Asymptotic Queuing Analysis for Dynamic Spectrum Access Networks in the Presence of Heavy Tails

Pu Wang, *Student Member, IEEE,* and Ian F. Akyildiz, *Fellow, IEEE*

***Abstract*—The heavy tailed nature exhibited in both primary and secondary users' traffic fundamentally challenges the performance limit of dynamic spectrum access (DSA) networks under the conventional light tailed assumptions. This paper provides an asymptotic analysis of the steady-state queue length distribution of secondary users (SUs) under the heavy tailed network environment. Specifically, two network scenarios are investigated. In the first scenario where each SU has its exclusive access to a primary user (PU) channel, it is shown that the heavy tailed nature of either the PU traffic or the SU traffic can make SUs experience heavy tailed queue length with unbounded moments. In the second scenario where multiple SUs share a single PU channel, the queuing performance under throughput optimal scheduling policies is studied. It is proven that if the PU traffic has a heavier tail than any SU traffic, the queue length of each SU is at least one order heavier than the PU traffic under any scheduling policy. Otherwise, if the traffic from at least one of the SUs has a heavier tail than the PU traffic, it is proven that the celebrated throughput-optimal maximum weight scheduling leads to the worst possible asymptotic queuing performance for SUs by letting each SU queue have the heaviest possible tail. On the contrary, it is shown that there always exists a feasible set of β parameters such that the maximum weight-β scheduling yields the best asymptotic performance for the SU queues by letting each queue have the lightest possible tail.**

***Index Terms*—Heavy tail, queuing analysis, dynamic spectrum access**

I. Introduction

DYNAMIC spectrum access (DSA) is an emerging technique that allows the secondary users (SUs) to share the spectrum in an opportunistic manner [1]. Using such scheme, the SUs can access the unoccupied spectrum during idle periods of the primary users (PUs), and stop transmissions when the PU channels become busy. The achievable Quality of Service (QoS) performance of cognitive users is significantly affected by the dynamically changing PU traffic and the access policies used by the SUs.

Heavy tailed distributions have been widely observed in the current data-oriented communication networks. Specifically, the file size on the Internet servers, the web access pattern, and the scene length distribution of VBR (variable bit rate) and MPEG video streams have shown heavy tailed statistical characteristics [10]. In the classical communication network paradigms, the heavy tail distributions have been shown to have a non-negligible impact on the network performance in terms of network throughput, queue stability, and system scalability.

So far, the majority of research in dynamic spectrum access networks focuses on the development of the resource allocation and spectrum management schemes under the assumption of the light tailed behavior of primary and secondary users. Contrary to this conventional assumption, significant empirical evidence establishes that both PU and SU traffic can actually exhibit the heavy tailed nature. As for the primary users, it is shown that the call holding time of mobile users in 3G cellular networks and the session duration of licensed users in WLANs show heavy tailed statistics [15][8]. On the other hand, the emerging applications such as mobile internet, multimedia surveillance, video conferencing, and online gaming require secondary users to support internet and multimedia traffic, which is inherently bursty and exhibits heavy tailed nature. In spite of its importance, the performance limits of DSA network in the presence of the heavy tailed traffic is still an under-explored area, which, however, can fundamentally challenge the applicability and effectiveness of DSA scheme. For example, recent research shows that such heavy tailed behavior not only has a significant impact on the spectrum sensing performance [15], but also induces unbounded transmission delay moments for the SUs, which is certainly unfavorable for many delay-sensitive applications [13].

In this paper, we analyze the asymptotic tail behavior of the queue length for the SUs in the DSA networks under the heavy tailed network environment. Towards this, we consider a cognitive radio network in which multiple SUs opportunistically exploit the spectrum holes of a PU channel. The PU channel is modeled by an alternating renewal process, which alternates between busy periods $\{B_i\}_{i\geq 1}$ and idle periods $\{I_i\}_{i\geq 1}$. Each SU is associated with an input queue and a message arrives to the queue at each time slot with a certain probability. Upon the arrival of a message with random size $L > 0$, the SU first splits it into multiple packets with constant size. At each time slot, one of the SUs can be scheduled to transmit one packet provided that the PU channel is currently detected idle. Apparently, under such generic settings, the queuing performance for the SUs has a close relationship with the

Manuscript received 15 April 2012; revised 1 September 2012. This work was supported by the US Army Research Office (ARO) as MURI (Multidisciplinary University Research Initiative) under Grant No. W911NF-08-1-0238. The preliminary version was presented at IEEE International Dynamic Spectrum Access Networks symposium (DySPAN 2011), [13].

Pu Wang and Ian F. Akyildiz are with the Broadband Wirless Networking Laboratory, School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, GA, 30332 (e-mail: pwang40@ece.gatech.edu and ian@ece.gatech.edu).

Digital Object Identifier 10.1109/JSAC.2013.130316.

message size, the PU channel availability, and the scheduling policies. For the detailed description of this model, see Section III.

We first study the queue length asymptotics of the SU when it has exclusive access to the PU channel without competing with other SUs. Specifically, it is shown that if either the busy time or message size is heavy tailed, then the steady state queue length is one order heavier than the one with the heavier tail. This result implies that the expected queue length of the SUs can be infinite even through both the SU's transmitting messages and the PU busy periods are of finite mean size. We next investigate the asymptotic performance of the queue length under the general work conserving scheduling policies, where all the detected idle time of the PU channel are occupied for the SUs' transmissions unless the SUs have empty queues. It is proven that if the PU busy time has a heavier tail than the input traffic of any SU, then the asymptotic performance for all the SU queues is insensitive to the choice of a scheduling policy. In this case, under any work conserving policy, the tail distribution of the queue length of any SU is always one order heavier than that of the PU busy time. In contrary, if at least one of the SU queues has the input traffic with a heavier tail than the PU busy time, only the queue fed by the traffic with the heaviest tail exhibits the asymptotic behavior independent of the choice of a scheduling policy.

As the subclasses of the work-conserving scheduling policies, the maximum-weight scheduling and many of its variants [11] are known to be throughput optimal by stabilizing the queuing system for every supportable set of traffic arrival rates. The maximum-weight scheduling makes scheduling decision based on queue lengths, while the maximum-weight-β scheduling associates each queue with a different parameter β and makes the scheduling decision based on the queue lengths raised to the β-th power. Our asymptotic queuing analysis shows that the maximum-weight scheduling leads to the worst possible asymptotic performance for the SU queues by letting each queue have the heaviest possible tail. In contrary, it is shown that there always exists a feasible set of β parameters such that the maximum weight-β scheduling yields the best asymptotic performance for the SU queues by letting each queue have the lightest possible tail. In this case, the maximum-weight-β schedule promises the optimality in terms of maximizing the orders of the finite moments of the queue length.

The rest of this paper is organized as follows. Section II summarizes the related work. Section III introduces system model and preliminaries. Section IV presents the asymptotic queue length analysis for the exclusive access policies. The asymptotic queuing performance of the maximum-weight-β scheduling is presented in Section V. Finally, Section VI concludes this paper.

II. Related Work

Although the queue length distribution is a key factor affecting the QoS performance in wireless networks, the asymptotic queuing analysis for cognitive radio networks is still scarce to the best of our knowledge. In [7] and [14], the queuing delay of SUs in a multi-channel cognitive network was investigated with different objectives. Specifically, using large deviation approximation, [7] aimed to analyze the stationary queue distribution of SUs under the Markov chain based PU traffic model. On the contrary, [14] studied the moments of the SUs' queue length under the PU traffic modeled as an alternating ON/OFF process, where the ON periods follow a general distribution and the OFF periods are exponentially distributed. Different form [7] and [14], which consider the queuing performance under the light tailed PU and SU traffic, we aim to investigate the asymptotic behavior of the SU's queue length in the presence of the heavy tailed PU and SU traffic and study the effectiveness of the queue-length based scheduling policies on the asymptotic queuing performance. To the best of our knowledge, little work on the analysis of such queue length performance has been done for cognitive radio networks.

It is worth to note that [6] and [9] are among the first research efforts to study the performance of the maximum weight-β in the queuing network and show that the maximum weight-β is effective to mitigate the impact of the queue with heavy-tailed traffic on the other queue with light-tailed traffic. Different from [6] which consider a queuing system of two users competing a single channel, we study the maximum weight-β scheduling for an arbitrary number of SUs which dynamically access a PU channel with heavy tailed behavior. In this case, SU can only access PU channel when it is detected idle and if a miss detection happens, SU needs to retransmit the collided/loss packet. Apparently, the existing literature on heavy tails do not consider such dynamic channel access schemes. What is more important, the existing work only consider the simple case with two queues and one server/channel. However, cognitive radio networks generally need to support much more SUs. This greatly complicates the queueing analysis because of the correlation and dependence of the queue length of multiple SUs. Thus, the existing results can not be applied.

III. System Model and Preliminaries

A. System Model

Consider N SUs sharing a PU channel, as shown in Fig.1. Time is slotted, with a unit slot length. Without loss of generality, we assume that the PU channel is of unit capacity and modeled by an alternating renewal process, which alternates between busy periods with length $\{B_i\}_{i\geq 1}$ and idle periods with length $\{I_i\}_{i\geq 1}$. $\{B_i\}_{i\geq 1}$ and $\{I_i\}_{i\geq 1}$ are mutually independent random sequences of i.i.d. random variables with distribution F_B and F_I, respectively. At each time slot, if the PU channel is detected idle, one of the SUs can be scheduled to transmit one packet per time slot. If the transmitted packet is collided with the PU transmission because of miss detection, the packet is retransmitted in the future. Assume that before the scheduling takes place, the PU channel detection result is available either through cooperative sensing or through the fusion center [1]. Let p_f denote the false alarm probability. By renewal theory, we have the service rate (throughput) of the PU channel as follows

$$\mu := \frac{(1-p_f)E[I_1]}{E[B_1]+E[I_1]}. \tag{1}$$

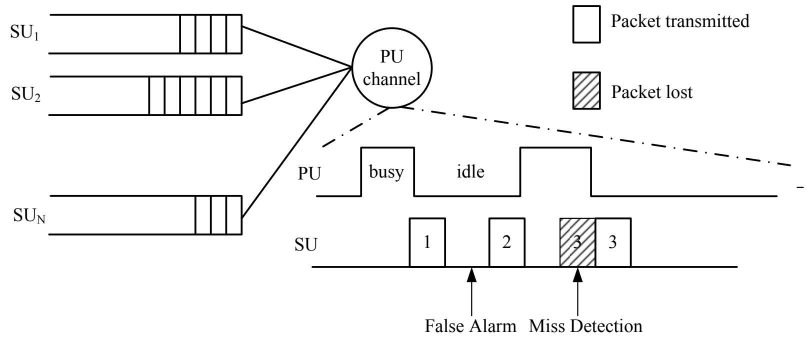

Fig. 1. System model.

Let q_i denote the queue associated with SU_i. In each time slot, a message arrives to the queue q_i with a probability λ_i. Let $L_i > 0$ denote the number of packets in the message that arrives to q_i. L_i is an independent and identically distributed (i.i.d.) random variable (r.v.) from slot-to-slot, and is independent of the channel states $\{B_i\}_{i\geq 1}$ and $\{I_i\}_{i\geq 1}$. Let $A_i(t)$ denote the number of packets that arrive during time slot t to q_i. Accordingly, the input rate Λ_i of the queue q_i is given by

$$\Lambda_i := E[A_i(t)] = \lambda_i E[L_i]. \tag{2}$$

We assume $\sum_{i=1}^{N} \Lambda_i < \mu$ so that the system is stable under any work-conserving scheduling policy, where every detected idle time slot of the PU channel is used for transmitting SUs' packets unless the SUs have empty queues. Let $Q_i(t)$ denote the queue length of q_i in time slot t. Let Q_i denote the steady-state queue length of q_i. The main objective of this paper is to study the asymptotic tail distribution of the steady-state length Q_i under throughput optimal scheduling polices.

B. Preliminaries

In this paper we use the following notations. For any two real functions $a(t)$ and $b(t)$, we let $a(t) \sim b(t)$ denote $\lim_{t\to\infty} a(t)/b(t) = 1$. We say that $a(t) \lesssim b(t)$ if $\limsup_{t\to\infty} a(t)/b(t) \leq 1$, and $a(t) \gtrsim b(t)$ if $\liminf_{t\to\infty} a(t)/b(t) \geq 1$. Furthermore, we say that $a(t) = o(b(t))$ if $\lim_{t\to\infty} a(t)/b(t) = 0$. In addition, for any two non-negative r.v.s X and Y, we say that $X \leq_{a.s.} Y$ if $X \leq Y$ almost surely, and $X \leq_{s.t.} Y$ if X is stochastically dominated by Y, i.e., $P(X > t) \leq P(Y > t)$ for all $t \geq 0$. We say $X \overset{d}{=} Y$ if X and Y are equal in distribution. Also, let $F_X(x) = P(X \leq x)$ denote the cumulative distribution function (cdf) of a non-negative r.v. X. Let $\overline{F}_X(x) = P(X > x)$ denote its tail distribution function.

Definition 1 *A r.v. X is heavy tailed (HT) if for all $\theta > 0$*

$$\lim_{x\to\infty} e^{\theta x}\overline{F}_X(x) = \infty, \tag{3}$$

or, equivalently, if for all $z > 0$

$$E[e^{zX}] = \infty. \tag{4}$$

Definition 2 *A r.v. X is light tailed (LT) if it is not heavy tailed or, equivalently, if there exists $z > 0$ such that*

$$E[e^{zX}] < \infty. \tag{5}$$

Remark 1 Generally speaking, a r.v. is HT if its tail distribution decreases slower than exponentially. On the contrary, a r.v. is LT if its tail distribution decreases exponentially or faster. Some typical HT distributions include Pareto and log-normal, while Some typical LT distributions cover exponential and Gamma.

Based on the existence of the moments, we define the tail index of a non-negative random variable.

Definition 3 *The tail index $\kappa(X)$ of a nonnegative random variable X is defined by*

$$\kappa(X) = \sup\{k \geq 0 : E[X^k] < \infty\}. \tag{6}$$

Remark 2 The tail index specifies the threshold order above which a random variable has infinite moments. Some HT distributions, such as Pareto, have finite tail index, which leads to infinite moments of certain orders, Some HT distribution, such as log-normal, have infinite tail index and therefore possesses finite moments of all orders. In this work, we focus on heavy tail distributed random variables with finite tail index because they can effectively characterize lots of network attributes such as the frame length of variable bit rate (VBR) traffic, the session duration of licensed users in WLANs, and files sizes on internet severs [8] [10].

The following Lemma presents the sufficient condition regarding the existence of finite tail index for a r.v. X [4].

Lemma 1 *A nonnegative r.v. X has $\kappa(X)$ if and only if the tail distribution of X satisfies*

$$\lim_{t\to\infty} \frac{\log[P(X > t)]}{\log t} = -\kappa(X). \tag{7}$$

An important subclass of HT distributions with finite tail index is the class of regularly varying distributions [2]. Its definition involves the slowly varying function which is defined as follows.

Definition 4 *A measurable positive function $\mathcal{L}(x)$ defined in some interval $[a, \infty)$ is called slowly varying if for all $y > 0$*

$$\lim_{x\to\infty} \frac{\mathcal{L}(yx)}{\mathcal{L}(x)} = 1. \tag{8}$$

For example, a constant and a logarithmic function are both slowly varying functions.

Definition 5 *A r.v. X is called regularly varying with tail index $\alpha > 0$, denoted by $X \in \mathcal{RV}(\alpha)$, if*

$$\overline{F}_X(x) \sim x^{-\alpha}\mathcal{L}(x), \tag{9}$$

where $L(x)$ is a slowly varying function.

Remark 3 Regularly varying distributions are a generalization of power law distributions. The tail index α indicates how heavy the tail distribution is, where smaller values of α imply heavier tail. Moreover, for a r.v. $X \in \mathcal{RV}(\alpha)$, the exact values of α determine whether the moments of X are bounded or not, that is if $X \in \mathcal{RV}(\alpha)$, then $\kappa(X) = \alpha$.

IV. Asymptotic Queue Length Analysis for Exclusive Access Policies

In this section, we first study the queue length asymptotics of the SU when there exists only one SU in the secondary network and thus it has exclusive access to the PU channel. Based on the derived queuing performance under this single SU scenario, we evaluate the queuing performance of multiple SUs accessing one PU channel (Theorem 2 - 4).

Let q_e denote the queue associated with the SU, $Q_e(t)$ the queue length at time slot t, and Q_e the steady-state queue length.

A. Main Theorem

Theorem 1 *Assume the SU message size $L \in \mathcal{RV}(\alpha_l)$ and the PU busy time $B_1 \in \mathcal{RV}(\alpha_b)$ and let $\alpha_l = \infty$ or $\alpha_b = \infty$ indicate that L or B_1 is light tailed. Then, the steady-state queue length Q_e of the SU satisfies*

$$\lim_{t\to\infty} \frac{\log[P(Q_e > t)]}{\log t} = -\min(\alpha_l, \alpha_b) + 1. \tag{10}$$

Remark 4 The preceding results establish the relationship between the tail asymptotics of the message size L, the PU busy time B_i, and the queue length Q_e. Specifically, if either the busy time or message size is heavy tailed, then the steady-state queue length is one order heavier than the one with the heavier tail. This result implies that the expected queue length of the SUs can be infinite even if both the SU's transmitting messages and PU busy periods are of finite mean size. For example, if the message size is LT and the PU busy time is heavy tailed with tail index $2 > \alpha_b > 1$, then both the message size and the PU busy time are finite. In this case, by Theorem 1, the steady-state queue length of the SU has a tail index $1 > \kappa(Q_e) > 0$, which implies that both the mean and variance of Q_e are infinite. Moreover, by Theorem 1, it is evident that the detection results of the PU channel have no impact on the asymptotic behavior of the queue length.

B. Proof of the Main Theorem

1) Fictitious Queues: To prove Theorem 1, we construct two fictitious queues, namely the slow queue $\widetilde{q_s}$ and the fast queue $\widetilde{q_f}$, which have the same packet arrivals, experience the same PU channel activities, and obtain the same PU channel detection results as queue q_e, but receive different services. Without loss of generality, for each queue, we assume that the first message arrives at the beginning of an idle period of the PU channel. As for the slow queue $\widetilde{q_s}$, the transmission of a new message always starts at the beginning of an idle period. This means that even if the transmission of the current message is over in the middle of an idle period, the transmission of the next message is not initiated until the next idle period begins. Thus, during the same time interval, less messages are served in the slow queue $\widetilde{q_s}$ than in the original queue q_e. This implies that

$$Q_s(t) \geq Q_e(t). \tag{11}$$

As for the fast queue $\widetilde{q_f}$, if the transmission of a message is finished in the middle of an idle period, we do not count this idle period in its service time so that each message waiting in the queue starts to be served from the beginning of the idle period during which the transmission of the previous message is finished. If a message arrives when the queue is empty, we consider two scenarios. (1) If it arrives at the beginning of an idle period, its service time will not include the idle period during which its transmission is finished. Otherwise, (2) if it arrives in the middle of an idle period, we treat this message as if it arrives at the beginning of the idle period. It is easy to verify that during the same time interval, more messages are served in the fast queue $\widetilde{q_f}$ than in the original queue q_e. This implies that

$$Q_f(t) \leq Q_e(t). \tag{12}$$

By (11) and (12), we obtain

$$P(Q_f(t) > t) \leq P(Q_e(t) > t) \leq P(Q_s(t) > t). \tag{13}$$

We will next prove Theorem 1 by showing that the lower and upper bounds in (13) asymptotically coincide. Towards this, we derive the tail asymptotics of the steady-state queue length for the slow queue $\widetilde{q_s}$ and the fast queue $\widetilde{q_f}$, respectively.

2) Queue Length Asymptotics of Queue $\widetilde{q_f}$:

Lemma 2 *Assume the SU message size $L \in \mathcal{RV}(\alpha_l)$ and the PU busy time $B_1 \in \mathcal{RV}(\alpha_b)$ and let $\alpha_l = \infty$ or $\alpha_b = \infty$ indicate that L or B_1 is light tailed. Then the steady-state queue length Q_f of the SU satisfies*

$$\lim_{t\to\infty} \frac{\log[P(Q_f > t)]}{\log t} = -\min(\alpha_l, \alpha_b) + 1. \tag{14}$$

To prove Lemma 2, we first define the transmission time of a message with size L in the fast queue $\widetilde{q_f}$. The construction of $\widetilde{q_f}$ indicates that the transmission attempt of a packet is always started at the beginning of an idle period. In addition, the last idle period during which the transmission is finished is excluded from the service time. Accordingly, we have the service time $T_f(L)$ of transmitting a message of size L in the queue $\widetilde{q_f}$ as follows. During an idle period I_i, let $e(j)$ denote the event that the PU channel is detected idle at time slot j and $\mathbf{1}_{e(j)}$ denote the indicator function of the event $e(j)$ where $\mathbf{1}_{e(j)} = 1$ iff the event $e(j)$ occurs.

Definition 6 *During an idle period with length I_i, the transmission time X_i of the SU is defined as*

$$X_i := \sum_{j=1}^{I_i} \boldsymbol{1}_{e(j)}, \tag{15}$$

the total number of idle periods the SU occupies for transmitting a message of size L,excluding the last idle period during which the transmission is finished, is defined as

$$M_f := \inf \left\{ m : \sum_{i=1}^{m} X_i \geq L \right\} - 1, \tag{16}$$

and for the fast queue $\widetilde{q}_f$, the total service (transmission) time $T_f(L)$ of a message of size L is defined as

$$T_f(L) := \sum_{i=1}^{M_f} \{I_i + B_i\}. \tag{17}$$

The tail asymptotics of the transmission time T_f is given by the following Lemma.

Lemma 3 *Assume that $B_i \in \mathcal{RV}(\alpha_b)$. If $L \in \mathrm{LT}$ or $L \in \mathcal{RV}(\alpha_l)$ with $\alpha_b < \alpha_l$, we have*

$$P(T_f(L) > t) \sim E[M_f]P(B_1 > t). \tag{18}$$

Assume that $L \in \mathcal{RV}(\alpha_l)$ and $E[L] < \infty$.

1) *If $B_i \in \mathrm{LT}$, we have*

$$P(T_f(L) > t) \sim P\left(L > \frac{E[X_1]}{E[I_1] + E[B_1]}t\right). \tag{19}$$

2) *If $B_i \in \mathcal{RV}(\alpha_b)$ with $\alpha_l < \alpha_b$, we have*

$$\begin{aligned} P(T_f(L) > t) \quad &\sim \quad \left(\frac{E[B_1]}{E[X_1]}\right)^{\alpha_l} P(L > t) \\ &\quad + \left(\frac{E[I_1]}{E[X_1]}\right)^{\alpha_l} P(L > t). \end{aligned} \tag{20}$$

3) *If $B_i \in \mathcal{RV}(\alpha_b)$ with $\alpha_l = \alpha_b$, we have*

$$\begin{aligned} P(T_f(L) > t) \quad &\sim \quad \left(\frac{E[I_1]}{E[X_1]}\right)^{\alpha_l} P(L > t) \\ &\quad + (E[B_1])^{\alpha_l} P(L > t) \\ &\quad + E[M_f]P(B_1 > t). \end{aligned} \tag{21}$$

Remark 5 From the above results, we see that the tail distribution of the message transmission time is as heavy as either the SU's message size or the PU busy time, whichever has the heavier tail.

Proof of Lemma 3: The proof is lengthy and relies on the large deviation theory. See [12] for details. ■

We are now ready to prove Lemma 2 regarding the tail asymptotics of queue length of the fast queue $\widetilde{q}_f$.

Proof of Lemma 2: Let Q_m denote the steady-state number of messages waiting in the queue. Thus, Q_m is actually the steady-state queue length of a $GI/G/1$ queue, with the message arrival rate λ and service time $T_s(L)$. Since each message i that arrives to the queue $\widetilde{q}_f$ consists of L_i packets, the steady-state queue length Q_f satisfies

$$\sum_{i=1}^{Q_m - 1} L_i \leq Q_f \leq \sum_{i=1}^{Q_m} L_i. \tag{22}$$

We next prove that the lower and upper bounds match asymptotically by considering the following three cases.

(1) If $B_i \in \mathcal{RV}(\alpha_b)$ and $L \in \mathrm{LT}$ or $L \in \mathcal{RV}(\alpha_l)$ with $\alpha_b < \alpha_l$, it follows by Lemma 3 that the service time $T_f(L) \in \mathcal{RV}(\alpha_b)$, which implies that $T_f(L)$ is subexponentially distributed. Let $\rho = \lambda E[T_f(L)]$ is the traffic intensity. By applying Theorem 1 in [3], the steady-state waiting time W_m of a message in the queue is given by

$$P(W_m > t) \sim \frac{\rho}{1-\rho} \int_t^{\infty} \frac{P(T_f(L) > x)}{E[T_f(L)]} dx, \tag{23}$$

which, by distributional Little's law and regular variation, yields

$$P(Q_m > t) \sim P(\lambda W_m > t) \sim \frac{\lambda^{\alpha_b + 1} E[M]}{(1-\rho)(\alpha_b - 1)} t P(B_1 > t). \tag{24}$$

This implies that $Q_m \in \mathcal{RV}(\alpha_b - 1)$. Combining (22) and (24), it follows from the sum property of random number of regularly varying random variables [5] that

$$P(Q_f > t) \sim \frac{\lambda^{\alpha_b + 1} E[M] E[L]^{\alpha_b}}{(1-\rho)(\alpha_b - 1)} t P(B_1 > t). \tag{25}$$

(2) If $L \in \mathcal{RV}(\alpha_l)$ with $E[L] < \infty$ and $B_i \in \mathrm{LT}$, by Lemma 3, we have $T_f(L) \in \mathcal{RV}(\alpha_l)$ and thus $T_f(L)$ is subexponentially distributed. By the similar arguments for the case (1), we have

$$P(Q_f > t) \sim \frac{\lambda^{\alpha_l + 1} \left((E[I_1] + E[B_1])E[L]\right)^{\alpha_l}}{(1-\rho)(\alpha_b - 1)E[X_1]^{\alpha_l}} t P(L > t). \tag{26}$$

(3) If $L \in \mathcal{RV}(\alpha_l)$ with $E[L] < \infty$ and $B_i \in \mathcal{RV}(\alpha_l)$ with $\alpha_l \leq \alpha_b$, this implies by Lemma 3 and the properties of slowly varying function that $T_f(L) \in \mathcal{RV}^{\alpha_l}$. By the similar arguments for the case (1) and (2), we have $T_f(L) \in \mathcal{RV}(\alpha_l - 1)$. ■

By the similar techniques, it can be shown that the steady-state queue length of slow queue Q_s has the same asymptotic performance as the fast queue Q_f. This indicates by (13) that the lower and upper bounds of the queue length Q_e coincide, which completes the proof of Theorem 1.

V. Asymptotic Queue Length Analysis for Throughput Optimal Scheduling Policies

In this section, we study the steady-state queue length asymptotics under throughput optimal scheduling policies. We first investigate the asymptotic queue length performance under the general work conserving scheduling policies. Our results show that the tailness of the PU traffic has a profound impact on the effectiveness of the scheduling policy. Then, we study the asymptotic queuing performance under maximum-weight-β scheduling and prove its optimality in terms of maximizing the orders of the finite moments of the queue length.

Theorem 2 *If $1 < \alpha_b < \min_{1 \leq i \leq N} \alpha_{l_i}$, then under any work conserving scheduling policy, the steady-state queue length Q_i of any SU $i \leq N$ is one order heavier than the PU busy period, i.e.,*

$$\kappa(Q_i) = \alpha_b - 1, \quad \forall 1 \leq i \leq N \tag{27}$$

Theorem 3 *Assume $\alpha_b \geq \min_{1 \leq i \leq N} \alpha_{l_i} > 1$. Let $\alpha^- := \min_{1 \leq i \leq N} \alpha_{l_i}$. Under any work-conserving policy, the steady-state queue length Q_i of the queue q_i with the smallest tail* *index $\alpha_{l_i} = \alpha^-$ follows*

$$\kappa(Q_i) = \alpha_{l_i} - 1 = \alpha^- - 1, \tag{28}$$

while the steady-state queue length Q_i of any other queue q_i with $\alpha_{l_i} > \alpha^-$ follows

$$\min(\alpha_{l_i}, \alpha_b) - 1 \geq \kappa(Q_i) \geq \alpha^- - 1. \tag{29}$$

Remark 6 We can see from Theorem 2 that if the PU busy time has a heavier tail than the input traffic of any SU, then the tail asymptotic of the queue length is insensitive to the choice of a scheduling policy. In this case, under any work conserving policy, the tail distribution of the SU queue length is always one order heavier than that of the PU busy time. In contrary, by Theorem 3, if at least one of the SU queues has the input traffic with a heavier tail than the PU busy time, only the queue fed by the traffic with the heaviest tail exhibits the asymptotic behavior independent of the choice of a scheduling policy, while the other queues have bounded asymptotic performance.

Proof of Theorem 2: Since the best scheduling scheme for a particular queue q_i is to let it receive service whenever the queue is not empty. In this case, q_i behaves as if it has exclusive access to the PU channel and no other queues compete for the service. Under this scheduling policy, q_i behaves like q_e. Thus, under any work conserving policy, we have $Q_i(t) > Q_e(t)$ and thus $P(Q_i > t) > P(Q_e > t)$, which, by Theorem 1 and the assumption $\alpha_b < \min_{1 \leq i \leq N} \alpha_{l_i}$, implies that the upper bound of the tail index of Q_i satisfies

$$\kappa(Q_i) \leq \min(\alpha_b, \alpha_{l_i}) - 1 \leq \alpha_b \ - 1. \tag{30}$$

Moreover, since $P(Q_i > t) \leq P(\sum_{i=1}^N Q_i > t)$, invoking Lemma 4, we have the lower bound of the tail index of Q_i, i.e.,

$$\kappa(Q_i) \geq \min(\min_{1 \leq i \leq N} \alpha_{l_i}, \alpha_b) - 1 \geq \alpha_b \ - 1, \tag{31}$$

which agrees with the upper bound and completes the proof. ■

Proof of Theorem 3: We first prove the asymptotic results in (28) regarding the queue q_i with the input process of the heaviest tail, i.e., the smallest tail index $\alpha_{l_i} = \arg\min_{1 \leq i \leq N} \alpha_{l_i}$. It is evident that the queue length Q_i is stochastically dominated by the composite queue length $\sum_{i=1}^N Q_i$, which, by Lemma 4 and the assumption $\alpha_{l_i} \leq \alpha_b$, proves the lower bound of (28), i.e., $\kappa(Q_i) \geq \alpha_{l_i} - 1$.

As to the upper bound, we consider the best scheduling policy for q_i, which allows q_i to receive the service whenever q_i is not empty. This policy yields the best asymptotic results for the queue q_i since q_i does not have to compete with other queues for the service and thus behaves like q_e. Invoking Theorem 1, it follows from the assumption $\alpha_{l_i} \leq \alpha_b$ that the lower bound in (28) holds, i.e., $\kappa(Q_i) \leq \alpha_{l_i} - 1$, which matches the upper lower and proves (28).

Using the similar arguments, we can prove (29) by showing that the tail asymptotics of the queue length Q_i are lower bounded by those of the composite queue length and upper bounded by those of the queue q_e. The details are omitted in the interest of brevity. ■

A. Maximum-Weight-β Scheduling

As shown in Theorem 2, if the PU busy time has a heavier tail than the input traffic of any SU, the asymptotic behavior of the SU queue is insensitive to the choice of the scheduling policy. More specifically, the queue length of each SU is one order heavier than the PU traffic under any scheduling policy. That is, it is impossible for any scheduling policy to either improve or deteriorate the tail asymptotic performance for the SUs.

Therefore, to investigate the effectiveness of the scheduling policy, in this section we assume that at least one of the SUs has the input traffic with a heavier tail than the PU busy time, i.e., $\alpha_b \geq \min_{1 \leq i \leq N} \alpha_{l_i}$.

In this section, we study the tail asymptotics of the steady-state queue length distribution under maximum-weight-β scheduling, which works as follows. For N queues $\{q_i\}_{1 \leq i \leq N}$, each queue q_i is assigned with a positive parameter β_i. During each time slot t, the queue q_i, which satisfies the condition

$$Q_i(t)^{\beta_i} = \max_{1 \leq j \leq N} Q_j(t)^{\beta_j} \tag{32}$$

wins the competition and one packet from this queue is served provided that the PU channel is detected idle. Ties are broken arbitrarily. If all parameters $\{\beta_i\}_{1 \leq i \leq N}$ are equivalent, maximum-weight-β scheduling becomes the conventional maximum-weight scheduling, where at each time slot, the largest queue is served.

The asymptotic analysis of the queue length distribution in this section shows that the well-known maximum-weight scheduling leads to the worst possible asymptotic behavior for the SU queues such that each queue can have a queue length with the heaviest possible tail, which indicates that as long as one SU queue has unbounded delay or variance, so do all SU queues. On the contrary, the maximum-weight-β scheduling is proven to yield the best asymptotic performance for the SU queues by letting each queue have the smallest possible tail. Consequently, the maximum-weight-β scheduling is asymptotically optimal because it can ensure the queue length has the same asymptotic performance as the exclusive access case, which is the best performance one can expect.

Theorem 4 *Let $\alpha_i := \min(\alpha_{l_i}, \alpha_b) > 1$, i.e., $E[L_i] < \infty$ and $E[B] < \infty$, and $\alpha^- := \min_{1 \leq i \leq N} \alpha_i$. Define*

$$\alpha_i^m = \min_{1 \leq j \leq N} \frac{\beta_i}{\beta_j}(\alpha_j - 1). \tag{33}$$

Under Maximum-Weight-β scheduling, the tail index of the steady-state queue length Q_i for SU i follows

$$\kappa(Q_i) = \max(\alpha_i^m, \alpha^- - 1). \tag{34}$$

Remark 7 (**Ineffectiveness of Maximum Weight Scheduling**) From the above results, we see that if all parameters $\{\beta\}_{1 \leq i \leq N}$ are equivalent, all queues have the same tail index as the heaviest queue which has the smallest tail index equal to $\min_{1 \leq j \leq N}(\alpha_j - 1)$. This implies that the maximum-weight scheduling leads to the worst possible tail asymptotics for the SU queues so that the queue length of each queue has the lowest orders of the finite moments. In this case, if among all queues, the queue q_i is fed by the traffic with the smallest tail index $1 < \alpha_{l_i} < 2$, then under maximum-weight scheduling, all the queues have the infinite mean steady state queue length.

Remark 8 (**Asymptotic Optimality of Maximum Weight-β Scheduling**) Theorem 4 indicates that by adjusting the parameters $\{\beta_i\}_{1\leq i\leq N}$, the maximum-weight-$\beta$ scheduling can lead to the best possible asymptotic queue length performance which is as good as that under the case where the queue has the exclusive access to the PU channel. To see this, recalling Theorem 3, the best tail performance (the largest tail index) of the queue length Q_i is that $\kappa(Q_i) = \alpha_i - 1 = \min(\alpha_{l_i}, \alpha_b) - 1$. Our objective is as follows.

$$\begin{aligned} &\text{Find} && \{\beta_i\}_{1\leq i\leq N} \\ &\text{Such that} && \alpha_i - 1 = \min_{1\leq j\leq N} \frac{\beta_i}{\beta_j}(\alpha_j - 1) \quad \forall 1 \leq i \leq N \end{aligned}$$

One feasible solution to the above optimization problem is given by

$$\beta_i = \frac{\alpha_i - 1}{\alpha^- - 1}, \quad \forall 1 \leq i \leq N. \tag{35}$$

The feasibility of this solution can be easily verified by inserting (35) into (33).

The proof of Theorem 4 relies on Lemma 4, 5 and 6, which we state and prove first.

Lemma 4 *Define $\alpha^- := \min_{1\leq i\leq N} \alpha_{l_i}$. We have*

$$\kappa\left(\sum_{i=1}^{N} Q_i\right) = \min(\alpha_b, \alpha^-) - 1. \tag{36}$$

Proof of Lemma 4: Consider a fictitious queue q_v which has the arrival process $A_v(t) = \sum_{i=1}^{N} A_i(t)$ and experiences the same PU channel as the original queuing system. Since $A_i(t) \in \mathcal{RV}(\alpha_{l_i})$, it implies by regular variation that the arrival process $A_v(t) \in \mathcal{RV}(\alpha^-)$, where $\alpha^- = \min_{1\leq i\leq N} \alpha_{l_i}$. Let Q_v denote the steady state queue length of q_v, It follows by Theorem 1 that $\kappa(Q_v) = \alpha^- - 1$. Let $Q_v(t)$ denote the queue length of q_v at time t. Under any work conserving policy in the original queuing system, we have $Q_v(t) = \sum_{i=1}^{N} Q_i(t)$. This implies that $Q_v = \sum_{i=1}^{N} Q_i$ and thus $\kappa(\sum_{i=1}^{N} Q_i) = \alpha^- - 1$. This completes the proof. ∎

Lemma 5 *Under Maximum-Weight-β scheduling, the tail index $\kappa(Q_i)$ of the steady-state queue length Q_i is lower bounded by*

$$\kappa(Q_i) \geq \alpha_i^m. \tag{37}$$

Proof of Lemma 5: For any $1 > \delta > 0$, we have

$$\begin{aligned} P(Q_i > t) &= P\left(Q_i > t \wedge \left\{\bigcap_{j\neq i} Q_j^{\frac{\beta_j}{\beta_i}} < \delta t\right\}\right) \\ &\quad + P\left(Q_i > t \wedge \left\{\bigcup_{j\neq i} Q_j^{\frac{\beta_j}{\beta_i}} \geq \delta t\right\}\right) \\ &:= I + II. \end{aligned} \tag{38}$$

As to the term I, it denotes the probability that the queue q_i has a queue length Q_i larger than t, when all the other queues have a queue length less than $(\delta t)^{\beta_i/\beta_j}$, i.e., $Q_j < (\delta t)^{\beta_i/\beta_j}$, $\forall j \neq i$. Without loss of generality, we assume that this event occurs at time 0, which means $Q_j(0)^{\beta_i/\beta_j} < \delta t$. Let $-\tau$ denote the last time when some of the queues $j \neq i$ receive service. We have two implications. (1) $Q_i(-\tau)^{\beta_i} < Q_j(-\tau)^{\beta_j}, \forall j \neq i$ since q_i did not receive service at time $-\tau$. (2) $Q_j(-\tau)^{\beta_j/\beta_i} < \delta t, \forall j \neq i$ since q_j did not receive service during the time interval $[-\tau + 1, 0]$. The two implications imply that $Q_i(-\tau) < \delta t$. Thus, to ensure $Q_i(0) > t$, the number of packets accumulated in q_i during the time interval $[-\tau + 1, 0]$ is at least larger than $(1 - \delta)t$, i.e., $\sum_{n=-\tau+1}^{0}(A_i(n) - C_i(n)) > (1 - \delta)t$, where $C_i(n)$ denotes the number of packets that depart from q_i at time n. Thus, we obtain the upper bound of I

$$\begin{aligned} I &\leq P\left(\sum_{n=-\tau+1}^{0}(A_i(n) - C_i(n)) > (1 - \delta)t, \exists\tau \geq 0\right) \\ &= P\left(\sup_{\tau\geq 0} S_\tau > (1 - \delta)t\right) = P\left(Q_e > (1 - \delta)t\right), \end{aligned}$$

where $S_\tau := \sum_{n=-\tau+1}^{0}(A_i(n) - C_i(n))$. The last equality holds since $\sup_{\tau\geq 0} S_\tau$ is actually the event that a single server queue (a queue having exclusive access to the PU channel) has a queue length beyond $(1 - \delta)t$ at time 0. It follows by Theorem 1 that

$$\kappa(Q_e) = \min(\alpha_{l_i}, \alpha_b) - 1 = \alpha_i - 1. \tag{39}$$

As to the term II, it follows by union bound that

$$\begin{aligned} II &= P\left(\bigcup_{j\neq i}\left\{Q_i > t \wedge Q_j^{\frac{\beta_j}{\beta_i}} \geq \delta t\right\}\right) \\ &\leq \sum_{j\neq i} P\left(Q_i > t \wedge Q_j^{\frac{\beta_j}{\beta_i}} \geq \delta t\right) \\ &\leq \sum_{j\neq i} P\left(Q_i + Q_j \geq (\delta t)^{\frac{\beta_i}{\beta_j}}\right). \end{aligned} \tag{40}$$

Invoking Lemma 4, we have $\kappa(Q_i + Q_j) = \min(\alpha_i, \alpha_j) - 1$. It follows from Lemma 1 that

$$\kappa((Q_i + Q_j)^{\beta_j/\beta_i}) = \min\left(\frac{\beta_i}{\beta_j}(\alpha_i - 1), \frac{\beta_i}{\beta_j}(\alpha_j - 1)\right), \tag{41}$$

which, by Lemma 1, implies that

$$\sum_{j\neq i} P\left((Q_i + Q_j)^{\frac{\beta_j}{\beta_i}} \geq (\delta t)\right) \sim P\left((Q_i + Q_{j^*})^{\frac{\beta_{j^*}}{\beta_i}} \geq (\delta t)\right) \tag{42}$$

and

$$\kappa((Q_i + Q_{j^*})^{\frac{\beta_{j^*}}{\beta_i}}) = \min_{\{1\leq j\leq N\}} \frac{\beta_i}{\beta_j}(\alpha_j - 1) = \alpha_i^m. \tag{43}$$

Combining with (38), (39), (40), and (42) yields

$$P(Q_i > t) \lesssim P\left(Q_e > (1 - \delta)t\right) + P\left((Q_i + Q_{j^*})^{\frac{\beta_{j^*}}{\beta_i}} \geq (\delta t)\right) \tag{44}$$

by which, we obtain the upper bound of the steady-state queue length Q_i under two cases.

(1) If $\alpha_i - 1 < \alpha_i^m$, it follows by (39) and (43) that $\kappa(Q_e) < \kappa((Q_i + Q_{j^*})^{\frac{\beta_{j^*}}{\beta_i}})$, which, by Lemma 1, implies that

$$\sum_{j\neq i} P\left((Q_i + Q_j)^{\frac{\beta_j}{\beta_i}} \geq (\delta t)\right) = o(P\left(Q_e > (1 - \delta)t\right)) \tag{45}$$

by which we obtain from (44) that

$$\limsup_{t\to\infty} \frac{[\log(P(Q_i > t))]}{\log t} \le \limsup_{t\to\infty} \frac{\log[P(Q_e > (1-\delta)t)}{\log t}. \tag{46}$$

This implies from (39) that

$$\kappa(Q_i) \ge \kappa(Q_e) = \alpha_i - 1. \tag{47}$$

(2) If $\alpha_i - 1 \ge \alpha_i^m$, we have $\kappa(Q_e) < \kappa((Q_i + Q_{j^*})^{\frac{\beta_{j^*}}{\beta_i}})$. It follows by Lemma 1 that

$$P(Q_e > (1-\delta)t) = o\left(\sum_{j\neq i} P\left((Q_i + Q_j)^{\frac{\beta_j}{\beta_i}} \ge (\delta t)\right)\right). \tag{48}$$

This, combining (43) and (44), yields

$$\kappa(Q_i) \ge \kappa((Q_i + Q_{j^*})^{\beta_{j^*}/\beta_i}) = \alpha_i^m, \tag{49}$$

which, in conjunction with (47), completes the proof. ■

Lemma 6 *Under Maximum-Weight-β scheduling, the tail index $\kappa(Q_i)$ of the steady-state queue length Q_i is upper bounded by*

$$\kappa(Q_i) \le \alpha_i^m. \tag{50}$$

The proof of Lemma 6 depends on Lemma 7 [5]. Let $\{Y_i\}_{i\ge 1}$ be non-negative i.i.d. random variables independent of the non-negative random variable N. Define $S_N := \sum_{i=1}^N Y_i$.

Lemma 7 1) *Assume $Y_1 \in \mathcal{RV}(\alpha)$, $E[N] < \infty$ and $P(N > t) = o(P(Y_1 > t))$. Then,*

$$P(S_N > t) \sim E[N]P(Y_1 > t).$$

2) *Assume $N \in \mathcal{RV}(\alpha)$, $E[Y_1] < \infty$, and $P(Y_1 > t) = o(P(N > t))$. Moreover, assume that $E[N] < \infty$ if $\alpha = 1$. Then,*

$$P(S_N > t) \sim P(N > (E[Y_1])^{-1}t).$$

Proof of Lemma 6: To prove Lemma 6, we construct a fictitious queuing system, which consists of N queues $\{\overline{q}_i\}_{1\le i\le N}$. Each queue $\overline{q}_i$ has the same input process as q_i and is associated with a dedicated PU channel i. All PU channels have the same PU activities and the same channel detection results as the PU channel in the original system. Each queue follows the same regulations as the fast queue $\widetilde{q}_f$.

Consider a particular queue $\overline{q}_i$. We let all the queues $\{\overline{q}_j\}_{j\neq i}$ except $\overline{q}_i$ have the exclusive access to their own dedicated PU channel j without competing with each other. The queue $\overline{q}_i$ receives service if and only if $\overline{Q}_i^{\beta_i} = \max_{1\le j\le N} \overline{Q}_j^{\beta_j}$. In such a system, it is easy to prove that the fictitious queue $\overline{q}_i$ has shorter queue length than the queue q_i in the original system, i.e.,

$$Q_i(t) \ge \overline{Q}_i(t). \tag{51}$$

We assume that the fictitious system is in the steady state. Let p_j denote the probability that the queue $\overline{q}_{j\neq i}$ is not empty, i.e., $p_j := P(Q_j > 0)$. Let $\mathcal{E}_j$ denote the event where $\overline{q}_j$ is not empty and all other queues are empty, i.e.,

$$\mathcal{E}_j := \left\{\overline{Q}_j \neq 0 \wedge \bigcap_{k\neq i,j} \overline{Q}_k = 0\right\} \tag{52}$$

and $P(\mathcal{E}_j) := p_j \prod_{k\neq i,j}(1-p_k)$. Thus, by (51), we have the lower bound of moments of Q_i with any order d

$$E[Q_i^d] \ge \sum_{j\neq i} P(\mathcal{E}_j) E\left[\overline{Q}_i^d | \mathcal{E}_j\right]. \tag{53}$$

In the rest of the proof, we will derive the lower bound of the conditional moments $E[\overline{Q}_i^d|\mathcal{E}_j]$. We first define the following denotations for the queue $\overline{q}_j$. Assume that the event $\mathcal{E}_j$ occurs at time t. Let $L_j^r(t)$ denote the residual length of the message currently in service, which is the number of packets that belongs to this message but still remain in the queue at time t. Let $\widetilde{L}_j^r(t)$ denote the residual length of the message currently in service if the queue is served at every time slot of the PU channel. Since actually, the queue can be only served at the idle time periods of the PU channel, this implies that

$$L_j^r(t) > \widetilde{L}_j^r(t). \tag{54}$$

Let $L_j^s(t)$ denote the age of the message currently in service, which is the number of packets from this message that are already served. Let $T_j^r(t)$ and $T_j^s(t)$ denote respectively the residual and the expanded service time of the message currently in service. From renewal theory and Lemma 3, we have

$$\kappa(\widetilde{L}_j^r(t)) = \alpha_{l_j} - 1 \tag{55}$$

and if $B_1 \in \mathcal{RV}(\alpha_b)$ and $P(L_j > t) = o(P(B_1 > t))$, then

$$P(T_j^s(t) > t) \sim C_1 t P(B_1 > t), \tag{56}$$

where C_1 is a constants. Otherwise, if $L_j \in \mathcal{RV}(\alpha_{l_j})$ and $P(B_1 > t) = o(P(L_j > t))$, then

$$C_2 t P(L_j > t) \lesssim P(T_j^s(t) > t) \lesssim C_3 t P(L_j > t), \tag{57}$$

where C_2 and C_3 are some constants. By renewal theory, it follows from (56) and (57) that

$$\kappa(T_j^s(t)) = \kappa(T_j^r(t)) = \min(\alpha_{l_j}, \alpha_b) - 1 = \alpha_j - 1. \tag{58}$$

We are now ready to prove the lower bound of the conditional moments $E[\overline{Q}_i^d|\mathcal{E}_j]$. If the event $\mathcal{E}_j$ occurs, then two possible events, $\Gamma(t)$ and $\Gamma^c(t)$, occur to $\overline{Q}_i(t)$. Define $\Gamma(t) = \{\overline{Q}_i(t)^{\beta_i} \ge T_j^r(t)^{\beta_j}\}$ and its complement $\Gamma^c(t)$. If $\Gamma(t)$ occurs, we have

$$\overline{Q}_i(t) \ge T_j^r(t)^{\frac{\beta_j}{\beta_i}}. \tag{59}$$

Otherwise, if $\Gamma^c(t)$ occurs, there are two possibilities including (1) $L_j^r(t)^{\beta_j} < \overline{Q}_i(t)^{\beta_i} < T_j^r(t)^{\beta_j}$, and (2) $\overline{Q}_i(t)^{\beta_i} < L_j^r(t)^{\beta_j}$. In the case (1), it implies from (54) that

$$\overline{Q}_i(t) \ge L_j^r(t)^{\frac{\beta_j}{\beta_i}} \ge \widetilde{L}_j^r(t)^{\frac{\beta_j}{\beta_i}}. \tag{60}$$

In the case (2), let τ denote the last time before t that $\overline{q}_i$ receives service. This means that $\overline{Q}_j(\tau)^{\beta_j} < \overline{Q}_i(\tau)^{\beta_i}$. This, combining with the fact that $\overline{Q}_j(t)^{\beta_j} \ge L_j^r(t)^{\beta_j} > \overline{Q}_i(t)^{\beta_i} > \overline{Q}_i(\tau)^{\beta_i}$, implies that the burst being served at time t did not begin to receive service at τ, i.e., $t - \tau > T_j^s(t)$. This implies that

$$\overline{Q}_i(t) = \sum_{k=1}^{t-\tau} A_i(k) + \overline{Q}_i(\tau) \ge \sum_{k=1}^{T_j^s(t)} A_i(k). \tag{61}$$

Let $S_{T_j^s} := \sum_{k=1}^{T_j^s(t)} A_i(k)$. Applying Lemma 7, it follows from (56) and (57) and

$$\kappa(S_{T_j^s}) = \min(\alpha_{l_j}, \alpha_b) - 1 = \alpha_j - 1. \quad (62)$$

Let $p_\Gamma = P(\Gamma(t))$ and $p_{\Gamma^c} = P(\Gamma^c(t))$. Combining (53), (59), (60) and (61), we obtain

$$\begin{aligned} E[Q_i^d] \;\geq\; & \sum_{j\neq i} P(\mathcal{E}_j)\left(p_\Gamma E\left[T_j^r(t)^{\frac{d\beta_j}{\beta_i}}\right]\right. \\ & \left. + p_{\Gamma^c} E\left[\min(\widetilde{L}_j^r(t)^{\frac{d\beta_j}{\beta_i}}, (S_{T_j^s})^d)\right]\right) \\ \geq\; & \sum_{j\neq i} P(\mathcal{E}_j)\left(p_\Gamma E\left[T_j^r(t)^{\frac{d\beta_j}{\beta_i}}\right]\right. \\ & \left. + p_{\Gamma^c} \min(E\left[\widetilde{L}_j^r(t)^{\frac{d\beta_j}{\beta_i}}\right], E\left[(S_{T_j^s})^d\right])\right). \quad (63) \end{aligned}$$

This, combining with (55), (58), and (62), implies that if the order of the moments $d \geq \min_{j\neq i} \frac{\beta_i}{\beta_j}(\alpha_j - 1)$, then at least one of the terms on the right hand of (63), is infinite, which implies

$$\kappa(Q_i) \leq \min_{j\neq i} \frac{\beta_i}{\beta_j}(\alpha_j - 1). \quad (64)$$

Moreover, since under any working conserving scheduling policy, Q_i is lowered bounded by Q_e. This implies that

$$\kappa(Q_i) \leq \frac{\beta_i}{\beta_i}(\alpha_i - 1), \quad (65)$$

which, combining with (64) completes the proof. ■

Proof of Theorem 4: By Lemma 6 and 5, it follows that the upper and lower bounds of $\kappa(Q_i)$ matches, This, combining the fact that $\kappa(Q_i) \geq \alpha^- - 1$ by Theorem 3, completes the proof. ■

VI. Conclusions

This paper provides an asymptotic analysis of the steady-state queue length distribution of secondary users (SUs) for Dynamic Spectrum Access (DSA) networks. For the single-user single-channel case, it is shown that if either the busy time or message size is heavy tailed, then the steady state queue length is one order heavier than either the busy time or message size, whichever has the heavier tail. For the multi-user single-channel case, we study the throughput-optimal scheduling policies. Specifically, it is shown that the celebrated maximum-weight scheduling yields the worst possible queuing performance by making the tail of the SU queue length as heavy as possible. On the contrary, the maximum-weight-α scheduling can lead to the best queueing performance in terms of inducing the lightest possible tail distribution of the SU queue length.

References

[1] I. F. Akyildiz, W. Y. Lee, and K. Chowdhury. Crahns: Cognitive radio ad hoc networks. *Ad Hoc Networks (Elsevier) Journal*, 7(5):810–836, July. 2009.

[2] N. H. Bingham, C. M. Goldie, and J. L. Teugels. *Regular Variation*. Cambridge University Press, 1989.

[3] J. W. Cohen. Some results on regular variation for distributions in queueing and fluctuation theory. *J. Appl. Probab.*, 10:343C353, 1973.

[4] D. J. Daley. The moment index of minima 2. *J. Appl. Probab.*, 38(A):33–36, 2001.

[5] G. Fay, B. Gonzalez-Arevalo, T. Mikosch, and G. Samorodnitsky. Modeling teletraffic arrivals by a poisson cluster process. *Queueing Systems: Theory and Applications*, 4(2):121–140, 2006.

[6] K. Jagannathan, M. Markakis, E. Modiano, and J. N. Tsitsiklis. Queue length asymptotics for generalized max-weight scheduling in the presence of heavy-tailed traffic. In *Proc. IEEE INFOCOM*, Apr. 2011.

[7] A. Laourine, S. Chen, and L. Tong. Queuing analysis in multichannel cognitive spectrum access: A large deviation approach. In *Proc. IEEE INFOCOM*, Mar. 2010.

[8] S. Luo, J. Li, K. Park, and R. Levy. Proc. exploiting heavy-tailed statistics for predictable qos routing in ad hoc wireless networks. In *IEEE INFOCOM'08*, Mar. 2008.

[9] M. Markakis, E. Modiano, and J. N. Tsitsiklis. Scheduling policies for single-hop networks with heavy-tailed traffic. In *Proc. 47th Annual Allerton Conference on Communication, Control, and Computing*, 2009.

[10] K. Park and W. Willinger. *Self-Similar Network Traffic and Performance Evaluation*. A Wiley-Interscience Pubilication, 2000.

[11] L. Tassiulas and A. Ephremides. Stability properties of constrained queueing systems and scheduling policies for maximum throughput in multihop radio networks. *IEEE Transactions on Automatic Control*, 37(12):1936C1948, 1992.

[12] P. Wang and I. F. Akyildiz. Asymptotic queueing analysis for dynamic spectrum access networks. Technical report, Available: http://users.ece.gatech.edu/~pwang40/publications.html.

[13] P. Wang and I. F. Akyildiz. Can dynamic spectrum access induce heavy tailed delay? In *Proc. IEEE DySPAN 2011*, May 2011.

[14] S. Wang, J. Zhang, and L. Tong. Delay analysis for cognitive radio networks with random access: A fluid queue view. In *Proc. IEEE INFOCOM*, Mar. 2010.

[15] M. Wellens, J. Riihijarvi, and P. Mahonen. Empirical time and frequency domain models of spectrum use. *Physical Communication*, 2(1-2):10–32, 2009.

Pu Wang received the B.S. degree in Electrical Engineering from the Beijing Institute of Technology, Beijing, China, in 2003 and the M.Eng. degree in Computer Engineering from the Memorial University of Newfoundland, St. Johns, NL, Canada, in 2008. Currently, he is a graduate research assistant in Broadband Wireless Networking Laboratory (BWN Lab), School of Electrical and Computer Engineering, Georgia Institute of Technology. He is currently working toward the Ph.D. degree in Electrical and Computer Engineering under the supervision of Prof. Ian F. Akyildiz. His research interests include wireless sensor networks, cognitive radio networks, Internet of Things and Cyber-Physical systems.

Ian F. Akyildiz (M'86-SM'89-F'96) received the B.S., M.S., and Ph.D. degrees in Computer Engineering from the University of Erlangen-Nrnberg, Germany, in 1978, 1981 and 1984, respectively. Currently, he is the Ken Byers Chair Professor in Telecommunications with the School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, the Director of the Broadband Wireless Networking Laboratory and Chair of the Telecommunication Group at Georgia Tech. Dr. Akyildiz is an honorary professor with the School of Electrical Engineering at Universitat Politcnica de Catalunya (UPC) in Barcelona, Catalunya, Spain and the founder of N3Cat (NaNoNetworking Center in Catalunya). Since 2011, he is a Consulting Chair Professor at the Department of Information Technology, King Abdulaziz University (KAU) in Jeddah, Saudi Arabia. Since January 2013, Dr. Akyildiz is also a FiDiPro Professor (Finland Distinguished Professor Program (FiDiPro) supported by the Academy of Finland) at Tampere University of Technology, Department of Communications Engineering, Finland. He is the Editor-in-Chief of Computer Networks (Elsevier) Journal, and the founding Editor-in-Chief of the Ad Hoc Networks (Elsevier) Journal, the Physical Communication (Elsevier) Journal and the Nano Communication Networks (Elsevier) Journal. He is an IEEE Fellow (1996) and an ACM Fellow (1997). He received numerous awards from IEEE and ACM. His current research interests are in nanonetworks, Long Term Evolution (LTE) advanced networks, cognitive radio networks and wireless sensor networks.