Timing Capacity of Queues with Random Arrival and Modified Service Times

Guido C. Ferrante°⋆, Tony Q. S. Quek° and Moe Z. Win⋆
°Singapore University of Technology and Design, Singapore
⋆Massachusetts Institute of Technology, MA, USA

***Abstract*—A queue timing channel with random arrival and service times is investigated. The message is encoded into a sequence of additional delays that packets are subject to before they depart from the queue. We derive upper and lower bounds of the channel capacity for general arrival and service processes, and general load. We establish the channel capacity for the queue with exponential server and no load constraint by keeping the queue stable. We discuss the consequences of this result and describe a possible application where the timing channel is used to send covert information.**

I. Introduction

A. Background and Motivation

In the seminal paper "Bits through Queues" (BTQ), Anantharam and Verdú [1] introduced the concept of capacity of a queue. A queue was regarded as a mechanism that blurs the timing information of the stream of packets at its input. In particular, a message is encoded into time intervals between packets, and the decoder infers the message upon observation of the sequence of interdeparture times, that are random due to the random service time and the queueing processes. The capacity of the exponential server queue with arrival rate λ and service rate μ is $C(\lambda) = \lambda \log(\mu/\lambda)$ nats/s. Moreover, for any fixed μ, the arrival rate $e^{-1}\mu$ maximizes $C(\lambda)$ and allows to achieve the capacity of the queue, $C = e^{-1}\mu$ nats/s.

Notable contributions that followed include the works of Bedekar and Azizoglu [2], where analog results were derived for the discrete-time setting, Arikan [3], where upper (sphere-packing) and lower (random-coding) bounds for the reliability function were provided, Sundaresan and Verdú [4], [5], where the maximum-likelihood decoding and a rigorous point-process formulation of queue's capacity were investigated, Prabhakar and Gallager [6], where the entropy rates of arrival and departure processes and the timing capacity of discrete-time queues with Poisson inputs and Poisson outputs were studied, and Wagner and Anantharam [7], where the zero-rate reliability function for the queue with exponential server was found.

It is of practical and theoretical interest the case where not only the service process but also the arrival process is random: Is it still possible to transmit information at a nonzero rate over a queue timing channel when both interarrival and service times are random?

B. Contributions

In this paper we propose a scheme where the service time process is modified by the introduction of an additional delay sequence that encodes the message. The service process $(S_i)_{i\geqslant 1}$ is thus formed by the *nominal* service process $(B_i)_{i\geqslant 1}$, that is defined by the model of the queue, and the *delay* process $(T_i)_{i\geqslant 1}$, that encodes the message to be transmitted. We show that the capacity of the queue with random arrivals and random exponential service times is equal to (in a specific sense) the capacity of the queue with random exponential service times only (as in [1]). Therefore, even when the arrival process is random, there is no loss in capacity as long as the transmitter can encode a message in the delay process. We discuss in detail the consequences of such a result.

C. Notation

The exponential distribution with mean $1/\lambda$ is denoted e_λ. In accordance with Kendall's notation, we succinctly refer to the model with additional delays as the $\cdot/\cdot+\cdot/1$ queue, where we use the plus ('+') sign to indicate that the service time consists of two components and the dot ('$\cdot$') to indicate a distribution. Moreover, by following Anantharam-Verdú's notation, we leave the dot ('$\cdot$') in place of the (unknown) input distribution, while all other distributions are specified in accordance with Kendall's notation. For example, when interarrival times are general and nominal service times are exponentially distributed, we refer to the model as the $G/M+\cdot/1$ queue. We use the above Kendall-Anantharam-Verdú's notation throughout the paper.

II. Model

A. System Model

The system model is depicted in Fig. 1. A message m is transmitted over a timing channel as follows. A queue is fed by packets that are generated with random interarrival times $A^n := (A_1, \ldots, A_n)$, that is, packet i arrives at time $\sum_{k=1}^{i} A_k$. Each packet waits in the queue before it is served. The nomial service time of packet i is denoted B_i. Once it is served, packet i is additionally delayed by a time T_i. As a matter of fact, the service time is $S_i := B_i + T_i$, and consists of two components, one that is random, that is B_i, and one that is intentionally introduced, that is T_i. The message m is encoded in the sequence of delays $T^n := (T_1, \ldots, T_n)$. We will refer to B^n, T^n and S^n as nominal service time, delay, and service time sequences, respectively. Packet i departs, therefore, at time $\sum_{k=1}^{i} U_k = \sum_{k=1}^{i} A_k + S_k$. Finally, the decoder is fed by U^n, on the basis of which it infers the transmitted message m.

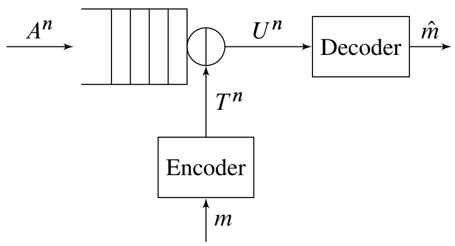

FIG. 1: System model of the timing channel. Packets are generated and enter the queue with random interarrivals A^n. The timing encoder maps a message m into a sequence of delays $T^n = T^n(m)$. Each packet experiences a service time that depends on the nominal service time, due to the queue, and the additional delay, due to the encoding (we emphasized this by representing the server as a circle with a diameter). A decoder observes the interdeparture times U^n, upon which the message m is inferred.

There are two main differences with respect to the model in BTQ: first, the message is encoded in the sequence of delays that contribute to the service times, rather than in the sequence of interarrival times; second, there are two sources of noise, namely the random arrivals A^n and random nominal service times B^n, rather than service time B^n only.

B. Problem Statement

Definitions of code and capacity are adapted from [1].

Definition 1. (Code) An $(n, M, \mathcal{T}, \epsilon)$-code consists of a codebook of M codewords, each of which is a vector of n nonnegative delays $(t_1, \ldots, t_n)$; a decoder which upon estimation of all n departures selects the correct codeword with probability greater than $1-\epsilon$, assuming that the queue is initially empty; and the nth departure from the queue occurs on average no later than $\mathcal{T}$. The rate of an $(n, M, \mathcal{T}, \epsilon)$-code is defined as $(\log M)/\mathcal{T}$. □

Since in any stable queue the interdeparture rate is equal to the interarrival rate, we have $\mathcal{T} = n/\lambda$. Definitions of capacity C, ϵ-achievable rate, and capacity $C(\lambda)$ at output rate λ, are identical to those in [1], and are not reported here.

The problem is that of finding the capacity $C(\lambda)$ at output rate λ and the capacity C.

III. RESULTS

A. Preliminaries

Interdeparture times U^n are related to interarrival times A^n through service times S^n and idle (or free) times[1] F^n (see Fig. 2). As mentioned in the previous section, the sequence of service times S^n is split into a random term B^n and a delay term T^n, the latter of which encodes the message. The encoder can choose the distribution of the nonnegative random variables T^n, that in turn modifies interdeparture times U^n, while both A^n and B^n are sources of noise. Interval F_i is the time interval the queue is idle before packet i arrival, that is, it is the time between packet $i-1$ departure and packet i arrival if this time is positive, otherwise it is nil [1]:

$$F_i = \mathsf{f}_i(A^i, U^{i-1}) := \left(\sum_{k=1}^{i} A_k - \sum_{k=1}^{i-1} U_k \right)^+, \qquad (1)$$

where $(x)^+ := \max\{0, x\}$. Note that F_i is a deterministic function of A^i and U^{i-1}. From the above, the interdeparture time U_i is:

$$U_i = F_i + S_i = \mathsf{f}_i(A^i, U^{i-1}) + B_i + T_i. \qquad (2)$$

For stable queues with packet interarrival rate λ, *i.e.*, $\mathbb{E}[A_i] = 1/\lambda$, it results also $\mathbb{E}[U_i] = 1/\lambda$. Denote $\mathbb{E}[B_i] = 1/\nu$, $\mathbb{E}[T_i] = 1/\beta$, and $\mathbb{E}[S_i] = 1/\mu$. Stability holds iff $\lambda < \mu$. This is usually expressed by writing an equivalent constraint on the load $\rho := \lambda/\mu$, that is, $\rho < 1$. In general, when the load is equal to ρ, it results

$$\frac{1}{\beta} = \mathbb{E}[T_i] = \mathbb{E}[S_i] - \mathbb{E}[B_i] = \frac{1}{\mu} - \frac{1}{\nu} = \frac{\rho}{\lambda} - \frac{1}{\nu}. \qquad (3)$$

B. Upper Bounds

We present two converse results valid for any load $\rho \in [0, 1)$: the first, Lemma 1, is valid for the general server, and the second, Theorem 1, is valid for the exponential server.

Lemma 1 (Converse, $G/G{+}\cdot/1$). The $G/G{+}\cdot/1$ queue with output rate λ satisfies

$$C(\lambda) \leqslant \lambda[1 - \log\lambda - h(\mathbb{B})] \text{ nats/s} \qquad (4)$$

irrespective of $\rho \in [0, 1)$, where $h(\mathbb{B})$ is the average entropy rate of the service time process, *i.e.*,

$$h(\mathbb{B}) = \lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} h(B_i), \qquad (5)$$

when it exists.

Proof: Fano's inequality and the data processing lemma imply that:

$$\log M \leqslant \frac{1}{1-\epsilon}[I(T^n; U^n) + \log 2]. \qquad (6)$$

We bound the mutual information term by applying the chain rule:

$$\begin{aligned} I(T^n; U^n) &= \sum_{i=1}^{n} I(U_i; T^n, U^{i-1}) - I(U_i; U^{i-1}) \\ &\overset{\text{(a)}}{\leqslant} \sum_{i=1}^{n} I(U_i; T^n, U^{i-1}) \\ &\overset{\text{(b)}}{=} \sum_{i=1}^{n} I(U_i; T^i, U^{i-1}) \\ &\overset{\text{(c)}}{=} \sum_{i=1}^{n} I(U_i; T^i, U^{i-1}, F_i) - I(U_i; F_i | T^i, U^{i-1}) \end{aligned}$$

[1] We denote idle times with F^n instead of the more natural I^n in order to maintain I denoting mutual information only. We can think of F_i as the "free" time before the arrival of packet i, whence the F.

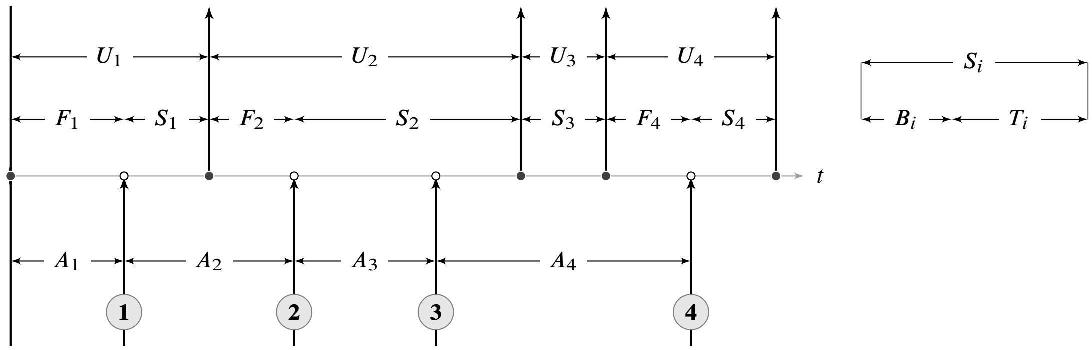

FIG. 2: Example of evolution of a queue. Interarrival and interdeparture times are denoted A^n and U^n, respectively. Packet i arrives at time $\sum_{k=1}^{i} A_k$ and departs at time $\sum_{k=1}^{i} U_k$. Service time S_i is formed by two components, namely the nominal service time B_i and the delay T_i. Idling time before packet i arrival is denoted F_i.

$$\begin{aligned}
&\overset{(d)}{\leqslant} \sum_{i=1}^{n} I(U_i; T_i, F_i) \\
&\overset{(e)}{=} \sum_{i=1}^{n} h(U_i) - h(B_i|T_i, F_i) \\
&\overset{(f)}{\leqslant} \sum_{i=1}^{n} \sup_{\substack{U_i \sim P_U \\ \mathbb{E}[U_i]=1/\lambda}} h(U_i) - h(B_i) \\
&\overset{(g)}{=} n(1-\log\lambda) - \sum_{i=1}^{n} h(B_i), \qquad (7)
\end{aligned}$$

where (a) follows by discarding a mutual information term and since mutual information is nonnegative, (b) is implied by $U_i \perp\!\!\!\perp T_{i+1}^n$, *i.e.*, independence of U_i from future delays $T_{i+1}^n = (T_{i+1}, \dots, T_n)$, (c) follows by Kolmogorov's identity, (d) follows by discarding the second mutual information term in (c) and since F_i is a sufficient statistic of $[T^{i-1}, U^{i-1}, F_i]$ for U_i, (e) follows since entropy is invariant under translations, (f) results by taking the supremum over distributions of nonnegative random variables with given mean and from independence of B_i on any other random variable, and (g) follows by evaluating the supremum (that is attained by the exponential distribution with mean $1/\lambda$). Introducing (7) in (6) yields

$$\frac{\log M}{n} \leqslant \frac{1}{1-\epsilon}\left((1-\log\lambda) - \frac{1}{n}\sum_{i=1}^{n} h(B_i) + \frac{\log 2}{n}\right), \quad (8)$$

that yields the statement by letting $n \to \infty$ and $\epsilon \to 0$. □

The following theorem results from Lemma 1 by specializing the result to the $G/M+\cdot/1$ queue.

Theorem 1 (Converse, $G/M+\cdot/1$, output rate λ). The $G/M+\cdot/1$ queue with output rate λ satisfies, irrespective of $\rho \in [0, 1)$,

$$C(\lambda) \leqslant \lambda \log \frac{\nu}{\lambda} \text{ nats/s.} \quad (9)$$

Proof: Since $(B_i)_{i\geqslant 1}$ is an i.i.d. process with marginal $P_{B_i} = e_\nu$, $\forall i$, it follows that $h(B_i) = 1 - \log\nu$ and $h(\mathbb{B}) = h(B_i)$. The result then follows by using (4). □

We note that the exponential server reaches the minimum of (4) and, therefore, it yields the tightest converse. Moreover, the two above converse results provide an upper bound for the capacity of the $G/G+\cdot/1$ and $G/M+\cdot/1$ queues: by introducing F_i as an input, the channel $P_{U_i|F_i,T_i}$ becomes memoryless and the only remaining source of noise is B_i, which cannot be in any way inferred. Therefore, the upper bound corresponds to the capacity of a memoryless additive exponential channel with fixed mean.

C. Lower Bounds

We provide two achievable rates. The following Lemma 2 is based on the choice of an auxiliary channel and provides an achievable rate for the $G/G+\cdot/1$ queue with fixed λ and ρ. Then we present in Theorem 2 the tightest bound by optimizing the load.

Lemma 2 (Direct, $G/G+\cdot/1$, output rate λ, load ρ). The $G/G+\cdot/1$ queue with output rate λ and fixed load $\rho \in [0, 1)$ satisfies

$$C(\lambda) \geqslant \lambda \log \frac{\frac{1}{\lambda}}{\frac{1-\rho}{\lambda} + \frac{1}{\nu}} \text{ nats/s.} \quad (10)$$

Proof: We start by lower bounding the mutual information as follows:

$$\begin{aligned}
I(T^n; U^n) &\overset{(a)}{=} \sum_{i=1}^{n} I(T_i; U^n, T^{i-1}) \\
&\overset{(b)}{\geqslant} \sum_{i=1}^{n} I(T_i; U_i), \qquad (11)
\end{aligned}$$

where (a) follows from the chain rule and by assuming i.i.d. inputs T^n, and (b) follows because removing data (in this case $[U_1^{i-1}, U_{i+1}^n, T^{i-1}]$) cannot increase mutual information. We proceed by lower bounding each term $I(T_i; U_i)$ by means of an auxiliary channel. From [8, Theorem (Auxiliary-Channel Lower Bound)] it results

$$I(T_i; U_i) \geqslant \mathbb{E}[\log Q_{U_i|T_i}] - \mathbb{E}[\log Q_{U_i}], \quad (12)$$

where Q_{T_i} is any input distribution that maintains the stability of the queue ($\lambda < \mu$), $Q_{U_i} = Q_{T_i} \circ Q_{U_i|T_i}$ denotes the output distribution, and the expectation is carried with respect to the original (true) channel. In particular, we choose the memoryless channel

$$Q_{U^n|T^n=t^n}(u^n) = \prod_{i=1}^{n} Q_{U_i|T_i=t_i}(u_i) = \prod_{i=1}^{n} e_\zeta(u_i - t_i), \quad (13)$$

where $Q_{U_i|T_i} = e_\zeta$ is the exponential distribution with mean

$$\frac{1}{\zeta} = \mathbb{E}[\, Z_i \,] = \mathbb{E}[\, F_i + B_i \,] = \frac{1-\rho}{\lambda} + \frac{1}{\nu} = \frac{1}{\lambda} - \frac{1}{\mu} + \frac{1}{\nu}, \quad (14)$$

that implies

$$\frac{1}{\beta} = \mathbb{E}[\, U_i \,] - \mathbb{E}[\, Z_i \,] = \frac{1}{\lambda} - \frac{1}{\zeta} = \frac{\rho}{\lambda} - \frac{1}{\nu}. \quad (15)$$

In other words, our choice consists in regarding $\{Z_i := F_i + B_i\}_{i=1}^n$ as a set of i.i.d. exponentially distributed random variables. We choose Q_{T_i} in order to attain the maximum mutual information of the *auxiliary* channel so to tighten (12) as much as possible. Denote $a = \mathbb{E}[\, T_i \,] = 1/\beta$ and $b = \mathbb{E}[\, Z_i \,] = 1/\zeta$ for brevitiy. Based on [1, Theorem 3], the optimum Q_{T_i} is as follows,

$$Q_{T_i} = \frac{b}{a+b}\delta_0 + \frac{a}{a+b} e_{1/(a+b)}, \quad (16)$$

where δ_0 is the point mass at 0. Denote e_{λ_1,λ_2} the hypo-exponential distribution with rates λ_1 and λ_2, that is the distribution of the sum of two independent exponentially distributed random variables with rates λ_1 and λ_2. It is formally equal to

$$e_{\lambda_1,\lambda_2} = \frac{\lambda_2}{\lambda_2 - \lambda_1} e_{\lambda_1} - \frac{\lambda_1}{\lambda_2 - \lambda_1} e_{\lambda_2}. \quad (17)$$

The output distribution of the auxiliary channel when the input distribution is (16) is, therefore, given by

$$\begin{aligned} Q_{U_i} &\overset{(a)}{=} \frac{b}{a+b} e_{1/b} + \frac{a}{a+b} e_{1/b,1/(a+b)} \\ &\overset{(b)}{=} \frac{b}{a+b} e_{1/b} + \frac{a}{a+b} \left(\frac{\frac{1}{a+b}}{\frac{1}{a+b} - \frac{1}{b}} e_{1/b} - \frac{\frac{1}{b}}{\frac{1}{a+b} - \frac{1}{b}} e_{1/(a+b)} \right) \\ &\overset{(c)}{=} e_{1/(a+b)} = e_\lambda, \end{aligned} \quad (18)$$

where (a) follows from (16), the definition of the auxiliary channel $Q_{U_i|T_i}$, and $e_\zeta = e_{1/b}$, (b) results from (17), and (c) follows from straightforward computation. Introducing (16) in (12) and using (18) yields

$$I(T_i; U_i) \geqslant \log \frac{\zeta}{\lambda} = \log \frac{\frac{1}{\lambda}}{\frac{1}{\lambda} - \frac{1}{\beta}}. \quad (19)$$

The result follows by using (19) in (11) and using (15). □

Fig. 3 shows the upper bound given in Theorem 1 (cf. (9)) and the lower bound in Lemma 2 (cf. (10)) as a function of the ratio ν/λ, for a fixed load (on figure it is shown $\rho = 0.75$). Let us rewrite the two bounds in nats/use as follows:

$$\log \frac{\frac{\nu}{\lambda}}{(1-\rho)\frac{\nu}{\lambda} + 1} \leqslant C(\lambda) \leqslant \log \frac{\nu}{\lambda} \text{ nats/use.} \quad (20)$$

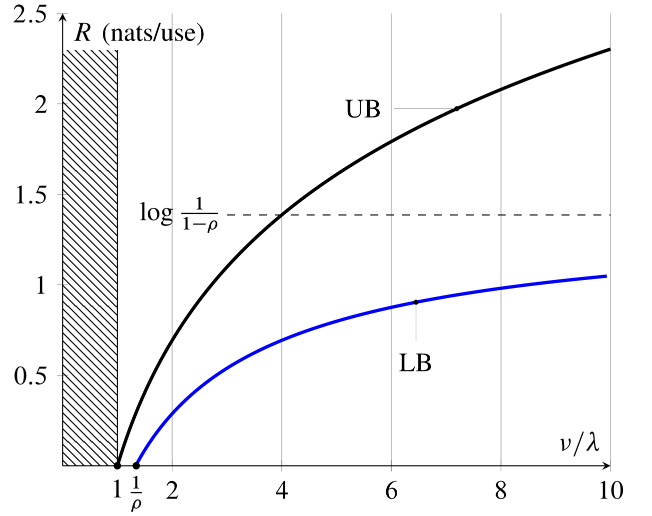

FIG. 3: Rate bounds for the $G/M{+}\cdot/1$ queue. Upper and lower bounds refer to (9)-(10). The particular value of the load is $\rho = 0.75$.

The lower bound tends to $\log \frac{1}{1-\rho}$ nats/use as $\nu/\lambda \to \infty$ (dashed line on figure), and it is nonnegative for $\nu/\lambda \geqslant 1/\rho$. The upper bound is unbounded as $\nu/\lambda \to \infty$, and it is nonnegative for $\nu/\lambda \geqslant 1$ (stability condition). Upper and lower bound match when $\rho \to 1^-$ as discussed in the next section.

Lemma 2 allows to derive the following achievability bound by optimizing the load.

Theorem 2 (Direct, $G/G{+}\cdot/1$, output rate λ, optimum load). The $G/G{+}\cdot/1$ queue with output rate λ satisfies

$$C(\lambda) \geqslant \lambda \log \frac{\nu}{\lambda} \text{ nats/s.} \quad (21)$$

Proof: The result follows from Lemma 2 by taking the supremum over $\rho \in [0, 1)$, which is achieved as $\rho \to 1^-$. □

The direct lemma above provides an achievable rate for the $G/G{+}\cdot/1$ queue with fixed load ρ by using i.i.d. inputs and by replacing the channel with memory with a worse additive channel, namely a memoryless additive exponential channel, with same first-order statistic. In this case, the lower bound is maximized as $\rho \to 1^-$ (cf. (10)), *i.e.*, in the *heavy-traffic regime*: in words, the maximum rate is achieved when there is no "free" (idle) time. The intuition is that, in the mentioned regime, it results $F_i \xrightarrow{d} \delta_0$ and thus $Z_i := F_i + B_i \xrightarrow{d} P_B$: when B_i is exponentially distributed, then the selected auxiliary channel (additive exponential channel) tends to the true channel. The resulting bound in Theorem 2 turns out to be tight for the $G/M{+}\cdot/1$ queue with the optimum load, for which capacity is found.

D. Capacity

In the below Theorems 3 and 4 we provide capacity results for the $G/M{+}\cdot/1$ queue.

Theorem 3 (Capacity of $G/M{+}\cdot/1$, output rate λ). The capacity of a $G/M{+}\cdot/1$ queue with output rate λ and nominal service rate ν is

$$C(\lambda) = \lambda \log \frac{\nu}{\lambda} \text{ nats/s.} \tag{22}$$

Proof: The result follows by (9) and (21) from Theorem 1 and 2, respectively. □

The result in Theorem 3 is very similar to the capacity of the $\cdot/M/1$ queue found in BTQ [1], provided that we identify the nominal service rate ν with the service rate μ in [1]. The difference that lies in the two models, however, imply that for the $G/M+\cdot/1$ queue, the total service rate tends to λ because the delay process $(T_i)_{i\geqslant 1}$ tends to saturate the queue, $\rho \to 1^-$. Nonetheless, given a queue with nominal service time that is exponentially distributed with rate ν, the capacity of both schemes is given by (22). In other words, one can achieve a rate equal to $\lambda \log(\nu/\lambda)$ nats/s with the exponential server queue by either feeding the queue with a Poisson process as in [1] or delaying output packets according to a suitable process (cf. (16)) as in the present paper. In both cases, the resulting queue is the $M/M/1$ queue. It follows without surprise that the maximum of the capacity with respect to λ is same as in [1], as shown in the below theorem:

Theorem 4 (Capacity of $G/M+\cdot/1$). The capacity of a $G/M+\cdot/1$ queue is

$$C = e^{-1}\nu \text{ nats/s.} \tag{23}$$

Proof: The result follows by taking the maximum of (22) with respect to λ, that is achieved for $\lambda = e^{-1}\nu$. □

IV. Discussion

In this paper we presented a scheme to transmit a message by means of a queue timing channel. In particular, we assumed that both arrival and service times were random. The message was encoded in a sequence of additional delays introduced as part of the service time process. Each service time was thus formed by a nominal service time, due to the server, and a delay, due to the encoding. This scheme is different from that in BTQ, where arrival times are used to encode the message.

We derived upper and lower bounds for general queue models (general arrival and service processes) and general load. We established the channel capacity for the exponential server queue without load constraint, and showed that it is equal (in a specific sense) to that derived in BTQ. From a practical standpoint, the result in [1] requires that an infinite set of packets is ready to be transmitted, while the model in this paper does not require this assumption. This comes at the price that the capacity is achieved by saturating the queue, and therefore it can be applied to infinite-buffer queue only. In a future work we will extend the above analysis to finite-buffer queues.

We conclude by describing a possible application of the model discussed in this paper, that differs from that possible with the model in BTQ. In general, point-to-point timing channels can be used to either convey covert information or "piggyback" timing information on packets to increase the capacity (or the energy efficiency) of the "compound channel,"

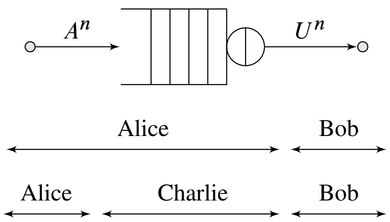

Fig. 4: Possible scenario. Case 1: Alice transmits a message to Bob over the content channel and another message over the timing channel. Case 2: Alice transmits a message to Bob over the content channel; Alice's packets are collected and forwarded to Bob by Charie, who transmits another message to Bob using the timing channel.

that is that using the timing channel on top of the traditional content channel. Consider the following practical scenario. Alice intends to send a message to Bob through a stream of packets. The model in BTQ assumes that Alice generates and introduces a stream of packets in her own queue according to a specific arrival process: this is shown in Fig. 4 (Case 1). Contrarily, suppose now that the communication occurs through a third party, Charlie, that forwards Alice's packets to Bob: this is shown in Fig. 4 (Case 2). Suppose that Charlie cannot modify the content or the order of packets, and uses a timing channel only to transmit a message to Bob. Alice may or may not be aware of the presence of Charlie: in both cases, suppose there is no coordination between them. Differently from Alice, Bob always knows that Charlie is forwarding the packets and is using a timing channel to send him a message. When Alice transmits her packets, Charlie collects the stream in his queue (it is natural to assume that the arrival process is general or Poisson). Results in this paper apply to the capacity of the timing channel between Charlie and Bob, that is made possible thanks to Alice's packet stream.

References

[1] V. Anantharam and S. Verdú, "Bits through queues," *IEEE Trans. Inf. Theory*, vol. 42, no. 1, pp. 4–18, 1996.

[2] A. Bedekar and M. Azizoglu, "The information-theoretic capacity of discrete-time queues," *IEEE Trans. Inf. Theory*, vol. 44, no. 2, pp. 446–461, 1998.

[3] E. Arikan, "On the reliability exponent of the exponential timing channel," *IEEE Trans. Inf. Theory*, vol. 48, no. 6, pp. 1681–1689, 2002.

[4] R. Sundaresan and S. Verdu, "Robust decoding for timing channels," *IEEE Trans. Inf. Theory*, vol. 46, no. 2, pp. 405–419, 2000.

[5] R. Sundaresan and S. Verdú, "Capacity of queues via point-process channels," *IEEE Trans. Inf. Theory*, vol. 52, no. 6, pp. 2697–2709, 2006.

[6] B. Prabhakar and R. Gallager, "Entropy and the timing capacity of discrete queues," *IEEE Trans. Inf. Theory*, vol. 49, no. 2, pp. 357–370, 2003.

[7] A. Wagner and V. Anantharam, "Zero-rate reliability of the exponential-server timing channel," *IEEE Trans. Inf. Theory*, vol. 51, no. 2, pp. 447–465, 2005.

[8] D. M. Arnold, H.-A. Loeliger, P. O. Vontobel, A. Kavcic, and W. Zeng, "Simulation-based computation of information rates for channels with memory," *IEEE Trans. Inf. Theory*, vol. 52, no. 8, pp. 3498–3508, 1998.