

Finite Block-Length Achievable Rates for Queuing Timing Channels

Thomas J. Riedl, Todd P. Coleman and Andrew C. Singer

University of Illinois at Urbana-Champaign

Urbana, IL 61801, USA

Email: {triedl2,colemant,acsinger}@illinois.edu

Abstract—The exponential server timing channel is known to be the simplest, and in some sense canonical, queuing timing channel. The capacity of this infinite-memory channel is known. Here, we discuss practical finite-length restrictions on the codewords and attempt to understand the amount of maximal rate that can be achieved for a target error probability. By using Markov chain analysis, we prove a lower bound on the maximal channel coding rate achievable at blocklength n and error probability ϵ is approximated by $C - n^{-1/2}\sigma Q^{-1}(\epsilon)$ where Q denotes the Q-function and σ^2 is the asymptotic variance of the underlying Markov chain. A closed form expression for σ^2 is given.

Index Terms—Timing channel, finite block-length, achievability, geometric ergodicity, asymptotic variance

I. INTRODUCTION

While most communication systems convey information by controlling the amplitudes of signals at each time instant, information can also be sent by controlling the timing at which events occur. For example it is widely believed that neurons exchange information by sending spike trains [1], where information is contained in the random lengths of the interspike intervals. Another example are packet switching networks, where forwarding moves packets from their source toward their ultimate destination. The sources can choose when to send packets, but a queuing mechanism in the forwarding nodes obscures the timing information.

The landmark paper “Bits through Queues” by Anantharam & Verdú [2] characterizes such channels. Suppose the “packets” are identical and only their arrival time carries information. The times at which the sender puts packets on the network encodes a message. The packets go through a first-come, first-serve single-server queue with exponential service times. The decoder observes when the packets depart from the queue and then chooses one of the possible messages. For an

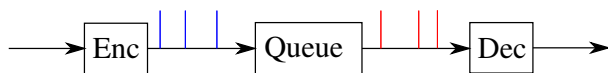


Fig. 1. Conveying information through packet timings in a queuing system

arrival process constrained to be of rate λ packets per second,

This work was supported in part by the department of the Navy, Office of Naval Research, under grants ONR MURI N00014-07-1-0738 and ONR N00014-07-1-0311.

it was demonstrated [2] that for an exponential service time distribution of rate $\mu > \lambda$ the capacity $C(\lambda)$ is given by

$$C(\lambda) = \lambda \log_2 \frac{\mu}{\lambda} \text{ nats/s,}$$

A point process viewpoint version of the problem with the same fundamental limits was considered in [3], [4]. Rather than considering n inter-arrival times and n inter-departure times of the queue, the time axis was fixed to be $[0, T_n]$ at the encoder and $[0, T_n]$ at the decoder. In [5], Bedekar and Azizoglu considered a discrete time analog to the continuous-time model studied in [2] where packets arrive to and depart from a discrete-time single-server queue with i.i.d. geometrically distributed service times. For an arrival process constrained to be of rate λ packets per time slot, it was demonstrated [5] that for queue with service times of rate $\mu > \lambda$ the capacity $C(\lambda)$ is given by

$$C(\lambda) = H(\lambda) - \frac{\lambda}{\mu} H(\mu) \text{ nats/slot,}$$

where $H(\cdot)$ denotes the binary entropy function.

The timing channels with memoryless service times (i.e. exponential in the continuous case and geometric in the discrete case), are known to be the simplest, and in some sense canonical, queuing timing channels. This paper focuses on the discrete time model with memoryless service times and discusses the maximal achievable rate of communication when there is a finite-length restriction on the codewords.

When each codeword corresponds to the timing of packets in n time units and the probability of error may not exceed ϵ , the maximal achievable rate can be substantially less than capacity. By using Markov chain analysis, we prove a lower bound on the maximal channel coding rate achievable at blocklength n and error probability ϵ . We shall show that the maximal channel coding rate is lower bounded by

$$C(\lambda) - n^{-1/2}\sigma Q^{-1}(\epsilon) - \frac{\log n}{2n} + O(n^{-1})$$

where $C(\lambda)$ is the channel capacity whose closed form expression is given above, $Q(\cdot)$ denotes the Q-function and σ^2 is the asymptotic variance of the underlying Markov chain for which we give a closed form expression below. Dropping the last two terms in this expression yields a good approximation

which in turn can be used to anticipate the achievable rate on this channel in the finite block length regime.

Asymptotic expansions of the above form were studied extensively in the 60s for the case of memoryless channels [6]–[9]. The strongest results were given by Strassen in [9]. Recently, this research was rediscovered and extended to Gaussian channels with an input constraint [10]. The major components of the proof can be traced back to the work of Strassen who combined hypothesis testing arguments (Neyman-Pearson lemma), Feinstein’s lemma [11] and bounds on the convergence rate of the central limit theorem [12], [13].

The queuing timing channel considered here has memory and previous results therefore do not apply. We show that the bound provided by the Berry-Esseen theorem in the memoryless channel case still holds and then prove the asymptotic result above by use of Feinstein’s lemma. Further, as mentioned before, we obtain a closed form expression for the asymptotic variance σ^2 . Finding such an expression for a given Markov chain is generally hard and significant research in the area of steady-state stochastic simulation [14]–[16] yields a closed form solution only for the class of homogeneous birth-death processes.

II. BASIC DEFINITIONS AND CONVENTIONS

- For $x \in [0, 1]$, denote $\bar{x} \triangleq 1 - x$.
- Denote $\text{Bern}(p)$ to be the Bernoulli distribution with parameter p and $\text{geom}(p)$ to be the geometric distribution with parameter p :

$$P(k) = p(1-p)^k, \quad k \geq 0.$$

- Denote the binary entropy function $H(p) = -p \log p - \bar{p} \log \bar{p}$.
- X denotes a random variable, $\mathbb{E}[X]$ denotes an expectation, and x denotes a realization.
- \mathbf{x} denotes a vector (x_1, x_2, \dots, x_n)
- A random process $\Phi = (\Phi_1, \Phi_2, \dots)$ on a probability space (Ω, \mathcal{F}, P) is a *Markov process* if for any n ,

$$\begin{aligned} & P(\Phi_1 \in A_1, \Phi_2 \in A_2, \dots, \Phi_n \in A_n) \\ &= P(\Phi_1 \in A_1) \prod_{i=2}^n P(\Phi_i \in A_i | \Phi_{i-1} \in A_{i-1}). \end{aligned}$$

- \mathbb{Z} is the set of all integers and $\mathbb{Z}_+ = \{z \in \mathbb{Z} : z \geq 0\}$.
- Denote $[n]_j = \{j, \dots, n\}$ with $[n] \equiv [n]_1$.
- Denote X^n to be the sequence of counting functions on $[n]_0$, i.e. the set of functions x_0, \dots, x_n for which $x_i \in \mathbb{Z}_+$ and $x_i \geq x_{i-1}$. Denote Y^n to be the set of counting functions y on $[n]_0$ for which $y_0 = 0$.
- For a sequence of input sets and output sets $\{X^n, Y^n : n \geq 1\}$, a channel is a sequence of conditional distributions $\{P_{Y^n|X^n}(\cdot|\mathbf{x}^n) : \mathbf{x}^n \in X^n, n \geq 1\}$.
- Given a distribution P_{X^n} on X^n and channel $P_{Y^n|X^n}(\cdot|\mathbf{x}^n)$, denote P_{Y^n} as the induced output distribution.
- Denote the information density as $i(\mathbf{x}^n, \mathbf{y}^n) \triangleq \log \frac{P_{Y^n|X^n}(\mathbf{y}^n|\mathbf{x}^n)}{P_{Y^n}(\mathbf{y}^n)}$.

- For any $\epsilon \in (0, 1)$, an (M, n, ϵ) code is a sequence $\{(\mathbf{x}^{(i)}, \mathbf{D}^{(i)}), i = 1, \dots, M\}$ where $\mathbf{x}^{(i)} \in X^n$ and $\{\mathbf{D}^{(i)}\}$ are mutually disjoint with $P(\mathbf{D}^{(i)}|\mathbf{x}^{(i)}) > 1 - \epsilon \forall i$.
- The rate of an (M, n, ϵ) code is denoted by $R = \frac{\log M}{n}$.
- Let $N(\epsilon, n, \lambda)$ be the supremum of the integers M such that an (M, n, ϵ) -code exists and $\mathbb{E}[X_n/n] = \lambda$.
- Denote the rate-constrained capacity as $C(\lambda) \triangleq \lim_{\epsilon \rightarrow 0} \lim_{n \rightarrow \infty} \frac{\log N(\epsilon, n, \lambda)}{n}$.
- We drop subscripts whenever they are clear from the context. For example $P_{Y^n|X^n}(\mathbf{y}^n|\mathbf{x}^n) = P(\mathbf{y}^n|\mathbf{x}^n)$.

III. SYSTEM DESCRIPTION AND PRELIMINARIES

Throughout this document, we consider a discrete-time point process version of the problem, analogous to [3], [4]. The communication channel we consider is an interesting example of a channel with memory. It is essentially a probabilistic single server queuing system with the length of the queue being the memory of channel. At each discrete time instance i the random variable $\tilde{X}_i, i \in [n-1]$, indicates if there was an arrival at the back of the queue at time i . The initial length of the queue $Q_0 \equiv X_0$ is a non-negative integer-valued random variable with distribution $P_{Q_0} = P_{X_0}$. Denote $\tilde{X}_i = X_i - X_{i-1}$ and denote \tilde{Y}_i analogously. Denote the queue state for time $i \geq 1$ as $Q_i = X_i - Y_{i-1}$. Then note that we have:

$$X_i = X_0 + \sum_{l=1}^i \tilde{X}_l \quad (1)$$

$$Y_i = 0 + \sum_{l=1}^i \tilde{Y}_l \quad (2)$$

$$Q_i = Q_0 + X_i - Y_{i-1} = Q_{i-1} + \tilde{X}_i - \tilde{Y}_{i-1} \quad (3)$$

Note that there is a bijection between $(Q_0, \tilde{X}_1, \dots, \tilde{X}_n)$ and the channel input, $X^n \triangleq (X_0, X_1, \dots, X_n)$. For geometrically distributed service times, the binary random variables \tilde{Y}_i are conditionally independent given Q_i and are distributed according to the conditional law pertaining to a Z channel:

$$P_{\tilde{Y}_i|Q_i}(\tilde{Y}_i|Q_i) = \begin{cases} 1; & \tilde{Y}_i = 0, Q_i = 0 \\ \bar{\mu}; & \tilde{Y}_i = 0, Q_i > 0 \\ 0; & \tilde{Y}_i = 1, Q_i = 0 \\ \mu; & \tilde{Y}_i = 1, Q_i > 0 \end{cases} \quad (4)$$

The vector $\mathbf{Y}^n \in Y^n$ is the channel output vector and there is a bijection between $\mathbf{Y}^n \in Y^n$ and $(\tilde{Y}_i : i \in [n])$. With this the channel law reads

$$P(\mathbf{y}^n|\mathbf{x}^n) = \prod_{i=0}^{n-1} P(\tilde{y}_i|q_i). \quad (5)$$

Clearly, X_i (Y_i) counts the total number of arrivals (departures), Q_i denotes the length of the queue at time i and \tilde{Y}_i indicates if there was a departure from the front of the queue at time i . This is illustrated in Figure 2.

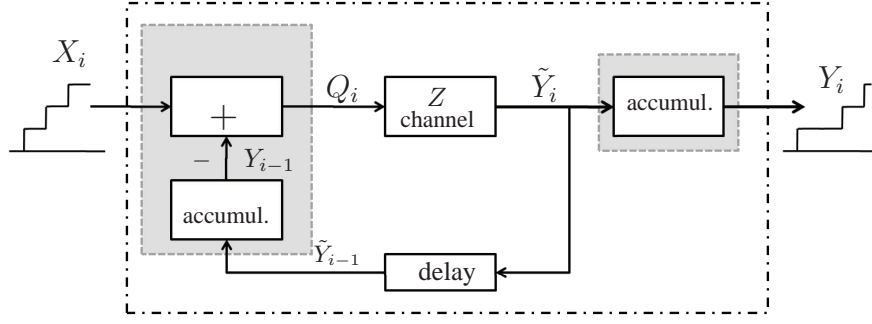


Fig. 2. A simple time-invariant description of the queuing timing channel.

We assume the queue to be stable and hence the arrival rate $\lambda = \frac{\mathbb{E}[X_n]}{n}$ to be smaller than the serving rate μ . We now state the following theorem

Theorem 1. [5]: For the queueing timing channel of rate μ given by (5),

$$C(\lambda) = H(\lambda) - \frac{\lambda}{\mu} H(\mu). \quad (6)$$

The optimal $P_{X^n}^*$ is given by \tilde{X}_i drawn i.i.d. with Bern(λ) distribution and Q_0 independently drawn with π_Q given by

$$\pi_Q(q) = \begin{cases} \frac{\lambda\mu - \lambda\bar{\mu}}{\lambda\mu}; & q = 0 \\ \frac{\lambda\mu - \lambda\bar{\mu}}{\bar{\mu}\mu} \rho^q; & q > 0 \end{cases} \quad (7)$$

where $\rho \triangleq \frac{\lambda\bar{\mu}}{\lambda\mu}$.

Note that $\rho < 1$ if and only if $\lambda < \mu$. We will also exploit how under $P_{X^n}^*$, the output \tilde{Y}_i 's are i.i.d.:

Theorem 2. (Burke's Theorem) For any n , for the channel given by (5) and $P_{X^n} = P_{X^n}^*$ given in Theorem 1, the outputs \tilde{Y}_i 's are i.i.d. with Bern(λ) distribution.

Proof: The proof is similar to the one for continuous time queues and can be found in [17]. A concise proof is also given in [18]. ■

Throughout the remainder of this paper, we assume that $P_{X^n} = P_{X^n}^*$.

By the above theorem

$$P(\mathbf{y}^n) = \prod_{i=0}^{n-1} P(\tilde{y}_i) \quad (8)$$

It is also well-known from Burke's theorem that under $P_{X^n}^*$, the $(Q_i : i \geq 0)$ form a Markov chain and likewise for the random process $((Q_i, \tilde{Y}_i) : i \geq 0)$. The transition probabilities for the Markov chain pertaining to $(Q_i : i \geq 0)$ are given by

$$P_{Q_{i+1}|Q_i}(q_{i+1}|q_i) = \begin{cases} \lambda; & q_{i+1} = q_i + 1, q_i = 0 \\ \bar{\lambda}; & q_{i+1} = q_i, q_i = 0 \\ \bar{\lambda}\mu; & q_{i+1} = q_i - 1, q_i > 0 \\ \lambda\bar{\mu}; & q_{i+1} = q_i + 1, q_i > 0 \\ 1 - \lambda\bar{\mu} - \bar{\lambda}\mu; & q_{i+1} = q_i, q_i > 0 \end{cases} \quad (9)$$

If and only if $\lambda < \mu$, there exists a probability measure π_Q on \mathbb{N}_0 that solves the system of equations

$$\sum_{q_i \in \mathbb{N}_0} \pi_Q(q_i) P_{Q_{i+1}|Q_i}(q_{i+1}|q_i) = \pi_Q(q_{i+1}) \quad (10)$$

for all $q_{i+1} \in \mathbb{N}_0$ and this measure is called the invariant measure. Note that for irreducible Markov chains the existence of such a probability measure is equivalent to positive recurrence. For the transition probabilities given it can be checked that $\pi_Q(q_i)$ as defined in Theorem 1 is the solution.

The following lemma uses arguments introduced by Feinstein [11] to give a lower bound on $N(\epsilon, n, \lambda)$.

Lemma 1. (Feinstein) For any distribution P_{X^n} and any $\theta \in \mathbb{R}$ there exists an (M, n, ϵ) code such that

$$M \geq e^\theta \{\epsilon - P(i(\mathbf{x}^n, \mathbf{y}^n) \leq \theta)\}. \quad (11)$$

Proof: The proof is short and elegant and reproduced in [18]. ■

IV. FINITE-LENGTH SCALING

Recall that by (5) and (8) the distributions $P(\mathbf{y}^n|\mathbf{x}^n)$ and $P(\mathbf{y}^n)$ factor and hence

$$i(\mathbf{x}^n, \mathbf{y}^n) = \sum_{i=0}^{n-1} \log \frac{P(\tilde{y}_i|q_i)}{P(\tilde{y}_i)} = \sum_{i=0}^{n-1} f(\tilde{y}_i, q_i) \quad (12)$$

where

$$f(\tilde{y}_i, q_i) \triangleq \log \frac{P(\tilde{y}_i|q_i)}{P(\tilde{y}_i)}. \quad (13)$$

The composed state $\psi_i = (\tilde{y}_i, q_i)$ again forms a positive recurrent Markov chain whose transition probabilities are illustrated in Figure 3. The invariant measure π_ψ for this chain is only a slight extension to π_Q :

$$\pi_\psi(\tilde{y}, q) = P_{\tilde{Y}_i|Q_i}(\tilde{y}|q) \pi_Q(q) \quad (14)$$

The proof of the following theorem is one of the main contributions of this paper because it can be used to proof bounds and an asymptotic on the quantity $N(\epsilon, n, \lambda)$.

Theorem 3. The asymptotic variance

$$\sigma^2 = \lim_{n \rightarrow \infty} \frac{1}{n} \text{Var}(i(\mathbf{x}^n, \mathbf{y}^n)) \quad (15)$$

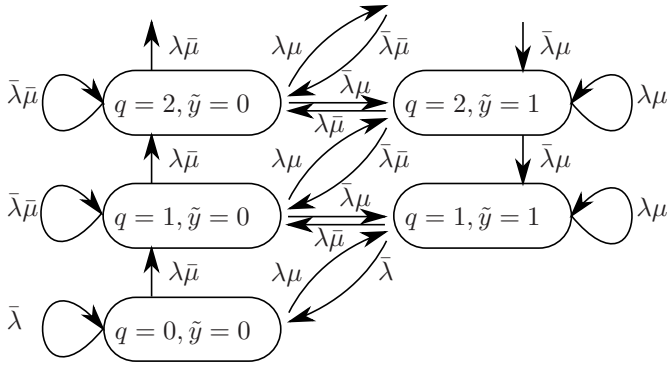


Fig. 3. Possible Transitions in the Markov Chain (\tilde{Y}, Q)

is well defined, positive and finite, and

$$\sigma^2 = \text{Var}(f(\Psi_0)) + 2 \sum_{i=1}^{\infty} \text{Cov}(f(\Psi_0), f(\Psi_i)). \quad (16)$$

Further the following Berry-Esseen type bound holds:

$$\sup_{\xi \in \mathbb{R}} \left| P \left(\frac{i(\mathbf{x}^n, \mathbf{y}^n) - nC(\lambda)}{\sigma \sqrt{n}} \leq \xi \right) - \Phi(\xi) \right| \leq \mathcal{O}(n^{-1/2}) \quad (17)$$

Proof: A detailed proof can be found in [18]. We only give a sketch here. The Markov Chain Ψ is aperiodic and irreducible. The state space of Ψ_i can be chosen to be $\mathbb{X} = \{0, 1\} \times \mathbb{N} \cup \{(0, 0)\}$. First we verify that there exists a Lyapunov function $V : \mathbb{X} \rightarrow (0, \infty]$, finite at some $\psi_0 \in \mathbb{X}$, a finite set $\mathbb{S} \subset \mathbb{X}$, and $b < \infty$ such that

$$\mathbb{E}[V(\Psi_{i+1}) - V(\Psi_i) | \Psi_i = \psi] \leq -1 + b \mathbf{1}_{\mathbb{S}}(\psi), \quad \psi \in \mathbb{X} \quad (18)$$

The chain is skip-free and the found Lyapunov function is linear and hence also Lipschitz. These properties imply that the chain is geometric ergodic [19], [20] and the bound in (17) hence holds by arguments made in [21]. ■

Remark 1. An explicit solution to the asymptotic variance of a general irreducible positive recurrent Markov chain is not available.

Significant research in the area of steady-state stochastic simulation has focused on obtaining an expression for the asymptotic variance [14]–[16] and has yielded a closed form solution only for the class of homogeneous birth-death processes when $f(\psi_i)$ simply returns the integer valued state itself.

We build up on an idea introduced in [22] to give an explicit closed form solution to the asymptotic variance in (15).

Theorem 4. The asymptotic variance defined in (15) has a closed form solution:

$$\sigma^2 = -\text{Var}(f(\Psi_0)) + 2 \sum_{i=0}^{\infty} \text{Cov}(f(\Psi_0), f(\Psi_i)) \quad (19)$$

where

$$\begin{aligned} \text{Var}(f(\Psi_0)) &= \log^2\left(\frac{1}{\lambda}\right) \pi_Q(0) + \log^2\left(\frac{\mu}{\lambda}\right) \overline{\mu \pi_Q(0)} \\ &\quad + \log^2\left(\frac{\bar{\mu}}{\lambda}\right) \overline{\bar{\mu} \pi_Q(0)} - C^2 \end{aligned} \quad (20)$$

$$\begin{aligned} \sum_{i=0}^{\infty} \text{Cov}(f(\Psi_0), f(\Psi_i)) &= \log \frac{1}{\lambda} (-c_{\bar{M}} \frac{\rho}{1-\rho} - c_{M0} \rho) \\ &\quad + \log \frac{\mu}{\lambda} c_{M0} \frac{\rho}{1-\rho} + \log \frac{\bar{\mu}}{\lambda} \frac{\rho}{1-\rho} (c_{\bar{M}} - \rho c_{M0}) \end{aligned} \quad (21)$$

and we defined

$$c_{M0} = \frac{\bar{\lambda}}{\bar{\mu}} \left(\mu \log\left(\frac{\mu}{\lambda}\right) + \bar{\mu} \log\left(\frac{\bar{\mu}}{\lambda}\right) - C \right) \quad (22)$$

$$c_{\bar{M}} = \left\{ \frac{c_{M0}}{\mu} + \left(C - \log \frac{\mu}{\lambda} \right) \frac{\pi_Q(0)}{\bar{\mu}} \right\} \quad (23)$$

Proof: Again we only sketch the proof here and refer to [18] for a detailed version. For the computation of the sum $\sum_{i=0}^{\infty} \text{Cov}(f(\Psi_0), f(\Psi_i))$ we will setup and solve a recursion.

We define

$$r(\psi, i) = \sum_{\psi' \in \mathbb{X}} (f(\psi') - C) \pi_{\Psi}(\psi') p_{\Psi_i | \Psi_0}(\psi | \psi') \quad (24)$$

Clearly

$$r(\psi, 0) = (f(\psi) - C) \pi_{\Psi}(\psi) \quad (25)$$

and

$$\text{Cov}(f(\Psi_0), f(\Psi_i)) = \sum_{\psi \in \mathbb{X}} (f(\psi) - C) r(\psi, i) \quad (26)$$

$$= \sum_{\psi \in \mathbb{X}} f(\psi) r(\psi, i) \quad (27)$$

Note however that for the computation of the asymptotic variance we actually do not even need to know this covariance for each i . It is sufficient to know its sum. So we define

$$R(\psi) = \sum_{i=0}^{\infty} r(\psi, i) \quad (28)$$

exchange limits

$$\sum_{i=0}^{\infty} \text{Cov}(f(\Psi_0), f(\Psi_i)) = \sum_{\psi \in \mathbb{X}} f(\psi) R(\psi) \quad (29)$$

and derive and solve a recursion for the sequence $R(\psi)$. ■

Using the result stated in Theorem 3 we can prove the final contribution of this paper:

Theorem 5.

$$\log N(n, \epsilon, \lambda) \geq nC(\lambda) - \sqrt{n} \sigma Q^{-1}(\epsilon) - \frac{1}{2} \log n + O(1) \quad (30)$$

where $C(\lambda)$ is given by (6) and σ is defined as in Theorem 3.

Proof: By Theorem 3 $\exists A > 0$:

$$|P((i(\mathbf{x}, \mathbf{y}) - nC)/\sqrt{n\sigma^2} \leq \xi_1) - \Phi(\xi_1)| \leq \frac{A}{\sqrt{n}} \quad \forall \xi_1 \in \mathbb{R} \quad (31)$$

Let $B > A$ and $\xi_1 = \Phi^{-1}(\epsilon - \frac{B}{\sqrt{n}}) < \Phi^{-1}(\epsilon) = \xi_0$. Set $\theta = \sqrt{n}\sigma\xi_1 + nC$ and the application of Feinstein's Lemma [11], [18] yields

$$\begin{aligned} & \log N(n, \epsilon, \lambda) - nC - \sqrt{n}\sigma\xi_0 \\ & \geq \log \left(\epsilon - P \left(\frac{i(\mathbf{x}, \mathbf{y}) - nC}{\sqrt{n}\sigma} \leq \xi_1 \right) \right) + \sqrt{n}\sigma(\xi_1 - \xi_0) \end{aligned} \quad (32)$$

$$\geq \log \left(\epsilon - \Phi(\xi_1) - \frac{A}{\sqrt{n}} \right) + \sqrt{n}\sigma O\left(\frac{1}{\sqrt{n}}\right) \quad (33)$$

Remark 2. We believe that the above result can be strengthened by dropping the term $\frac{1}{2} \log n$ from the right hand side of the inequality. By use of hypothesis testing arguments Strassen [9] was able to prove the above bound without the $\frac{1}{2} \log n$ term for the class of discrete memoryless channels. The used arguments can probably be extended to hold for the non-memoryless channel considered here as well since the information density factors and by Theorem 3 a Berry-Esseen type bound holds.

Theorem 5 confirms that $C(\lambda)$ is the operational capacity of the channel and any rate $R < C(\lambda)$ is achievable. The real beauty of this theorem is, however, that it proposes to use the asymptotic

$$\frac{\log N(n, \epsilon, \lambda)}{n} \sim C(\lambda) - n^{-1/2} \sigma Q^{-1}(\epsilon) \quad (34)$$

as an approximation to the channel coding rate $\frac{\log N(n, \epsilon, \lambda)}{n}$ which anticipates the achievable rate on this channel in the finite block length regime. For illustration we plotted this asymptotic for blocklengths ranging between 50 and 3000, various values for ϵ and the example values $\lambda = 0.2$, $\mu = 0.8$ in Figure 4.

ACKNOWLEDGEMENTS

The authors thank Rüdiger Urbanke for hosting one of the authors at his lab at the École Polytechnique Fédérale de Lausanne for several month and providing great guidance. The authors thank Sean Meyn for introducing us to advanced ergodic theory which was crucially important in proving the above Berry-Essen type result.

REFERENCES

- [1] S. Ikeda and J. H. Manton, "Capacity of a single spiking neuron channel," *Neural Computation*, vol. 21, pp. 1714–1748, 2009.
- [2] V. Anantharam and S. Verdú, "Bits through queues," *IEEE Trans. Inform. Theory*, vol. 42, pp. 4–18, 1996.
- [3] R. Sundaresan and S. Verdú, "Capacity of queues via verdú-process channels," *Information Theory, IEEE Transactions on*, vol. 52, no. 6, pp. 2697–2709, 2006.

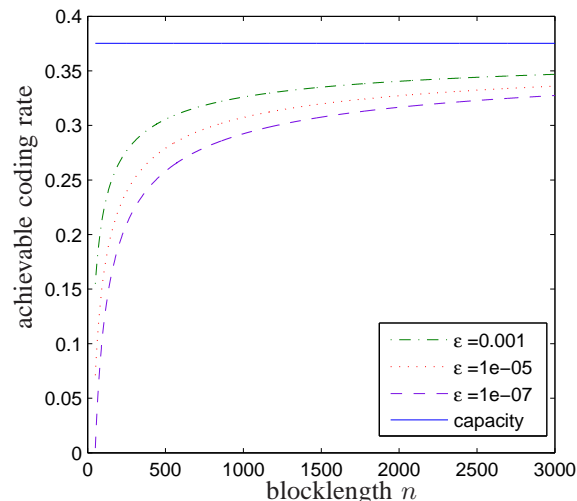


Fig. 4. Channel Coding Rate in the Finite Block-Length Regime

- [4] T. P. Coleman, "A simple memoryless proof of the capacity of the exponential server timing channel," in *Proceedings of the IEEE Information Theory Workshop (ITW)*, Volos, Greece, 2009.
- [5] A. S. Bedekar and M. Azizoglu, "On the information-theoretic capacity of discrete-time queues," *IEEE Trans. Inform. Theory*, vol. 44, pp. 446–461, 1998.
- [6] L. Weiss, "On the strong converse of the coding theorem for symmetric channels without memory," *Quart. Appl. Math.*, vol. 18, no. 3, 1960.
- [7] J. Wolfowitz, "Coding theorems of information theory," *Ergebnisse der Mathematik und ihrer Grenzgebiete*, 1961.
- [8] R. L. Dobruschin, "Mathematical problems in the shannon theory of optimal coding of information," *Fourth Berkeley Symposium*, 1961.
- [9] V. Strassen, "Asymptotische Abschätzungen in Shannon's Informationstheorie," *Trans. Third Prague Conf. Information Theory*, pp. 689–723, 1962.
- [10] Y. Polyanskiy, H. V. Poor, and S. Verdú, "Channel coding rate in the finite blocklength regime," *IEEE Transactions on Information Theory*, vol. 56, pp. 2307–2359, 2010.
- [11] A. Feinstein, "A new basic theorem of information theory," *IRE Trans. Inform. Theory*, vol. PGIT-4, pp. 2–22, 1954.
- [12] W. Feller, "Generalization of a probability limit theorem of cramer," *Transactions of The American Mathematical Society*, vol. 54, pp. 361–361, 1943.
- [13] C. G. Esseen, "Fourier analysis of distribution functions," *Acta Math.*, vol. 77, pp. 1–125, 1945.
- [14] W. Whitt, "Asymptotic formulas for markov processes with applications to simulation," *Operations Research*, vol. 40, pp. 279–291, 1992.
- [15] D. Y. Burman, "A functional central limit theorem for birth and death processes," *ORSA/TIMS Conference*.
- [16] J. G. Kemeny and J. L. Snell, *Finite Markov Chains*. Princeton, N.J., USA: Van-Nostrand, 1960.
- [17] H. Takagi, *Queueing Analysis: A Foundation of Performance Evaluation*. Amsterdam, The Netherlands: North Holland, 1993, vol. 3.
- [18] T. J. Riedl, T. P. Coleman, and A. C. Singer, "Finite block-length achievable rates for queuing timing channels," *IEEE Transactions on Information Theory*, 2011, to be submitted.
- [19] F. M. Spieksma and R. L. Tweedie, "Strengthening ergodicity to geometric ergodicity for markov chains," *Stochastic Models*, vol. 10, pp. 45–75, 1994.
- [20] S. Meyn, *Control Techniques for Complex Networks*, 1st ed. New York, NY, USA: Cambridge University Press, 2007.
- [21] I. Kontoyiannis and S. Meyn, "Spectral theory and limit theorems for geometrically ergodic markov processes," in *the 2001 INFORMS Applied Probability Conference*, 2001, pp. 304–362.
- [22] W. Grassmann, "The asymptotic variance of a time average in a birth-death process," *Annals of Operations Research*, vol. 8, pp. 165–174, 1987.