Average-delay optimal policies for reliable point-to-point communication with variable length block codes

Vineeth Bala Sukumaran and Utpal Mukherji

***Abstract*—A noisy point-to-point channel with random arrival of message symbols to the channel's transmitter is considered in this paper. These message symbols are buffered at the transmitter and are required to be reliably transmitted to the channel's receiver with low average delay. For reliability, the transmitter does channel coding and jointly encodes multiple message symbols using finite length block codes. For reducing the average delay, the transmitter dynamically schedules the number of message symbols jointly encoded. This paper considers the design of transmitter scheduling policies which minimize the average delay when the transmitter uses adaptive coding schemes that ensure the reliability of every transmission. An important feature of such policies and coding schemes is that the transmission durations for jointly encoded message symbols are not fixed. The duration of each transmission is a function $\tau(.)$ of the number s of message symbols that are jointly encoded and transmitted. The design of scheduling policies is considered by modelling the point-to-point channel as a discrete time batch-arrival batch-service queueing model with the batch service time being a function $\tau(s)$ of batch service size s. Approximately optimal dynamic control policies and performance bounds for typical forms of $\tau(s)$ that arises in point-to-point communication are obtained using a semi-Markov decision process framework. The typical forms for $\tau(s)$ are obtained from the Gallager's random coding upper bound as well as the Polyanskiy's normal approximation for the codeword error as a function of the codeword length and rate. For high reliability it is observed that a policy that serves exhaustively is close to optimal. The paper also considers the higher-layer tradeoff between reliability and the average queueing delay as an extension of the traditional tradeoff characterization of reliability and codeword length.**

I. INTRODUCTION

We consider a noisy point-to-point channel with random arrival of message symbols to the channel's transmitter. The message symbols are buffered at the transmitter. The transmitter's objective is to reliably transmit these message symbols with low average delay to the receiver. We assume that each message symbol has to be transmitted to the receiver with a reliability of at least $1-P_{e,r}$ or a probability of error of at most $P_{e,r}$. For ensuring reliability, the transmitter jointly encodes and transmits batches of message symbols using finite length block codewords. Even though the reliability requirement is on message symbols, we assume that the transmitter uses a block coding scheme so that the probability of codeword error of every transmission of a batch of message symbols is at most $P_{e,r}$. We note that this ensures the probability of error of every message symbol to be at most $P_{e,r}$.

The first (corresponding) author is with the Department of Avionics, Indian Institute of Space Science and Technology, Trivandrum. The second author is with the Department of Electrical Communication Engineering, Indian Institute of Science, Bangalore - 560012. Email: vineethbs@gmail.com,utpal@iisc.ac.in

Our motivation for considering such a scenario stems from the use of adaptive modulation and coding (AMC) methods in wireless communication systems such as WiFi [1], [19]. AMC methods lead to an improved and efficient use of wireless communication resources by adapting the transmitter rate using modulation and coding in response to dynamic channel conditions as well as network load. We note that AMC is required to provide a bit error rate or symbol error rate guarantee to the upper layer while adapting the modulation and coding schemes to the channel or load conditions, which forms the motivation for the reliability requirement stated above.

An important question which arises is how AMC should dynamically choose the modulation and coding scheme as a function of the time varying channel conditions and network load so as to optimize quality of service performance metrics, such as average delay, for this fixed reliability guarantee. This policy design problem has been considered by a number of authors, such as [5], [6], [26], [45], [46] (wherein it was assumed that any reliability could be achieved in a fixed amount of channel usage time). Performance analysis of specific policies have been considered in [28], [18] (wherein the performance of specific coding schemes were considered). In this paper, we consider this question for the case of a non-fading point-to-point link with adaptive coding where the dependence of the block codeword length on reliability and code rate are explicitly modelled using asymptotic [16] and non-asymptotic [30] information theoretic relationships.

We now observe that the dynamic choice of the coding scheme can be modelled as the control of a queue with service dependent batch service times. The transmission rate is dynamically controlled by changing the size of a batch of message symbols which is mapped into a block codeword at every transmission. Then, the length of the block codeword, i.e., the transmission time, for the batch of message symbols is an increasing function of the number of message symbols in the batch, and depends on the fixed reliability. If s is the number of message symbols transmitted in a batch, then the transmission time is a function $\tau(s)$ which is increasing in s. Since the transmitter also has low average delay as an objective, intuitively, the transmitter should transmit the message symbols at the maximum rate, i.e., it should choose

s such that the rate $\frac{s}{\tau(s)}$ is the maximum possible. However, we note that the total delay experienced by a message symbol is the sum of its waiting time in the transmitter buffer followed by the time taken to transmit the symbol as part of a block codeword to the receiver. So, if the rate of transmission is increased then the transmission time increases, with a possible reduction in the waiting delay of the message symbols in the buffer. It is then not clear if the total delay is minimized by a transmission policy that always transmits at the maximum rate possible. So the problem of designing a average delay optimal policy under per-transmission fixed reliability constraint is not direct; there are tradeoffs between different components of the average delay which needs to be taken into consideration for minimizing the average delay. In this paper, we consider this policy design problem and characterization of the minimum average delay as a function of reliability. We propose heuristic policies and analyze their performance using average delay upper and lower bounds subject to a reliability constraint. We show that a policy that serves exhaustively is *near-optimal* when the required reliability is *large*.

We also note that average delay and average error probability are basic cross-layer performance measures for a non-fading noisy point-to-point channel, when queueing of randomly arriving message symbols at the transmitter is also taken into consideration. The above performance bounds are useful for quantifying the tradeoff between these cross-layer performance measures. Such a tradeoff characterization builds on the information theoretic tradeoffs that quantify the asymptotic [16] or non-asymptotic [30] dependence of the average codeword error probability on codeword block length.

A. Prior work

We note that transmitter rate control policy design has been considered by many authors, a small sample of which is [45], [46], [6], [5], and [15]. However, they assume that any rate in the capacity region of the communication system can be instantaneously achieved in a large enough slot time. A more information theoretically motivated model should also consider the change in codeword block lengths which are required to achieve that rate at a particular reliability level. Such information theoretically motivated models were first considered by Telatar and Gallager in [39]. Telatar and Gallager [39] introduced the method of viewing a point-to-point channel as a queueing system by interpreting transmission requirements (such as probability of error) as the service requirement of a message or 'customer' arriving into the transmitter, and the error exponent of the channel as the service rate. Sayee and Mukherji, in [33] and subsequent papers, have extended these ideas to the multiple access channel and the degraded broadcast channel and obtained a discrete time processor sharing queueing model for both the independent and joint decoding receivers. Our modelling of the reliability as a function of power, rate, and codeword length is motivated by these approaches.

Recently, ultra reliable and low latency communication (URLLC) use cases in 5G have motivated the study of performance of systems using short codes [10], which are used to provide reliability with low latency. URLLC require transmissions to have reliability of 99.999% and meet delay deadlines (rather than average delay) of 1 ms [38]. Devassy et. al. [14] and Schiessl et. al. [34] have considered point-to-point queueing models where reliability dependence on the codeword length is modelled using non-asymptotic results such as those in [30]. We note that specific transmission rate control policies were considered in this paper rather than optimal control. Anand et al. [4] consider a variable length coding problem in the context of scheduling radio resource units in the time-frequency domain to URLLC traffic. Packets arrive into a multiclass system and radio esource units are allocated both in time and frequency. The number of units in time and frequency dynamically varies according to the traffic class. The queueing behaviour of point-to-point channels with specific codes such as the BCH code have been studied by Parag et al. [28] and Hamidi-Sepehr et. al [18], but without the use of dynamic control. Zhao et. al. [47] considers the tradeoff of average power and delay for variable length coding schemes for URLLC use cases in 5G. The variable length coding is done along resource units at different frequencies rather than across time, so that the variable service time feature of our model does not arise here. We note that our study of average delay cannot be directly applied to the performance measures which are studied here, although such an analysis is part of future work.

We note that the tradeoff of performance measures such as average delay and average power for a point-to-point channel with fixed length codewords has been studied in [5] and exhaustively covered in [26]. Berry and Gallager [5] characterize the minimum transmission power as a function of average delay constraints for a fading point to point channel and can be thought of as a characterization of the fundamental power-delay tradeoff for a fixed reliability. We note that in [5] the transmission power expended in each slot is modelled as a function of the channel state as well as the rate of service. For the special case of a non-fading system (such as ours), replacing the power function with an average error probability function leads to a characterization of the tradeoff between average error probability and average delay for a fixed power. We note that the average error probability is allowed to change from slot to slot, where each slot corresponds to the time duration to transmit a fixed length codeword. Thus, [5] obtains an asymptotic characterization of the tradeoff of delay, power, and reliability, either in the delay-power plane for a fixed reliability, or the delay-reliability plane for a fixed power in the regime of large delays. The control that the transmitter has is on the rate of the transmission for fixed length codewords.

Neely [26] considers the problem of tradeoff of average power and average delay in a general stochastic network setting and proposes drift-plus-penalty algorithms that can achieve the asymptotic tradeoffs obtained in [5]. Again the control that the transmitter(s) has is on the rate of transmission for fixed length codewords.

Biyikoglu et. al. [42] consider a related problem of minimizing the total transmission energy required to transmit multiple packets within a deadline. Their scheme uses variable length coding of each packet in order to tradeoff the transmission energy with the delay. An optimal offline scheduling algorithm has been proposed for minimizing the total transmission energy for a finite number of packets and extended to the case of infinite number of packets. We again note that if the transmission energy function is replaced by an average error probability function, we have an optimal scheduling algorithm for minimizing the average error probability for packets with a delay deadline being met. The control used is the length of the codeword, but multiple packets are not jointly encoded together.

We also note that the queueing model that we study in this paper, which is a batch service queue with batch size dependent service times, has only been considered for some special cases in the queueing theory literature. Neuts [27] provides examples where such batch size dependent service time queueing models are applied in the study of car ferries, elevator scheduling, and traffic scheduling. Claeys [7, Chapter 2] has also applied such queueing models to the study of the waiting time, as well as other performance measures, for dynamic group testing policies, under any general group-subgroup screening strategy. We observe that our system can be modelled as a discrete time queue with batch service, and therefore we now review the literature on the study of queueing models with batch service. Neuts [27] considered a continuous time queueing model with Poisson arrivals where customers are served in batches, according to the following threshold policy. At every decision epoch, if the number of customers is less than a threshold, then none are served. On the other hand, if the number of customers are greater than or equal to the threshold but less than the maximum number which can be served, then all the customers are served, and otherwise the maximum number is served. Probability generating functions (PGF) of various performance measures such as the number of customers at every departure and the length of the busy period were obtained for this policy. The service times of each batch could be random, with the distribution depending on the batch size used. Claeys et al. [8] and [9] considers a discrete time version of Neut's queueing model. In [7] Claeys et al. consider a policy which is similar to Neut's except that at a decision epoch, if the number of customers is less than a threshold, then service is initiated with some probability. For this policy, for independent and identically distributed (IID) batch arrival processes, the PGF of the queue length in every slot is used to obtain moments and tail probabilities of the queue length numerically. Heavy and light traffic approximations of the average queue length for the policy are then obtained which are compared with the numerically obtained exact results. Furthermore, the PGF analysis is also used to obtain higher moments of the average queue length and customer delay. The analysis is also extended to correlated arrival processes.

To the best of our knowledge, the study of dynamic control of batch service queues to minimize the average queue length was initiated by Deb and Serfozo in [12]. Deb and Serfozo considered a continuous time queueing model with Poisson arrivals. At a scheduling epoch, a batch of customers, with batch size at most S_{max}, can be served. The batch service times are assumed to be independent of the batch size, and IID. Using a Markov decision theoretic approach, Deb and Serfozo show that a threshold policy (which is the same as that of Neut's) minimizes the average queue length. Deb also considers the problem, of minimizing the sum of the average queue length and the average switching cost, when non-negative costs are incurred in switching the server on and off, in [11], wherein the optimal policy is shown to be a similar threshold policy. Control problems for batch service queues, with service times that were IID and independent of the batch size, were also considered by Powell and Humblet [31], Weiss [44], and Aalto [3],[2]. Gallisch [17] considers a continuous time M/G/1 queueing system, where the service time distribution of each customer can be chosen by the scheduler at the time of service initiation from a given set (which is stochastically ordered). Using a Markov decision theoretic approach, Gallisch obtains that the optimal policy which minimizes the average queue length, chooses service time distributions which are monotone stochastically lesser as a function of the queue length. We note that in most of the prior work on optimal control of batch service queues for minimizing the average queue length, either the batch service time was independent of the batch size or only the service time was controllable. The model that we analyze in this paper is a generalization of the models considered by [12], [31], [44], [3], and [2], as we consider dynamic control of batch service queues, where the batch service time is a function of the service batch size. Musy and Telatar in [24] consider a continuous time batch service queue with Poisson arrivals. They assume that the server can serve any number of customers in a batch, but the batch service time is an affine function of the batch size. Upper and lower bounds to the minimum average queue are obtained. The model that we consider is a discrete time version of the model considered by Musy and Telatar. Thus our study would be of independent interest and the results that we have obtained can be applied to other applications too.

B. Outline of the paper and contributions

We discuss the discrete time queueing model with batch-service size dependent batch-service time for the point-to-point channel with per-transmission reliability in Section II. The average delay minimization problem is also discussed in Section II for the queueing model. We also discuss the average delay optimal policies for two simple forms of $\tau(s)$ in the same section. The forms of $\tau(s)$ that arise in point-to-point channels for block coding schemes are also discussed. In Section III, we show that for the problem considered in this paper, it is enough to consider stationary deterministic policies. We show this by formulating the problem as a semi-Markov decision problem. We give new sufficient conditions for the existence of the average cost optimality equation and the existence of a stationary deterministic policy. We then propose heuristic stationary deterministic policies for point-to-point channels in Sections IV and V. We obtain upper and lower bounds on the performance of these heuristic policies and compare their performance with the optimal policy in the same sections. Finally, we consider the fundamental tradeoff of average error probability, average power, and average delay for the case of the point-to-point channel with variable length coding in Section VI.

In our view, the main contributions of this paper can be summarized as follows:

1) we obtain that a policy that serves exhaustively at a transmission instant is close to optimal for a high reliability regime,
2) we obtain analytical performance bounds for the above policy,
3) our results are applied to characterize the fundamental tradeoff of average probability of error and average delay for non-fading noisy point-to-point channels.

C. Notation and conventions

The set of all real numbers is denoted by $\mathbb{R}$, the set of all integers by $\mathbb{Z}$, the set of non-negative integers by $\mathbb{R}_+$, and the set of all non-negative integers by $\mathbb{Z}_+$. The probability of an event E is denoted as $\Pr\{E\}$. We use capital letters to denote random variables, e.g., X is a random variable. We denote the expectation of a random variable X by $\mathbb{E}X$ and the variance by $var(X)$. If the expectation is taken with respect to a particular distribution π or the distribution induced by a policy γ, then that is denoted by $\mathbb{E}_\pi$ or $\mathbb{E}_\gamma$ respectively. We say that $X \sim Y$ and $X \sim f$ to denote that X and Y have the same distribution and X has the distribution f respectively. We note that a Binomial arrival process is a discrete time IID random process in which at every discrete time the random variable is a Binomial random variable with standard parameters n and p. The rate of the arrival process is $\lambda = np$.

II. System model and the average delay minimization problem

We assume that the system evolves in discrete time with the slots indexed by $n \in \mathbb{Z}_+$. We assume that the message symbols arrive in a random process into the transmitter and are queued in an infinite length buffer. The random number of message symbols that arrive into the transmitter in the n^{th} slot is denoted as $A_s[n]$. We assume that these arrivals happen just before the end of the n^{th} slot. We assume that $A_s[n] \in \{0, \dots, A_{max}\}$, with $\mathbb{E}A_s[1] = \lambda$ and $var(A_s[1]) = \sigma^2$. For technical reasons[1] we also assume that $\Pr\{A_s[1] = a\} > 0$ for all $a \in \{0, \dots, A_{max}\}$. We also assume that the arrival random process $(A_s[n])$ is independently and identically distributed (IID).

A slot in which there is no transmission on the channel and the transmitter is idle is called an idle slot. The transmitter scheduler makes decisions about the number of message symbols to be transmitted at the beginning of a slot just after every end of every transmission or at the beginning of a slot just after an idle slot. These epochs are called decision epochs. At a decision epoch, the transmitter may choose to either: (i) idle for one slot (batch size $s = 0$), or (ii) choose to encode and transmit s message symbols in a batch taking a transmission time of $\tau(s)$ slots, if the number of message symbols in the queue is at least s (here s is a positive integer).

We define the state of the system as the number of message symbols in the transmitter queue at every decision epoch. This queue length state is denoted by $Q[m]$, where $m \in \mathbb{Z}_+$ indexes the decision epochs. The size of the batch of message symbols, which starts transmission at the m^{th} decision epoch, is denoted by $S[m]$. Then, the queue length $Q[m]$ at the decision epochs m evolves as follows for every m (see Figure 1).

$$Q[m+1] = Q[m] - S[m] + A[m], \tag{1}$$

where the batch size $S[m] \leq Q[m]$. We note that when the queue length is q, the set of batch sizes which can be used for service is denoted by $\mathcal{S}(q) = \{s : s \in \mathbb{Z}_+, s \leq q\}$. We note that the queue length in every slot n, denoted by $Q_s[n]$ evolves in the following manner. Whenever n is a decision epoch, e.g., the m^{th} decision epoch, then $Q_s[n] = Q[m]$. In between decision epochs, $Q_s[n]$ is non-decreasing since there are arrivals into the system. We note that the state evolution is embedded at the decision epochs in the queue length evolution process $Q_s[n]$. We note that $S[m]$ symbols take $\tau(S[m])$ slots for transmission after which we have the next decision epoch. The cumulative number of message symbols which arrive in these $\tau(S[m])$ slots is denoted as $A[m]$.

Suppose the initial number of message symbols in the queue be denoted by q_0; then $Q[0] = Q_s[0] = q_0$. A service policy γ is then defined as a sequence of batch sizes $(S[0], S[1], S[2], \cdots)$ where $S[m]$ could be chosen dependent on the history $(Q[0], S[0], \cdots, Q[m-1], S[m-1], Q[m])$. The set of all policies is denoted by Γ. We assume that any $\gamma \in \Gamma$ transmits message symbols in first-in first-out (FIFO) order. A stationary deterministic policy is such that $S[m]$ is a deterministic function $s(.)$ of $Q[m]$; i.e. $S[m] = s(Q[m])$. A stationary deterministic policy allows for a simpler specification of the operating policy compared to a general service policy. The set of stationary deterministic policies is denoted by Γ_s. We note that the process $Q[m]$ under a policy $\gamma \in \Gamma_s$ is a Markov chain with state space $\mathbb{Z}_+$. The evolution of the embedded Markov chain $(Q[m], m \geq 0)$ is given by (1) and the transition time from the m^{th} to the $(m+1)^{th}$ embedded epoch by $\tau(S[m])$. We note that the transition probabilities are determined by the service policy and the arrival statistics. The evolution of the queue is illustrated in Figure 1.

From Little's law the average queue length for a policy γ is proportional to the average delay for a fixed arrival rate. Therefore, we consider the average queue length for a policy $\gamma \in \Gamma$. We note that the average queue length can be expressed in two ways as in Ross [32]. The time average queue length

[1]We need these assumptions for irreducibility and aperiodicity properties to hold for a Markov chain model for the system.

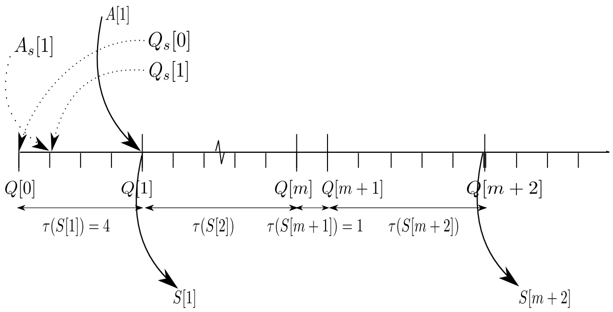

Fig. 1: Evolution of the queue state for an example policy γ

for a policy $\gamma \in \Gamma$ starting in the state q_0 is defined as

$$\phi^{\gamma}(q_0) = \limsup_{N \to \infty} \mathbb{E}_{\gamma} \left[\frac{\sum_{n=0}^{N-1} Q_s[n]}{N} \middle| Q_s[0] = q_0 \right], \quad (2)$$

where $\mathbb{E}_{\gamma}$ is used to denote the expectation with respect to the joint probability distribution of $(Q_s[n], n \geq 0)$ under γ. It is also possible to define the average queue length in terms of the embedded process $Q[m]$. For this we define the following cost function:

$$h(q, s) = q\tau(s) + \frac{1}{2}\lambda\tau(s)(\tau(s) - 1). \quad (3)$$

The function $h(Q[m], S[m])$ is the holding cost incurred at each epoch m of $(Q[m], m \geq 0)$. The holding cost $h(q, s)$ is the cumulative expected queue length during a period between two consecutive decision epochs, conditioned on the queue length being q at the initial decision epoch and s being the batch size chosen for transmission. For any policy γ an alternative definition of the average queue length is:

$$g^{\gamma}(q_0) = \limsup_{M \to \infty} \left[\frac{\mathbb{E}_{\gamma} \sum_{m=0}^{M-1} h(Q[m], S[m])}{\mathbb{E}_{\gamma} \sum_{m=0}^{M-1} \tau(S[m])} \middle| Q(0) = q_0 \right]. \quad (4)$$

In this paper, we find that for the average queue length minimization problem the policies of interest are such that $\phi^{\gamma}(q_0) = g^{\gamma}(q_0)$ (see Section III).

The average queue length minimization problem that we study in this paper is

$$\min_{\gamma \in \Gamma} \phi^{\gamma}(q_0), \forall q_0. \quad (5)$$

Any policy that is optimal for the above problem is denoted as γ^*.

We now discuss the tradeoffs that are present in the above average queue length (or equivalently average delay) minimization problem. We note that the delay that a message symbol experiences in the system is composed of the waiting time in the system followed by the batch service time or the transmission time. The waiting time in the system can be further subdivided as: (i) the waiting time from the time of arrival to the first service opportunity; the first service opportunity is the next decision epoch, and (ii) the waiting time if the symbol is not transmitted at the first service opportunity. An optimal policy should minimize the sum of these three components. However, we observe that there is a tradeoff that exists between these components depending on the form of $\tau(s)$.

For some simple forms of $\tau(s)$ the optimal policy and the minimum average queue length can be easily obtained since the tradeoff between the delay components discussed above becomes trivial. We discuss two such forms here.

a) The case where $\tau(s) = 1$: Of the three components of delay, we note that there is no change in the transmission time as a function of the batch size s and the waiting time till the first service epoch is always 1. Hence, we can increase the service rate in order to minimize the waiting time from the first service opportunity. We define the policy γ_{exh} such that $S[m] = Q[m]$; i.e., the policy exhaustively serves all the customers in the system in every slot. Let $\gamma \in \Gamma$ be any other policy. We consider the evolution of the system under γ_{exh} and γ, with the same arrival sample path and initial condition q_0. We see that the queue length $Q^{\gamma_{exh}}[m] = Q_s^{\gamma_{exh}}[n]$ is always less than or equal to $Q_s^{\gamma}[n]$. The minimum average delay is 1 and the minimum average queue length is λ for any arrival rate λ. We note that the $\tau(s) = 1$ form is equivalent to the assumption that any rate within the capacity region can be achieved with a large enough slot time.

b) The case where $\tau(s) = a.s$, for any positive integer a and $\tau(0) = 1$: Suppose we consider a service epoch m at which there are q message symbols in the queue. We define a policy γ_1 which is such that $S[m] = 1$ if $Q[m] > 0$ and $S[m] = 0$ if $Q[m] = 0$. Consider any other policy γ. Suppose at a decision epoch, we have q message symbols in the queue. We note that under FIFO service, the total time to service the q message symbols is aq for both γ and γ_1. For γ_1 we observe that the delay of the message symbols are $a, 2a, \ldots, qa$ which is less than the delay experienced by the message symbols under γ which uses batches of size more than one. Therefore, γ_1 has the minimum average delay. We note that the minimum average delay is finite for any $\lambda < \frac{1}{a}$ (this can be shown using standard Lyapunov drift arguments, but the intuitive reason is that the maximum service rate is $\frac{1}{a}$). We discuss an approximation to the minimum average delay and queue length for the above policy in Appendix II. We note that in this case the transmission time as well as the waiting time till the first epoch are function of the service batch sizes. However, for whatever non-zero s that we chose, the service rate is constant and is $\frac{1}{a}$ so that the waiting time after the first service opportunity may be same. So intuitively the optimal policy should have the smallest transmission time as well as the waiting time till the first service opportunity which is what γ_1 achieves.

We note that for both forms of $\tau(s)$ discussed above, the tradeoffs between the different components of delay were direct leading to complete characterizations of the respective optimal policies. For the point-to-point channel with per-transmission reliability requirements, we find that the codeword transmission time $\tau(s)$ is not one of the forms discussed above. The tradeoffs involved between the components of delay are then non-trivial. We now discuss the forms of $\tau(s)$ for the point-to-point channel. These forms are obtained using

Gallager's random coding bound [16] and Polyanskiy's normal approximation [30].

A. Batch service time $\tau(s)$ for point-to-point channels

In this paper we study specific forms of $\tau(s)$ which have been motivated by coding schemes for point-to-point channels. We assume that each message symbol has an average error probability requirement $P_{e,r}$. We recall that the coding schemes considered in this paper are assumed to jointly encode s message symbols into a block codeword of length τ slots. Suppose for such a coding scheme, the average codeword error probability is denoted as $P(s,\tau)$ which is a function of s and τ. We guarantee that the reliability requirement is met by ensuring that $P(s,\tau) \leq P_{e,r}$, for every transmission, since then every message symbol transmitted has an average error probability of at most $P_{e,r}$. To ensure $P(s,\tau) \leq P_{e,r}$, it is intuitive that the codeword length τ has to be chosen as a function $\tau(s)$ of the batch size s. An analytical upper bound for $\tau(s)$ can be obtained by using Gallager's random coding upper bound on $P(s,\tau)$. It turns out that $\tau(s) = \lceil as + b \rceil, s > 0$, where a and b are positive constants. A complete discussion containing the formulae for a and b is presented in Appendix I. We also note the above affine form for $\tau(s)$ also arises in the case if a convolutional code with memory $K = b$ is used as a block code (with terminating zeros).

We note that the Gallager random coding bound is tight (in the exponent) only for message symbol arrival rates λ close to the capacity of the channel and when large block lengths are used (i.e. when $\tau(s)$ is large). Motivated by the normal approximation for the codeword error probability [30], an approximation for the function $\tau(s)$, under the assumption that each message symbol requires an average probability of error guarantee of $P_{e,r}$ can be obtained. For any coding scheme that transmits s message symbols using a block length of $\tau(s)$ and with an average error probability $\leq P_{e,r}$, we have from [30] that $\tau(s) = \lceil cs + d + \sqrt{es + f} \rceil$ for $s > 0$. A complete discussion can again be found in Appendix I where the formulae for c, d, e, and f are given.

For these forms of $\tau(s)$ which arise for point-to-point channel coding schemes, the derivation of average delay optimal policies are not direct, as in the case of $\tau(s) = 1$ and $\tau(s) = as$ discussed earlier. For example, for $\tau(s) = as + b$ we find that an increase in s leads to an increase in the rate as well as the transmission time, leading to a tradeoff between transmission time and the two components of the waiting time. Therefore, for general forms of $\tau(s)$, we formulate the problem as a semi-Markov decision process in order to obtain operating policies and approximations for performance.

III. Optimality of stationary deterministic policies for minimizing average queue length

In this section, we simplify the minimization problem in (5) by showing that there exists an optimal policy γ^* in the set of stationary deterministic policies. If such an optimal policy γ^* exists, then (5) reduces to the simpler problem

$$\min_{\gamma \in \Gamma_s} \phi^\gamma(q_0), \forall q_0, \tag{6}$$

where we only optimize over only stationary deterministic policies, i.e., the set Γ_s.

The optimality of stationary deterministic policies for (5) is shown by formulating (5) as a semi-Markov decision theoretic problem (SMDP) and then showing that there exists optimal stationary deterministic policies for the SMDP. In Theorem III.1 we present the main result that will be used subsequently and present a complete discussion in Appendix VI.

Theorem III.1. *Suppose $r_{max} = \lim_{s\to\infty} \frac{s}{\tau(s)}$ is the maximum service rate. If the arrival rate $\lambda < r_{max}$ then*

1) *the optimal or minimum values of* (2) *and* (4) *denoted as $\phi^*(q_0)$ and $g^*(q_0)$ are finite, equal, and do not depend on q_0. This finite minimum value is called the optimal average queue length and denoted as g^*,*
2) *there exists an optimal stationary deterministic policy γ^*. The optimal policy γ^* prescribes an optimal batch size $s^*(q)$ as a deterministic function of q. The optimal batch size $s^*(q)$ is the minimizer for the following average cost optimality equation (ACOE):*

$$\begin{aligned} J(q) &= \min_{s\in\mathcal{S}(q)} \left\{ \left(q + \frac{(\tau(s)-1)\lambda}{2} - g^*\right)\tau(s) + \right. \\ &\qquad \left. \mathbb{E}J(q - s + A^{\tau(s)})\right\}, \forall q \in \mathbb{Z}_+, \end{aligned} \tag{7}$$

where $J(q)$ is the optimal relative value function.

The proof of this theorem can be found in Appendix VI. In Appendix VI, we also present new sufficient conditions for the existence of the above average cost optimality equation for the SMDP. These sufficient conditions are required since prior work [37] considers the case where $\tau(s)$ is bounded, while in general the transmission time $\tau(s)$ is an unbounded function of the batch size s.

We consider systems where $\lambda < r_{max}$ for which we have an optimal stationary deterministic policy which satisfies the ACOE. We note that as the optimal policy satisfies the ACOE, it satisfies the c-regularity property [22] (see Appendix III), with the c function being the function $h(q, s^*(q))$ where $s^*(q)$ is the optimal policy. From Appendix III, the optimal policy then belongs to the set of stationary deterministic policies which (a) are positive recurrent, (b) have finite average queue length, and (c) have that $\phi^\gamma(q_0) = g^\gamma(q_0), \forall q_0 \in \mathbb{Z}_+$. Therefore, we can consider a simpler problem

$$\min_{\gamma \in \Gamma_s} \phi^\gamma, \tag{8}$$

where the initial state q_0 has been dropped from the notation since there is no dependence on the initial state, and we redefine Γ_s to be the set of stationary deterministic policies which have the properties stated above. The optimal value of the above problem, is the minimum average queue length g^*.

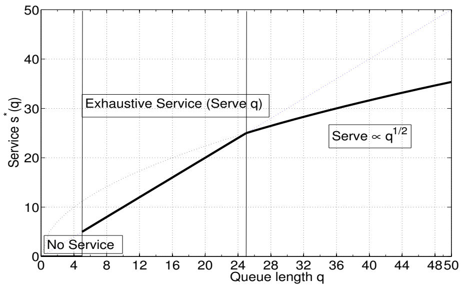

Fig. 2: General structure of optimal policies for affine $\tau(s)$ with $b \gg a$.

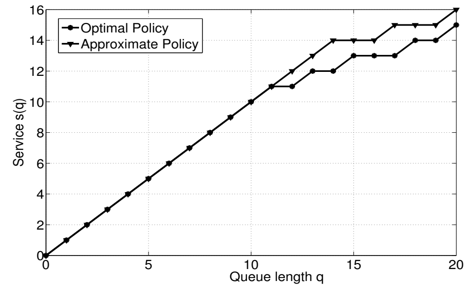

Fig. 3: Average-delay optimal policy and approximate policy for binomial arrival process with $\lambda = 0.19, a = 5, b = 1, n_s = 5$

In the following, we consider (8) for $\tau(s)$ obtained from Gallager's random coding bound and Polyanskiy's approximation using this SMDP formulation. For each form of $\tau(s)$ we propose approximately optimal stationary deterministic policies whose performance is studied through analytical bounds as well as simulations.

IV. Affine case where $\tau(s) = \lceil as + b \rceil ; s > 0$ and $\tau(0) = 1$

We recall that the *affine* $\tau(s)$ case where $\tau(s) = \lceil as + b \rceil ; s > 0$ and $\tau(0) = 1$ arises when we use Gallager's random coding bound to model the relationship between reliability and codeword length or if we use a convolutional coder.

In contrast to the usual approach in applying MDPs where structural properties of the optimal policy are derived using the optimality equations [20], for this affine form of $\tau(s)$ we do not have an analytical characterization of the structure of the optimal policy or that of the relative value function $J(q)$ from the SMDP analysis. We start with a numerical solution of the semi-Markov decision process using policy iteration [40] with a suitably truncated state space. We observe that in general the optimal policy for affine $\tau(s)$ exhibits the behaviour shown in Figure 2, for several choices of the parameters a, b, and arrival statistics. We note that since we are interested in the case where the reliability is high, the parameter choices are such that b is chosen to be much greater than a (see Appendix I). We observe three distinct regions of queue length values with different behaviours for the stationary deterministic optimal policy. For queue length q less than an initial threshold, the optimal policy does not serve. For queue length q greater than another larger threshold, the optimal policy uses a service batch size which is proportional to $\sqrt{q}$. In between these two thresholds the service is exhaustive. We note that in general the thresholds defining the behaviour of policy would be different from what is shown in Figure 2 and are dependent on the parameters of the system. Since we are not able to solve (7) for all q we solve (7) in the asymptotic regime as $q \to \infty$.

A. Asymptotic analysis of the ACOE (eq. (7)) for affine $\tau(s)$

We note that the utility of the following asymptotic analysis of the ACOE is that: (a) it leads to an approximate policy for the affine $\tau(s)$ case with performance close to that of optimal and (b) it provides an analytical reason for the $\sqrt{q}$ behaviour of the optimal policy observed in Figure 2. Before we go on to the derivation of the approximate policy, we give an example in Figure 3 where we compare the approximate policy with the optimal policy obtained from policy iteration. In Figure 3 we consider a system with $a = 5, b = 1$, and binomial arrival process with $\lambda = 0.19$ and batch size $n_s = 5$. We note the similarity in the computed optimal policy and approximate policy. We note that we choose $a > b$ in this example so that the difference between the approximate and optimal policy can be emphasized compared to cases for $b \gg a$. The policy iteration is done for state space truncated at 1000.

In this section, for our analysis we use the simple affine approximation $as + b$ for $\lceil as + b \rceil$ (i.e., without the ceiling operation). We solve the ACOE in an asymptotic regime as follows. We first relax the integer constraints on the number of symbols s and the queue length q. Then, we note that the minimization in (7) is over the interval $[0, q]$. Suppose there exists a function $J_a(q) : \mathbb{R}_+ \to \mathbb{R}_+$ and a non-negative real number g_a such that

$$\lim_{q \to \infty} \left| \min_{s \in [0,q]} \left\{ q\tau(s) + \frac{\tau(s)(\tau(s)-1)\lambda}{2} - g_a \tau(s) + \mathbb{E} J_a(q - s + A^{\tau(s)}) \right\} - J_a(q) \right| = 0. \tag{9}$$

Then, we define $J_a(q)$ and g_a to be an asymptotic solution to the ACOE. It is natural to ask whether obtaining any $J_a(q)$ and g_a would be beneficial in solving the original problem, since the asymptotic solution *solves* the ACOE only for large values of q and in steady state the evolution of the queue length may not reach such large values. As discussed previously, such a solution (see Figure 3) can be used to derive approximate policies with average queue length which are close to optimal.

An approach to obtaining a solution $(J_a(q), g_a)$ to (9) is to verify that a candidate pair $(J_a(q), g_a)$ satisfies (9). Here we use a candidate function $J_a(q)$ which is of the form of a linear combination of basis functions, i.e., $\sum_{i=1}^{n} c_i f_i(q)$ where $f_i(q)$ are basis functions which are chosen and c_i are the coefficients which have to be determined.

To determine what $f_i(q)$ should be, we will first do a preliminary analysis using two basis functions - q^2 and q^η, where

η will be determined in the following. We first differentiate the expression within the minimization in (9) in order to obtain the minimizing s. We assume that for sufficiently large q the minimizing s is non-zero and approximate $\tau(s)$ by $as+b$. Then we obtain that for large q, the minimizing s denoted by s^* has the form given by $k_1 q + k_2 q^{\eta-1}$, where $k_1 = (-2a+4c_1-4ac_1\lambda)$. From Figure 2 and other numerical evaluations we observe that for large q, $\frac{s^*(q)}{q} \to 0$. We note that this holds only if $k_1 = 0$ which implies that $c_1 = \frac{a}{2(1-a\lambda)}$. With this expression for c_1 it is seen that $s^* = \frac{(1-a\lambda)\eta c_2}{a} q^{\eta-1}$. On substituting $J_a(q) = c_1 q^2 + c_2 q^\eta$ and $s^* = \frac{(1-a\lambda)\eta c_2}{a} q^{\eta-1}$ we find that the coefficients of q^η (for $1 < \eta < 2$) on the LHS of (9) is $ak + c_2$, where $k = \frac{(1-a\lambda)\eta c_2}{a}$. We also find that there are terms which contain the $2\eta - 2$ power of q. Suppose $2\eta - 2 \neq 1$, then the coefficient of $q^{2\eta-2}$ in the LHS of (9) is

$$\frac{\lambda}{2}(ak)^2 + c_1(1-a\lambda)^2 k^2 - k(1-a\lambda)\eta c_\eta,$$

which should be zero for $J_a(q)$ to be an asymptotic solution. However, this can be simplified to $0 = -\frac{1}{2}$. This implies that our assumption $2\eta - 2 \neq 1$ is false. So $2\eta - 2 = 1$, or $\eta = \frac{3}{2}$. Then we have that

$$\begin{aligned} 0 &= b + \frac{\lambda}{2}(ak)^2 + \frac{a}{2}(1-a\lambda)k^2 + 2b\lambda c_1 - ak^2 \\ \frac{ak^2}{2} &= b(1+2\lambda c_1) = \frac{b}{1-a\lambda} \end{aligned}$$

which yields $c_2 = \frac{2\sqrt{2}\sqrt{ab}}{3(1-a\lambda)^{\frac{3}{2}}}$.

We now use $\eta = \frac{3}{2}$ obtained above for a new form for $J_a(q)$. We assume that $J_a(q) = c_1 q^2 + c_2 q^{\frac{3}{2}} + c_3 q + c_4 q^{\frac{1}{2}}$ as a candidate solution for (9). The additional basis functions q and $q^{\frac{1}{2}}$ are needed in order that (9) is satisfied. Substituting $J_a(q)$ in (9) and under the assumption that for large enough q, $s^*(q) \neq 0$, we have that

$$\begin{aligned} 0 = \min_{s\in[0,q]} \Bigg\{ & q(as+b) + \frac{(as+b)(as+b-1)\lambda}{2} \\ & -g(as+b) + c_1((q-s)^2 + \mathbb{E}(A^{\tau(s)})^2 + 2\mathbb{E}A^{\tau(s)}(q-s)) \\ & + c_2\Bigg(q^{\frac{3}{2}} - \frac{3}{2} q^{\frac{1}{2}} (s - \mathbb{E}A^{\tau(s)} \\ & + \frac{3}{8} q^{\frac{-1}{2}} (s^2 - 2s\mathbb{E}A^{\tau(s)} + \mathbb{E}(A^{\tau(s)})^2) \\ & + \frac{1}{16} q^{\frac{-3}{2}} (s^3 - 3s^2\mathbb{E}A^{\tau(s)} + 3s\mathbb{E}(A^{\tau(s)})^2 - \mathbb{E}(A^{\tau(s)})^3)) \Bigg) \\ & + c_3(q - s + \mathbb{E}A^{\tau(s)}) + c_4\Bigg(q^{\frac{1}{2}} - \frac{1}{2} q^{\frac{-1}{2}} (s - \mathbb{E}A^{\tau(s)}) \\ & - \frac{1}{8} q^{\frac{-3}{2}} (s^2 + \mathbb{E}(A^{\tau(s)})^2 - 2s\mathbb{E}A^{\tau(s)}) \Bigg) \Bigg\} \\ & -(c_1 q^2 + c_2 q^{\frac{3}{2}} + c_3 q + c_4 q^{\frac{1}{2}}) \end{aligned}$$

Differentiating the term within the minimization with respect to s, we obtain that for large q, the minimizer $s^*(q) = k_1\sqrt{q} + \frac{k_2}{\sqrt{q}} + \frac{k_3}{q} + \frac{k_4}{q^{\frac{3}{2}}} + k_5$, where $k_1 = \frac{3c_2(1-a\lambda)}{2a}$ and the rest are constants but with more complex expressions. Note that $c_1 = \frac{a}{2(1-a\lambda)}$ for $s^*(q)$ to have this form.

After substituting this $s^*(q)$ into the above equation we seek an asymptotic solution. The asymptotic solution is obtained by a choice of c_2, c_3, c_4, and g such that the coefficients of all non-negative powers of q in the above equation is zero. The non-negative powers are $1, \frac{1}{2}$ and 0. Thus we obtain that

$$c_2 = \frac{2\sqrt{2}\sqrt{ab}}{3(1-a\lambda)^{\frac{3}{2}}}$$

From forcing the coefficient of $q^{\frac{1}{2}}$ to zero, a linear relationship between c_3 and g is obtained. Also by forcing the constant to zero, a linear relationship between g and c_4 is obtained. We do not have unique solutions for c_3, c_4, and g. We choose any value for c_4, and then c_3 and g are determined in terms of c_4. Although we can obtain the exact expressions for c_3, c_4, and g, they are not presented here since for the purposes of obtaining an approximate policy, we use only c_1 and c_2 which have been uniquely defined.

We define the approximate unconstrained minimizer as $\sqrt{\frac{2b}{a(1-a\lambda)}}\sqrt{q}$. We then define the approximate policy (denoted as Approx.) as the stationary deterministic policy with batch service size

$$s_{approx}(q) = \begin{cases} q & \text{if } q \le \frac{2b}{a(1-a\lambda)} \\ \left\lfloor \sqrt{\frac{2b}{a(1-a\lambda)}}\sqrt{q} \right\rfloor & \text{otherwise} \end{cases} \quad (10)$$

We note that here q as well as the batch size $s(q)$ is integer valued, as this is a policy for our original system.

B. Heuristic policies

We propose two heuristic operating policies in this section. The first policy, exhaustive service policy or EXH, uses $s(q) = q$ for every $q \in \mathbb{Z}_+$. We note that for the point-to-point channel context, for small values of $P_{e,r}$ we observe that $b \gg a$. Then, $\tau(s)$ can be approximated as a constant or $\tau(s) = T$. The EXH policy is then motivated from our analysis for $\tau(s) = T$ in [43]. In [43] we obtained that a threshold policy that serves exhaustively only if the queue length is over a threshold was optimal. The exhaustive service policy is a threshold policy with a threshold of 0 and was found to have performance comparable with that of the optimal threshold policy.

We also define a heuristic policy QAP by considering a uniformized version of the system. The uniformized system is obtained using the uniformization technique from [40] elaborated in Appendix IV. From Appendix IV we also have a Markov decision process (MDP) associated with the uniformized system. This MDP has the ACOE:

$$\begin{aligned} \tilde{J}(q) = \min_{s\in\{0,1,\ldots,q\}} & \Big\{ q + \frac{(\tau(s)-1)\lambda}{2} - g^* + \\ & \mathbb{E}\tilde{J}\big(q - s + \tilde{A}^{\tau(s)}\big) \Big\}, \end{aligned}$$

where $\tilde{J}(q)$ is the optimal relative value function for the uniformized MDP, g^* is the optimal minimum cost, and $\tilde{A}^{\tau(s)}$ is the arrival random variable for the uniformized MDP. From

Appendix IV, we note that $\tilde{A}^{\tau(s)}$ is $A^{\tau(s)}$ with probability $\frac{1}{\tau(s)}$ and it is s with probability $1-\frac{1}{\tau(s)}$.

We have the following characterization for $\tilde{J}(q)$.

Lemma IV.1. *The optimal relative value function $\tilde{J}(q) = \mathcal{O}(q^2)$.*

The proof is given in Appendix VII-A. We note that this result holds for any general $\tau(s)$ and not just the affine form of $\tau(s)$ as discussed above.

The heuristic policy that we propose is based on a greedy minimization of the ACOE with the above quadratic approximation for $\tilde{J}(q)$. We note that finding out the minimizer in the ACOE above with $\tilde{J}(q) = c_1 q^2$ is equivalent to

$$\underset{s\in\{0,1,\dots,q\}}{\arg\min} \left\{\frac{(\tau(s)-1)\lambda}{2} + c_1\mathbb{E}\big(q-s+\tilde{A}^{\tau(s)}\big)^2\right\}.$$

By expanding out the square inside the expectation and substituting $\mathbb{E}\big(\tilde{A}^{\tau(s)}\big)^2 = \sigma^2 + \lambda^2\tau(s) + \big(1-\frac{1}{\tau(s)}\big)s^2$ and $\mathbb{E}\tilde{A}^{\tau(s)} = \lambda + s - \frac{s}{\tau(s)}$, we have that

$$\underset{s\in\{0,1,\dots,q\}}{\arg\min} \left\{\frac{c_1 s^2}{\tau(s)} - 2c_1 s\Big(\frac{q}{\tau(s)}+\lambda\Big) + c_1\lambda^2\tau(s) + \frac{\lambda\tau(s)}{2}\right\}.$$

We propose the heuristic quadratic approximation policy (QAP) as the function $s_{qap}(q)$ which is the batch service size that achieves the minimum in the above optimization problem for every q, with $c_1 = \frac{a}{2(1-a\lambda)}$ where $a = \lim_{s\uparrow\infty}\frac{\tau(s)}{s}$. We note that this policy can be computed offline for any general form of $\tau(.)$ using numerical methods after approximating s as well as q as being real valued.

C. Performance analysis

In this section, we present upper and lower bounds on the optimal average queue length g^*. These bounds are helpful in ascertaining whether the policies that we have proposed, i.e., Approx., EXH, and QAP have average queue length close to that of the minimum.

Proposition IV.2. *The minimum average queue length $g^* \geq \max\left\{\frac{b\lambda}{2(1-a\lambda)}(2+a\lambda), \frac{a\sigma^2}{2(1-a\lambda)} + \frac{a\lambda b\lambda}{2(1-a\lambda)}\right\}$.*

The proof is given in Appendix VII-B. We now present an upper bound to the minimum average queue length for the system with $\tau(0) = \lceil b \rceil$. We note that an upper bound for the system with $\tau(0) = \lceil b \rceil$ is a upper bound for the system with $\tau(0) = 1$ since the following comparison result holds.

Lemma IV.3. *Suppose there are two queueing systems (modelled as in Section II) denoted as A and B with batch service time functions given by $\tau_a(s)$ and $\tau_b(s)$ such that $\tau_a(s) \leq \tau_b(s), \forall s$. Suppose g_A^* and g_B^* are the optimal average queue lengths for A and B respectively. Then $g_A^* \leq g_B^*$.*

The proof is given in Appendix VII-C.

We note that an upper bound for the $\tau(0) = \lceil b \rceil$ system can be obtained by obtaining the average queue length for any policy. Therefore, we consider the EXH policy with $s(q) = q$. The special form of $s(q)$ enables us to easily obtain an analytical upper bound on the average queue length of the exhaustive policy, which leads to the following result.

Proposition IV.4. *The minimum average queue length $g^* \leq (b+1)\left(\frac{a\sigma^2}{2(1-a\lambda)} + \frac{3(b+1)\lambda}{2(1-a\lambda)} + \frac{a\sigma^2}{2(1-a^2\lambda^2)} - \frac{\lambda}{2}\right)$. For sufficiently large λ, the minimum average queue length is approximately bounded above by*

$$\left(\frac{b+1}{b}\right)\left(\frac{a\sigma^2}{2(1-a\lambda)} + \frac{3(b+1)\lambda}{2(1-a\lambda)} + \frac{a\sigma^2}{2(1-a^2\lambda^2)} - \frac{\lambda}{2}\right).$$

The proof can be found in Appendix VII-D.

We recall that Musy and Telatar [24] had considered a continuous time queueing model with Poisson arrivals and affine service times; i.e., if a batch size $s > 0$ is used, then the service time is $D + ks$ (we note that there is no ceiling applied to the service time since it is a continuous time model and $b = D$ and $a = k$ in our notation). Under an asymptotic regime where the slot durations are made infinitesimally small, we show that our bounds are the same as those in [24] in Appendix VII-E. We note that the performance bounds that we have derived hold for any IID batch arrival process for discrete time systems.

D. Numerical comparison of policies for affine $\tau(s)$

In this section we numerically evaluate the average queue length (or equivalently average delay for a fixed arrival rate) performance of the EXH, Approx., and QAP policies. We also use the performance bounds which were derived in the previous section for evaluating the performance of these policies. In Figure 4 we consider a representative case where $a = 1$ and $b = 1$ and compare the average queue length of EXH, QAP, and Approx. policies (which are obtained by simulation) with that of an (approximately) optimal policy obtained using policy iteration[2]. From Figure 4, we observe that the average queue length for the three proposed policies are essentially the same, and essentially the same as that of optimal policy. In Figure 5 we compare the average queue length of EXH, Approx., QAP for a system where $a = 2$ and $b = 397.87$ respectively, which models the code block length for a binary symmetric channel (BSC) with a crossover probability $\epsilon = 0.1$ when the $P_{e,r}$ is 10^{-6} (see Table I in Appendix I). We plot the upper and lower bounds on the optimal average queue length in the same figure. We observe that again the average queue lengths of EXH, Approx., and QAP are similar and furthermore they closely match the upper and lower bounds. We therefore conclude that EXH, Approx. as well as QAP perform close to optimal in this case.

In Figures 6, 7, and 8 we compare the average queue length of EXH, Approx., and QAP for additive white Gaussian noise (AWGN) channels with different SNR values and $P_{e,r}$

[2]The optimal policy is only approximately characterized since policy iteration is carried out for a semi-Markov decision process defined on a truncated state space.

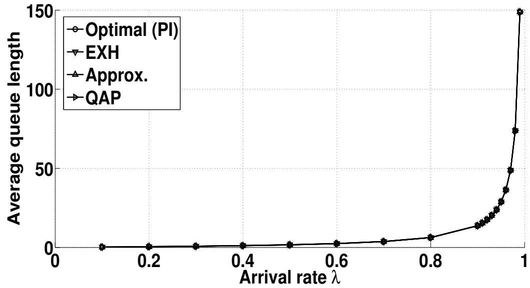

Fig. 4: Comparison of the average queue length for an optimal policy (computed using policy iteration), EXH, QAP, and Approx. policies the affine $\tau(s)$ case. The parameters a and b are 1 and 1 respectively.

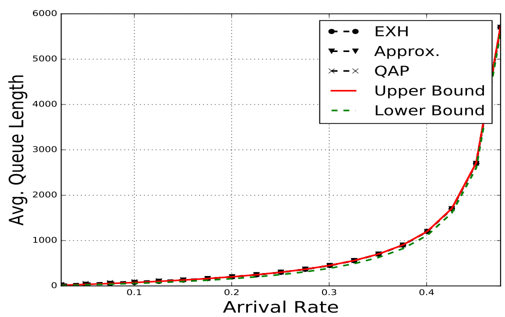

Fig. 5: Comparison of the average queue length for EXH, QAP, and Approx. policies for a BSC with $\epsilon = 0.1$ for the affine $\tau(s)$ case. The $P_{e,r}$ requirement is 10^{-6}. The parameters a and b are 2 and 397.87 respectively. The upper bound from Proposition (IV.4) and lower bound from Proposition (IV.2) are also shown.

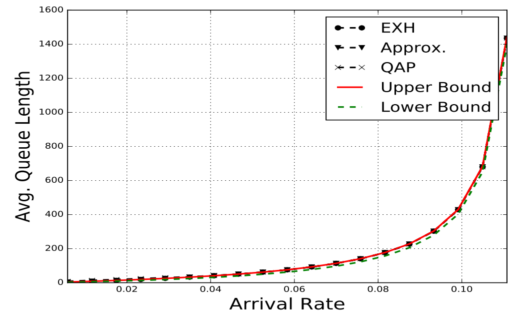

Fig. 6: Comparison of the average queue length for EXH, QAP, and Approx. policies for an AWGN channel with $SNR = 0$dB for the affine $\tau(s)$ case. The $P_{e,r}$ requirement is 10^{-6}. The parameters a and b are 8.58 and 427.31 respectively. The upper bound from Proposition (IV.4) and lower bound from Proposition (IV.2) are also shown.

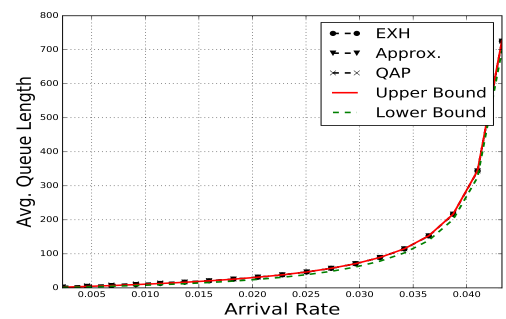

Fig. 7: Comparison of the average queue length for EXH, QAP, and Approx. policies for an AWGN channel with $SNR = -10$dB for the affine $\tau(s)$ case. The $P_{e,r}$ requirement is 10^{-3}. The parameters a and b are 21.94 and 546.74 respectively. The upper bound from Proposition (IV.4) and lower bound from Proposition (IV.2) are also shown.

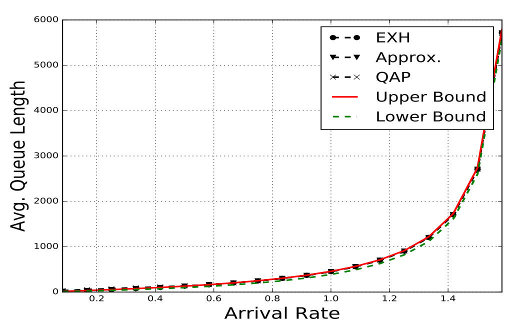

Fig. 8: Comparison of the average queue length for EXH, QAP, and Approx. policies for an AWGN channel with $SNR = 20$dB for the affine $\tau(s)$ case. The $P_{e,r}$ requirement is 10^{-6}. The parameters a and b are 0.60 and 119.53 respectively. The upper bound from Proposition (IV.4) and lower bound from Proposition (IV.2) are also shown.

requirements (see Table II in Appendix I). We again observe that the average queue lengths of EXH, Approx., and QAP are similar and furthermore they closely match the upper and lower bounds. We have observed similar behaviour in other simulations, in which $b \gg a$, which are not reported here. Therefore, we conclude that EXH, Approx., and QAP perform close to optimal in the case of affine $\tau(s)$ when $b \gg a$. We note that the case $b \gg a$ arises in the cases where $P_{e,r}$ is small. An interesting observation from these numerical results is that the EXH policy is the optimal policy for fixed slot duration models. So for the cases where the $P_{e,r}$ requirement is small, the question of the optimal policy seems to be answered by fixed slot size models. However, the average queue length and therefore delay values would be different compared to a fixed slot size model.

We consider two cases where $a \gg b$ in Figures 9 and 10. We observe that in this case the EXH policy has a higher average queue length than Approx. or QAP. It can be inferred that serving exhaustively at large values of queue length is not optimal since then the waiting time to first service opportunity for message symbols which arrive during a transmission time are larger. Therefore, the $\sqrt{q}$ nature of the optimal policy which is suggested in the Approx. policy gives us lower average queue length, since it minimizes the sum of the three components delay.

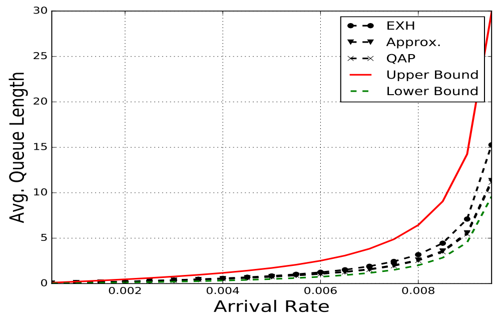

Fig. 9: Comparison of the average queue length for EXH, QAP, and Approx. policies for $a = 100$ and $b = 1$. The upper bound from Proposition (IV.4) and lower bound from Proposition (IV.2) are also shown.

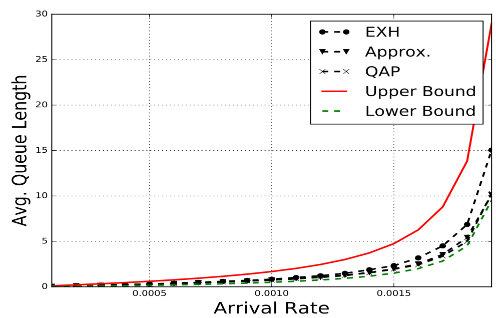

Fig. 10: Comparison of the average queue length for EXH, QAP, and Approx. policies for $a = 500$ and $b = 1$. The upper bound from Proposition (IV.4) and lower bound from Proposition (IV.2) are also shown.

V. The case where $\tau(s) = \lceil cs + d + \sqrt{es + f} \rceil ; s > 0$ and $\tau(0) = 1$

We note that for more general functional forms for $\tau(s)$, such as that obtained from Polyanskiy's approximation, the approach followed in the previous section for deriving the approximately optimal policy Approx. is even more tedious. For the case of $\tau(s) = \lceil cs + d + \sqrt{es + f} \rceil ; s > 0$ and $\tau(0) = 1$, which is obtained from Polyanksiy's approximation, we therefore propose to use just the EXH and QAP policy defined above as operating policies. We evaluate the performance of these policies for the Polyanskiy form of $\tau(s)$ in the following sections.

A. Performance analysis

We obtain an upper bound on the average queue length for the Polyanskiy form for $\tau(s)$ by first identifying a system (B) with $\tau_B(s) \geq cs + d + \sqrt{es + f}$ and τ_B being affine. We have from Lemma IV.3 that the average queue length for our system is bounded above by the upper bound for system (B). We use the upper bound from Proposition (IV.4) for system (B). Then, we have the following result.

Proposition V.1. *For any $s_0 \in \mathbb{Z}_+$, we define $\tilde{a} = c + \frac{e}{2\sqrt{f + es_0}}$ and $\tilde{b} = d + \frac{2f + es_0}{2\sqrt{f + es_0}}$. We also define $g_{ub,approx}(s_0)$ to be*

$$\left(\frac{\tilde{b}+1}{\tilde{b}}\right)\left(\frac{\tilde{a}\sigma^2}{2(1-\tilde{a}\lambda)} + \frac{3(\tilde{b}+1)\lambda}{2(1-\tilde{a}\lambda)} + \frac{\tilde{a}\sigma^2}{2(1-\tilde{a}^2\lambda^2)} - \frac{\lambda}{2}\right).$$

The minimum average queue length g^* has the approximate upper bound:

$$g^* \leq \min_{\{s_0 : g_{ub}(s_0) \geq 0\}} g_{ub,approx}(s_0).$$

The proof is given in Appendix VIII-A. We obtain a lower bound on the average queue length by identifying a system (A) with $\tau_A(s) \leq cs + d + \sqrt{es + f}$. We then derive a lower bound which is similar to that in Proposition (IV.2) and use Lemma IV.3 to obtain our desired lower bound.

Proposition V.2. *For any $s_0 \in \mathbb{Z}_+$ we define $\tau_A(s) = \min(\tau_1(s), \tau_2(s))$, where*

$$\begin{aligned} \tau_1(s) &= \left(c + \frac{\sqrt{es_0 + f} - \sqrt{f}}{s_0}\right)s + \left(d + \sqrt{f}\right), \\ \tau_2(s) &= cs + \left(d + \sqrt{es_0 + f}\right). \end{aligned}$$

Now for any s_0 such that $s_0/\tau_A(s_0) < \lambda$, we define $\tilde{C}$ as

$$\lambda\left(d + \sqrt{f} + \frac{\left(\sqrt{es_0 + f} - \sqrt{f}\right)\left(\lambda - \frac{s_0}{\tau_A(s_0)}\right)}{\frac{1}{c} - \frac{s_0}{\tau_A(s_0)}}\right)$$

Then we have that

$$g^* \geq \max_{\{s_0 : s_0/\tau_A(s_0) < \lambda\}} \left\{\frac{\tilde{C}(1 + c\lambda/2)}{1 - c\lambda} + \frac{\tilde{C}}{2} - \frac{\lambda}{2}(d + \sqrt{f})\right\}.$$

The proof is given in Appendix VIII-B.

B. Numerical comparison of policies for general $\tau(s)$

In this section we numerically evaluate the average queue length (or equivalently average delay for a fixed arrival rate) of the EXH and QAP policies. We also use the performance bounds which were derived in the previous section for evaluating the performance of these policies. In Figure 11 we consider a representative case where $c = 1, d = 1, e = 0.5$, and $f = 0.5$ and compare the average queue length of EXH and QAP policies (which are obtained by simulation) with that of an (approximately) optimal policy obtained using policy iteration. The optimal policy is approximate since policy iteration is carried out for a semi-Markov decision process defined on a truncated state space. From Figure 11, we observe that EXH and QAP have approximately the same average queue length, and their average queue length is almost the same as that of the optimal.

In Figure 12 we compare the average queue length of EXH and QAP for a system where c, d, e, and f are $1.88, 36.23, 153.93$, and 1480.85 respectively, which models the code block length for a BSC with a crossover probability $\epsilon = 0.1$ when the

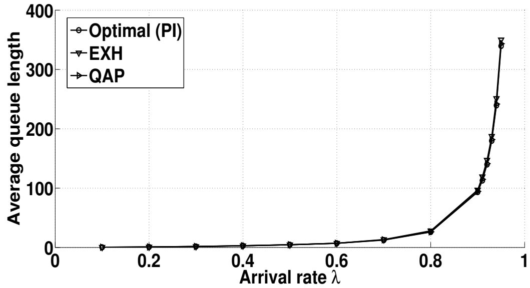

Fig. 11: Comparison of the average queue length for an optimal policy (computed using policy iteration), EXH, and QAP policies for the $\tau(s)$ obtained from Polyanskiy's approximation. The parameters c, d, e, and f are $1, 1, 0.5$, and 0.5 respectively.

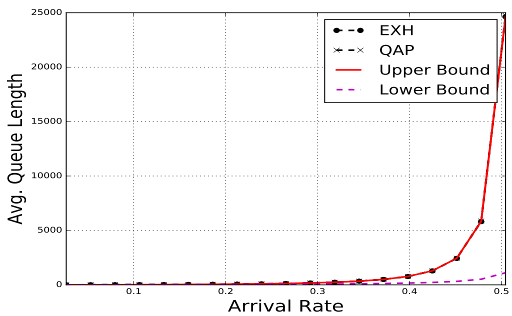

Fig. 12: Comparison of the average queue length for EXH and QAP policies for a BSC with $\epsilon = 0.1$ for the $\tau(s)$ from Polyanskiy's approximation. The $P_{e,r}$ requirement is 10^{-6}. The parameters c, d, e, and f are $1.88, 36.23, 153.93$, and 1480.85 respectively. The upper bound from Proposition (V.1) and lower bound from Proposition (V.2) are also shown.

$P_{e,r}$ is 10^{-6} (see Table III in Appendix I). We plot the upper and lower bounds on the optimal average queue length in the same figure. We observe that again the average queue lengths of EXH and QAP are similar. We also observe that the approximate upper bound from Proposition V.1 is tight. However, we observe that the lower bound from Proposition V.2 is not tight. Our conjecture is that the lower bound is not tight but EXH and QAP has performance close to that of the optimal.

In Figures 13, 14, and 15 we compare the average queue length of EXH and QAP for AWGN channels with different SNR values and $P_{e,r}$ requirements. The parameters used can be seen to correspond to different SNR values from Table IV in Appendix I. We again observe that the average queue lengths of EXH and QAP are similar and furthermore they closely match the upper bound. We have observed similar behaviour in other simulations, in which $P_{e,r}$ is small. We note that in the cases where $P_{e,r}$ is small we have that $d \gg c$ and $f \gg e$. So in the high reliability scenario the EXH policy is close to optimal, and EXH is the optimal policy for a model with a service duration that is fixed and not a function of the batch size. We note that although the fixed service duration model gives the same optimal policy, the average queue length behaviour would be different for this refined model.

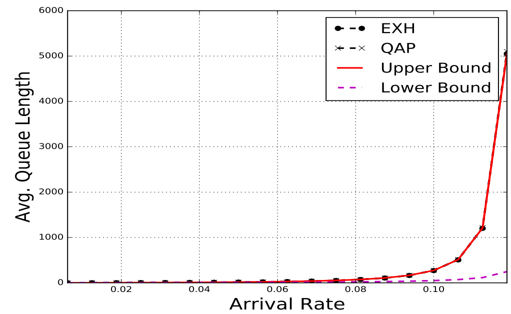

Fig. 13: Comparison of the average queue length for EXH and QAP policies for an AWGN channel with $SNR = 0$dB for the $\tau(s)$ from Polyanskiy's approximation. The $P_{e,r}$ requirement is 10^{-6}. The parameters c, d, e, and f are $8, 35.27, 564.34$, and 1244.08 respectively. The upper bound from Proposition (V.1) and lower bound from Proposition (V.2) are also shown.

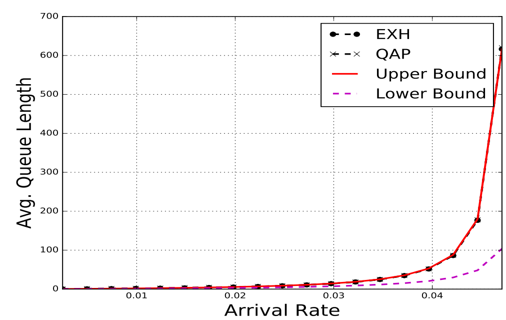

Fig. 14: Comparison of the average queue length for EXH and QAP policies for an AWGN channel with $SNR = -10$dB for the $\tau(s)$ from Polyanskiy's approximation. The $P_{e,r}$ requirement is 10^{-3}. The parameters c, d, e, and f are $20.18, 53.48, 339.18$, and 449.38 respectively. The upper bound from Proposition (V.1) and lower bound from Proposition (V.2) are also shown.

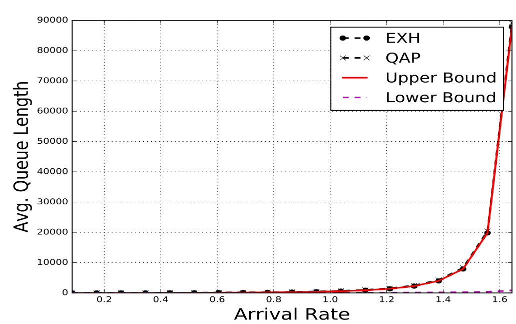

Fig. 15: Comparison of the average queue length for EXH, QAP, and Approx. policies for an AWGN channel with $SNR = 20$dB for the $\tau(s)$ from Polyanskiy's approximation. The $P_{e,r}$ requirement is 10^{-6}. The parameters c, d, e, and f are $0.58, 3.90, 53.93$, and 181.76 respectively. The upper bound from Proposition (V.1) and lower bound from Proposition (V.2) are also shown.

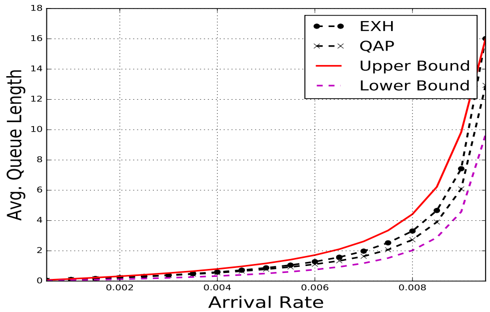

Fig. 16: Comparison of the average queue length for EXH and QAP policies for $c = 100, d = 1, e = 1$, and $f = 1$. The upper bound from Proposition (V.1) and lower bound from Proposition (V.2) are also shown.

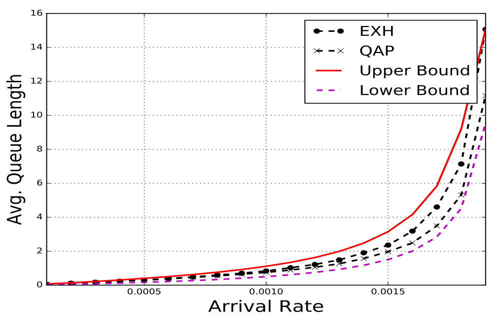

Fig. 17: Comparison of the average queue length for EXH and QAP policies for $c = 500, d = 1, e = 1$, and $f = 1$.

We consider two cases where $c \gg d$ in Figures 16 and 17. We observe that in this case the EXH policy has a higher average queue length than QAP. We note that such cases are not common in the scenarios that we are considering in the context of error-probability delay tradeoff.

VI. Cross layer tradeoffs for the noisy point-to-point channel

In this section, we study the tradeoff between average queue length and probability of message symbol error for BSC and AWGN channels, which is a cross layer performance tradeoff. In Figures 18, 19, 20 we plot the tradeoff for AWGN channels with different SNR values, while in Figures 21 and 22 we plot the tradeoff for BSCs with different cross over probabilities. In each case we plot the average queue length versus $\log(P_{e,r})$ which is an achievable tradeoff between average queue length and average message symbol error. We observe that compared to the $\tau(s)$ derived from Polyanskiy's approximation, the $\tau(s)$ derived from Gallager's random coding exponent overestimates the codeword length and this overestimation is amplified when considering the average queue length as a function of $P_{e,r}$ (the plots show 0.1-th of the actual average queue length). We also note that the tradeoff shown using Polyanskiy's approximation is a non-asymptotic characterization of the tradeoff. We observe that the tradeoff performance of the EXH, QAP, and Approx. policies are similar.

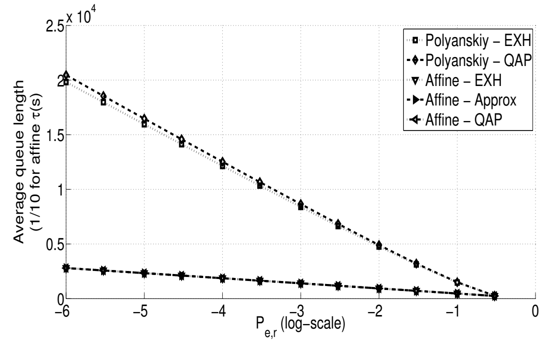

Fig. 18: Comparison of the average queue length as a function of $log_{10}(P_{e,r})$ for EXH, Approx, and QAP policies for affine $\tau(s)$, and EXP and QAP policies for the $\tau(s)$ from Polyanskiy's approximation for a AWGN channel with SNR = $20dB$. The arrival rate is 1.5567 (90% of the capacity which is 1.7297).

In the regime of $P_{e,r} \downarrow 0$ (or when $-\log(P_{e,r}) \uparrow \infty$) we also obtain an approximate characterization of the above tradeoff. This characterization is in the form of an achievable error exponent of the message symbol error with respect to the average queue length (or equivalently with respect to the average delay). From Proposition IV.4 we have that for a achievable message symbol error probability of $P_{e,r}$ the minimum average delay (say $\overline{d}^*(P_{e,r})$) is approximately bounded above by

$$\left(\frac{b+1}{b\lambda}\right)\left(\frac{a\sigma^2}{2(1-a\lambda)} + \frac{3(b+1)\lambda}{2(1-a\lambda)} + \frac{a\sigma^2}{2(1-a^2\lambda^2)} - \frac{\lambda}{2}\right),$$

where $b = -\log\left(P_{e,r}\right)/E_0(\rho, \mathcal{Q})$ (see Appendix I). The error exponent with respect to the average delay is defined as

$$\lim_{P_{e,r}\to 0} \frac{-\log\left(P_{e,r}\right)}{\overline{d}^*(P_{e,r})}.$$

Then, from the above expression we obtain an achievable error exponent with respect to average delay as

$$\frac{2}{3}\big(E_0(\rho, \mathcal{Q}) - \rho\lambda\log(\mathcal{M})\big).$$

We note that this is $2/3$ of the Gallager random coding exponent with respect to the codeword length. Therefore, for small $P_{e,r}$ we have that $P_{e,r} \approx e^{-2/3(\text{Delay})\left(E_0(\rho,\mathcal{Q})-\rho\lambda\log(\mathcal{M})\right)}$, which says that the time required for coding would be two-thirds of a given average delay requirement for a given requirement on the average probability of message symbol error.

VII. Conclusion and Future work

In this paper, we considered the problem of minimizing the average queue length (or equivalently the average delay) by dynamic scheduling of message symbols over a point-to-point link. The point-to-point link uses variable length coding schemes for transmitting batches of message symbols with per-transmission reliability constraints. We identified two examples of the relationship between the length of the code and the batch size from Gallager's random coding exponent

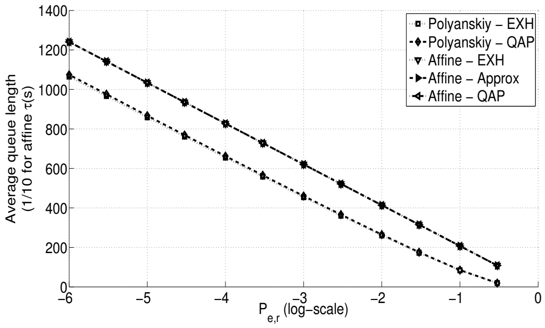

Fig. 19: Comparison of the average queue length as a function of $log_{10}(P_{e,r})$ for EXH, Approx, and QAP policies for affine $\tau(s)$, and EXP and QAP policies for the $\tau(s)$ from Polyanskiy's approximation for a AWGN channel with SNR = $0dB$. The arrival rate is 0.4 (80% of the capacity which is 0.5).

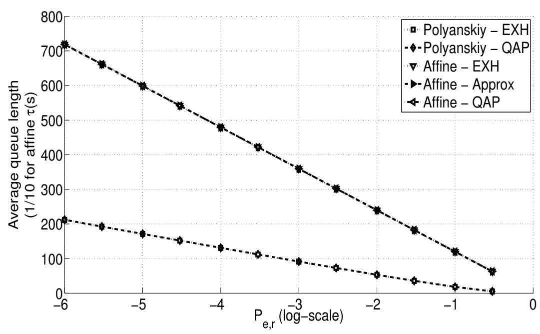

Fig. 20: Comparison of the average queue length as a function of $log_{10}(P_{e,r})$ for EXH, Approx, and QAP policies for affine $\tau(s)$, and EXP and QAP policies for the $\tau(s)$ from Polyanskiy's approximation for a AWGN channel with SNR = $-10dB$. The arrival rate is 0.1387 (70% of the capacity which is 0.1982).

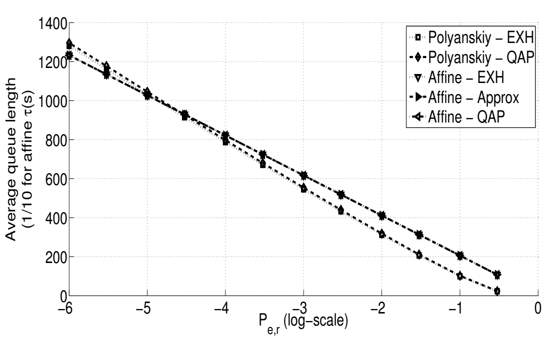

Fig. 21: Comparison of the average queue length as a function of $log_{10}(P_{e,r})$ for EXH, Approx, and QAP policies for affine $\tau(s)$, and EXP and QAP policies for the $\tau(s)$ from Polyanskiy's approximation for a BSC channel with $\epsilon = 0.1$. The arrival rate is 0.4248 (80% of the capacity which is 0.5310).

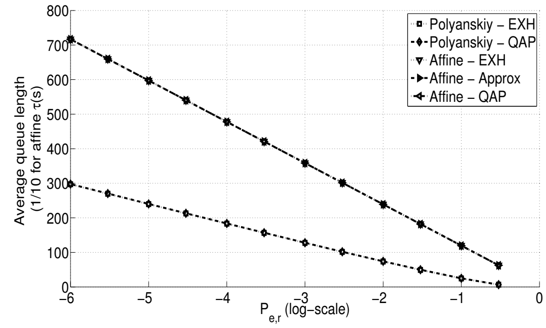

Fig. 22: Comparison of the average queue length as a function of $log_{10}(P_{e,r})$ for EXH, Approx, and QAP policies for affine $\tau(s)$, and EXP and QAP policies for the $\tau(s)$ from Polyanskiy's approximation for a BSC channel with $\epsilon = 0.2$. The arrival rate is 0.1947 (80% of the capacity which is 0.2781).

and Polyanskiy's approximation. Motivated from a SMDP analysis, we proposed three operating policies EXH, QAP, and Approx. for the case of the affine transmission time and two policies EXH and QAP for the Polyanskiy form for the transmission time. We showed that EXH was close to optimal whenever the reliability requirement was large. We also characterized the tradeoff between the average queue length and reliability numerically. For small $P_{e,r}$ we observed that the time required for coding would be two-thirds of a given average delay requirement for a given requirement on the average probability of message symbol error.

An important extension in the future is to consider the case of block fading wireless channels. We have presented these scheduling policies for a queueing system with variable batch service times which are motivated from point-to-point communication schemes. However, such queueing systems have independent interest and it is important to see if insights can be obtained for a general $\tau(s)$.

References

[1] *IEEE 802.11n: Standard for WLAN MAC and PHY Specifications Amendment 5 : Enhancements for higher throughput*. 2009.

[2] Samuli Aalto. Optimal control of batch service queues with compound poisson arrivals and finite service capacity. *Mathematical Methods of Operations Research*, 48:317–335, 1998.

[3] Samuli Aalto. Optimal control of batch service queues with finite service capacity and linear holding costs. *Mathematical Methods of Operations Research*, 51:263–285, 2000.

[4] Arjun Anand and Gustavo de Veciana. Resource allocation and harq optimization for urllc traffic in 5g wireless networks. *IEEE Journal on Selected Areas in Communications*, 36(11):2411–2421, 2018.

[5] Randall A Berry and Robert G Gallager. Communication over fading channels with delay constraints. *IEEE Transactions on Information Theory*, 48(5):1135–1149, 2002.

[6] Ido Bettesh and SS Shamai. Optimal power and rate control for minimal average delay: The single-user case. *IEEE Transactions on Information Theory*, 52(9):4115–4141, 2006.

[7] Dieter Claeys. *Analysis of Queueing Models with Batch Service*. PhD thesis, Gent University, 2011.

[8] Dieter Claeys, Joris Walraevens, Koenraad Laevens, and Herwig Bruneel. A batch server with service times dependent on the number of served customers. *European Simulation and Modelling Conference 2007*, pages 214–219, 2007.

[9] Dieter Claeys, Joris Walraevens, Koenraad Laevens, and Herwig Bruneel. Analysis of threshold-based batch-service queueing systems with batch arrivals and general service times. *Performance Evaluation*, 68(6):528 – 549, 2011.

[10] Mustafa Cemil Coşkun, Giuseppe Durisi, Thomas Jerkovits, Gianluigi Liva, William Ryan, Brian Stein, and Fabian Steiner. Efficient error-correcting codes in the short blocklength regime. *Physical Communication*, 34:66–79, 2019.

[11] Rajat K. Deb. Optimal control of batch service queues with switching costs. *Advances in Applied Probability*, 8(1):pp. 177–194, 1976.

[12] Rajat K. Deb and Richard F. Serfozo. Optimal control of batch service queues. *Advances in Applied Probability*, 5(2):pp. 340–361, 1973.

[13] Dee Denteneer, Augustus Josephus Elizabeth Maria Janssen, and JSH van Leeuwaarden. Moment inequalities for the discrete-time bulk service queue. *Mathematical Methods of Operations Research*, 61(1):85–108, 2005.

[14] Rahul Devassy, Giuseppe Durisi, Guido Carlo Ferrante, Osvaldo Simeone, and Elif Uysal. Reliable transmission of short packets through queues and noisy channels under latency and peak-age violation guarantees. *IEEE Journal on Selected Areas in Communications*, 37(4):721–734, 2019.

[15] Navid Ehsan and Tara Javidi. Delay optimal transmission policy in a wireless multiaccess channel. *IEEE Transactions on Information Theory*, 54(8):3745–3751, 2008.

[16] Robert G Gallager. *Information theory and reliable communication*, volume 2. Springer, 1968.

[17] Eckhardt Gallisch. On monotone optimal policies in a queueing model of m/g/1 type with controllable service time distribution. *Advances in Applied Probability*, 11(4):pp. 870–887, 1979.

[18] Fatemeh Hamidi-Sepehr, Henry D Pfister, and Jean-François Chamberland. Delay-sensitive communication over fading channels: Queueing behavior and code parameter selection. *IEEE Transactions on Vehicular Technology*, 64(9):3957–3970, 2014.

[19] Evgeny Khorov, Anton Kiryanov, Andrey Lyakhov, and Giuseppe Bianchi. A tutorial on ieee 802.11 ax high efficiency wlans. *IEEE Communications Surveys & Tutorials*, 21(1):197–216, 2018.

[20] Ger Koole. *Monotonicity in Markov Reward and Decision Chains : Theory and Applications*. Foundations and Trends in Stochastic Systems : NOW Publishers.

[21] Sean P Meyn. Algorithms for optimization and stabilization of controlled Markov chains. *Sadhana*, 24(4-5):339–367, 1999.

[22] Sean P Meyn. Stability, performance evaluation, and optimization. In *Handbook of Markov decision processes*, pages 305–346. Springer, 2002.

[23] Sean P. Meyn and Richard Tweedie. *Markov Chains and Stochastic Stability*. Springer Verlag, 1993.

[24] S. Musy and E. Telatar. On the transmission of bursty sources. In *2006 IEEE International Symposium on Information Theory*, pages 2899–2903, July 2006.

[25] Stéphane Musy. Delay and coding in multiple-user communications. 2007.

[26] Michael J Neely. Stochastic network optimization with application to communication and queueing systems. *Synthesis Lectures on Communication Networks*, 3(1):1–211, 2010.

[27] Marcel F. Neuts. A general class of bulk queues with poisson input. *The Annals of Mathematical Statistics*, 38(3):pp. 759–770, 1967.

[28] Parimal Parag, Jean-Francois Chamberland, Henry D Pfister, and Krishna Narayanan. Code-rate selection, queueing behavior, and the correlated erasure channel. *IEEE Transactions on Information Theory*, 59(1):397–407, 2013.

[29] Y. Polyanskiy, H.V. Poor, and S. Verdu. Dispersion of gaussian channels. In *IEEE International Symposium on Information Theory*, pages 2204 –2208, July 2009.

[30] Y. Polyanskiy, H.V. Poor, and S. Verdu. Channel coding rate in the finite blocklength regime. *IEEE Transactions on Information Theory*, 56(5):2307 –2359, May 2010.

[31] Warren B. Powell and Pierre Humblet. The bulk service queue with a general control strategy: Theoretical analysis and a new computational procedure. *Operations Research*, 34(2):pp. 267–275, 1986.

[32] Sheldon M. Ross. Average cost semi-Markov decision processes. *Journal of Applied Probability*, 7(3):649–656, 1970.

[33] KCV Kalyanarama Sesha Sayee and Utpal Mukherji. Stability of scheduled multi-access communication over quasi-static flat fading channels with random coding and independent decoding. *Information Theory, 2005 International Symposium on*, pages 2261–2265, Sept. 2005.

[34] Sebastian Schiessl, James Gross, and Hussein Al-Zubaidy. Delay analysis for wireless fading channels with finite blocklength channel coding. In *Proceedings of the 18th ACM International Conference on Modeling, Analysis and Simulation of Wireless and Mobile Systems*, pages 13–22. ACM, 2015.

[35] L. I. Sennott. *Stochastic dynamic programming and the control of queues*. Wiley- IEEE, 1999.

[36] Linn I. Sennott. Average cost optimal stationary policies in infinite state Markov decision processes with unbounded costs. *Operations Research*, 37(4):pp. 626–633, 1989.

[37] Linn I. Sennott. Average cost semi-Markov decision processes and the control of queueing systems. *Probability in the Engineering and Informational Sciences*, 3:247–272, 1989.

[38] M Series. Imt vision–framework and overall objectives of the future development of imt for 2020 and beyond. *Recommendation ITU*, pages 2083–0, 2015.

[39] I. Emre Telatar and Robert G. Gallager. Combining queueing theory with information theory for multiaccess. *IEEE Journal on Selected Areas in Communications*, 13(6):963–969, 1995.

[40] C. H. Tijms. *A first course in stochastic models*. Wiley, 2003.

[41] R.L. Tweedie. Hitting times of Markov chains, with applications to state dependent queues. *Bulletin of the Australian Mathematical Society*, 17(1):97–107, 2009.

[42] Elif Uysal-Biyikoglu, Balaji Prabhakar, and Abbas El Gamal. Energy-efficient packet transmission over a wireless link. *IEEE/ACM Transactions on Networking (TON)*, 10(4):487–499, 2002.

[43] B. S. Vineeth and U. Mukherji. Average-delay optimal policies for the point-to-point channel. In *7th International Symposium on Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks, WiOpt*, pages 1 –10, 2009.

[44] Howard J. Weiss. The computation of optimal control limits for a queue with batch services. *Management Science*, 25(4):pp. 320–328, 1979.

[45] E.M. Yeh and A.S. Cohen. Throughput and delay optimal resource allocation in multiaccess fading channels. In *IEEE International Symposium on Information Theory*, page 245, june-4 july 2003.

[46] E.M. Yeh and A.S. Cohen. Throughput optimal power and rate control for queued multiaccess and broadcast communications. In *International Symposium on Information Theory*, page 112, june-2 july 2004.

[47] Xiaoyu Zhao and Wei Chen. Queue-aware variable-length coding for ultra reliable low latency communications with random arrival. *CoRR*, abs/1903.05804, 2019.

APPENDIX I
BATCH SERVICE TIMES FOR POINT-TO-POINT CHANNEL CODES

1) Affine form for $\tau(.)$ for random block coding: We consider a memoryless point-to-point channel, with transition probability function $P_{Y|X}$, and with input and output alphabets $\mathcal{X}$ and $\mathcal{Y}$. The random variables $X \in \mathcal{X}$ and $Y \in \mathcal{Y}$ denote the input to and the output from the channel respectively. Assume that each message symbol has an average error probability requirement $P_{e,r}$. We guarantee that this requirement is met by ensuring that the average probability of codeword error $P(s, \tau) \leq P_{e,r}$, for every transmission, since then every message symbol transmitted has an average error probability of at most $P_{e,r}$.

To ensure $P(s, \tau) \leq P_{e,r}$, it is intuitive that the codeword length τ has to be chosen as a function $\tau(s)$ of the batch size s. In the following, we obtain an analytical form for $\tau(s)$ by using Gallager's random coding upper bound on $P(s, \tau)$. The message alphabet size is $\mathcal{M}$, therefore the alphabet size of the batch of s message symbols is $\mathcal{M}^s$. Therefore, the rate in bits per channel use is $R = \frac{s \log \mathcal{M}}{\tau(s)}$. We consider a randomly generated codebook, in which each codeword symbol is chosen independently according to the distribution $\mathcal{Q}$ on the input alphabet $\mathcal{X}$. The receiver is assumed to do

ϵ	$P_{e,r}$	a	b
0.1	10^{-3}	2.00	198.94
0.1	10^{-6}	2.00	397.88
0.4	10^{-3}	37.80	3766.97
0.4	10^{-6}	37.80	7533.93

TABLE I: Values for a and b computed for a BSC with crossover probability ϵ, $M = 2$, and probability of error requirement $P_{e,r}$. The parameter ρ is assumed to be 0.01.

SNRdB	$P_{e,r}$	M	a	b
-10	10^{-3}	16	21.94	546.74
0	10^{-6}	16	8.58	427.31
20	10^{-6}	2	0.60	119.53

TABLE II: Values for a and b computed for an AWGN channel with SNR SNR, symbol size M, and probability of error requirement $P_{e,r}$. The parameter ρ is assumed to be 0.01.

maximum likelihood decoding of the joint message. From [16, Theorem 5.6.2], we have that the average probability of codeword error is bounded above as:

$$\begin{aligned} P(s,\tau) &\leq e^{-\tau\{-\rho R + E_0(\rho,\mathcal{Q})\}}, \\ \text{or } \ln P(s,\tau) &\leq \rho s \log \mathcal{M} - \tau E_0(\rho,\mathcal{Q}), \end{aligned}$$

where $\rho \in [0,1]$ and

$$E_0(\rho,\mathcal{Q}) = -\log \int_{y\in\mathcal{Y}} \left[\int_{x\in\mathcal{X}} \mathcal{Q}(x) P_{Y|X}(y|x)^{\left(\frac{1}{1+\rho}\right)} dx\right]^{1+\rho} dy.$$

We note that for discrete input and output channels (such as the binary symmetric channel) the integrals are replaced by sums and the probability density functions by probability mass functions in the above definition.

To guarantee the average error probability requirement for every transmission, we constrain $P(s,\tau)$ to be $\leq P_{e,r}$. If

$$e^{-\tau\{-\rho R + E_0(\rho,\mathcal{Q})\}} \leq P_{e,r},$$

then $P(s,\tau) \leq P_{e,r}$. For every s, if τ is chosen as a function $\tau(s)$ to satisfy the above inequality, we have that

$$\frac{-\log P_{e,r}}{E_0(\rho,\mathcal{Q})} + \frac{\rho s \log \mathcal{M}}{E_0(\rho,\mathcal{Q})} \leq \tau(s).$$

Thus $\tau(s)$ has to be chosen as the smallest integer greater than or equal to $as + b$, where $a = \frac{\rho \log M}{E_0(\rho,\mathcal{Q})}$ and $b = \frac{-\log P_{e,r}}{E_0(\rho,\mathcal{Q})}$. Therefore $\tau(s) = \lceil as + b \rceil, s > 0$, so that $P(s,\tau) \leq P_{e,r}$. For any policy γ, for which $\tau(s)$ is chosen as above, we have that the average probability of error is less than or equal to $P_{e.r}$. For this affine case, we have that $\tau(0) = 1$ and $\tau(s) = \lceil as + b \rceil, s > 0$.

We now discuss some specific channels and typical a and b values. For a binary symmetric channel (BSC) with crossover probability ϵ we have that

$$E_0(\rho, Q) = \rho \log(2) - (1+\rho)\big(\epsilon^{1/(1+\rho)} + (1-\epsilon)^{1/(1+\rho)}\big).$$

Some typical values for the parameters a and b for the AWGN channel are shown in Table I. For an AWGN channel with receiver SNR denoted by SNR we have that

$$E_0(\rho, Q) = \frac{\rho}{2} \log\big(1 + SNR/(1+\rho)\big).$$

Some typical values for the parameters a and b for the AWGN channel are shown in Table II.

2) Form for $\tau(s)$ using Polyanskiy's bounds for error probability of block codes: We note that the Gallager random coding bound is tight (in the exponent) only for arrival rates λ close to the capacity of the channel and when large block lengths are used (i.e. when $\tau(s)$ is large). We will now use the normal approximation for the codeword error probability, derived by Polyanskiy et al. [30], to obtain an approximation for the function $\tau(s)$. We again assume that each message symbol requires an average probability of error guarantee of $P_{e,r}$. This is again guaranteed by choosing $P(s,\tau) \leq P_{e,r}$. For any coding scheme that transmits s message symbols using a block length of $\tau(s)$ and with an average error probability $\leq P_{e,r}$, we have from Polyanskiy's approximation that

$$s \log_2 \mathcal{M} \lessapprox \tau(s)C - \sqrt{\tau(s)V}\mathbb{Q}^{-1}(P_{e,r}),$$

where C is the capacity of the channel, V is the channel dispersion [30], and $\mathbb{Q}$ is the Gaussian Q function. From this we obtain the approximation that

$$\begin{aligned} \tau(s) &= \Bigg\lceil \frac{s\log\mathcal{M}}{C} + \frac{V(\mathbb{Q}^{-1}(P_{e,r}))^2}{2C^2} + \\ &\quad \frac{\sqrt{V}\mathbb{Q}^{-1}(P_{e,r})}{C}\sqrt{4Cs\log\mathcal{M} + V(\mathbb{Q}^{-1}(P_{e,r}))^2}\Bigg\rceil. \end{aligned}$$

For brevity, we will use $\tau(s) = \lceil cs + d + \sqrt{es+f} \rceil$ in the following to represent the above functional form for $\tau(s)$, where

$$\begin{aligned} c &= \frac{\log\mathcal{M}}{C}, \\ d &= \frac{V(\mathbb{Q}^{-1}(P_{e,r}))^2}{2C^2}, \\ e &= \left(\frac{\sqrt{V}\mathbb{Q}^{-1}(P_{e,r})}{C}\right)^2 4C\log\mathcal{M}, \text{ and} \\ f &= \left(\frac{\sqrt{V}\mathbb{Q}^{-1}(P_{e,r})}{C}\right)^2 V(\mathbb{Q}^{-1}(P_{e,r}))^2. \end{aligned}$$

For this example, we have that $\tau(0) = 1$ and $\tau(s) = \lceil cs + d + \sqrt{es+f} \rceil, s > 0$.

We now discuss some specific channels and typical c, d, e, and f values. For a binary symmetric channel (BSC) with crossover probability ϵ we have that

$$\begin{aligned} C &= 1 + \epsilon\log(\epsilon) + (1-\epsilon)\log(1-\epsilon) \\ V &= \epsilon(1-\epsilon)\big(\log\big(\frac{1-\epsilon}{\epsilon}\big)\big)^2. \end{aligned}$$

Some typical values for the parameters c, d, e, and f for

ϵ	$P_{e,r}$	c	d	e	f
0.1	10^{-3}	1.88	15.31	65.06	264.51
0.1	10^{-6}	1.88	36.23	153.93	1480.85
0.4	10^{-3}	34.42	464.67	107.99	728.83
0.4	10^{-6}	34.42	1099.45	255.50	4080.25

TABLE III: Values for c, d, e, and f computed for a BSC with crossover probability ϵ, $M = 2$, and probability of error requirement $P_{e,r}$.

SNRdB	$P_{e,r}$	M	c	d	e	f
-10	10^{-3}	16	20.18	53.48	339.18	449.38
0	10^{-6}	16	8.00	35.27	564.34	1244.08
20	10^{-6}	2	0.58	3.90	53.93	181.76

TABLE IV: Values for c, d, e, and f computed for an AWGN channel with SNR SNR, symbol size M, and probability of error requirement $P_{e,r}$.

the BSC are shown in Table III. For an AWGN channel with receiver SNR denoted by SNR we have that [29]

$$C = \frac{1}{2}\log\left(1 + SNR\right)$$
$$V = \frac{1}{2}\left(1 - 1/(1 + SNR)^2\right).$$

Some typical values for the parameters c, d, e, and f for the AWGN channel are shown in Table IV.

Appendix II
The minimum average delay for $\tau(s) = as$

Under the policy γ_1 the system idles for one slot when there are no message symbols to be transmitted. Otherwise, each symbol takes a slots to be transmitted. For obtaining an approximate expression for the minimum average queue length and delay we consider a system which idles for a slots when there are no message symbols to be transmitted. Then, we have an equivalent discrete time queueing system where the slot is equivalent to a slots for the original queueing system. For the equivalent system, the arrival process is IID in the slots with arrival rate and variance being $a\lambda$ and $a\sigma^2$ instead of λ and σ^2. For the equivalent system, a single message symbol is transmitted in every slot. From [13] we have that the average queue length for the equivalent system is

$$\frac{a\sigma^2}{2(1 - a\lambda)} + \frac{a\lambda}{2}.$$

From Little's law the average delay is

$$\frac{a\sigma^2}{2a\lambda(1 - a\lambda)} + \frac{1}{2}.$$

We propose this an approximate expression for the average delay for γ_1 for the original system. We expect that as the arrival rate $\lambda \uparrow \frac{1}{a}$ the above approximation would be tight.

Appendix III
c-regularity of scheduling policies

We present the following material from [21], [22] and [23] as background for c-regularity for scheduling policies, since the c-regularity property is used as a building block for many of the proofs in this paper. We note that these results are stated in the context of countable state Markov chains which arise in our queueing models rather than for the general cases considered in [23]. The Markov chain that we consider is the embedded queueing process $(Q[m], m \geq 0)$. From Section II we have that $(Q[m], m \geq 0)$ is a Markov chain if the policy γ is a stationary deterministic policy. In this discussion, the state of a Markov refers to the queue length value for $Q[m]$.

We first define ψ-irreducibility. A Markov chain is ψ-irreducibile if there is a single communicating class which is reachable with positive probability from any initial state. Under the assumption that there is positive probability of zero arrivals in a slot and for policies which serve at least in some state, such a class can be obtained in our model. We also note that since at state 0 there is no service, we have a transition to 0 with positive probability since $\Pr\left\{A_s[m] = 0\right\} > 0$. Then, we have that the Markov chain $(Q[m])$ is aperiodic too.

Suppose $c : \mathbb{Z}_+ \to \mathbb{R}_+$, is a cost function defined on the integer valued queue length. From [21], the Markov chain $(Q[m], m \geq 0)$ is c-regular if for some $\theta \in \mathbb{Z}_+$ and every $q \in \mathbb{Z}_+$ we have that

$$\mathbb{E}_q\Big[\sum_{n=0}^{\tau_\theta - 1} c(Q[m]) | Q[0] = q\Big] < \infty,$$

where τ_θ is the first return time to θ.

For a countable state space Markov chain, like the one considered here, a c-regular Markov chain is positive recurrent, having a stationary probability π which is such that $\mathbb{E}_\pi c(Q) < \infty$, i.e., finite stationary average cost.

A scheduling policy γ is said to be c-regular if the Markov chain $(Q[m], m \geq 0)$ under γ is c-regular. From our discussion above, for a c-regular policy, we have that $(Q[m], m \geq 0)$ is a positive recurrent Markov chain with $\mathbb{E}_\pi c(Q) < \infty$, where π is the stationary distribution of the Markov chain induced by the policy.

For the system considered in this paper, we now verify that a particular candidate scheduling policy is c-regular for the functions $h(q, s(q))$ and $\tau(s(q))$ (or h-regular and τ-regular) using a Lyapunov drift criterion. The candidate policy γ, serves a batch size $s(q) = \min(q, s_1)$, where s_1 is any integer (say, the smallest) such that $\lambda < \frac{s_1}{\tau(s_1)}$. We will use the following theorem [21, Theorem 2.2] or [22, Theorem 10.3] to verify that γ is both h-regular as well as τ-regular.

Theorem III.1. *Sufficient condition for c-regularity: Suppose there exists a Lyapunov function $L : \mathbb{Z}_+ \to \mathbb{R}_+$, the function c is near monotone, (i.e., $\{q \in \mathbb{Z}_+, c(q) \leq \eta\}$ is finite for any $\eta < \sup_q c(q)$), and there exists a constant $J < \sup_q c(q)$ such that*

$$\mathbb{E}[L(Q[m+1]|Q[m] = q] - L(q) \leq -c(q) + J.$$

Then $(Q[m], m \geq 0)$ is a c-regular Markov chain.

We note that h-regularity and τ-regularity are verified by replacing c in the above theorem with $h(q, s(q))$ and $\tau(s(q))$ respectively.

Proposition III.2. *For the candidate policy γ, the Markov chain $(Q[m], m \geq 0)$ is both h-regular and τ-regular.*

Proof. We define the Lyapunov function $L(q) = q^2$. The evolution of the queue for γ is given by

$$Q[m+1] = \begin{cases} A^1 & \text{if } Q[m] = 0, \\ A^{\tau(Q[m])} & \text{if } Q[m] \leq s_1, \\ Q[m] - s_1 + A^{\tau(s_1)} & \text{if } Q[m] > s_1. \end{cases}$$

Define $\Delta(q) = \mathbb{E}[L(Q[m+1]) - L(Q[m]) | Q[m] = q]$. Then

$$\Delta(q) = \begin{cases} \mathbb{E}[(A^1)^2] & \text{if } q = 0, \\ \mathbb{E}[((q-s_1)^+ + A^{\tau(q)})^2] - q^2 & \text{if } q \leq s_1, \\ \mathbb{E}[(q - s_1 + A^{\tau(s_1)} + q)(A^{\tau(s_1)} - s_1)] & \text{if } q > s_1. \end{cases}$$

If $q = 0$, then

$$\Delta(q) = \mathbb{E}[(A^1)^2] \leq 2q(\mathbb{E}A^{\tau(s_1)} - s_1) + \mathbb{E}[(A^{\tau(s_1)})^2] + s_1^2.$$

And if $q \leq s_1$, then

$$\Delta(q) = \mathbb{E}[(q - q + A^{\tau(q)} + q)(A^{\tau(q)} - q)] \leq \\ 2q(\mathbb{E}A^{\tau(q)} - q) + \mathbb{E}[(A^{\tau(q)})^2] + s_1^2.$$

And if $q > s_1$, then

$$\Delta(q) = \mathbb{E}[(q - s_1 + A^{\tau(s_1)} + q)(A^{\tau(s_1)} - s_1)] \leq \\ 2q(\mathbb{E}A^{\tau(s_1)} - s_1) + \mathbb{E}[(A^{\tau(s_1)})^2] + s_1^2.$$

We consider the second case (i.e., $q \leq s_1$) in the above set of equations. As $q \geq s_1 - \delta_q$ for some $0 \leq \delta_q = s_1 - q \leq s_1$ we have that $2q(\mathbb{E}A^{\tau(q)} - q) \leq 2q(\mathbb{E}A^{\tau(q)} - s_1) + 2q\delta_q$. We note that $q \leq s_1$ and $\delta_q \leq s_1$. Thus, $2q(\mathbb{E}A^{\tau(q)} - s_1) + 2q\delta_q \leq 2q(\mathbb{E}A^{\tau(q)} - s_1) + 2s_1^2$. We also note that $\mathbb{E}A^{\tau(q)} \leq \mathbb{E}A^{\tau(s_1)}$ and $\mathbb{E}(A^{\tau(q)})^2 \leq \mathbb{E}(A^{\tau(s_1)})^2$. So $\Delta(q) \leq 2q(EA^{\tau(s_1)} - s_1) + \mathbb{E}(A^{\tau(s_1)})^2 + 3s_1^2, \forall q$. We note that $\mathbb{E}A^{\tau(s_1)} = \tau(s_1)\lambda$. We had assumed that $\lambda < \frac{s_1}{\tau(s_1)}$. Thus we have that $\Delta(q) \leq -(1 + 2q(s_1 - \lambda\tau(s_1))) + J$, where $J = \mathbb{E}(A^{\tau(s_1)})^2 + 3s_1^2 + 1$. As $1 + 2q(s_1 - \lambda\tau(s_1))$ is near monotone and $J < \infty$, the Markov chain induced by this policy is c-regular with $c(q) = 1 + 2q(s_1 - \lambda\tau(s_1))$. We note that the single stage cost under γ is $h(q, s(q)) \leq q\tau(s_1) + \tau(s_1)(\tau(s_1) - 1)\lambda/2$. We also note that c-regularity implies that the average cost with single-stage cost $h(q, s)$ is also finite (from [22, Theorem 10.2]). Under the above policy, all transition times are bounded above by $\tau(s_1)$. Again from [22, Theorem 10.2], c-regularity implies that the system is τ-regular. □

Appendix IV
Uniformization of the semi-Markov chain $(Q[m], m \geq 0)$

In this section we use the uniformization method from [40, pp. 282-284] to obtain a Markov chain $(\tilde{Q}[m], m \geq 0)$ from the Markov chain $(Q[m], m \geq 0)$ embedded in the queue evolution process. The consideration of the uniformized Markov chain $(\tilde{Q}[m], m \geq 0)$ turns out to be useful in Section III, for obtaining the quadratic approximation policy (QAP), and for deriving bounds on the average queue length.

The uniformized Markov chain is constructed such that the stationary distribution $\tilde{\pi}$ of the Markov chain $(\tilde{Q}[m])$, $\tilde{\pi}(q)$ is the same as that of embedded Markov chain $(Q[m])$, i.e., $\tilde{\pi}(q) = \pi(q), \forall q \in \mathbb{Z}_+$. By redefining the cost for the uniformized chain, the average queue length for any stationary deterministic policy γ for the original system (modelled by the embedded process $(Q[m], m \geq 0)$) can be obtained from the analysing $(\tilde{Q}[m])$.

We define the uniformized chain $\tilde{Q}[m]$ by defining its evolution as

$$\tilde{Q}[m+1] = \tilde{Q}[m] - s(\tilde{Q}[m]) + \tilde{A}[m+1]. \tag{11}$$

Here the function $s(q)$ is specified by the policy $\gamma \in \Gamma_s$ for our original system, which serves $s(q)$ when in state q. Suppose $\tilde{Q}[m] = q$ and $s = s(q)$, then we let $\tilde{A}[m+1]$ be distributed as $\tilde{p}_a(s) = Pr\{\tilde{A}[m+1] = a\}$, where

$$\tilde{p}_a(s) = \frac{Pr\{A^{\tau(s)} = a\}}{\tau(s)} \text{ if } a \neq s, \text{ and,}$$
$$\tilde{p}_a(s) = \frac{Pr\{A^{\tau(s)} = a\}}{\tau(s)} + \left(1 - \frac{1}{\tau(s)}\right) \text{ if } a = s.$$

Equivalently, the evolution of $\tilde{Q}[m]$ can also be specified by defining its transition probability $\tilde{p}_{q,q'}(s) = Pr\{\tilde{Q}[m+1] = q' \Big| \tilde{Q}[m] = q, \tilde{S}[m+1] = s\}$ in terms of the transition probability of $Q[m]$, $p_{q,q'}(s) = Pr\{Q[m+1] = q' \Big| Q[m] = q, S[m+1] = s\}$. We note that for $\gamma \in \Gamma_s$, the transition probabilities of both $Q[m]$ and $\tilde{Q}[m]$ are functions of the service batch size s. We define

$$\tilde{p}_{q,q'}(s) = \frac{p_{q,q'}(s)}{\tau(s)} \text{ for } q \neq q', \text{ and,}$$
$$\tilde{p}_{q,q}(s) = \frac{p_{q,q}(s)}{\tau(s)} + \big(1 - \frac{1}{\tau(s)}\big).$$

We have that for any $\gamma \in \Gamma_s$, for this choice of $\tilde{p}_{q,q'}$, $\tilde{\pi}(q) = \pi(q), \forall q \in \mathbb{Z}_+$. For the uniformized system, we define the single stage cost as

$$\tilde{h}(q, s) = h(q, s)/\tau(s) = q + \lambda(\tau(s) - 1)/2. \tag{12}$$

Then we have that

$$\mathbb{E}_{\tilde{\pi}}\tilde{h}(\tilde{Q}, s(\tilde{Q})) = \frac{\mathbb{E}_{\pi}h(Q, s(Q))}{\mathbb{E}_{\pi}\tau(s(Q))} = g^{\gamma}.$$

We also note that $\tilde{\pi}_s(s) = \pi_s(s)$.

We note that a Markov decision process (MDP) can be defined with (i) transition probabilities $\tilde{p}_{q,q'}(s)$ and (ii) single stage cost $\tilde{h}(q, s)$ at state q and action s. If $\lambda < r_{max}$, then we have that the following ACOE exists for the MDP:

$$\tilde{J}(q) = \min_{s \in \mathcal{S}(q)} \left\{ \tilde{h}(q, s) - g^* + \mathbb{E}\tilde{J}\big(q - s + \tilde{A}^s\big) \right\}, \tag{13}$$

where $\tilde{A}^s \sim \tilde{p}_a(s)$ and $\tilde{J}(q)$ is the optimal relative value function. It is known that the optimal stationary deterministic policy, obtained from the above ACOE, minimizes the average queue length g^{γ}, provided that the Markov chain $\tilde{Q}[m]$ induced by the optimal policy is unichain. We also note that $\tilde{J}(q) = J(q)$.

Appendix V
Relationships implied by (1) and (11) for the stationary moments associated with $\gamma \in \Gamma_s$

In this section we list some relationships which are implied by (1) and (11) for the stationary moments of the queue length $Q[m]$ associated with policies $\gamma \in \Gamma_s$. We note that Γ_s consists of stationary deterministic policies which are positive recurrent, have a stationary distribution π, and have finite average queue length. These relationships are useful in obtaining bounds for performance analysis.

We consider a policy γ in Γ_s, which prescribes a batch size $s(q)$ to be used when the queue length is q. We also use the same policy γ for the uniformized form $\tilde{Q}[m]$ discussed in Appendix IV. We let Q denote the random variable having the same stationary distribution as $Q[m]$ and let $\tilde{Q}$ have the same stationary distribution as $\tilde{Q}[m]$. Then $Q \sim \pi$ and $\tilde{Q} \sim \tilde{\pi} = \pi$. We also have that $\mathbb{E}_{\pi} Q = \mathbb{E}_{\tilde{\pi}} \tilde{Q} < \infty$. In the following, the probability distribution with respect to which the expectation is taken will be indicated by the random variable itself.

From [23, Theorem 14.0.1] we have that $\mathbb{E}Q[m] \to \mathbb{E}Q$ and $\mathbb{E}\tilde{Q}[m] \to \mathbb{E}\tilde{Q}$. Taking expectations with respect to the stationary distribution π on both LHS and RHS of (1) we have that

$$\begin{aligned} \mathbb{E}Q[m+1] &= \mathbb{E}Q[m] - \mathbb{E}s(Q[m]) + \mathbb{E}A^{\tau(s(Q[m]))}, \\ \text{as } m \to \infty, \mathbb{E}Q &= \mathbb{E}Q - \mathbb{E}s(Q) + \mathbb{E}A^{\tau(s(Q))}, \\ \mathbb{E}s(Q) &= \mathbb{E}A^{\tau(s(Q))}. \end{aligned} \tag{14}$$

Similarly for the uniformized chain we obtain that

$$\mathbb{E}s(\tilde{Q}) = \mathbb{E}\tilde{A}^{s(\tilde{Q})}, \tag{15}$$

where $\tilde{A}^{s(\tilde{Q})} \sim \tilde{p}_a(s(\tilde{Q}))$.

Suppose the policy γ is such that $\mathbb{E}Q^2 = \mathbb{E}\tilde{Q}^2 < \infty$. Then again from [23, Theorem 14.0.1] we have that $\mathbb{E}(Q[m])^2 \to \mathbb{E}Q^2$ and $\mathbb{E}(\tilde{Q}[m])^2 = \mathbb{E}\tilde{Q}^2$. Squaring (1) and taking expectations with respect to the stationary distribution π on both LHS and RHS, we have that

$$\mathbb{E}(Q[m+1])^2 = \mathbb{E}Q[m]^2 - \mathbb{E}s(Q[m])^2 + \mathbb{E}\big(A^{\tau(s(Q[m]))}\big)^2,$$

and as $m \to \infty$, we have that

$$\begin{aligned} \mathbb{E}Q^2 &= \mathbb{E}Q^2 + \mathbb{E}s(Q)^2 + \mathbb{E}\big(A^{\tau(s(Q))}\big)^2 + \\ &\quad 2\big[\mathbb{E}QA^{\tau(s(Q))} - \mathbb{E}Qs(Q) - \mathbb{E}s(Q)A^{\tau(s(Q))}\big], \end{aligned}$$

Or we have that

$$\begin{aligned} 0 &= \mathbb{E}s(Q)^2 + \mathbb{E}\big(A^{\tau(s(Q))}\big)^2 + \\ &\quad 2\big[\mathbb{E}QA^{\tau(s(Q))} - \mathbb{E}Qs(Q) - \mathbb{E}s(Q)A^{\tau(s(Q))}\big]. \end{aligned} \tag{16}$$

We note that a similar equation can be derived for the uniformized chain

$$\begin{aligned} 0 &= \mathbb{E}s(\tilde{Q})^2 + \mathbb{E}\big(\tilde{A}^{\tau(s(\tilde{Q}))}\big)^2 + \\ &\quad 2\big[\mathbb{E}\tilde{Q}A^{\tau(s(\tilde{Q}))} - \mathbb{E}\tilde{Q}s(\tilde{Q}) - \mathbb{E}s(\tilde{Q})A^{\tau(s(\tilde{Q}))}\big]. \end{aligned} \tag{17}$$

We find use for the relationships in (14), (15), (16), and (17) in the performance bounds that we derive in this paper.

Appendix VI
Average queue length optimality of stationary deterministic policies

In this section, we present in detail how the minimization problem in (5) can be simplified to

$$\min_{\gamma \in \Gamma_s} \phi^{\gamma}(q_0), \forall q_0, \tag{18}$$

where we only optimize over only stationary deterministic policies, i.e., the set Γ_s.

The optimality of stationary deterministic policies for (5) is shown by formulating (5) as a semi-Markov decision theoretic problem (SMDP). We show that there exists a stationary deterministic optimal policy which minimizes the average queue length. We present new sufficient conditions for the existence of an average cost optimality equation for the SMDP. These sufficient conditions are required since prior work [37] considers the case where $\tau(s)$ is bounded while $\tau(s)$ is not bounded in our case.

We obtain the above average queue length optimal policy through a sequence of optimal control problems as explained below. Each optimal control problem's objective is to minimize the total discounted queue length for a discount factor α. We get a sequence of optimal control problems by considering a sequence of discount factors α increasing to 1. We note that in the development that follows, we will use the terms cost and queue length interchangeably.

A. The discounted cost problem

Since we will consider multiple policies we will use notation which explicitly states what policy is being used. We denote the state of the system under policy γ sampled at the m^{th} decision epoch as $Q^{\gamma}[m]$. For a discount factor $\alpha \in (0,1)$ the expected total discounted cost for γ is

$$V_{\alpha}^{\gamma}(q) = \lim_{M \to \infty} \mathbb{E}_{\gamma}\Big[\sum_{m=0}^{M-1} \alpha^{T[m]} h(Q^{\gamma}[m], S^{\gamma}[m]) \Big| Q^{\gamma}[0] = q\Big],$$

where $T[m]$ is the total number of slots elapsed until the m^{th} decision epoch, and q is the queue length at time 0. As function $h(q,s)$ is non-negative, $V_{\alpha}^{\gamma}(q)$ is always well defined (but could be infinity). The optimal total discounted cost is defined as

$$V_{\alpha}(q) = \inf_{\gamma \in \Gamma} V_{\alpha}^{\gamma}(q), \forall q$$

A policy γ is α-discount optimal if $V_{\alpha}^{\gamma}(q) = V_{\alpha}(q), \forall q \in \mathbb{Z}_+$. From [37, Theorem 1], if $V_{\alpha}(q) < \infty, \forall q$, then the optimal total discounted cost $V_{\alpha}(q)$ satisfies the discounted cost optimality equation (DCOE):

$$V_{\alpha}(q) = \min_{s \in \mathcal{S}(q)} \Big\{ h(q,s) + \alpha^{\tau(s)} \mathbb{E} V_{\alpha}(q - s + A^{\tau(s)}) \Big\}, \tag{19}$$

$\forall q \in \mathbb{Z}_+$, and $A^{\tau(s)}$ has the same distribution as the random number of message symbols which arrive in $\tau(s)$ slots. Any policy γ, prescribing an action $s(q)$ which attains the minimum in the RHS of the equation above is a stationary deterministic α-discount optimal policy.

For our system we have that:

Lemma VI.1. *If $\lambda < r_{max}$ then $V_\alpha(q)$ is finite for every q, and hence $V_\alpha(q)$ satisfies the DCOE* (19).

Proof. We first note that from the definition $V_\alpha^\gamma(q) \leq \tilde{V}_\alpha^\gamma(q)$, where

$$\tilde{V}_\alpha^\gamma(q) = \mathbb{E}\big[\sum_{m=0}^{\infty} \alpha^m h(Q^\gamma[m], s(Q^\gamma[m]))\Big|Q^\gamma[0] = q\big],$$

since $\alpha \in [0, 1)$ and the total number of slots elapsed till the m^{th} decision epoch should be at least m.

We consider the deterministic policy γ which serves a batch size $s(q) = \min(q, s_1)$, where s_1 is a positive batch service size. In Appendix III we have shown that the policy γ is $h(q, s(q))$-regular if $\lambda < r_{max}$ and hence for γ, the average cost

$$\lim_{M\to\infty} \frac{1}{M}\mathbb{E}\big[\sum_{m=0}^{\infty} h(Q^\gamma[m], s(Q^\gamma[m]))\Big|Q^\gamma[0] = q\big] < \infty.$$

We note that the policy γ is aperiodic and irreducible. Then from [36, Proposition 5], we have that $\tilde{V}_\alpha^\gamma(q) < \infty, \forall q$ and therefore $V_\alpha(q) < \infty, \forall q$ since $V_\alpha(q) \leq V_\alpha^\gamma(q) \leq \tilde{V}_\alpha^\gamma(q)$. □

We now present a non-decreasing property of $V_\alpha(q)$ which will be used subsequently.

Lemma VI.2. *$V_\alpha(q)$ is a non-decreasing function of $q \in \mathbb{Z}_+$.*

Proof. For the proof of this property, we need an intermediate result for which we define the M stage total discounted cost for a policy γ as:

$$V_{\alpha,M}^\gamma(q) = \mathbb{E}_\gamma\big[\sum_{m=0}^{M-1} \alpha^{T[m]} h(Q^\gamma[m], S^\gamma[m])\Big|Q^\gamma[0] = q\big]. \tag{20}$$

The optimal M stage total discounted cost is $V_{\alpha,M}(q) = \inf_{\gamma\in\Gamma} V_{\alpha,M}^\gamma$. From Bellman's optimality principle, for $M \geq 1$, we have that

$$V_{\alpha,M}(q) = \min_{s\in\mathcal{S}(q)} \left\{h(q, s) + \alpha^{\tau(s)}\mathbb{E}V_{\alpha,M-1}(q - s + A^{\tau(s)})\right\}. \tag{21}$$

We define $V_{\alpha,0}(q) = 0, \forall q$. We have that:

Lemma VI.3. *$V_{\alpha,M}(q)$ is non-decreasing in M, i.e., $V_{\alpha,M}(q) \geq V_{\alpha,M-1}(q), \forall q, M \geq 1$.*

Proof. By definition $V_{\alpha,1}(q) \geq V_{\alpha,0}(q)$. Assume that $V_{\alpha,k}(q) \geq V_{\alpha,k-1}(q), \forall k \in \{2, \cdots, M\}$. We have that

$$V_{\alpha,M}(q) = \min_{s\in\mathcal{S}(q)} \{h(q, s) + \alpha^{\tau(s)}\mathbb{E}V_{\alpha,M-1}(q - s + A^{\tau(s)})\}$$

Suppose the optimal action for state q at the $(M+1)^{th}$ stage is s^*. Then we have the following

$$\begin{aligned} V_{\alpha,M+1}(q) &= h(q, s^*) + \alpha^{\tau(s^*)}\mathbb{E}V_{\alpha,M}(q - s^* + A^{\tau(s^*)}), \\ V_{\alpha,M}(q) &\leq h(q, s^*) + \alpha^{\tau(s^*)}\mathbb{E}V_{\alpha,M-1}(q - s^* + A^{\tau(s^*)}), \end{aligned}$$

and therefore $V_{\alpha,M+1}(q) - V_{\alpha,M}(q)$

$$\begin{aligned} &\geq \alpha^{\tau(s^*)}[\mathbb{E}V_{\alpha,M}(q - s^* + A^{\tau(s^*)}) - \\ &\quad \mathbb{E}V_{\alpha,M-1}(q - s^* + A^{\tau(s^*)})] \geq 0. \end{aligned}$$

Thus $V_{\alpha,M}(q)$ is non-decreasing in M. □

We now prove the non-decreasing property of $V_\alpha(q)$ in two steps : 1) we prove that $V_{\alpha,M}(q)$ is non-decreasing in q, for every M, and 2) we prove that $\lim_{M\to\infty} V_{\alpha,M}(q) = V_\alpha(q)$ and hence $V_\alpha(q)$ is also non-decreasing in q. In both of these steps, we will use Lemma VI.3.

From (21) we have that $V_{\alpha,1}(q+1) = h(q+1, 0) \geq h(q, 0) = V_{\alpha,1}(q), \forall q$. We assume that $V_{\alpha,m}(q + 1) \geq V_{\alpha,m}(q), \forall m \in \{1, \cdots, M\}$. We have that $V_{\alpha,M+1}(q) =$

$$\min_{s\in\mathcal{S}(q)} \left\{h(q, s) + \alpha^{\tau(s)}\mathbb{E}V_{\alpha,M}(q - s + A^{\tau(s)})\right\},$$

and $V_{\alpha,M+1}(q + 1) =$

$$\min_{s\in\mathcal{S}(q+1)} \left\{h(q + 1, s) + \alpha^{\tau(s)}\mathbb{E}V_{\alpha,M}(q + 1 - s + A^{\tau(s)})\right\}.$$

Assume that the $(M + 1)^{th}$ stage optimal action for the state $q+1$ is s^*. We first consider the case where s^* is feasible for q. Then

$$\begin{aligned} V_{\alpha,M+1}(q + 1) &= h(q + 1, s^*) + \\ &\quad \alpha^{\tau(s^*)}\mathbb{E}V_{\alpha,M}(q + 1 - s^* + A^{\tau(s^*)}), \text{ and,} \\ V_{\alpha,M+1}(q) &\leq h(q, s^*) + \alpha^{\tau(s^*)}\mathbb{E}V_{\alpha,M}(q - s^* + A^{\tau(s^*)}), \end{aligned}$$

implying that $V_{\alpha,M+1}(q + 1) - V_{\alpha,M+1}(q)$

$$\begin{aligned} &\geq h(q + 1, s^*) - h(q, s^*) + \\ &\quad \alpha^{\tau(s^*)}\Big[\mathbb{E}V_{\alpha,M}(q + 1 - s^* + A^{\tau(s^*)}) - \\ &\quad \mathbb{E}V_{\alpha,M}(q - s^* + A^{\tau(s^*)})\Big] \geq 0. \end{aligned}$$

Now let us consider the case when s^* is not feasible for the state q. Then by definition $s^* = q+1$. We have for state $(q+1)$ that $V_{\alpha,M+1}(q + 1)$ is then:

$$= h(q + 1, q + 1) + \alpha^{\tau(q+1)}\mathbb{E}V_{\alpha,M}(A^{\tau(q+1)}).$$

In state q, with $(M + \tau(q + 1))$ stages to go, we use the non-stationary policy γ^{ns} which (1) idles for $\tau(q+1) - \tau(q)$ slots, (2) serves q message symbols with a batch-service time of $\tau(q)$ slots, and (3) follows the actions prescribed by the discount optimal policy for the remaining M stages. Therefore we have that $V_{\alpha,M+\tau(q+1)}(q)$

$$\begin{aligned} &\leq \mathbb{E}_{\gamma_{ns}}\big[\sum_{k=0}^{\tau(q+1)-1} h(Q^{\gamma_{ns}}[k], S^{\gamma_{ns}}[k])|Q^{\gamma_{ns}}[0] = q\big] + \\ &\quad \alpha^{\tau(q+1)}\mathbb{E}V_{\alpha,M}(A^{\tau(q+1)}). \end{aligned}$$

and,

$$V_{\alpha,M+1}(q) \leq V_{\alpha,M+\tau(q+1)}(q).$$

By an abuse of notation, we write $h(q, q + 1) = \mathbb{E}[\sum_{k=0}^{\tau(q+1)-1} h(Q^{\gamma_{ns}}[k], S^{\gamma_{ns}}[k])|Q^{\gamma_{ns}}[0] = q]$ so that

$$V_{\alpha,M+1}(q) \leq h(q, q + 1) + \alpha^{\tau(q+1)}\mathbb{E}V_{\alpha,M}(A^{\tau(q+1)}).$$

Thus we have that $V_{\alpha,M+1}(q+1) - V_{\alpha,M+1}(q) \geq h(q+1,q+1) - h(q,q+1) \geq 0$. So $V_{\alpha,M+1}(q)$ is non-decreasing in q.

We note that from Lemma VI.3, we have that $V_{\alpha,M}(q)$ is non-decreasing in M. Therefore $\lim_{M\to\infty} V_{\alpha,M}(q) = W(q)$, where $W(q)$ could potentially be infinity. The rest of the proof proceeds as in [35, Proposition 4.3.1], and following the proof leads to $W(q) = V_\alpha(q)$. Hence $V_\alpha(q)$ is non-decreasing in q, for all q. □

B. The average cost problem

We first present a theorem which will enable us to verify that a stationary deterministic optimal policy exists for (5). In literature, e.g. [37], sufficient conditions for the existence of a stationary deterministic optimal policy for SMDPs include an assumption of an uniform upper bound on $\tau(s), \forall s \in \mathbb{Z}_+$. We note that for our system model there is no uniform upper bound for $\tau(s), s \in \mathbb{Z}_+$. This necessitates the development of the following theorem, which relaxes the uniform upper bound requirement on $\tau(s)$ in [37], but follows the development in [37] specialized to the model considered in this paper.

Theorem VI.4. *If the following hold:*

1) $J_\alpha(q) := V_\alpha(q) - V_\alpha(0)$ *satisfies* $0 \overset{(a)}{\leq} J_\alpha(q) \overset{(b)}{\leq} M(q)$, *where* $M(q)$ *is a finite non-negative real number, and,*
2) *for every* $s \in \{0,\cdots,q\}, \mathbb{E}M(q-s+A^{\tau(s)}) < \infty$,

then

1) *the optimal or minimum values of* (2) *and* (4) *denoted as* $\phi^*(q_0)$ *and* $g^*(q_0)$ *are finite, equal, and do not depend on* q_0. *This finite minimum value is called the optimal average queue length and denoted as* g^*,
2) *there exists an optimal stationary deterministic policy* γ^*. *The optimal policy* γ^* *prescribes an optimal batch size* $s^*(q)$ *as a deterministic function of* q. *The optimal batch size* $s^*(q)$ *is the minimizer for the following average cost optimality equation (ACOE):*

$$J(q) = \min_{s\in\mathcal{S}(q)} \left\{ \left(q + \frac{(\tau(s)-1)\lambda}{2} - g^*\right)\tau(s) + \mathbb{E}J(q-s+A^{\tau(s)})\right\}, \forall q \in \mathbb{Z}_+, \tag{22}$$

where $J(q)$ *is the optimal relative value function.*

The proof follows the approach and the same notation as in Sennott [37], with ideas borrowed from Sennott [36], Ross [32], and Meyn [22], [23]. We specialize the proof in Sennott [37] to the model considered in this paper. But we relax Assumption 2 of Sennott [37]; instead of requiring that the transition time is uniformly bounded above for every state and action, we only require it to be finite for every state and action (which is in fact true in our model).

Proof. As in [37] we first construct a candidate stationary optimal policy for the average cost problem via discounted cost problems. Consider any sequence of discount factors $\alpha_n \uparrow 1$. Let f_{α_n} be the stationary deterministic α_n discount optimal policy (as determined by (19)). Note that $f_{\alpha_n} \in \Pi_q\mathcal{S}(q)$. From the [36, Lemma], we obtain that $\Pi_q\mathcal{S}(q)$ is a compact set as it is the product of compact sets. As f_{α_n} is a sequence in a compact set, there exists a subsequence β_n such that $f_{\beta_n} \to f$. This f is the candidate stationary deterministic optimal policy for the average cost problem.

We now consider the sequence β_n and define $J_{\beta_n}(q) = V_{\beta_n}(q) - V_{\beta_n}(0)$. From the first hypothesis of Theorem VI.4 we have that $J_{\beta_n} \in \times_q[0, M(q)]$ which is again a compact set (see the Appendix of [36]). Thus there exists a further subsequence δ_n of β_n such that $J_{\delta_n} \to J$. This J is the candidate optimal relative value function for the average cost problem.

Let $p_{q,q'}(f(q))$ be defined as

$$Pr\{Q[m+1] = q' | Q[m] = q, S[m] = f(q)\}.$$

Then, from the DCOE (19) we have that:

$$V_{\delta_n}(q) = h(q,f(q)) + \delta_n^{\tau(f(q))}\sum_{q'} p_{q,q'}(f(q))V_{\delta_n}(q'),$$

from which we have that $V_{\delta_n}(q)(1-\delta_n^{\tau(f(q))})$

$$\begin{aligned} &= h(q,f(q)) + \\ &\quad \delta_n^{\tau(f(q))}\sum_{q'} p_{q,q'}(f(q))J_{\delta_n}(q') - \delta_n^{\tau(f(q))}J_{\delta_n}(q) && (23)\\ &\leq h(q,f(q)) + \delta_n^{\tau(f(q))}\sum_{q'} p_{q,q'}(f(q))M(q'). && (24)\end{aligned}$$

We note that here we have used the finiteness of $V_\alpha(q)$, for $\alpha \in [0,1)$, obtained in Lemma VI.1.

We note that the RHS of (24) is a number say d_q. Again using that fact that $V_{\delta_n}(.)(1-\delta_n^{\tau(f(.))}) \in \times_q[0,d_q]$ which is compact, we obtain that there is subsequence, say ϵ_n of δ_n such that $\lim_{n\to\infty} V_{\epsilon_n}(q)(1-\epsilon_n^{\tau(f(q))}) = g(q)$, for some $g(q) \geq 0$. We note that $\epsilon_n \uparrow 1$. Let $\nu_n = 1-\epsilon_n$. Hence we have that $\lim_{n\to\infty} V_{\epsilon_n}(q)\nu_n = \frac{g(q)}{\tau(f(q))}$. Now as in [36] we show that $\frac{g(q)}{\tau(f(q))} = g, \forall q$. We have that

$$\begin{aligned} 0 &\leq \left|\frac{g(q)}{\tau(f(q))} - \frac{g(0)}{\tau(f(0))}\right|, \text{ and therefore,}\\ &\leq \lim_{n\to 0} \nu_n |J_{\epsilon_n}(q)| = 0,\end{aligned}$$

since $J_\alpha(q) \leq M(q)$ by the stated assumption. Hence $\tau(f(q))\lim_{n\to\infty} V_{\epsilon_n}(q)\nu_n = g\tau(f(q))$. Now consider (24), for the sequence $\epsilon_n = 1-\nu_n$, we have that $\lim_{n\to\infty} V_{\epsilon_n}(q)(1-\epsilon_n^{\tau(f(q))})$

$$\begin{aligned} &= h(q,f(q)) + \lim_{n\to\infty} \epsilon_n^{\tau(f(q))} \times \\ &\quad \Big(\sum_{q'} p_{q,q'}(f(q))J_{\epsilon_n}(q') - \epsilon_n^{\tau(f(q))}J_{\epsilon_n}(q)\Big), \quad (25)\end{aligned}$$

and $g\tau(f(q)) + J(q)$

$$= h(q,f(q)) + \lim_{n\to\infty}\sum_{q'} p_{q,q'}(f(q))J_{\epsilon_n}(q'). \tag{26}$$

Now using the second hypothesis and applying dominated convergence theorem, we obtain that

$$g\tau(f(q)) + J(q) = h(q, f(q)) + \sum_{q'} p_{q,q'}(f(q))J(q'). \quad (27)$$

Thus from the assumptions, we have obtained a constant g, functions $f(q)$ and $J(q)$ such that the above equation is satisfied. We note that $\forall q, 1 \leq \tau(f(q)) < \infty$. Therefore (27) can be written as

$$g + \frac{J(q)}{\tau(f(q))} = \frac{h(q, f(q))}{\tau(f(q))} + \frac{\sum_{q'} p_{q,q'} J(q')}{\tau(f(q))}, \quad (28)$$

or $g + J(q) =$

$$\frac{h(q, f(q))}{\tau(f(q))} + \frac{\sum_{q'} p_{q,q'} J(q')}{\tau(f(q))} + \left(1 - \frac{1}{\tau(f(q))}\right) J(q). \quad (29)$$

For the stationary policy f, note that $Q[m]$ is a Markov chain. We note that using the procedure in Tijms [40], which is explained in Section IV, we can construct an equivalent Markov chain $\tilde{Q}[m]$ with transition probabilities

$$\tilde{p}_{q,q'}(f(q)) = \begin{cases} \frac{p_{q,q'}(f(q))}{\tau(f(q))}, \text{ if } q \neq q', \\ \frac{p_{q,q'}(f(q))}{\tau(f(q))} + \left(1 - \frac{1}{\tau(f(q))}\right), \text{ if } q = q'. \end{cases}$$

With a modified single stage cost $\tilde{h}(q, f(q)) = \frac{h(q,f(q))}{\tau(f(q))}$, the average queue length for the uniformized Markov chain, starting in state q is

$$\tilde{g}^f(q) = \lim_{M\to\infty} \frac{1}{M} \mathbb{E}\big[\sum_{m=0}^{M-1} \tilde{h}(\tilde{Q}[m], f(\tilde{Q}[m])) \Big| \tilde{Q}[0] = q \big].$$

From Appendix IV we have that

$$\tilde{g}^f(q) = g^f(q).$$

Therefore, we can obtain the average queue length defined as $g^f(q)$ from $\tilde{Q}[m]$. Now we note that Lemma A.1 of [36] can be applied to (29), to obtain

$$g \geq \tilde{g}^f(q) = g^f(q). \quad (30)$$

We now note that (29) verifies the Lyapunov condition (10.13) of [22] for $Q[m]$ which implies that f induces $Q[m]$ which is h-regular and positive recurrent. Then from [40] and Theorem 1 of Ross [32] we have that

$$\begin{aligned} g^f(q) &= \lim_{N\to\infty} \mathbb{E}\big[\frac{\sum_{m=0}^{M-1} h(Q[m], f(Q[m]))}{\sum_{m=0}^{M-1} \tau(f(Q[m]))} \Big| Q[0] = q \big] \\ &= \lim_{N\to\infty} \mathbb{E}\big[\frac{\sum_{n=0}^{N-1} Q_s[n]}{N} \Big| Q_s[0] = q \big] = \phi^f(q). \quad (31) \end{aligned}$$

We note that Proposition 1 of [37] does not assume that $\tau(s)$ is bounded. Then from [37, Proposition 1] we have that $g \leq \phi^\gamma(q)$ for any policy γ. Hence, for the policy f, $g^f(q) = g = \phi^f(q) \leq \phi^\gamma(q)$ for any policy γ. Therefore f, a stationary policy, is average cost optimal, with average cost g.

Now consider (22), note that for any policy $\tilde{f}$ which minimizes the RHS, we have that $g = g^{\tilde{f}}(q) = \phi^{\tilde{f}}(q) \leq \phi^\gamma(q)$ which shows that $\tilde{f}$ is average cost optimal. Thus we have shown that (i) the average queue length is finite and does not depend on q_0, furthermore the two definitions of the average queue length coincide for the optimal policy, and ii) the optimal policy prescribes a batch size which is the minimizer for the RHS of the ACOE. □

The properties of $V_\alpha(q)$ are used, to verify the hypotheses in the above theorem in the following proposition.

Proposition VI.5. *If $\lambda < r_{max}$, then the sufficient conditions in Theorem VI.4 hold true.*

Proof. A sufficient condition for (1) is that $V_\alpha(q)$ is non-decreasing in q. Therefore, from Lemma VI.2, (1) holds.

A sufficient condition for (2) to hold is that for some policy γ, starting from any state q, the expected total cost and the expected total time until state 0 is reached are both finite and $A_s[n] \leq A_{max}$. Let us denote the service batch size prescribed by the policy γ by $s(q)$. If the policy γ is c-regular, with the c function being $h(q, s(q))$, then starting from any state q, the expected total cost and the expected total time until state 0 is reached are both finite. In Appendix III, we have proved that if $\lambda < r_{max}$ then there exists a candidate policy γ, which serves $s(q) = \min(q, s_1)$, which is c-regular. Therefore we have verified that (2) holds. Thus we have proved that if $\lambda < r_{max}$ then the hypotheses in Theorem VI.4 hold. □

Appendix VII
Proofs for Section IV

A. Proof of Lemma IV.1

We will use some of the ideas from the proof of Proposition III.2. Let $L(q) = q^2$, and consider the candidate policy γ, with batch size $s(q) = \min(q, s_1)$, where $\frac{s_1}{\tau(s_1)} > \lambda$. However, here we will consider the uniformized Markov chain $\tilde{Q}[m]$ (see Appendix IV), with single stage cost $\tilde{h}(q, s) = q + \big(\tau(s(q)) - 1\big)\lambda/2$. Then as in Proposition III.2, it can be shown that for the policy γ,

$$\mathbb{E}\big[L(\tilde{Q}[m+1]) - L(\tilde{Q}[m]) \Big| \tilde{Q}[m] = q\big] \leq -\tilde{h}(q, s(q)) + K_2,$$

where K_2 is a finite constant. Taking expectations conditional on $\tilde{Q}[0] = q_0$, we have that

$$\begin{aligned} &\mathbb{E}[L(\tilde{Q}[m+1])|\tilde{Q}[0] = q_0] - \mathbb{E}[L(\tilde{Q}[m])|\tilde{Q}[0] = q_0] \\ &\quad \leq -\mathbb{E}\big[\tilde{h}(\tilde{Q}[m], s(\tilde{Q}[m]) \Big| \tilde{Q}[0] = q_0\big] + K_2. \quad (32) \end{aligned}$$

For the policy γ, let T be the random time of first hitting the state 0 for the policy γ starting from q_0. Consider the inequalities generated from (32) by varying m from 0 to $T-1$. Addition of these inequalities yields

$$\begin{aligned} &\mathbb{E}\big[L(\tilde{Q}(T)) \Big| \tilde{Q}[0] = q_0\big] - L(q_0) \leq \\ &-\sum_{m=0}^{T-1} \mathbb{E}\big[\tilde{h}(\tilde{Q}[m], s(\tilde{Q}[m]) \Big| \tilde{Q}[0] = q_0\big] + TK_2. \end{aligned}$$

Taking expectation with respect to the random variable T, we have that

$$\mathbb{E}\Big[\sum_{m=0}^{T-1}\mathbb{E}\big[\tilde{h}(\tilde{Q}[m],s(\tilde{Q}[m]))\Big|\tilde{Q}[0]=q_0\big]\Big]\leq L(q_0)+K_2\mathbb{E}T.$$

We denote the above expectation by $c_{q_0,0}$ and note that $c_{q_0,0}$ is the expected total cost starting from q_0, until 0 is reached for the first time.

We now verify that the hypotheses of [41, Theorem 2] to show that $\mathbb{E}T = O(q_0)$. The random time T is the hitting time of the set of states $\Delta = \{0\}$. We choose the function $g(q)$, in [41, Theorem 2] as q. For the candidate policy we then have that $\mathbb{E}\big[g(Q[m+1]) - g(Q[m])|Q[m] = q\big] < -\epsilon$ for all $q \geq s_1$. Furthermore for any $q < s_1$, we have that $\mathbb{E}\big[g(Q[m+1]) - g(Q[m])|Q[m] = q\big] \leq \lambda\tau(s_1) < \infty$. Also note if there is positive probability of zero arrivals, then $Pr\{Q[m+1] = 0|Q[m] = q\} > 0$, for any $q < s_1$. Thus all three hypotheses in [41, Theorem 2] hold true, and therefore $\mathbb{E}T = O(q_0)$.

Now we note that the optimal relative value function $\tilde{J}(q)$ is the minimum total cost starting from any state q till the state 0 is reached, but for a MDP with a single stage cost $\tilde{h}(q,s)-g^*$. Therefore $\tilde{J}(q) \leq c_{q_0,0} - g^*\mathbb{E}T \leq L(q)+\mathbb{E}T\big(K_2-g^*\big)$. Hence $\tilde{J}(q) = \mathcal{O}(q^2)$, since $L(q) = q^2$.

B. Proof of Proposition IV.2

For deriving this lower bound we consider the uniformized system described in Appendix IV. We consider the average queue length for any policy γ serving a batch of size $s(q)$ at a queue length q. The average queue length $g^\gamma = \mathbb{E}\tilde{h}(\tilde{Q}, s(\tilde{Q}))$, where $\tilde{h}(q, s(q))$ is defined in (12). We note that $\tilde{\pi}(q) = \frac{\pi(q)\tau(s(q))}{\mathbb{E}_\pi\tau(s(Q))}$. The time average number of customers in the transmitter is always greater than the time average number of customers in service, so that $\mathbb{E}\tilde{Q} \geq \mathbb{E}s(\tilde{Q})$. For brevity, let $\tilde{S} = s(\tilde{Q})$. From (15) we obtain that $\mathbb{E}\frac{\tilde{S}}{\tau(\tilde{S})} = \lambda$. Note that at $s = 0$, $\frac{s}{\tau(s)} = 0$ for any value for $\tau(0)$. Also $\tau(s) \geq as + b$. Hence we have that $\mathbb{E}\frac{\tilde{S}}{a\tilde{S}+b} \geq \lambda$. Now we note that the function $\frac{s}{as+b}$ is a concave function of s. Using Jensen's inequality we obtain that $\frac{\mathbb{E}\tilde{S}}{a\mathbb{E}\tilde{S}+b} \geq \lambda$, or $\mathbb{E}\tilde{S} \geq \frac{b\lambda}{1-a\lambda}$. Now we note that $\mathbb{E}\tau(\tilde{S}) \geq \mathbb{E}(a\tilde{S}+b) - \tilde{\pi}_s(0)(b-1)$, where $\tilde{\pi}_s(0)$ can be interpreted as the fraction of time the transmitter does not transmit any messages. We note that $\tilde{\pi}_s(0) \leq 1$, so that $\mathbb{E}\tau(\tilde{S}) \geq \mathbb{E}(a\tilde{S} + b) - (b-1) = a\mathbb{E}\tilde{S} + 1$. Hence $\mathbb{E}\tau(\tilde{S}) \geq \frac{ab\lambda}{1-a\lambda} + 1$. Now since $\mathbb{E}\tilde{h}(\tilde{Q}, s(\tilde{Q})) = \mathbb{E}\big[\tilde{Q} + \frac{\lambda}{2}\big(\tau(\tilde{S}) - 1\big)\big]$, using the above lower bound and $\mathbb{E}\tilde{S} \geq \frac{b\lambda}{1-a\lambda}$ we obtain the lower bound $\frac{b\lambda}{2(1-a\lambda)}(2 + a\lambda)$ on g^γ.

To obtain the second term in the $\max$ expression, we use $\mathbb{E}\tau(\tilde{S}) \geq \frac{ab\lambda}{1-a\lambda}+1$. But instead of bounding $\mathbb{E}\tilde{Q}$ by $\mathbb{E}\tilde{S}$ below, we will use (17), and then show that $\mathbb{E}\tilde{Q} \geq \frac{a\sigma^2}{2(1-a\lambda)}$. Hence we obtain that $g^* \geq \frac{a\sigma^2}{2(1-a\lambda)} + \frac{ab\lambda^2}{2(1-a\lambda)}$. Combining the two lower bounds we obtain that $g^* \geq \max\big\{\frac{b\lambda}{2(1-a\lambda)}(2+a\lambda), \frac{a\sigma^2}{2(1-a\lambda)} + \frac{ab\lambda^2}{2(1-a\lambda)}\big\}$.

C. Proof of Lemma IV.3

Consider any stationary deterministic policy for system (B) (note that there is an optimal policy within the set of stationary deterministic policies for system (B)). Let us denote this policy by π. We now find an equivalent policy π_{eq} for system (A) which is non-stationary and which has average queue length value for system (A) which is at most what π has for system (B). At slot 0 with the same initial queue length for systems (A) and (B), π_{eq} chooses the same batch size s_0 as π. Note that since $\tau_A(s_0) \leq \tau_B(s_0)$ the first batch for system (A) would finish service earlier than the first batch for system (B). Policy π_{eq} however then idles for an extra $\tau_B(s_0) - \tau_A(s_0)$ slots. Then we note that the next decision epoch in both systems are the same. Also the queue evolution in the system is the same under the same sample path of arrivals into both systems (A) and (B). We repeat the above procedure for π_{eq} at every decision epoch. Then it is clear that the average queue length for π_{eq} is less than that of π. We note that the minimum average queue length for system (A), i.e., g^*_A is less or equal to that of π_{eq}. It follows that if π is the optimal policy for system (B), then the minimum average queue length for (A) is less than or equal to that of (B).

D. Proof of Proposition IV.4

For exhaustive service policy, we have that $Q = s(Q)$ and $\tau(s) = \lceil as+b \rceil, \forall s$. We note that $\tau(s) \leq as + b + 1$ As $\mathbb{E}Q < \infty$, we have that $\mathbb{E}Q = \lambda\mathbb{E}\tau(Q)$ from (14). Then, we obtain that $\mathbb{E}Q \leq \frac{(b+1)\lambda}{1-a\lambda}$. From (16) we have that

$$\begin{aligned}
\mathbb{E}Q^2 &= \sigma^2\mathbb{E}\tau(Q) + \lambda^2\mathbb{E}\tau(Q)^2, \\
&\leq \sigma^2(a\mathbb{E}Q + b + 1) + a^2\lambda^2\mathbb{E}Q^2 + \\
&\quad \lambda^2\big((b+1)^2 + 2a(b+1)\mathbb{E}Q\big) \text{ or,} \\
\mathbb{E}Q^2 &\leq \frac{1}{(1-a^2\lambda^2)}\big\{(b+1)\sigma^2 + (b+1)^2\lambda^2 + \\
&\quad \frac{(b+1)\lambda}{1-a\lambda}(a\sigma^2 + 2a(b+1)\lambda^2)\big\}.
\end{aligned}$$

We have that, for EXH, the average queue length is bounded above by $\frac{\mathbb{E}Q(aQ+b+1)+\mathbb{E}\frac{(aQ+b)(aQ+b+1)\lambda}{2}}{\mathbb{E}\tau(Q)}$. We substitute the upper bounds obtained above for $\mathbb{E}Q^2$, and $\mathbb{E}Q$ in the following upper bound and proceed to lower bound the $\mathbb{E}\tau(Q)$ in the denominator as follows. We note that $\mathbb{E}\tau(Q) \geq aQ + 1$. Then, $\mathbb{E}Q \geq \lambda a\mathbb{E}Q + 1$, or $\mathbb{E}Q \geq 1/(1-a\lambda)$. Therefore, $\mathbb{E}\tau(Q) \geq 1/(1-a\lambda)$. Then we obtain the upper bound

$$(b+1)\Big(\frac{a\sigma^2}{2(1-a\lambda)} + \frac{3(b+1)\lambda}{2(1-a\lambda)} + \frac{a\sigma^2}{2(1-a^2\lambda^2)} - \frac{\lambda}{2}\Big).$$

We note that the above lower bound for $\mathbb{E}\tau(Q)$ in the above expression can be motivated as follows. Since $\tau(q) \geq aq + b, \forall q > 0$ we can write $\mathbb{E}\tau(Q) \geq a\mathbb{E}Q + b - \pi_0(b-1)$ where π_0 is the stationary probability of $Q = 0$. Assuming that $b \geq 1$, we use $\pi_0 \leq 1$ in order to obtain $\mathbb{E}\tau(Q) \geq a\mathbb{E}Q + 1$. But, for sufficiently large λ we approximate $\tau(s) \leq as + b$ since $\pi_0 \downarrow 0$. Then, $as + b \leq \tau(s)$ we obtain that, $\mathbb{E}Q \geq \frac{b\lambda}{1-a\lambda}$ and $\mathbb{E}\tau(Q) \geq \frac{b}{1-a\lambda}$. This is the approximation that is used in deriving the second expression in Proposition IV.4.

E. Relationship to bounds derived in [24]

We now show that under a limiting regime where the slot durations are made infinitesimally small, our bounds are the same as those in [24]. We assume that D and k are integers. Let δ be such that $b\delta = D$ and $a\delta = k$, where a and b are integers. Then if we assume that each slot in our model is of δ duration, then the time taken for service of s customers in our model and the model in [24] is the same. In the following, we compare the bounds obtained in Propositions IV.2 and IV.4, with the bounds in [24], when $\delta \downarrow 0$. We denote the arrival rate of the Poisson message symbol arrival process in Musy's model by λ_m. Now assume that the arrival process in our model is Bernoulli with an arrival probability (rate) of $\lambda = \lambda_m\delta$ in each slot. We note that if a and b are integers, then $\tau(s) = as + b$ for every $s > 0$. Then the upper bound in Proposition IV.4 can be shown to simplify to

$$\frac{a\sigma^2}{2(1-a\lambda)} + \frac{3b\lambda}{2(1-a\lambda)} + \frac{a\sigma^2}{2(1-a^2\lambda^2)} - \frac{\lambda}{2}. \tag{33}$$

We note that $\sigma^2 = \lambda - \lambda^2$ for a Bernoulli arrival process. Substituting $\lambda = \lambda_m\delta$, $a = \frac{k}{\delta}$ and $b = \frac{D}{\delta}$ and taking the limit as $\delta \downarrow 0$ (along a sequence such that a and b are integers) we obtain that the average queue length in the limit is

$$\frac{3D\lambda_m}{2(1-\lambda_m k)} + \frac{\lambda_m k(2+\lambda_m k)}{2(1-\lambda_m k)}. \tag{34}$$

We note that this is the same as that obtained by Musy in Section 2.3.2 of his thesis [25].

Appendix VIII
Proofs for Section V

A. Proof of Proposition V.1

We obtain an upper bound by identifying a system (B) with a $\tau_B(s) \geq cs + d + \sqrt{es+f}$ and with $\tau_B(s)$ being affine. We note that for any integer s_0 the tangent to the curve $cs + d + \sqrt{es+f}$ is an upper bound to the curve. For any s_0 this tangent is chosen as a $\tau_B(s)$; we have that

$$\tau_B(s) = \tilde{a}s + \tilde{b},$$

where $\tilde{a} = c + \frac{e}{2\sqrt{f+es_0}}$ and $\tilde{b} = d + \frac{2f+es_0}{2\sqrt{f+es_0}}$. From Lemma IV.3 we then have that the minimum average queue length is bounded above by the minimum average queue length for system (B) and further bounded above by the bound given for the EXH policy in Proposition IV.4. This leads us to define $g_{ub}(s_0)$ as

$$\left(\frac{\tilde{b}+1}{\tilde{b}}\right)\left(\frac{\tilde{a}\sigma^2}{2(1-\tilde{a}\lambda)} + \frac{3(\tilde{b}+1)\lambda}{2(1-\tilde{a}\lambda)} + \frac{\tilde{a}\sigma^2}{2(1-\tilde{a}^2\lambda^2)} - \frac{\lambda}{2}\right).$$

The minimum average queue length is the least amongst all such B-systems, which leads to

$$g^* \leq \min_{\left\{s_0 : g_{ub}(s_0) \geq 0\right\}} g_{ub}(s_0).$$

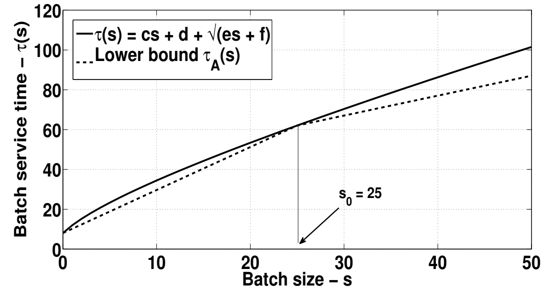

Fig. 23: An illustration of $\tau(s) = cs + d + \sqrt{es+f}$ and the lower bound $\tau_A(s)$ for $c = 1, d = 1, e = 50$, and $f = 50$. The parameter $s_0 = 25$.

B. Proof of Proposition V.2

We obtain a lower bound by first identifying a system (A) with $\tau_A(s) \leq cs + d + \sqrt{es+f}$. We define $\tau_A(s)$ (see Figure 23) as the minimum of two lines that bound $cs + d + \sqrt{es+f}$ from below.

For any $s_0 \in \mathbb{Z}_+$ we define $\tau_A(s) = \min(\tau_1(s), \tau_2(s))$, where

$$\begin{aligned}
\tau_1(s) &= \big(c + \frac{\sqrt{es_0+f} - \sqrt{f}}{s_0}\big)s + \big(d + \sqrt{f}\big), s \leq s_0 \\
\tau_2(s) &= cs + \big(d + \sqrt{es_0+f}\big), s > s_0.
\end{aligned}$$

Now consider the uniformized queueing system with $\tau_A(s)$ as defined above - but with an additional ceiling operation. We note that if a λ is stabilizable under the original system then it is stabilizable for system A. We therefore have that $\mathbb{E}\big[\frac{\tilde{S}}{\tau_A(\tilde{S})}\big] \geq \lambda$ (since $\tau_A(s) \leq \lceil \tau_A(s) \rceil$). Since $\frac{s}{\tau_A(s)}$ is concave we have that $\frac{\mathbb{E}\tilde{S}}{\mathbb{E}\tau_A(\tilde{S})} \geq \lambda$, or that $\mathbb{E}\tilde{S} \geq \lambda\mathbb{E}\tau_A(\tilde{S})$. We have that

$$\begin{aligned}
\mathbb{E}\tau_A(\tilde{S}) &\geq \Pr\big\{\tilde{S} \leq s_0\big\}\big[\mathbb{E}\big[\tau_1(\tilde{S})|\tilde{S} \leq s_0\big]\big] + \\
&\quad \Pr\big\{\tilde{S} > s_0\big\}\big[\mathbb{E}\big[\tau_2(\tilde{S})|\tilde{S} > s_0\big]\big], \\
&\geq c\mathbb{E}\tilde{S} + d + \sqrt{f} + \Pr\big\{\tilde{S} > s_0\big\}\big(\sqrt{es_0+f} - \sqrt{f}\big).
\end{aligned}$$

We now obtain a lower bound on $\Pr\big\{\tilde{S} > s_0\big\}$. Since $\mathbb{E}\big[\frac{\tilde{S}}{\tau_A(\tilde{S})}\big] \geq \lambda$, we have that

$$\begin{aligned}
Pr\big\{\tilde{S} \leq s_0\big\}\frac{s_0}{\tau(s_0)} + \frac{1}{c}Pr\big\{\tilde{S} > s_0\big\} &\geq \lambda, \\
\frac{s_0}{\tau(s_0)} + Pr\big\{\tilde{S} > s_0\big\}\big(\frac{1}{c} - \frac{s_0}{\tau(s_0)}\big) &\geq \lambda, \\
Pr\big\{\tilde{S} > s_0\big\} &\geq \frac{\lambda - s_0/\tau(s_0)}{1/c - s_0/\tau(s_0)},
\end{aligned}$$

which gives a positive lower bound for any s_0 such that $\lambda > s_0/\tau(s_0)$.

So for any s_0 such that $\lambda > s_0/\tau(s_0)$, we have that

$$\begin{aligned}
\mathbb{E}\tau_A(\tilde{S}) &\geq c\mathbb{E}\tilde{S} + d + \sqrt{f} + \\
&\quad \big(\sqrt{es_0+f} - \sqrt{f}\big)\frac{\lambda - s_0/\tau(s_0)}{1/c - s_0/\tau(s_0)}.
\end{aligned}$$

Since $\mathbb{E}\tilde{S} \geq \lambda\mathbb{E}\tau_A(\tilde{S})$ we have that

$$\mathbb{E}\tilde{S} \geq \lambda c\mathbb{E}\tilde{S} + \tilde{C},$$

where

$$\tilde{C} = \lambda\big(d + \sqrt{f} + \frac{\big(\sqrt{es_0 + f} - \sqrt{f}\big)\big(\lambda - \frac{s_0}{\tau_A(s_0)}\big)}{\frac{1}{c} - \frac{s_0}{\tau_A(s_0)}}\big)$$

Therefore, we have that

$$\begin{aligned} \mathbb{E}\tilde{S} &\geq \frac{\tilde{C}}{1 - c\lambda}, \text{ and,} \\ \mathbb{E}\tau_A(\tilde{S}) &\geq \frac{c\tilde{C}}{1 - c\lambda} + \frac{\tilde{C}}{\lambda}. \end{aligned}$$

Now, as in the proof of Proposition IV.2 we have that the minimum average queue length is bounded below by

$$\mathbb{E}\tilde{S} + \frac{\lambda}{2}\big(\mathbb{E}\tau_A(\tilde{S}) - 1 - (d + \sqrt{f} - 1)\big).$$

Substituting we have the lower bound

$$\begin{aligned} & \frac{\tilde{C}}{1 - c\lambda} + \frac{\lambda}{2}\big(\frac{c\tilde{C}}{1 - c\lambda} + \frac{\tilde{C}}{\lambda} - 1 - (d + \sqrt{f} - 1)\big), \\ = \; & \frac{\tilde{C}(1 + c\lambda/2)}{1 - c\lambda} + \frac{\tilde{C}}{2} - \frac{\lambda}{2}(d + \sqrt{f}). \end{aligned}$$

Since the above lower bound holds for any s_0 such that $s_0/\tau_A(s_0) < \lambda$ we therefore have that

$$g^* \geq \max_{\{s_0 : s_0/\tau_A(s_0) < \lambda\}} \big\{\frac{\tilde{C}(1 + c\lambda/2)}{1 - c\lambda} + \frac{\tilde{C}}{2} - \frac{\lambda}{2}(d + \sqrt{f})\big\}.$$