is the normalizing factor and it equals the probability that a codeword chosen at random based on the law $\prod_{k=1}^{n} f_X(x_k)$ satisfies the "shell constraint" $\phi(\boldsymbol{x}) = 1$. A straightforward application of the central limit theorem shows that $\alpha = e^{-o_n(1)}$ (Henceforth we write $o_n(1)$ to denote an unspecified positive-valued function that goes to zero as n goes to infinity. Furthermore, we write $o(n)$ to denote a function such that $o(n)/n = o_n(1)$). Since our final goal is to obtain an achievable exponent with n-length codes (for the TSC) either via fixed transmission-time or random transmission-time strategies, we shall consider throughout an ensemble of n-length codes the components of which meet a "strict" fixed transmission-time "shell constraint".

A.1. Random coding error exponent

This section is based on Gallager's result on computing the reliability exponent for continuous input/output channels [12, Chapter 7]. Gallager's bound on the ensemble average of the probability of ML decoding error, when the mth codeword is transmitted, is given by

$$\overline{P}_{er'} \le M^{\rho} \int d\boldsymbol{y} \left[\int d\boldsymbol{x} q(\boldsymbol{x}) p(\boldsymbol{y}|\boldsymbol{x})^{\frac{1}{1+\rho}} \right]^{1+\rho} \ , \tag{23}$$

with ρ chosen arbitrarily within the interval $0 < \rho \le 1$.

Since our feedback code renders the components of the random vector Y^n independent of each other we may write

$$p(\boldsymbol{y}|\boldsymbol{x}) = \prod_{k=1}^{n} e_{\mu}(y_k - x_k) \ . \tag{24}$$

Next, we note that for any $\theta \geq 0$

$$q(\boldsymbol{x}) \leq \alpha^{-1} e^{\theta\left(\sum_{k=1}^n x_k - \beta n + \sqrt{n}\right)} \prod_{k=1}^n f_X(x_k) \ . \qquad (25)$$

Thus, the inner integral in (23) may be evaluated to yield

$$\begin{aligned}\int d\boldsymbol{x} q(\boldsymbol{x}) p(\boldsymbol{y}|\boldsymbol{x})^{\frac{1}{1+\rho}} &\leq \alpha^{-1} e^{\theta\sqrt{n}} \prod_{k=1}^n \left\{ e^{-\theta\beta} \int e^{\theta x_k} f_X(x_k) \left[e_\mu(y_k - x_k)\right]^{\frac{1}{1+\rho}} dx_k \right\} \\ &= e^{o(n)} \prod_{k=1}^n e^{-\theta\beta} \frac{\left(\mu e^{-\mu y_k}\right)^{\frac{1}{1+\rho}}}{\left(\theta + \frac{\mu}{1+\rho} - B\right)} \\ &\quad \cdot \left[(1-a) B e^{\left(\theta + \frac{\mu}{1+\rho} - B\right) y_k} + a\left(\theta + \frac{\mu}{1+\rho}\right) - B\right] \ .\end{aligned}$$

Substituting the last expression into (23), we obtain

$$\begin{aligned}\overline{P}_{er'} \leq M^\rho e^{o(n)} & \frac{e^{-n\theta\beta(1+\rho)}}{\left(\theta + \frac{\mu}{1+\rho} - B\right)^{n(1+\rho)}} \\ & \cdot \left\{ \mu \int e^{-\mu y} \left[(1-a) B e^{\left(\theta + \frac{\mu}{1+\rho} - B\right) y} + a\left(\theta + \frac{\mu}{1+\rho}\right) - B\right]^{1+\rho} dy \right\}^n \ , \qquad (26)\end{aligned}$$

where $0 \leq \theta < B$.

As it is difficult to obtain a closed form expression for the integral in (26), we consider the upper bound resulting from the particular choice (which is the optimal when $\rho = 1$)

$$B = a\left(\theta + \frac{\mu}{1+\rho}\right) \ , \qquad (27)$$

while it is further assumed that $0 \leq \theta < a\mu / \left[(1-a)(1+\rho)\right]$.

With this choice the upper bound (26) yields

$$\begin{aligned}\overline{P}_{er'} &\leq e^{o(n)} M^\rho \left(\frac{e^{-\theta\beta} B(1-a)}{\theta + \frac{\mu}{1+\rho} - B}\right)^{n(1+\rho)} \left(\mu \int e^{-(B-\theta)(1+\rho) y} dy\right)^n \\ &= e^{o(n)} M^\rho \left[\frac{a^{1+\rho} \mu e^{-\frac{\theta(1-a)(1+\rho)}{a\left(\theta + \frac{\mu}{1+\rho}\right)}}}{a\mu - \theta(1-a)(1+\rho)}\right]^n\end{aligned}$$

The next step is to convert this bound on an ensemble average to a bound that applies to a specific code. For this, one considers an ensemble of codes with $2M$ codewords and uses expurgation (see, e.g., [12, Section 5.6]). The final result is

$$P_{er'} \leq e^{o(n)} M^\rho \left[\frac{a^{(1+\rho)} e^{-\frac{\theta_\mu(1-a)(1+\rho)}{a(\theta_\mu + 1/(1+\rho))}}}{a - \theta_\mu(1-a)(1+\rho)}\right]^n \ , \quad \theta_\mu \triangleq \frac{\theta}{\mu} \ , \qquad (28)$$

for any message $m \in \{1 \dots, M\}$.

A.2. An achievable exponent with a specific decoder

In this section we analyze the performance of a decoder that follows the decoding rule (4)-(5) when $\tilde{k} = n$. Although this decoding rule is the Maximum Likelihood decoding rule [1], following the reasoning of Wyner in [10], we do not exploit this fact here. Our analysis gives rise to an upper bound on the average probability of error and this bound is shown to coincide with the upper bound obtained via Gallager's random coding expression (23).

Suppose that the codeword $\boldsymbol{x}(1)$ from a given codebook $\{\boldsymbol{x}(1), \dots, \boldsymbol{x}(M)\}$ is transmitted and as result the decoder receives the sequence of interdeparture times $\boldsymbol{y}$. A decoder that implements the above decoding rule may be split into the following two stages:

Reduction stage: To each recorded interdeparture we assign a "survivors set" — a subset of $\{1, \dots, M\}$ denoted as $S^{(k)}$, of cardinality $s^{(k)}$. In this stage the survivors set is defined inductively with respect to the index k, of the recorded interdeparture y_k, as follows. Set $S^{(0)} = \{1, \dots, M\}$, and $s^{(0)} = \left|S^{(0)}\right| = M$. Assume that the set $S^{(k-1)}$ has already been defined. Upon receiving the kth interdeparture y_k, set

$$S^{(k)} = \left\{ i \in S^{(k-1)} : x_k(i) \leq y_k \right\} \ .$$

Consequently, after observing the nth interdeparture, the decoder forms the set $S^{(n)}$, defined as

$$S^{(n)} = \{ i \in \{1, \dots, M\} : x_k(i) \leq y_k \ , \ 1 \leq k \leq n \} \ .$$

Nearest Neighbor Decoding Stage: The decoder, having formed $S^{(n)}$, chooses the message $i \in S^{(n)}$ that minimizes the sum of service times, that is

$$\psi_d(\boldsymbol{y}) = \arg \min_{i \in S^{(n)}} \sum_{k=1}^{n} y_k - x_k(i) \ . \tag{29}$$

In fact, the decoder may implement just the nearest neighbor decoding stage (29) as it subsumes the reduction stage. Nevertheless, since the reduction stage may be implemented successively per each interdeparture this defines the exact list of ambiguous messages the decoder has upon observing the sequence y^{k-1}. The encoder is thus permitted to choose the next interarrival x_k based on $\left(x^{k-1}, y^{k-1}\right)$ – i.e., depending on the current list of survivors and the past sequence of interdepartures. In that case, the code would exploit all available degrees of freedom analogously to the code used by Lapidoth in [11] in order to derive the reliability function of the ideal Poisson channel with noiseless feedback. It should be emphasized, however, that as opposed to the code in [11] which exhibits a deterministic reduction factor in the survivors set per each photon arrival, our code exhibits a *random* reduction factor in the survivors set per each observed interdepature. The following analysis of the error exponent takes into account this fact.

Let $P_s(\boldsymbol{y})$ be the conditional probability that the codeword $\boldsymbol{x}(m), m \neq 1$ belongs to $S^{(n)}$ given the received sequence $\boldsymbol{y}$, i.e.

$$P_s(\boldsymbol{y}) = \Pr\left\{\boldsymbol{x}(m) \in S^{(n)}, m \neq 1 | \boldsymbol{y}\right\} .$$

Since the codewords are drawn with the same law $q(\boldsymbol{x})$ and independently of each other, $P_s(\boldsymbol{y})$ is not identified with a specific index m. Furthermore, the conditional probability, $P_s^{(l)}(\boldsymbol{y})$, that $S^{(n)}$ contains, apart from the transmitted codeword (which a.s. appears in $S^{(n)}$), exactly $l, 0 \leq l \leq M-1$ more codewords, is

$$P_s^{(l)}(\boldsymbol{y}) = \Pr\left\{\left|S^{(n)}\right| = l+1 | \boldsymbol{y}\right\} = \binom{M-1}{l} P_s^l(\boldsymbol{y}) \left(1 - P_s(\boldsymbol{y})\right)^{M-(l+1)} . \tag{30}$$

Next, let $P_e^{(m)}(\boldsymbol{x}(1), \boldsymbol{y})$ be defined as

$$P_e^{(m)}(\boldsymbol{x}(1), \boldsymbol{y}) = \Pr\left(\left\{\sum_{k=1}^{n} x_k(m) \geq \sum_{k=1}^{n} x_k(1)\right\} \mid \boldsymbol{x}(m) \in S^{(n)}, \boldsymbol{x}(1), \boldsymbol{y}\right) ,$$

that is, $P_e^{(m)}(\boldsymbol{x}(1), \boldsymbol{y})$ is the conditional probability that a surviving codeword $\boldsymbol{x}(m)$ won't be "farther" than $\boldsymbol{x}(1)$ from $\boldsymbol{y}$.

Thus, the conditional probability of error, given $s^{(n)} = l+1$ and $\boldsymbol{y}$, is

$$P_e(l, \boldsymbol{x}(1), \boldsymbol{y}) = 1 - \left(1 - P_e^{(m)}(\boldsymbol{x}(1), \boldsymbol{y})\right)^l . \tag{31}$$

Combining (30) with (31) it follows that

$$\begin{aligned}\sum_{l=0}^{M-1} P_s^{(l)}(\boldsymbol{y}) P_e(l, \boldsymbol{x}(1), \boldsymbol{y}) &= 1 - \left[1 - P_s(\boldsymbol{y}) P_e^{(m)}(\boldsymbol{x}(1), \boldsymbol{y})\right]^{M-1} \\ &= \Pr\left(\bigcup_{m=2}^{M}\left\{x_k(m) \leq y_k, 1 \leq k \leq n\right\} \cap \left\{\sum_{k=1}^{n} x_k(m) \geq \sum_{k=1}^{n} x_k(1)\right\}\right) .\end{aligned}$$

We may now express the average decoding error probability when using this scheme as

$$\begin{aligned}\overline{P}_{ed'} &= \int d\boldsymbol{x} q(\boldsymbol{x}) \int d\boldsymbol{y} p(\boldsymbol{y}|\boldsymbol{x}) \\ &\quad \cdot \Pr\left(\bigcup_{m=2}^{M}\left\{x_k(m) \leq y_k, 1 \leq k \leq n\right\} \cap \left\{\sum_{k=1}^{n} x_k(m) \geq \sum_{k=1}^{n} x_k(1)\right\}\right)\end{aligned}$$

The union of events bound then yields for any $0 \le \rho \le 1$

$$\begin{aligned}
\overline{P}_{ed'} &\le \int d\boldsymbol{x} q(\boldsymbol{x}) \int d\boldsymbol{y} p(\boldsymbol{y}|\boldsymbol{x}) \left[(M-1) P_s(\boldsymbol{y}) P_e^{(m)}(\boldsymbol{x}(1), \boldsymbol{y})\right]^\rho \\
&\le M^\rho \int d\boldsymbol{x} q(\boldsymbol{x}) \int d\boldsymbol{y} p(\boldsymbol{y}|\boldsymbol{x}) \left[P_s(\boldsymbol{y}) P_e^{(m)}(\boldsymbol{x}(1), \boldsymbol{y})\right]^\rho . \qquad (32)
\end{aligned}$$

In order to evaluate $P_e^{(m)}(\boldsymbol{x}(1), \boldsymbol{y})$ observe that the reduction stage picks a code $\{\boldsymbol{x}(1), \boldsymbol{x}(2), \ldots, \boldsymbol{x}(M)\}$ from the original ensemble and returns a subcode $\{\boldsymbol{x}(1), \boldsymbol{x}(j_2), \ldots, \boldsymbol{x}(j_{s^{(n)}})\}$ having $s^{(n)}$ codewords. Let us denote this codebook as $\{\boldsymbol{z}(1), \boldsymbol{z}(2), \ldots, \boldsymbol{z}(j_{s^{(n)}})\}$. Thus, the probability $q\left(\boldsymbol{x}(m)\right)$ that has been originally associated with $\boldsymbol{x}(m)$ is modified to

$$p\left(\boldsymbol{z}\right) = \frac{1}{P_s(\boldsymbol{y})} \tilde{\phi}(\boldsymbol{z}, \boldsymbol{y}) q(\boldsymbol{z}) \ , \qquad (33)$$

with

$$\tilde{\phi}(\boldsymbol{z}, \boldsymbol{y}) = \begin{cases} 1, & z_k \le y_k \ , \ 1 \le k \le n \\ 0, & \text{otherwise} \end{cases}$$

Next, consider the conditional probability $P_e^{(m)}(\boldsymbol{x}(1), \boldsymbol{y})$ that a surviving codeword $\boldsymbol{x}(m) \triangleq \boldsymbol{z}$ will be "closer", to the received interdeparture sequence $\boldsymbol{y}$, than the correct sequence $\boldsymbol{x}(1)$. By Markov's inequality, we have for any $s \ge 0$

$$\begin{aligned}
P_e^{(m)}(\boldsymbol{x}(1), \boldsymbol{y}) &= \Pr\left\{\sum_{k=1}^n Z_k \ge \sum_{k=1}^n X_k\right\} \\
&= \Pr\left\{\exp\left[s\left(\sum_{k=1}^n Z_k - \sum_{k=1}^n X_k\right)\right] \ge 1\right\} \\
&\le e^{-s\sum_{k=1}^n x_k} \cdot \mathsf{E}\left(e^{s\sum_{k=1}^n Z_k}\right) ,
\end{aligned}$$

where E is the expectation operator.

Using (33) we obtain

$$P_e^{(m)}(\boldsymbol{x}(1), \boldsymbol{y}) \le e^{-s\sum_{k=1}^n x_k} P_s^{-1}(\boldsymbol{y}) \int d\boldsymbol{z} \tilde{\phi}(\boldsymbol{z}, \boldsymbol{y}) q(\boldsymbol{z}) e^{s\sum_{k=1}^n z_k} \qquad (34)$$

Substituting (34) into (32), we obtain

$$\overline{P}_{ed'} \le M^\rho \int d\boldsymbol{x} q(\boldsymbol{x}) e^{-s\rho\sum_{k=1}^n x_k} \int d\boldsymbol{y} p(\boldsymbol{y}|\boldsymbol{x}) \left[\int d\boldsymbol{z} \tilde{\phi}(\boldsymbol{z}, \boldsymbol{y}) q(\boldsymbol{z}) e^{s\sum_{k=1}^n z_k}\right]^\rho \qquad (35)$$

By (25) we get for any $\theta \ge 0$,

$$\begin{aligned}
\int d\boldsymbol{z} \tilde{\phi}(\boldsymbol{z}, \boldsymbol{y}) q(\boldsymbol{z}) e^{s\sum_{k=1}^n z_k} &= \alpha^{-1} e^{\theta\sqrt{n}} \prod_{k=1}^n \left[e^{-\theta\beta} \int_{z_k=0}^{y_k} e^{(\theta+s)z_k} f_X(z_k) dz_k\right] \\
&= \alpha^{-1} e^{\theta\sqrt{n}} \prod_{k=1}^n e^{-\theta\beta} \frac{\left[(1-a)Be^{(\theta+s-B)y_i} + a\left(\theta+s\right) - B\right]}{(\theta+s-B)} .
\end{aligned}$$

Substituting (24)-(25), and the last expression into (35) we obtain

$$\begin{aligned}
\overline{P}_{ed'} &\leq M^{\rho} e^{o(n)} \prod_{k=1}^{n} \frac{e^{-\theta\beta(1+\rho)}}{(\theta+s-B)^{\rho}} \int_{x_k=0}^{\infty} e^{(\theta-s\rho)x_k} f_X(x_k) dx_k \\
&\quad \cdot \int_{y_k=x_k}^{\infty} \mu e^{-\mu(y_k-x_k)} \left[(1-a)Be^{(\theta+s-B)y_k} + a\,(\theta+s)-B\right]^{\rho} dy_k \\
&= M^{\rho} e^{o(n)} \prod_{k=1}^{n} \frac{e^{-\theta\beta(1+\rho)}}{(\theta+s-B)^{\rho}} \int_{y_k=0}^{\infty} \mu e^{-\mu y_k} \left[(1-a)Be^{(\theta+s-B)y_i} + a\,(\theta+s)-B\right]^{\rho} dy_k \\
&\quad \cdot \int_{x_k=0}^{y_k} e^{(\theta+\mu-s\rho)x_k} f_X(x_k) dx_k \ . \qquad (36)
\end{aligned}$$

The inner integral in (36) evaluates to

$$\begin{aligned}
&\int_{x_k=0}^{y_k} f_X(x_k) e^{(\theta+\mu-s\rho)x_k} dx_k = \\
&\quad \frac{\left[(1-a)Be^{(\theta+\mu-s\rho-B)y_k} + a\,(\theta+\mu-s\rho)-B\right]}{(\theta+\mu-s\rho-B)} \qquad (37)
\end{aligned}$$

Choosing $s = \frac{\mu}{1+\rho}$ in (37) and substituting this expression into (36) we obtain exactly the same bound as that obtained via Gallager's random coding argument in (26).

A.3. Expurgated error exponent

This section is based on Gallager's result on computing the expurgated exponent for continuous input/output channels [12, Chapter 7]. Using expurgation on an ensemble of codebooks each with $M' = 2M-1$ codewords, one may infer that there exists a size M code the error probability of which, when the mth codeword is transmitted, is bounded as follows.

$$P_{ex'} \;\leq\; M^{\rho} \left[\int d\boldsymbol{x} d\boldsymbol{x}' q(\boldsymbol{x}) q(\boldsymbol{x}') \left(\int d\boldsymbol{y} \sqrt{p(\boldsymbol{y}|\boldsymbol{x}) p(\boldsymbol{y}|\boldsymbol{x}')} \right)^{\frac{1}{\rho}} \right]^{\rho} \qquad (38)$$

with ρ chosen arbitrarily subject to the constraint $\rho \geq 1$.

First observe that using (24)

$$\begin{aligned}
\int d\boldsymbol{y} \sqrt{p(\boldsymbol{y}|\boldsymbol{x}) p(\boldsymbol{y}|\boldsymbol{x}')} &= \prod_{k=1}^{n} \int \sqrt{e_{\mu}(y_k - x_k) e_{\mu}(y_k - x'_k)} \; dy_k \\
&= \prod_{k=1}^{n} e^{-\frac{\mu|x_k - x'_k|}{2}} \ . \qquad (39)
\end{aligned}$$

Thus, substituting (39) and (25) into (38), we obtain

$$P_{ex'} \leq cM^{\rho} \prod_{k=1}^{n} \left[\int e^{\theta x_k} f_X(x_k) dx_k \int e^{\theta x'_k} e^{-\frac{\mu|x_k - x'_k|}{2\rho}} f_X(x'_k) dx'_k \right]^{\rho} , \qquad (40)$$

where $c = \alpha^{-2\rho} e^{2\rho\theta\sqrt{n}} e^{-2\rho n\theta\beta} = e^{o(n)} e^{-2\rho n\theta\beta}$.

Defining $\theta_B \triangleq \theta/B$ and $B_\mu \triangleq B/\mu$ the bracketed integrals in (40) may be evaluated to obtain

$$P_{ex'} \quad \leq \quad e^{o(n)} M^\rho \left\{ \frac{e^{-2\theta_B(1-a)} \left[\frac{a^2}{2\rho} + B_\mu \frac{(1-a\theta_B)^2}{1-\theta_B}\right]}{B_\mu(1-\theta_B) + \frac{1}{2\rho}} \right\}^{n\rho} \quad , \quad 0 \leq \theta_B < 1 \ . \tag{41}$$

A.4. Random transmission-time exponents

In order to obtain an achievable random transmission-time exponent with n-length codes recall that the code $\{\boldsymbol{x}(m)\}_{m=1}^{M}$ is selected such that each codeword satisfies the shell constraint (21), that is

$$\sum_{k=1}^{n} x_k(m) < n \left(\beta + \frac{1}{\sqrt{n}}\right) \quad , \quad 1 \leq m \leq M \ ,$$

while the average service time for n customers is n/μ.

Let us choose $T = n\,(1/\lambda + 1/\sqrt{n})$, where $\frac{1}{\lambda} = \frac{1}{\mu} + \beta$. Using the fact that $o(n) = o(T)$ and $\beta = (1-a)/B$ we conclude that the expurgated upper bound on the error probability, for a code with average transmission-time T, can be expressed as

$$P_{ex'} \leq e^{-TE_{ex'}(R) + o(T)} \ ,$$

where $E_{ex'}(R)$ is given in (17) and $\lambda = \frac{B_\mu \mu}{B_\mu + (1-a)}$.

Similarly, the random coding upper bound on the error probability, for n-length codes with average transmission-time T, can be expressed as

$$P_{er'} \leq e^{-TE_{r'}(R) + o(T)} \ ,$$

where $E_{r'}(R)$ is given in (17) and here the choice (27) implies that $\lambda = \frac{a[\theta_\mu(1+\rho)+1]\mu}{a[\theta_\mu(1+\rho)+1]+(1-a)(1+\rho)}$.

A.5. Fixed transmission-time exponents

In order to obtain an achievable fixed transmission-time exponent with n-length codes let P_{et} denote the probability that the nth departure occurs beyond time T, that is

$$P_{et} = \Pr\left\{\sum_{k=1}^{n} Y_k > T\right\} \ .$$

Using this definition, the probability of error is

$$P_e = \Pr\left(\{D \neq \psi(Y_1, Y_2, \ldots, Y_n)\} \bigcap \left\{ \sum_{k=1}^{n} Y_k \leq T \right\} \right) + P_{et} \, .$$

Using the union of events bound, the probability of error can be upper bounded by

$$P_e \leq \Pr\{D \neq \psi(Y_1, Y_2, \ldots, Y_n)\} + P_{et} \, . \tag{42}$$

Since each codeword satisfies (21) we may express P_{et} by

$$\begin{aligned} P_{et} &= \Pr\left\{ \sum_{k=1}^{n} X_k + \sum_{k=1}^{n} S_k > T \right\} \\ &\leq \Pr\left\{ \sum_{k=1}^{n} S_k > T - \beta n - \sqrt{n} \right\} \, , \end{aligned} \tag{43}$$

with $s_k = y_k - x_k$ denoting the service time of the kth customer.

The components of the random vector $(S_1, S_2, \ldots, S_n)$ are independent exponentially distributed (with parameter μ) random variables. Thus, the Chernoff bound can be used to upper bound the probability that their sum deviates from a specified value T_d as follows

$$\Pr\left\{ \sum_{k=1}^{n} S_k > T_d \right\} \leq \left(\frac{\mu}{\mu - \gamma} \right)^n e^{-T_d \gamma} \, ,$$

where γ is some nonnegative number, and $\gamma < \mu$.

Setting $T_d = T - \beta n - \sqrt{n}$ we can upper bound the r.h.s of (43) by

$$P_{et} \leq e^{o(n)} e^{-\gamma(T - \beta n)} \left(\frac{\mu}{\mu - \gamma} \right)^n \, . \tag{44}$$

Assuming $\mu \geq n/(T - \beta n)$, and substituting $\gamma = \mu - n/(T - \beta n)$ in (44) while ignoring the $o(n)$ term we obtain

$$P_{et} \leq e^{-\mu T} e^{n(1+\beta\mu)} \left[\frac{\mu(T - \beta n)}{n} \right]^n \, , \quad n \leq \frac{\mu T}{1 + \beta\mu}$$

For our convenience, we introduce the parameter $0 < r \leq 1$ and define $n = \frac{r B_\mu \mu T}{B_\mu + (1-a)}$, to obtain

$$P_{et} \;\leq\; e^{-\mu T} e^{r\mu T} \left[\frac{B_\mu + (1-a)(1-r)}{r B_\mu} \right]^{\frac{r B_\mu \cdot \mu T}{B_\mu + (1-a)}} \triangleq e^{-T \tilde{E}_{ex}^{(1)}} \, , \tag{45}$$

where $\tilde{E}_{ex}^{(1)}$ is defined in (14).

Similarly, substituting the expression for n into (41) we obtain

$$P_{ex'} \le e^{-T\tilde{E}_{ex}^{(0)}(R)} , \tag{46}$$

where $\tilde{E}_{ex}^{(0)}(R)$ is defined in (14). The combination of (42), (45) and (46) then yields the exponent $\tilde{E}_{ex}^{(f)}(R)$ in (13).

For the random coding exponent we use the relation (27) to get $n = \frac{ra[\theta_\mu(1+\rho)+1]\mu T}{a[\theta_\mu(1+\rho)+1]+(1-a)(1+\rho)}$ and obtain

$$\begin{aligned} P_{et} &\le e^{-\mu T} e^{r\mu T} \left\{ \frac{a\left[\theta_\mu(1+\rho)+1\right]+(1-a)(1-r)(1+\rho)}{ra\left[\theta_\mu(1+\rho)+1\right]} \right\}^{\frac{ra[\theta_\mu(1+\rho)+1]\mu T}{a[\theta_\mu(1+\rho)+1]+(1-a)(1+\rho)}} \\ &\triangleq = e^{-T\tilde{E}_r^{(1)}} , \end{aligned} \tag{47}$$

where $\tilde{E}_r^{(1)}$ is defined in (12).

Similarly, substituting the expression for n into (28) we obtain

$$P_{er'} \le e^{-T\tilde{E}_r^{(0)}(R)} , \tag{48}$$

where $\tilde{E}_r^{(0)}(R)$ is defined in (12). The combination of (42), (47) and (48) then yields the exponent $\tilde{E}_r^{(f)}(R)$ in (11).

B. Variable Block-Length Schemes

This section considers the performance of both fixed and random transmission-time block codes, as per Definitions 4 and 7, on the TSC. First we define explicitly the conditional channel law for the variable block-length fixed transmission-time code. Then using this expression, we compute both Gallager's ML decoding random coding exponent, and the expurgated exponent for these codes. Finally, we analyze the case of the random transmission time with random block-length.

B.1. Fixed transmission-time channel

Since the encoder observes only those departures that occur during the interval $[0, T]$, the length k of Y^k is a random variable. Fix some n and let us be given an n-length codebook $x^n(1), x^n(2), \ldots, x^n(M)$. In what follows we shall consider the performance of such a code taking into account that only a partial length $k \le n$ of a codeword has been transmitted over the TSC.

Given a partial sequence of interarrivals $x^{k+1}, k = 0, \ldots, n-1$ the conditional probability density of observing the interdeparture sequence y^k is

$$p\left(y^k|x^{k+1}\right) = \begin{cases} \int_T^\infty \chi(0) p(y_1|x_1) dy_1, & k = 0 \\ \int_{T-\sum_{i=1}^k y_i}^\infty \chi(k) \prod_{i=1}^{k+1} p(y_i|x_i) dy_{k+1}, & 1 \le k < n \end{cases} \tag{49}$$

while

$$p(y^n|x^n) = \chi(n) \prod_{i=1}^{n} p(y_i|x_i) \ . \tag{50}$$

Here the indicator $\chi(\cdot)$ is defined as

$$\chi(k) \quad = \quad \begin{cases} 1, & \left\{\sum_{i=1}^{k} Y_i \leq T\right\} \cap \left\{\sum_{i=1}^{k+1} Y_i > T\right\} \\ 0, & \text{otherwise} \end{cases} \qquad 0 < k < n$$

$$\chi(0) \quad = \quad \begin{cases} 1, & \{Y_1 > T\} \\ 0, & \text{otherwise} \end{cases} \qquad k = 0$$

$$\chi(n) = \begin{cases} 1, & \{\sum_{i=1}^{n} Y_i \leq T\} \\ 0, & \text{otherwise} \end{cases} \qquad k = n$$

and one can verify that

$$\sum_{k=0}^{n} \left[\int p\left(y^k|x^{k+1}\right) dy^k \right] = 1 \ .$$

Next, using the traditional bound, one has for any $\theta_k \geq 0$,

$$\chi(k) \quad \leq \quad \begin{cases} e^{-\theta_k \left(T - \sum_{i=1}^{k+1} y_i\right)} & 0 \leq k < n \\ 1 & k = n \end{cases}$$

The conditional law in (49) may be bounded as follows

$$p\left(y^k|x^{k+1}\right) \quad \leq \quad e^{-\theta_k T} \left[\prod_{i=1}^{k} p(y_i|x_i) e^{\theta_k y_i} \right] \int_{T - \sum_{i=1}^{k} y_i}^{\infty} p(y_{k+1}|x_{k+1}) e^{\theta_k y_{k+1}} dy_{k+1} \ , \quad 0 < k < n$$

Recalling that the service time $s_{k+1} = y_{k+1} - x_{k+1}$ is exponentially distributed with mean $1/\mu$, $p\left(y^k|x^{k+1}\right)$ may be further upper bounded when $0 \leq \theta_k < \mu$ by

$$p\left(y^k|x^{k+1}\right) \quad \leq \quad \begin{cases} \frac{\mu}{\mu - \theta_0} e^{-\theta_0 T} e^{\theta_0 x_1}, & k = 0 \\ \frac{\mu}{\mu - \theta_k} e^{-\theta_k T} e^{\theta_k x_{k+1}} \left[\prod_{i=1}^{k} p(y_i|x_i) e^{\theta_k y_i} \right], & 0 < k < n \\ \prod_{i=1}^{n} p(y_i|x_i), & k = n \end{cases} \tag{51}$$

thus $\theta_n = 0$.

B.2. Random coding error exponent

In this section we apply Gallager's result on computing the reliability exponent for continuous input/output channels [12, Chapter 7] to the conditional channel law defined previously.

Gallager's bound on the ensemble average of the probability of ML decoding error, when the mth codeword is transmitted, is given by

$$\overline{P}_{er'}^{(vl)} \leq M^\rho \sum_{k=0}^{n} \int dy^k \left[\int dx^{k+1} q(x^{k+1}) p\left(y^k|x^{k+1}\right)^{\frac{1}{1+\rho}} \right]^{1+\rho} \quad , \quad 0 < \rho \leq 1 \ . \tag{52}$$

Choosing $\theta_k = \mu - \frac{1}{T}$ for $k = 0$, makes both the bound for $p\left(y^0|x^1\right)$ given by (51), and the expression for $k = 0$ in the sum on the r.h.s. of (52), in the order of $e^{-\mu T + o(T)}$.

Next choosing n such that $n = e^{o(T)}$, (52) becomes

$$\overline{P}_{er'}^{(vl)} \leq e^{o(T)} M^\rho \max_{0<k\leq n} \left\{ \int dy^k \left[\int dx^{k+1} q(x^{k+1}) p\left(y^k|x^{k+1}\right)^{\frac{1}{1+\rho}} \right]^{1+\rho} \right\} \ . \tag{53}$$

With the choice of the input law

$$q\left(x^k\right) = \prod_{i=1}^{k} f_X(x_i) \ , \tag{54}$$

and together with the bound on the conditional channel law given by (51), one may verify that

$$\left[\int dx^{k+1} q\left(x^{k+1}\right) p\left(y^k|x^{k+1}\right)^{\frac{1}{1+\rho}} \right]^{1+\rho} \leq \frac{\mu}{\mu - \theta_k} e^{-\theta_k T}$$
$$\cdot \left[\int f_X(x_{k+1}) e^{\frac{\theta_k x_{k+1}}{1+\rho}} dx_{k+1} \right]^{1+\rho} \prod_{i=1}^{k} \mu e^{-(\mu-\theta_k) y_i} \left[\int f_X(x_i) e^{\frac{\mu x_i}{1+\rho}} U(y_i - x_i) dx_i \right]^{1+\rho} \ .$$

When choosing $B - \frac{\theta_k}{1+\rho} > 0$, the left square brackets on the r.h.s. of the last inequality are in the order of $e^{o(T)}$. Thus, similarly to section A.1 we obtain

$$\left[\int dx^{k+1} q\left(x^{k+1}\right) p\left(y^k|x^{k+1}\right)^{\frac{1}{1+\rho}} \right]^{1+\rho} \leq$$
$$e^{o(T)} \frac{\mu}{\mu - \theta_k} e^{-\theta_k T} \prod_{i=1}^{k} \mu e^{-(\mu-\theta_k) y_i} \left[\frac{(1-a) B e^{\left(\frac{\mu}{1+\rho} - B\right) y_i} + \frac{a\mu}{1+\rho} - B}{\frac{\mu}{1+\rho} - B} \right]^{1+\rho} \ .$$

The substitution of $k = rn \ , \ 0 < r \leq 1$ in the last expression yields for (53)

$$\overline{P}_{er'}^{(vl)} \leq e^{o(T)} M^\rho \sup_{0<r\leq 1} \frac{\mu}{\mu - \theta_r} e^{-\theta_r T} \frac{1}{\left(\frac{\mu}{1+\rho} - B\right)^{rn(1+\rho)}}$$
$$\cdot \left[\mu \int e^{-(\mu-\theta_r) y} \left[(1-a) B e^{\left(\frac{\mu}{1+\rho} - B\right) y} + \frac{a\mu}{1+\rho} - B \right]^{1+\rho} dy \right]^{rn} \tag{55}$$

Since the expression in (55) is difficult to calculate, we consider now the upper bound resulting from the particular choice $B = a\mu/(1+\rho)$, while θ_r is restricted such that $\frac{\mu}{\mu-\theta_r} \le e^{o(T)}$ (In the case $\rho = 1$, the calculation of (55) is easily done and the above choice of B is found to be optimum). With this choice, the last upper bound reduces to

$$\begin{aligned}\overline{P}_{er'}^{(vl)} &\le e^{o(T)} M^{\rho} \sup_{0<r\le 1} e^{-\theta_r T} \left(\frac{B(1-a)}{\frac{\mu}{1+\rho} - B}\right)^{rn(1+\rho)} \left(\mu \int e^{-[B(1+\rho)-\theta_r]y} dy\right)^{rn} \\ &= e^{o(T)} M^{\rho} \sup_{0<r\le 1} e^{-\theta_r T} \left[\frac{a^{1+\rho}\mu}{a\mu - \theta_r}\right]^{rn} , \quad a\mu - \theta_r > 0 \ .\end{aligned}$$

Again, to convert this bound on an ensemble average to a bound that applies to a specific code, one considers an ensemble of codes with $2M$ codewords and uses expurgation (see, e.g., [12, Section 5.6]). The final result is

$$P_{er'}^{(vl)} \le e^{o(T)} M^{\rho} \sup_{0<r\le 1} e^{-\theta_\mu \mu T} \left[\frac{a^{(1+\rho)}}{a - \theta_\mu}\right]^{rn} , \quad \theta_\mu \triangleq \frac{\theta_r}{\mu} ,$$

for any message $m \in \{1 \ldots, M\}$.

Expressing $P_{er'}^{(vl)}$ as $P_{er'}^{(vl)} \le e^{-\hat{E}_r^{(f)}(R)\cdot T + o(T)}$, and substituting $n_r \triangleq \frac{r\cdot n}{\mu T}$ we obtain

$$\hat{E}_r^{(f)}(R) = \min\left\{ \inf_{0<n_r<\frac{n}{\mu T}} \mu\theta_\mu - \rho R - n_r \mu \ln\left(\frac{a^{(1+\rho)}}{a-\theta_\mu}\right), -\frac{\rho n}{T}\ln(a) - \rho R \right\} \tag{56}$$

since in (51) $\theta_n = 0$ and therefore when $r = 1$, $\theta_\mu = 0$. In order to avoid the direct dependence of the l.h.s of (56) on n_r — the fraction of samples based on which the ML decoder makes its decision, we choose θ_μ so that $\ln\left(\frac{a^{(1+\rho)}}{a-\theta_\mu}\right) = 0$ which in turn implies that the relation $\theta_\mu = a(1-a^\rho)$ must hold.

As a result of that

$$\hat{E}_r^{(f)}(R) = \min\left\{\mu\theta_\mu - \rho R, -\frac{\rho n}{T}\ln(a) - \rho R\right\} . \tag{57}$$

We are now free to choose the parameter a in order to maximize the l.h.s of (57). Indeed, $\hat{E}_r^{(f)}(R)$ is maximized when $a = (1+\rho)^{-\frac{1}{\rho}}$ and then

$$\hat{E}_r^{(f)}(R) = \min\left\{ \sup_{0<\rho\le 1} \left[\frac{\rho\mu}{(1+\rho)^{(1+\rho)/\rho}} - \rho R\right], \frac{n\ln(1+\rho)}{T} - \rho R\right\} .$$

Remark : The average waiting time dictated by the codewords' components in this coding scheme, denoted as β, equals $(1-a)/B$. Consequently, $1/\lambda_{vl} = \beta + 1/\mu$, with λ_{vl} denoting the average departure rate. Thus, we obtain

$$\lambda_{vl} = \frac{\mu}{(1+\rho)/a - \rho} = \frac{\mu}{(1+\rho)^{(1+\rho)/\rho} - \rho} .$$

Choosing n large enough and particularly $n \geq \alpha_r(\rho) \cdot T$, $\alpha_r(\rho) = \frac{\rho\mu}{(1+\rho)^{(1+\rho)/\rho}\ln(1+\rho)}$ where $\alpha_r(\rho) < \lim_{\rho\to 0}\alpha_r(\rho) = C$, we obtain

$$\hat{E}_r^{(f)}(R) = \sup_{0<\rho\leq 1}\left[\frac{\rho\mu}{(1+\rho)^{(1+\rho)/\rho}} - \rho R\right] .$$

The last expression is identical to that obtained by Arikan in [8, equation 52] for the one-way window-code used over the exponential-server timing channel.

In fact, the same conclusion holds, even if n is chosen to be larger than $e^{o(T)}$ from (53) and on, this can be shown via the foregoing analysis in section B.4.

In addition to demonstrating that, at least $E_r(R)$ is achievable via variable block-length codes even when queuing is avoided, this result demonstrates that shell constraint is superfluous as long as variable block-length codes are considered. Indeed, if the preceding analysis is carried out with a choice of an input law which takes into account a shell constraint that is applied sub-blockwise to $x^n(m)$ the same exponent (57) is obtained.

B.3. Expurgated error exponent

As in section A.3, this section is based on Gallager's result regarding the expurgated exponent for continuous input/output channels [12, Chapter 7]. Using expurgation on an ensemble of codebooks each with $M' = 2M-1$ codewords, one may infer that there exists a size M code the error probability of which, when the mth codeword is transmitted, is bounded as follows.

$$P_{ex'}^{(vl)} \leq M^\rho\left[\int dx^n dx_1^n q\left(x^n\right) q\left(x_1^n\right)\left(\sum_{k=0}^{n}\int dy^k\sqrt{p\left(y^k|x^{k+1}\right)p\left(y^k|x_1^{k+1}\right)}\right)^{\frac{1}{\rho}}\right]^\rho ,$$

with ρ chosen arbitrarily subject to the constraint $\rho \geq 1$.

Since $\left(\sum_i x_i\right)^\lambda \leq \sum_i x_i^\lambda$ for all nonnegative x_i and $0<\lambda<1$, it follows that

$$P_{ex'}^{(vl)} \leq M^\rho\left[\sum_{k=0}^{n}\int dx^{k+1}dx_1^{k+1}q\left(x^{k+1}\right)q\left(x_1^{k+1}\right)\left(\int dy^k\sqrt{p\left(y^k|x^{k+1}\right)p\left(y^k|x_1^{k+1}\right)}\right)^{\frac{1}{\rho}}\right]^\rho .$$

Consequently, following similar steps as in B.2

$$P_{ex'}^{(vl)} \leq e^{o(T)}M^\rho\max_{0<k\leq n}\left[\int dx^{k+1}dx_1^{k+1}q\left(x^{k+1}\right)q\left(x_1^{k+1}\right)\left(\int dy^k\sqrt{p\left(y^k|x^{k+1}\right)p\left(y^k|x_1^{k+1}\right)}\right)^{\frac{1}{\rho}}\right]^\rho \quad (58)$$

Using (51) we may write

$$\int dy^k\sqrt{p\left(y^k|x^{k+1}\right)p\left(y^k|x_1^{k+1}\right)} \leq$$

$$e^{\frac{\theta_k\left(x_{k+1}+x_{1,k+1}\right)}{2}}\left(\frac{\mu}{\mu-\theta_k}\right)^{k+1}e^{-\theta_k T}\prod_{i=1}^{k}e^{\frac{\theta_k\left(x_i+x_{1,i}\right)}{2}}e^{\frac{-(\mu-\theta_k)\left|x_i-x_{1,i}\right|}{2}}\ . \tag{59}$$

Substituting (59) into (58) we obtain

$$P_{ex'}^{(vl)}\leq e^{o(T)}M^{\rho}\max_{0<k\leq n}\left[\int f_X\left(x_{k+1}\right)f_X\left(x_{1,k+1}\right)e^{\frac{\theta_k\left(x_{k+1}+x_{1,k+1}\right)}{2\rho}}dx_{k+1}dx_{1,k+1}\right]^{\rho}\cdot$$
$$\left(\frac{\mu}{\mu-\theta_k}\right)^{k+1}e^{-\theta_k T}\left[\prod_{i=1}^{k}\int f_X\left(x_i\right)f_X\left(x_{1,i}\right)e^{\frac{\theta_k\left(x_i+x_{1,i}\right)}{2\rho}}e^{\frac{-(\mu-\theta_k)\left|x_i-x_{1,i}\right|}{2\rho}}dx_i dx_{1,i}\right]^{\rho} \tag{60}$$

When choosing $B-\frac{\theta_k}{2\rho}>0$, the left square brackets are in the order of $e^{o(T)}$, then similarly to section A.3 we obtain

$$P_{ex'}^{(vl)}\leq$$
$$e^{o(T)}M^{\rho}\max_{0<k\leq n}\left(\frac{\mu}{\mu-\theta_k}\right)^{k}e^{-\theta_k T}\left[\frac{\frac{a^2(\mu-\theta_k)}{2\rho}+\frac{B^2-\frac{\theta_k}{2\rho}a\left(2B-\frac{\theta_k}{2\rho}a\right)}{B-\frac{\theta_k}{2\rho}}}{\frac{\mu-2\theta_k}{2\rho}+B}\right]^{k\rho}\ ,\ 0\leq\theta_k<\mu\ .$$

Substituting $k=rn\ ,\ 0<r\leq 1$ in the last expression, we consider now the upper bound resulting from the particular choice $B=a\mu/(2\rho)$. With this choice, the last upper bound becomes

$$P_{ex'}^{(vl)}\leq$$
$$e^{o(T)}M^{\rho}\sup_{0<r\leq 1}e^{-\theta_r T}\left(\frac{\mu}{\mu-\theta_r}\right)^{rn}\left[\frac{a^2(\mu-\theta_r)}{a\mu-\theta_r}\right]^{rn\rho}\ ,\ 0\leq\theta_r<a\mu\ .$$

Expressing $P_{ex'}^{(vl)}$ as $P_{ex'}^{(vl)}\leq e^{-\hat{E}_{ex}^{(f)}(R)\cdot T+o(T)}$, and substituting $n_r\triangleq\frac{r\cdot n}{\mu T}$, and $\theta_\mu\triangleq\frac{\theta_r}{\mu}$ one obtains

$$\hat{E}_{ex}^{(f)}(R)\ =\ \min\left\{\inf_{0<n_r<\frac{n}{\mu T}}\mu\theta_\mu-\rho R-n_r\mu\left[\ln\left(\frac{1}{1-\theta_\mu}\right)+\rho\ln\left(\frac{a^2(1-\theta_\mu)}{a-\theta_\mu}\right)\right]\ ,\right.$$
$$\left.-\frac{\rho n}{T}\ln(a)-\rho R\right\}\ . \tag{61}$$

In order to avoid the direct dependence of the l.h.s of (61) on n_r, we choose θ_μ so that $\ln\left(\frac{1}{1-\theta_\mu}\right)+\rho\ln\left(\frac{a^2(1-\theta_\mu)}{a-\theta_\mu}\right)=0$, and we also let $\frac{\partial\theta_\mu}{\partial a}=0$ which in turn implies that $\theta_\mu=a/2$ must hold. As a result we obtain

$$\hat{E}_{ex}^{(f)}(R)=\min\left\{\sup_{\rho\geq 1}\left[\mu a/2-\rho R\right],-\frac{\rho n}{T}\ln(a)-\rho R\right\}. \tag{62}$$

where a is the solution of $2a^\rho(2-a)^{\rho-1}=1$. Choosing $n \geq \alpha_{ex}(\rho,a)\cdot T$, $\alpha_{ex}(\rho,a)=\frac{\mu a}{-2\rho\ln(a)}$ where $\alpha_{ex}(\rho,a) \leq \alpha_{ex}(1,1/2)=\frac{\mu}{4\ln(2)}$ we obtain the final result presented in (7) and (8).

B.4. Random transmission-time with zero error probability

In this section we consider a variable block-length code that consumes a random transmission time. Specifically, we consider a decoder which performs just the reduction process described in section A.2 and stops once it is left with a single survivor. Since the surviving message is a.s. the correct one, zero error is established with a transmission rate depending on the mean transmission time, and $C_{ze}^{(r)} \geq \ln M/\left(\mathsf{E}[T_{ze}]\right)$.

Let $p(k|\boldsymbol{y})$ be the probability that a single survivor is left after the departure of exactly k customers, then (with $s^{(k)}$ denoting the cardinality of the survivors set after k departures)

$$p(k|\boldsymbol{y}) = \begin{cases} P\left(s^{(k)}=1, s^{(k-1)}>1|\boldsymbol{y}\right), & k>1 \\ P\left(s^{(1)}=1|\boldsymbol{y}\right), & k=1 \ . \end{cases}$$

The mean transmission time $\mathsf{E}[T_{ze}] \triangleq \overline{T}_{ze}$ is

$$\overline{T}_{ze} = \int p(\boldsymbol{y})\sum_{k=1}^{\infty}\left(p(k|\boldsymbol{y})\sum_{i=1}^{k} y_i\right)d\boldsymbol{y} = \sum_{k=1}^{\infty}\int\left(p(y^k)p(k|y^k)\sum_{i=1}^{k}y_i\right)dy^k \ . \tag{63}$$

We choose a critical $k_c>1$ the value of which shall be determined later on, and divide the sum into two parts; upper bounding them differently as follows

$$\overline{T}_{ze} \leq \sum_{k=1}^{k_c-1}\int\left(p(y^{k_c-1})p(k|y^{k_c-1})\sum_{i=1}^{k_c-1}y_i\right)dy^{k_c-1} + \sum_{k=k_c}^{\infty}\int\left(p(y^k)p(k|y^k)\sum_{i=1}^{k}y_i\right)dy^k \ . \tag{64}$$

Next, using (30), we upper bound $p(k|y^k)$ as follows

$$\begin{aligned} p(k|y^k) &\leq \left[1-P\left(s^{(k-1)}=1|y^{k-1}\right)\right]^\rho = \left[1-\left(1-\prod_{j=1}^{k-1}F_X(y_j)\right)^{M-1}\right]^\rho \\ &\leq M^\rho\prod_{j=1}^{k-1}F_X^\rho(y_j) \ , \quad k>1 \ , \ 0\leq\rho\leq 1 \ . \end{aligned}$$

Using this inequality (64) becomes

$$\begin{aligned} \overline{T}_{ze} &\leq \int\left(\sum_{k=1}^{k_c-1}p(k|y^{k_c-1})\right)p(y^{k_c-1})\sum_{i=1}^{k_c-1}y_i\ dy^{k_c-1} \\ &\quad + M^\rho\sum_{k=k_c}^{\infty}\int\left(p(y^k)\prod_{j=1}^{k-1}F_X^\rho(y_j)\sum_{i=1}^{k}y_i\right)dy^k \\ &\leq \int p(y^{k_c-1})\sum_{i=1}^{k_c-1}y_i\ dy^{k_c-1} + M^\rho\sum_{k=k_c}^{\infty}\sum_{i=1}^{k}\int\left(p(y^k)\prod_{j=1}^{k-1}F_X^\rho(y_j)y_i\right)dy^k \ . \end{aligned}$$

Since the random variables Y_i , $i = 1, 2, \ldots$ are IID, let Y be a random variable having the same law as Y_i, namely $f_Y(y) = \int_0^y f_X(x) e_\mu(y-x)dx$. Using this notation one may verify that

$$\begin{aligned} \int p(y^{k_c-1}) \sum_{i=1}^{k_c-1} y_i \; dy^{k_c-1} &= (k_c-1)\mathsf{E}[Y] = \frac{k_c-1}{\lambda} \\ \int \left(p(y^k) \prod_{j=1}^{k-1} F_X^\rho(y_j) y_i \right) dy^k &\leq \frac{1}{\lambda} \left(\mathsf{E}\left[F_X^\rho(Y)\right]\right)^{k-2} , \end{aligned}$$

where the expectation is taken w.r.t. the law of Y. Thus, we conclude that

$$\begin{aligned} \overline{T}_{ze} &\leq (k_c-1)/\lambda + M^\rho \sum_{k=k_c}^{\infty} \frac{k}{\lambda} \left(\overline{F_X^\rho(Y)} \right)^{k-2} \\ &\leq k_c/\lambda + \gamma k_c M^\rho \left(\overline{F_X^\rho(Y)} \right)^{k_c} \\ &= k_c/\lambda + \gamma k_c \exp\left\{ -\rho k_c \left[-\frac{1}{\rho} \ln \left(\overline{F_X^\rho(Y)} \right) - \frac{\ln M}{k_c} \right] \right\} , \end{aligned} \tag{65}$$

where $\gamma = \frac{1}{\lambda \left(\overline{F_X^\rho(Y)}\right)^2 \left(1-\overline{F_X^\rho(Y)}\right)^2} \leq \frac{1}{\rho^2 \lambda \left(\overline{F_X(Y)}\right)^2 \left(1-\overline{F_X(Y)}\right)^2}$ is independent of k_c.

We now choose $-\frac{1}{\rho} \ln \left(\overline{F_X^\rho(Y)} \right) - \frac{\ln M}{k_c} = \epsilon/\lambda$, where, for example, $\epsilon = 1/k_c^{1/3}$ can be chosen. As a result we obtain

$$\begin{aligned} \overline{T}_{ze} &\leq k_c/\lambda + \epsilon_T = \frac{\ln M}{-\frac{\lambda}{\rho} \ln \left(\overline{F_X^\rho(Y)} \right) - \epsilon} + \epsilon_T \;, \quad \epsilon_T \geq \gamma k_c e^{-\frac{\rho k_c \epsilon}{\lambda}} \\ \frac{\ln M}{\overline{T}_{ze}} &\geq \left[-\frac{\lambda}{\rho} \ln \left(\overline{F_X^\rho(Y)} \right) - \epsilon \right] \left(1 - \frac{\epsilon_T}{\overline{T}_{ze}} \right) . \end{aligned} \tag{66}$$

Taking $\rho \to 0$, for example $\rho = 1/k_c^{1/3}$, and since $\epsilon \to 0$ and ϵ_T can be as close to zero as we want by choosing k_c sufficiently large, we finally obtain (18) using L'Hôpital's rule.

For the rest of the proof we choose the specific law $f_X(x)$ as per (22), in which case

$$\begin{aligned} F_X(t) &= \int_{0^-}^t f_X(x)dx = 1-(1-a)e^{-Bt} \\ f_Y(y) &= \int_0^y f_X(x) e_\mu(y-x)dx = \frac{\mu}{\mu - B} \left[(a\mu - B)e^{-\mu y} + (1-a)Be^{-By} \right] . \end{aligned} \tag{67}$$

Consequently,

$$\overline{\ln\left(F_X(Y)\right)} = \int_0^\infty \frac{\mu}{\mu - B} \left[(a\mu - B)e^{-\mu y} + (1-a)Be^{-By} \right] \ln\left[1-(1-a)e^{-By} \right] dy \; .$$

The substitution $u = 1-(1-a)e^{-By}$ yields

$$\overline{\ln\left(F_X(Y)\right)} = \left\{ \frac{a - B_\mu}{B_\mu (1-a)^{1/B_\mu}} \int_a^1 (1-u)^{(1/B_\mu - 1)} \ln u \; du - (1-a+a\ln a) \right\} / \left(1 - B_\mu\right) \; .$$

The last integral can be expressed via the Hypergeometric function. Using the relation $\lambda = B_\mu/(B_\mu + 1 - a)$ we obtain

$$C_{ze} \geq \frac{\mu B_\mu}{(1-B_\mu)(B_\mu + 1 - a)} \cdot \left[1 - a + B_\mu \ln a - \frac{(1-a)(a-B_\mu)B_\mu}{(1+B_\mu)} \cdot {}_2F_1\left(1 + \frac{1}{B_\mu}\ ,\ 1\ ;\ 2 + \frac{1}{B_\mu}\ ;\ 1-a\right)\right] ,$$

and the maximum is obtained when $B_\mu \approxeq 0.8192$ and $a \approxeq 0.1230$, thus the lower bound (19) follows.

C. Fixed transmission-time Sphere-Packing Bound

In this section we argue that the sphere-packing bound, obtained by Arikan in [8, Section III] for the one way exponential timing channel, is valid for fixed transmission-time codes over the TSC. To establish such a claim we begin with the following

Lemma 1: The coding capacity of the fixed transmission-time exponential TSC with service rate μ and departure rate λ is given by

$$C(\mu, \lambda) = \lambda \ln(\mu/\lambda)\ , \quad 0 \leq \lambda \leq \mu\ . \tag{68}$$

Proof-Converse: A converse theorem, stating that no rate larger than $C(\mu, \lambda)$ can be achieved with departure rate λ fixed block-length random transmission-time codes with average transmission-time of T, is given in [1, Theorem 8]. The interpretation of this theorem is as follows. $\forall \varepsilon > 0$ and $R \geq C(\mu, \lambda) + \varepsilon$, the ML decoding error is bounded away from zero by a function $g(\varepsilon, \mu, \lambda)$ which is independent of T. We shall prove by way of contradiction that this result is valid for fixed transmission-time variable block-length codes with the same departure rate λ. This, of course, would imply that the same is true for fixed transmission-time fixed block-length codes.

Consider a fixed transmission-time T departure rate λ variable block-length code with rate $R = C(\mu, \lambda) + \varepsilon$ and suppose that $\exists \varepsilon > 0$ such that $\lim_{T\to\infty} P_e = 0$. For such a code the number of departures k, observed during $[0, T]$, is a random variable with $\mathsf{E}[k] = \lambda T$. Fix $\tilde{T} = T + \sqrt{T}$, and $\tilde{n} = \lambda \tilde{T}$. Now, consider a modified genie-aided feedback transmission scheme which attempts to send each message via a fixed block-length codeword as follows.

A message is drawn uniformly over $\{1, \ldots, e^{RT}\}$ and sent as a variable block-length k codeword during a fixed transmission-time T over TSC number 1. In parallel, the same message is sent as a fixed block-length $\tilde{n}$ codeword over TSC number 2 having the same average service time and in which the first k service times per each transmitted message are identical to those of TSC 1 conditioned on $k < \tilde{n}$.

- If $k < \tilde{n}$, TSC number 2 transmits the same message as TSC number 1, with the same departures realization until time T. However, the TSC 2 encoder continues to transmit customers beyond time T, until the decoder observes exactly $\tilde{n}$ departures, and then makes its decision. Therefore, the error probability of the TSC 2 decoder is not larger than that of the TSC 1 decoder that makes its decision based on the first k departures.

- If $k \geq \tilde{n}$ (because $\sum_{i=1}^{\tilde{n}} y_i \leq T$) a genie, that observes both departure processes and finds out that for transmitting the current message TSC 1 consumed at least $\tilde{n}$ departures, indicates to the TSC 2 decoder to declare an error in the current message.

Clearly,

$$\lim_{\tilde{T}\to\infty} \tilde{P}_e \leq \lim_{T\to\infty} P_e + \lim_{T\to\infty} \Pr\left\{\sum_{i=1}^{\tilde{n}} y_i \leq T\right\} .$$

On the other hand, it can be shown, using an application of the central limit theorem, that

$$\lim_{T\to\infty} \Pr\left\{\sum_{i=1}^{\tilde{n}} y_i \leq T\right\} = 0 ,$$

and therefore this modified transmission scheme satisfies $\lim_{\tilde{T}\to\infty} \tilde{P}_e = 0$.

However, the performance of this genie-aided transmission scheme is not better than the performance of a regular departure rate λ fixed block-length random transmission-time with mean $\tilde{T}$ code, as per [1, Definition 4]. Therefore, by [1, Theorem 8] the error probability $\tilde{P}_e$ must be bounded away from zero by at least $g(\varepsilon, \mu, \lambda)$ which is independent of $\tilde{T}$.

Achievability: To show the achievability of the rate $R = \lambda \ln(\mu/\lambda) - \epsilon$, we consider one of our lower bounds on the attainable error exponent with fixed transmission-time codes — the variable length random coding exponent. From (57) we know that as the block-length n gets sufficiently large $\hat{E}_r^{(f)}(R) = \mu a\left(1 - a^\rho\right) - \rho R$. Taking $\rho \to 0$ and using the fact that $1 - a^\rho \to \rho \ln(1/a)$ and $\lambda = \frac{\mu B}{B+\mu(1-a)} = \frac{a\mu}{1+\rho-a\rho}$, we obtain

$$\begin{aligned} a &\to \lambda/\mu \\ \hat{E}_r^{(f)}(R) &\to \rho\left(a\mu \ln(1/a) - R\right) \to \rho\left(\lambda \ln(\mu/\lambda) - R\right) . \end{aligned}$$

Thus, (68) is proved.

The claim in Proposition 3 may be established as follows.

Following Arikan's proof in [8, Section III] for the one way channel, we observe that based on the converse coding theorem obtained in *Lemma 1*, the proof is also valid for fixed transmission-time codes over the TSC. In Arikan's proof, the channel is described by (1). Nevertheless, even if we put $w_k = f_k\left(m, y^{k-1}\right)$, where $f_k : \{1, \ldots, M\} \times \mathbb{R}_+^{k-1} \to \mathbb{R}_+$ is an arbitrary mapping, the proof still holds because it does not depend on the particular way based on which w_k is determined. In the TSC case $w_k = x_k$, and the presence of feedback implies that x_k might be a function of y^{k-1} which is known to the decoder as well.

Appendix I

We calculate the various expurgation exponents when $R \to 0$, and we show that random coding achieves $R \to C$.

For the fixed transmission-time variable block-length code, the optimum for $R \to 0$ is achieved when $\rho \to \infty$ in which case the solution of (8) is $a = 1$, and we obtain $\hat{E}_{ex}^{(f)}(0) = \frac{\mu}{2}$.

For the fixed transmission-time fixed block-length code, the optimal result is obtained when $\theta_B = 0$ and $\rho \to \infty$, (14) becomes

$$\tilde{E}_{ex}^{(0)}(R \to 0) = \mu \frac{r(1-a^2)}{2[B_\mu + (1-a)]} \ . \tag{69}$$

If we take $\frac{B_\mu}{(1-a)} \to 0$ and $a \to 1$ we get from (69) and from (14) the following

$$\begin{aligned} \tilde{E}_{ex}^{(0)}(R \to 0) &= \mu r \\ \tilde{E}_{ex}^{(1)} &= \mu(1-r) \end{aligned}$$

where the choice $r = 1/2$ yields that $\tilde{E}_{ex}^{(f)}(0) = \mu/2$.

When we examine the random transmission-time fixed block-length code, we see that $E_{ex}^{(f)}(R \to 0)$ is also optimized when $\theta_B = 0$ and $\rho \to \infty$, and (17) becomes

$$E_{ex'}(R \to 0) = \mu \frac{(1-a^2)}{2[B_\mu + (1-a)]} \ . \tag{70}$$

If we once again, take $\frac{B_\mu}{(1-a)} \to 0$ and $a \to 1$ we obtain in (70) $E_{ex}^{(f)}(0) = \mu$.

As for $R \to C$, in the fixed transmission-time fixed block-length code, the optimal result is obtained when $\theta_\mu = 0$, and the first exponent in (12) becomes

$$\tilde{E}_r^{(0)}(R \to C) = \rho \left[-\mu \frac{ra}{1+\rho(1-a)} \ln a - R \right] \ . \tag{71}$$

Under these conditions the second exponent in (12) becomes

$$\tilde{E}_r^{(1)} = \mu \left\{ (1-r) - \frac{ra}{1+\rho(1-a)} \ln \left[1 + \frac{(1-r)[1+\rho(1-a)]}{ra} \right] \right\}$$

Since $\ln(1+x) < x$, $\forall x > 0$, we obtain $\tilde{E}_r^{(1)} > 0$. Substituting $r \to 1$ and $a = e^{-1}$ into (71) we obtain

$$\tilde{E}_r^{(0)}(R \to C) \to \rho \left[\frac{\mu e^{-1}}{1+\rho(1-e^{-1})} - R \right] ,$$

which means that by taking $\rho \to 0$, R can come as close as we want to μe^{-1} while $\tilde{E}_r^{(f)}(R) > 0$.

For the random transmission-time fixed block-length case $E_r^{(f)}(R \to C)$ is also optimized when $\theta_\mu = 0$, and (17) becomes

$$E_{r'}(R \to C) = \rho \left[-\mu \frac{a}{1+\rho(1-a)} \ln a - R \right] .$$

Choosing $a = e^{-1}$ we obtain

$$E_{r'}(R \to C) = \rho \left[\frac{\mu e^{-1}}{1+\rho(1-e^{-1})} - R \right] ,$$

which means that by taking $\rho \to 0$, R can come as close as we want to μe^{-1} while $E_r^{(f)}(R) > 0$.

References

[1] V. Anantharam and S. Verdú, "Bits through queues," *IEEE Trans. Inform. Theory,* vol. IT-42, No. 1, pp. 4-18, January 1996.

[2] R. Sundaresan and S. Verdú, "Robust decoding for timing channels," *IEEE Trans. Inform. Theory,* vol. IT-46, No. 2, pp. 405-419, March 2000.

[3] R. Sundaresan and S. Verdú, "Sequential decoding for the exponential server timing channel," *IEEE Trans. Inform. Theory,* vol. IT-46, No. 2, pp. 705-709, March 2000.

[4] R. Sundaresan, "Point process channel and the capacity of the exponential server queue," *Proc. of ISIT 2000,* p. 50, Sorrento Italy, June 2000.

[5] S. Verdú, "The exponential distribution in information theory," *Problems Inform. Transmission,* vol. IT-32, No. 1, pp. 86-95, January-March 1996.

[6] A. Bedekar and M. Azizoğlu, "The information-theoretic capacity of discrete-time queues," *IEEE Trans. Inform. Theory,* vol. IT-44, No. 2, pp. 446-461, march 1998.

[7] J. Giles and B. Hajek, "An information-theoretic and game-theoretic study of timing channels," *IEEE Trans. Inform. Theory,* vol. IT-48, No. 9, pp. 2455-2477, Sep. 2002.

[8] E. Arikan, "On the reliability exponent of the exponential timing channel," *IEEE Trans. Inform. Theory,* vol. IT-48, No. 6, pp.1681-1689, June 2002.

[9] A. B. Wagner and V. Anantharam, “On the zero-rate error exponent of the exponential-server timing channel,” in *Proc. 40th Allerton Conf. Communication, Control and Computing,* pp. 1279-1287, Monticello, IL, October 2002.

[10] A. D. Wyner, “Capacity and error exponent for the direct detection photon channel – Part I,” *IEEE Trans. Inform. Theory,* vol. IT-34, No. 6, pp. 1449-1461, Nov. 1988.

[11] A. Lapidoth, “On the reliability function of the ideal Poisson channel with noiseless feedback,” *IEEE Trans. Inform. Theory,* vol. IT-39, No. 2, pp. 491-503, March 1993.

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

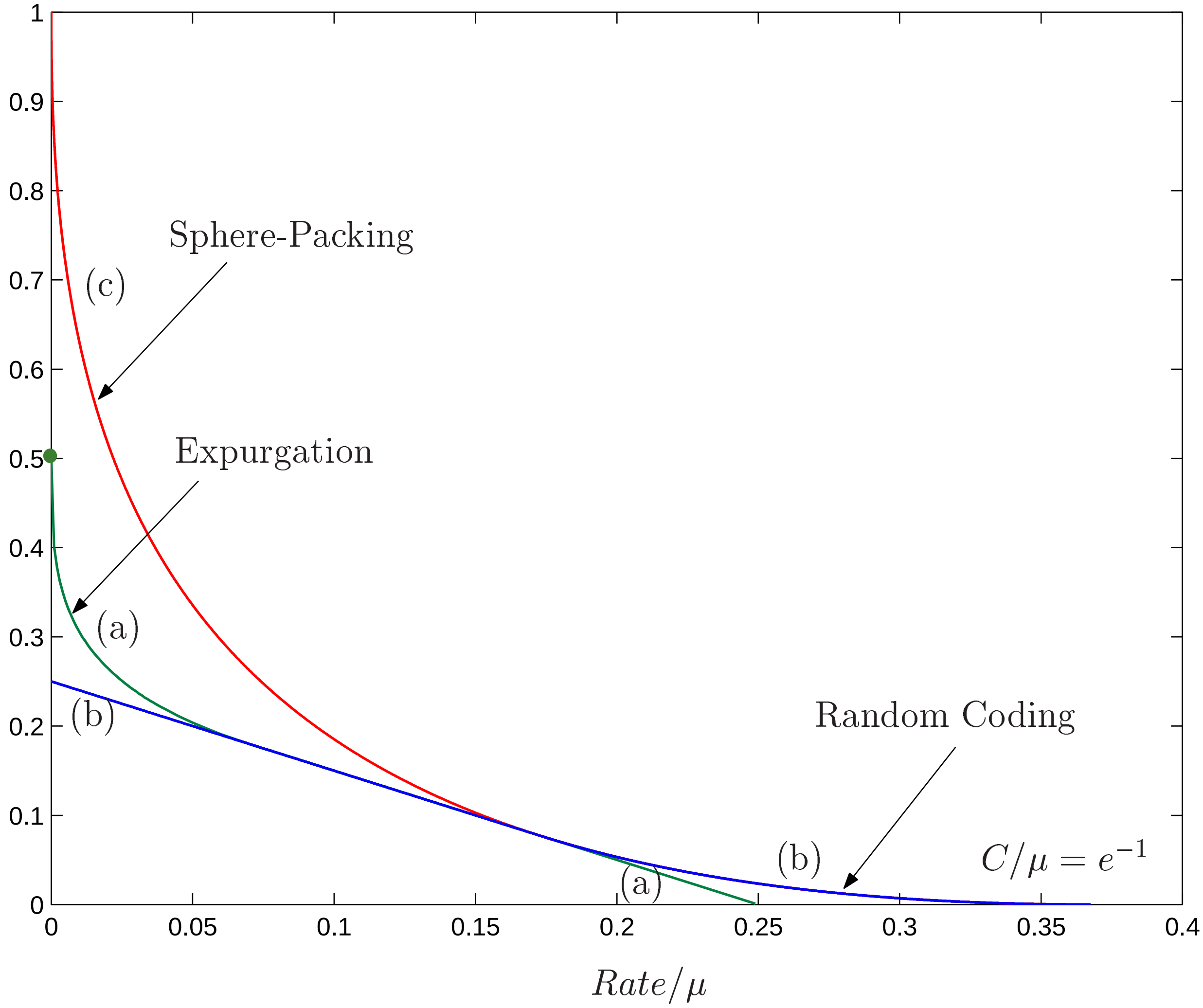

Figure 1: Achievable error exponents for fixed transmission-time variable block-length feedback codes. (a) $\hat{E}_{ex}^{(f)}(R)/\mu$, $\hat{E}_{ex}^{(f)}(0)/\mu = 1/2$. (b) $\hat{E}_{r}^{(f)}(R)/\mu = E_r(R)/\mu$. Sphere-packing exponent for fixed transmission-time feedback codes. (c) $\hat{E}_{sp}^{(f)}(R)/\mu = E_{sp}(R)/\mu$.

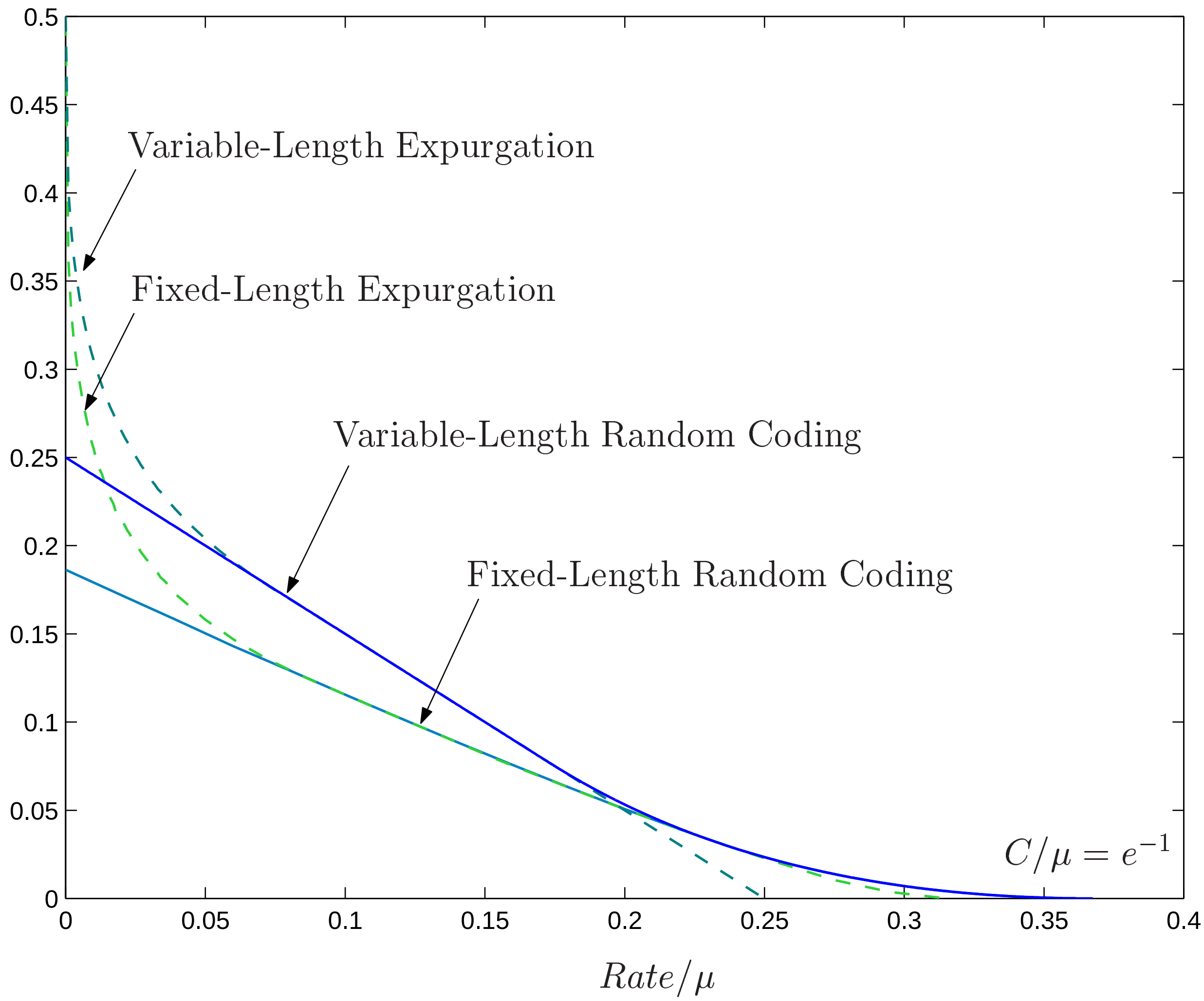

Figure 2: Achievable error exponents for fixed transmission-time, fixed & variable block-length feedback codes. $\hat{E}_{ex}^{(f)}(0)/\mu = \tilde{E}_{ex}^{(f)}(0)/\mu = 1/2$, $\hat{E}_{r}^{(f)}(R)/\mu \geq \tilde{E}_{r}^{(f)}(R)/\mu$.

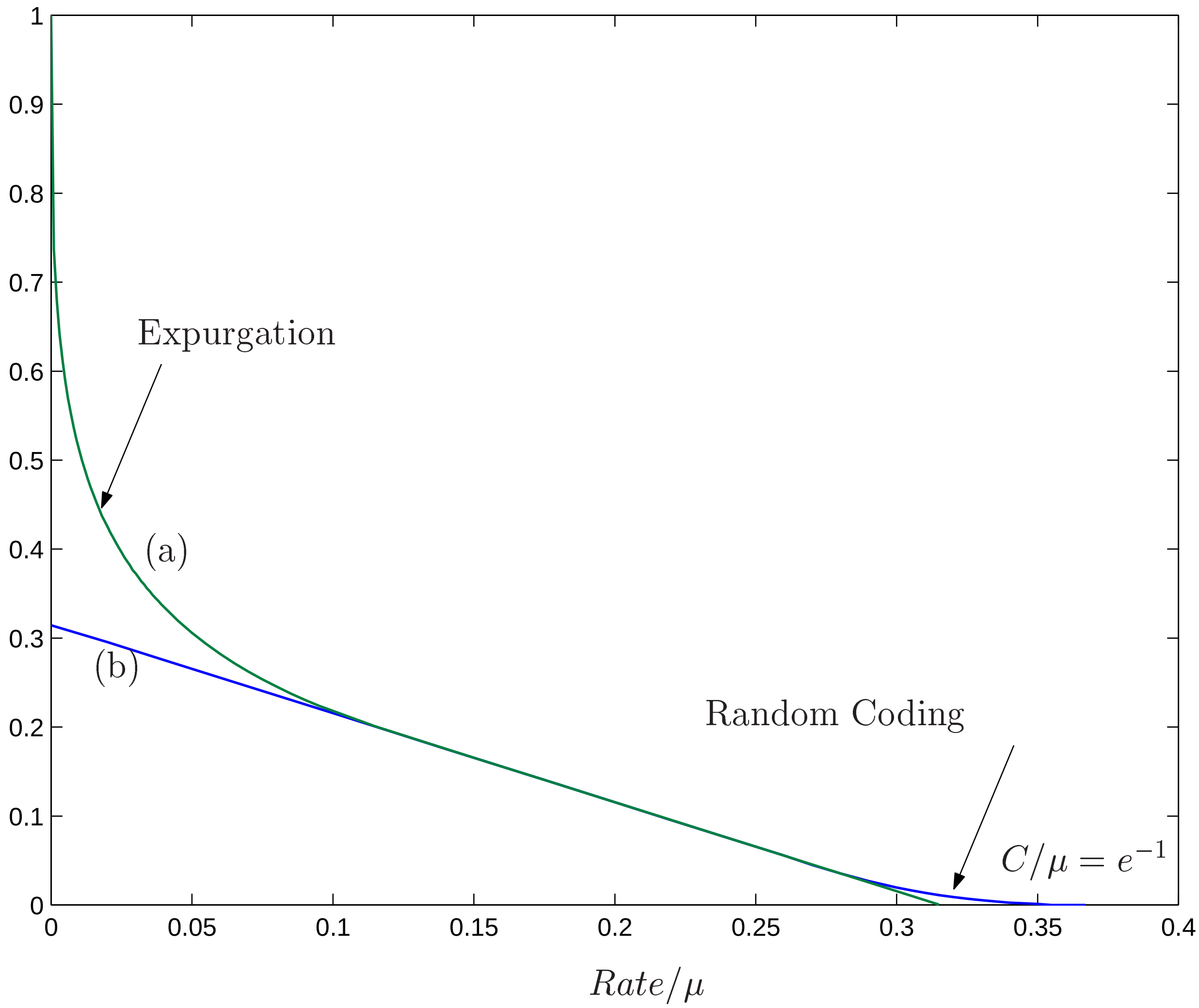

Figure 3: Achievable error exponents for random transmission-time fixed block-length feedback codes. (a) $E_{ex}^{(f)}(R)/\mu$, $E_{ex}^{(f)}(0)/\mu = 1$. (b) $E_r^{(f)}(R)/\mu$.

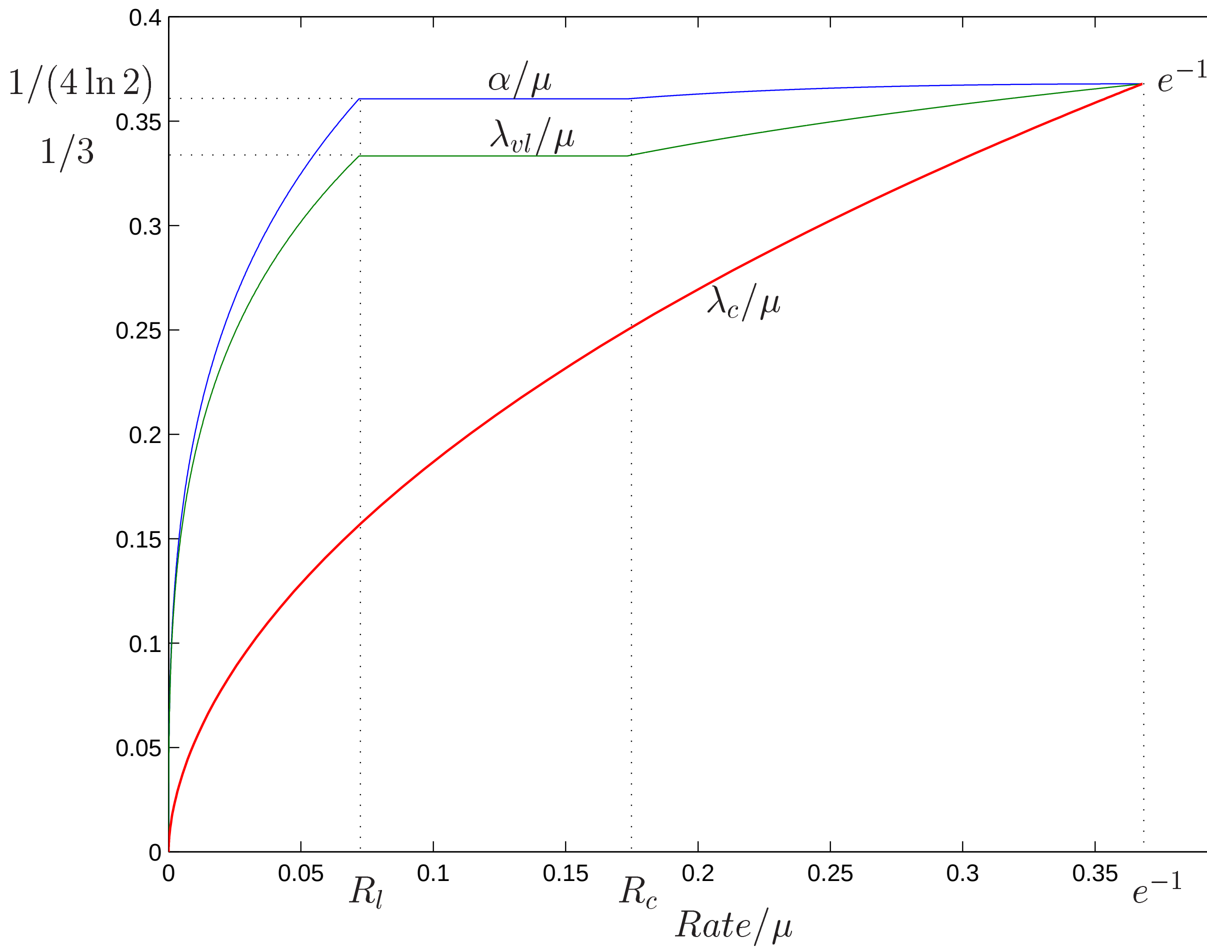

Figure 4: Fixed transmission-time variable block-length feedback codes parameters: $\alpha/\mu \geq \lambda_{vl}/\mu \geq \lambda_c/\mu$.