Delay-Optimal Opportunistic Scheduling and Approximations: The Log Rule

Bilal Sadiq, Seung Jun Baek, *Member, IEEE*, and Gustavo de Veciana, *Fellow, IEEE*

***Abstract*—This paper considers the design of multiuser opportunistic packet schedulers for users sharing a time-varying wireless channel from performance and robustness points of view. For a simplified model falling in the classical Markov decision process framework, we numerically compute and characterize mean-delay-optimal scheduling policies. The computed policies exhibit *radial sum-rate monotonicity*: As users' queues grow linearly, the scheduler allocates service in a manner that deemphasizes the *balancing of unequal queues* in favor of *maximizing current system throughput* (being opportunistic). This is in sharp contrast to previously proposed throughput-optimal policies, e.g., Exp rule and MaxWeight (with *any* positive exponent of queue length). In order to meet performance and robustness objectives, we propose a new class of policies, called the Log rule, that are radial sum-rate monotone (RSM) and provably throughput-optimal. In fact, it can also be shown that an RSM policy minimizes the asymptotic probability of sum-queue overflow. We use extensive simulations to explore various possible design objectives for opportunistic schedulers. When users see heterogenous channels, we find that emphasizing queue balancing, e.g., Exp rule and MaxWeight, may excessively compromise the overall delay. Finally, we discuss approaches to implement the proposed policies for scheduling and resource allocation in OFDMA-based multichannel systems.**

***Index Terms*—Delay/throughput optimality, Markov decision process, OFDMA resource allocation, opportunistic scheduling, radial sum-rate monotonicity (RSM).**

I. Introduction

THIS paper addresses the design of scheduling policies for a fixed number of users sharing a wireless channel. Each user's data arrives to a queue as a random stream where it awaits transmission. The wireless channel is time-varying in that the transmission rates supported for each user vary randomly over time. If the channel state is available, a policy can schedule users so as to exploit favorable channels, e.g., schedule the user that currently has the highest rate—this is referred to as opportunistic scheduling [1]–[3]. Our objective in this paper is to evaluate the design of queue-and-channel-aware schedulers both from the point of view of performance and robustness. By robustness, we informally mean a scheduler's ability to perform well for the majority of users under unpredicted/changing conditions and even transient "overloads" relative to the desired quality of service. Furthermore, if the system becomes temporarily overloaded, it is desirable for an opportunistic scheduler to exhibit graceful degradation of service. Though there has been a substantial amount of work on the schedulers, it is still unclear whether scheduler design should be guided by the objectives like minimizing mean delay and the asymptotic probability of sum-queue overflow, or instead by objectives like minimizing the asymptotic probability of max-queue overflow. In these considerations lies the motivation for this work and our efforts to leverage analysis, where possible, and simulation to reach a better understanding of this problem.

To put our work into context, we begin by summarizing some of the key related work in this area. Among many others, [3] considers opportunistic scheduling in a setting where users' queues are *infinitely backlogged*. They identify channel-aware opportunistic scheduling policies, which maximize the sum throughput under various types of fairness constraints. The missing element in this work is the impact of queueing dynamics. Recently, [4] showed that under a *constant* load, scheduling algorithms that are oblivious to queue state will incur an average delay that grows linearly in the number of users, whereas the channel-and-queue aware schedulers can achieve an average delay that is independent of the number of users. Even before this, it was immediately recognized that when queueing dynamics are introduced, opportunistic scheduling policies that are solely channel-aware may not be stable (i.e., keep the users' queues bounded) unless the policy is chosen carefully, e.g., using prior knowledge of mean arrival rates [5]. For this reason, a substantial focus was placed on designing schedulers that are both channel- and queue-aware and provably *throughput-optimal*, i.e., ensure the queues' stability without any knowledge of arrival and channel statistics if indeed stability can be achieved under any policy. Except for some degenerate cases, such policies must tradeoff *maximizing current transmission rate* (e.g., scheduling the queue with the best channel) versus *balancing unequal queues* (e.g., scheduling the longest queue). Balancing queues avoids empty queues, which enhances the ability to exploit high channel variations in the future. We will refer to this tradeoff many times in this paper. Two classes of policies known to be throughput-optimal are MaxWeight [6] (also known as Modified Largest Weighted Work/Delay First) and Exp rule [7]. Yet, stability is a weak form of performance optimality.

Thus, it is of interest to study opportunistic policies that are *delay-optimal*—e.g., polices that minimize the overall average

Manuscript received September 06, 2009; revised May 18, 2010; accepted July 22, 2010; approved by IEEE/ACM TRANSACTIONS ON NETWORKING Editor T. Bonald. Date of publication August 30, 2010; date of current version April 15, 2011. This work was supported in part by AFOSR Award FA9550-07-1-0428 and NSF Award CNS-0721532. S. Baek was supported in part by KEIT Contract No. 10035213-2010-01, KU K1011831, and a KU CIC Special Initiation Grant. This paper was presented in part at IEEE INFOCOM, Rio de Janeiro, Brazil, April 19–25, 2009.

B. Sadiq and G. de Veciana are with the Department of Electrical and Computer Engineering, The University of Texas at Austin, Austin, TX 78712 USA (e-mail: sadiq@ece.utexas.edu; gustavo@ece.utexas.edu).

S. Baek is with the College of Information and Communications, Korea University, Seoul 136-701, Korea (e-mail: sjbaek@korea.ac.kr).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TNET.2010.2068308

delay (per data unit) seen by the users—or policies that minimize the probability that either the sum-queue or the largest queue overflows a large buffer. These polices are harder to characterize for servers with time-varying capacity, but some results are available, which we briefly discuss next.

In [8] and [9], the Longest-Connected-Queue (LCQ) and Longest-Queue-Highest-Possible-Rate (LQHPR) policies are introduced. Strong results are shown for these policies; they stochastically minimize the max and sum queue process, and thus also the max and sum queue tails and mean delay. However, in addition to assuming certain symmetry conditions on arrival and channel statistics, [8] is limited to ON-OFF channels where only a single user can be scheduled per time slot, and [9] assumes that the scheduler can allocate service rates from the current information theoretic multiuser capacity region. In both cases, the above-mentioned tradeoff between queue balancing and throughput maximization is absent. Indeed in [8], all policies that pick a connected queue result in the same transmission rate, whereas, in the case of [9], all policies that pick a service vector from the maximal points of the current capacity region, i.e., points on the max-sum-rate face, result in the same overall transmission rate. Thus, one can achieve the *queue balancing* goal without ever compromising *throughput*. Not surprisingly, in both cases the optimal policy turns out to be greedy in that it allocates as much service rate as possible to the longest/longer queues.

A related server allocation problem is studied in [10]. The paper considers minimizing the mean delay in a two-queue system where each queue has a dedicated server and a third server can be dynamically shared between them. As a result, the two queues can be allocated service rates from a polymatroid capacity region, thus the objective of queue balancing can again be achieved without compromising the total service rate. However, without the underlying symmetry assumptions of [8] and [9], only the existence of a monotone increasing switching curve on the queue state space is shown; note that the switching curve under LCQ and LQHPR policies lies along the line where both queues are equal. For a system with a general compact, convex, and coordinate convex capacity region and any finite number of queues, [11] gives a large deviations principle (LDP) for transient queue process under MaxWeight scheduler. This LDP can be used to compute, e.g., the asymptotic probability of sum-queue or max-queue overflow, as well as the corresponding likely modes of overflow. Although the capacity region is not changing over time, the region is such that a scheduler must trade off maximizing total service rate with balancing unequal queues. Therefore, this result is insightful in relating the modes overflow to the tradeoff made by the MaxWeight scheduler. A more recent work [12] gives a many-sources large deviations result for the MaxWeight scheduler for a similar capacity region.

Finally, relaxing the symmetry assumptions of [8] and [9], the works in [13]–[15] consider the asymptotic probability of max-queue overflow. The server capacity in [13], though time-varying, is identical for all users at any given time, thus the need to trade off queue balancing versus service rate maximization is again absent. In fact, the sum-queue process in [13] is identical for all work conserving schedulers. However, [14] and [15] consider a server with *asynchronously* time-varying capacity across users. Reference [14] studies the asymptotic probability of max-queue overflow under MaxWeight scheduler and shows that as the exponent of queue length in the MaxWeight scheduler, α, becomes large, the asymptotic probability of max-queue overflow under MaxWeight approaches the minimum achievable under any other scheduler. A stronger result is shown in [15], that is, the Exp rule scheduler in fact minimizes the steady-state asymptotic probability of max-queue overflow. Indeed, the models in [14] and [15] accurately capture a wireless channel shared by heterogenous users and exhibit the tradeoff between queue balancing and service rate maximization. Existence of this tradeoff also implies that, unlike the LCQ and LQHPR policies, the asymptotic optimality of Exp rule does not translate to minimizing the asymptotic probability of sum-queue overflow or the mean delay. In fact, the policies that minimize mean delay and sum-queue overflow are very different, and we believe they are of practical interest.

In a related work on input-queued switches, [16] explains the conjecture that deemphasizing queue balancing improves mean delay of Maximum-Weight Matching algorithms.

1) Contributions: In this paper, we begin by characterizing mean-delay optimal opportunistic schedulers for heterogenous systems where the arrival and channel statistics are known. We consider a simple model falling in the classical Markov decision process framework, where we can numerically compute the optimal scheduling policy. Our first contribution is showing through numerical computation that mean-delay optimal policies exhibit *radial sum-rate monotonicity* (RSM), i.e., when user queues grow linearly (i.e., scaled up by a constant), the scheduler allocates service in a manner that deemphasizes the balancing of unequal queues in favor of maximizing current system throughput (being opportunistic). This is in sharp contrast to previously proposed policies, e.g., MaxWeight and Exp rules, which, nevertheless, have the advantage of being throughput-optimal. Our second contribution is to propose a new class of policies, called the Log rule, that are radial sum-rate monotone and provably throughput-optimal. These policies are favorable both in terms of mean delay and robustness. Our simulations for realistic wireless channels confirm the superiority of the Log rule, which achieves a 20%–75% reduction in the mean packet delay. The Log rule is proposed as a practical solution, but is not provably mean-delay optimal. However, in a companion paper [17], we use the approach of [15] to show that the candidate RSM policy indeed minimizes the asymptotic probability of *sum-queue* overflow.

Thus, we have at our disposal several opportunistic scheduling policies that are good for different objectives. The question remains: In designing an opportunistic scheduler, should one be guided by mean or asymptotic tail results, and should one focus on individual worst case or overall system criteria?

As the third contribution of this paper, we use extensive simulation to attempt to gain further insight on the question and evaluate the comparative effectiveness of various policies.

We also extend the proposed scheduling policies to multichannel systems supporting a large number of users, e.g., OFDMA-based WANs such as LTE/WiMAX networks, where bandwidth and power resources can be shared by multiple users over a scheduling/transmission time interval. Recognizing practical limits on the spectral granularity of channel feedback in multichannel systems, we suggest a queue-aware convex program formulation to realize opportunistic scheduling and resource allocation policies.

We make the following observations.

1) *Minimizing mean delay versus asymptotic tails:* Based on simulations, we observe that when users see heterogenous channels, policies such as Exp rule that aim to minimize the exponential decay rate of delay distribution tail of the *worst* user may excessively compromise average delay, in some cases penalizing the tail distributions of many of the users. Our simulations show that Log rule can achieve better mean delays (overall and on a per-user basis) and comparable or better distribution tails for many, if not all, the users under reasonably high loads.
2) *Graceful degradation:* Due to the uncertain and changing characteristics of wireless channels, precise resource allocation to meet quality-of-service (QoS) requirements for real-time or streaming flows is likely to be virtually impossible. As such, a desirable design objective is for a scheduler to gracefully degrade. If there is a change in the environment causing a temporary overload, then as many users as possible should meet their QoS requirements rather than all failing. Our simulation results show that Log rule compares favorably in this regard. In a system with unpredictable heterogenous channels, there will be a wider disparity in the performance users see under the Log rule, but a substantial number of users does very well. Hence, depending on the QoS objective and specific character of the change in user's channels, one could end up with no users seeing acceptable performance under the Exp rule while, say, half the users meet their QoS requirement under Log rule. Finally, we note that Log rule's underlying goal of minimizing mean packet delays might be a desirable objective from the point of view maximizing throughput seen by best-effort traffic.

2) Organization of This Paper: This paper is organized as follows. In Section II, the system model and definitions of "opportunistic capacity region" and "scheduling policy" are given. Optimality criterion is defined, which is slightly more general than mean packet delay. In Section III, a time-scale separation argument is used to formulate the optimal policy as a numerically tractable Markov decision process. In Section IV, "radial sum-rate monotonicity" is formally defined, and through numerical computations, the optimal policy is shown to be (weakly) RSM. By contrast, known heuristics (MaxWeight and Exp rule) are shown to differ. In Section V, a new class of scheduling policies, called the Log rule, is proposed which is both RSM and throughput-optimal. A large-deviations optimality result for an RSM policy is also stated. In Section VI, extensive simulation results are presented for an HDR-like downlink [18] to contrast various scheduling policies and underlying design objectives. In Section VII, implementation of the proposed schedulers for OFDMA-based multichannel systems is discussed.

II. System Model

Consider the following continuous time model for scheduling $\bar{n}$ users' traffic over a shared wireless channel. Each user $n \in \mathcal{N} \equiv \{1, 2, \ldots, \bar{n}\}$ is assigned a queue in which packets with independent and exponentially distributed sizes arrive as a Poisson stream with rate λ_n packets/s. At any time t, define the (random) vector $\boldsymbol{Q}(t) \equiv (Q_n(t), n \in \mathcal{N}) \in \mathbb{Z}_+^{\bar{n}}$, where $Q_n(t)$ denotes the number of packets in the nth queue at time t. The state of the users' wireless channels at time t is modeled by random $M(t)$, which can take values in the finite set $\mathcal{M} \equiv \{1, 2, \ldots, \bar{m}\}$. We assume that for all $t \neq t'$, the channel states $M(t)$ and $M(t')$ are independent and have the same distribution as a random variable M. Associated with each channel state $m \in \mathcal{M}$ is a vector $\boldsymbol{r}(m) = (r_n(m), n \in \mathcal{N})$, where $r_n(m)$ has the following interpretation: When the channel is in state m and dedicated to the nth user, then $r_n(m)$ is the instantaneous service rate in packets/second available to the nth user. We allow the channel to be split among users at any time instant according to a stochastic vector $\boldsymbol{\sigma}(t) = (\sigma_n(t), n \in \mathcal{N})$. Recall that a stochastic vector has nonnegative components that sum up to 1, in which case the service rate available to the nth user at time t will be $\sigma_n(t) r_n(M(t))$. We follow the convention that capital letters, e.g., $\boldsymbol{Q}(\cdot)$ and $M(\cdot)$, denote a random variable, whereas small letters, e.g., $\boldsymbol{q}(\cdot)$ and $m(\cdot)$, denote a particular realization. Moreover, we will make the natural distinction between "increasing" and "strictly increasing."

The problem of scheduling users for service is then to choose a vector $\boldsymbol{\sigma}(t)$ for each time instant t, such that a given optimality criterion is met. A scheduling policy is said to be static state-feedback if it chooses the vector $\boldsymbol{\sigma}(t)$ according to a *fixed* rule based solely on the *current* system state $(\boldsymbol{q}(t), m(t))$. More precisely, a static state-feedback scheduling policy is defined as a function $\boldsymbol{f}$, which takes the system state $(\boldsymbol{q}(t), m(t))$ at any time t into a stochastic vector $\boldsymbol{\sigma}(t)$

$$\boldsymbol{\sigma}(t) = \boldsymbol{f}(\boldsymbol{q}(t), m(t)). \tag{1}$$

Let $\mathcal{F}$ denote the set of static state-feedback policies. Given the optimality criterion described next, Poisson arrivals, exponentially distributed packet sizes, and i.i.d. channel state vectors, there is no loss of generality in restricting our attention to the policies in $\mathcal{F}$.

A. Optimality Criterion

Consider a system initiated at $t = 0$ in state $\boldsymbol{Q}(0) = \boldsymbol{q}(0)$, which evolves under scheduling policy $\boldsymbol{f}$. The expected long-run average queue for the nth user is given by

$$\overline{q}_n(\boldsymbol{f}) \equiv \limsup_{t \to \infty} \mathbb{E}^{\boldsymbol{f}}_{\boldsymbol{q}(0)} \left[\frac{1}{t} \int_0^t Q_n(\tau) d\tau \right] \tag{2}$$

where $\mathbb{E}^{\boldsymbol{f}}_{\boldsymbol{q}}$ denotes expectation under $\boldsymbol{f}$ conditional on $\boldsymbol{Q}(0) = \boldsymbol{q}$. We define the delay-optimal scheduling policy $\boldsymbol{f}^*$ as the one that minimizes the total weighted average queue length, if it exists, for a given weight vector $\boldsymbol{w} = (w_n, n \in \mathcal{N}) > \boldsymbol{0}$, i.e.,

$$\boldsymbol{f}^* \in \arg\min_{\boldsymbol{f} \in \mathcal{F}} \sum_{n \in \mathcal{N}} w_n \overline{q}_n(\boldsymbol{f}). \tag{3}$$

It follows from Little's Law that if the process $(\boldsymbol{Q}(t), t \geq 0)$ is stationary, this optimality criterion minimizes the overall (weighted) average packet delay seen by the n-users.

B. Stabilizability and Opportunistic Capacity Region

For $n \in \mathcal{N}$, let $\boldsymbol{e}_n \in \mathbb{R}_+^{\bar{n}}$ denote the nth standard basis vector in $\mathbb{R}_+^{\bar{n}}$. For each $m \in \mathcal{M}$, let $\mathcal{C}_m \in \mathbb{R}_+^{\bar{n}}$ denote the convex hull of origin and following $\bar{n}$ points:

$$r_1(m)\boldsymbol{e}_1, r_2(m)\boldsymbol{e}_2, \ldots, r_{\bar{n}}(m)\boldsymbol{e}_{\bar{n}}.$$

That is, $\mathcal{C}_m$ is the set of service rates that can be jointly offered to the $\bar{n}$ users, conditional on the channel being in state m. Define the opportunistic capacity region $\mathcal{C}$ of a channel as the set of long-run *average* service rates that can be jointly offered to the $\bar{n}$ users under all possible scheduling policies. The opportunistic capacity region associated with the distribution of M is given by the weighted Minkowski sum of regions $\mathcal{C}_m$, i.e.,

$$\begin{aligned}\mathcal{C} &\equiv \mathbb{P}(M=1)\mathcal{C}_1 \oplus \cdots \oplus \mathbb{P}(M=\bar{m})\mathcal{C}_{\bar{m}} \\ &= \left\{ \sum_{m\in\mathcal{M}} \mathbb{P}(M=m)\boldsymbol{u}(m) : \boldsymbol{u}(m) \in \mathcal{C}_m \right\}. \end{aligned} \tag{4}$$

The capacity region $\mathcal{C}$ is a compact, convex, coordinate-convex polyhedron in $\mathbb{R}_+^{\bar{n}}$ whose exact shape depends on the distribution of M[3]. Let $\mathcal{C}^v \equiv \{\boldsymbol{v}^{(1)}, \boldsymbol{v}^{(2)}, \ldots, \boldsymbol{v}^{(\bar{l})}\}$ denote the set of maximal vertices of $\mathcal{C}$, where $\bar{l}$ denotes the number of vertices. For any $\mathcal{N}' \subseteq \mathcal{N}$, define $\mathcal{C}(\mathcal{N}') \equiv \{\boldsymbol{u} \in \mathcal{C} : u_n = 0,\ \forall\, n \notin \mathcal{N}'\}$, where $\mathcal{C}(\mathcal{N}')$ is the channel capacity region when the channel is shared only among the users in $\mathcal{N}'$.

The following (restatement of) [19, Lemma 2.1] will be used in the subsequent sections.

Lemma 1: Assume all queues are infinitely backlogged. For any $\boldsymbol{\alpha} = (\alpha_n, n \in \mathcal{N}) \geq 0$, let $\boldsymbol{\beta}(\boldsymbol{\alpha}) \in \mathcal{C}$ denote the vector of average service rates seen by the queues under the policy that serves user n^* at time t if

$$n^* \in \arg\max_{n\in\mathcal{N}} \{\alpha_n r_n(m(t))\} \tag{5}$$

augmented with a tie-breaking rule; then

$$\boldsymbol{\beta}(\boldsymbol{\alpha}) \in \arg\max_{\boldsymbol{u}\in\mathcal{C}} \langle \boldsymbol{\alpha}, \boldsymbol{u}\rangle. \tag{6}$$

In other words, $\boldsymbol{\alpha}$ is an outer normal vector to the capacity region $\mathcal{C}$ at point $\boldsymbol{\beta}(\boldsymbol{\alpha})$ on the boundary.

As shown in [6], the system of $\bar{n}$ queues is stabilizable if and only if there exists a vector $\boldsymbol{u} = (u_n, n \in \mathcal{N}) \in \mathcal{C}$ such that for all $n \in \mathcal{N}$

$$\lambda_n < u_n. \tag{7}$$

We assume that the system under consideration is stabilizable, which implies that the weighted sum defined in (3) is bounded under at least one stationary policy.

III. Characterization of Delay-Optimal Policy

Consider the process $(\boldsymbol{Q}(t), t \geq 0)$ initiated in state $\boldsymbol{Q}(0) = \boldsymbol{q}(0)$ and evolving under a policy $\boldsymbol{f}$. Then, conditional on the process being in state $\boldsymbol{q}$, the nth queue is offered an average service rate of $\mu_n(\boldsymbol{q})$ given by

$$\mu_n(\boldsymbol{q}) = \mathbb{E}[r_n(M) f_n(\boldsymbol{q}, M)] \tag{8}$$

where the expectation is with respect to M. By definition of $\mathcal{C}$ in (4), the average service rate vector $\boldsymbol{\mu}(\boldsymbol{q}) = (\mu_n(\boldsymbol{q}), n \in \mathcal{N})$ lies[1] in $\mathcal{C}$. We assume that over an epoch, each queue $n \in \mathcal{N}$ is served constantly at rate $\mu_n(\boldsymbol{q})$, thus the set of n queues see state-dependent service rates chosen from $\mathcal{C}$. A rigorous justification of this can be found in [20] and relies on packet or file dynamics that are slow relative to channel variations, where the latter can be averaged. A similar assumption is made in [19] to obtain processor-sharing queueing model for a slotted time system where a packet (or file) typically takes many slots to process while the channel can change from slot to slot. Note that strictly speaking, it is shown in [21] that analysis under the assumption of infinitely fast channel variations leads to optimistic flow-level performance estimates.

[1]The map $\boldsymbol{f}(\boldsymbol{q}, \cdot) \to \boldsymbol{\mu}(\boldsymbol{q}) \in \mathcal{C}$ given by (8) is surjective.

Under these assumptions, the scheduling problem of finding the right function $\boldsymbol{f}(\boldsymbol{q}, \cdot)$ for each $\boldsymbol{q}$ such that the total (weighted) average queue length is minimized [see (3)] is one of finding the right service rate vector $\boldsymbol{\mu}(\boldsymbol{q}) \in \mathcal{C}$ for each $\boldsymbol{q}$. Using this, we redefine a scheduling policy as a function $\boldsymbol{\mu} : \mathbb{Z}_+^{\bar{n}} \to \mathcal{C}$ that takes a queue state vector in $\mathbb{Z}_+^{\bar{n}}$ to a service rate vector in $\mathcal{C}$, where $\boldsymbol{\mu}$ relates to $\boldsymbol{f}$ through (8).

Under a fixed policy $\boldsymbol{\mu}$, the process $(\boldsymbol{Q}(t), t \geq 0)$ forms a time-homogeneous Markov chain on $\mathbb{Z}_+^{\bar{n}}$ with state-dependent transition rates. For convenience, we shall uniformize $\boldsymbol{Q}(t)$. For any $\boldsymbol{q} \in \mathbb{Z}_+^{\bar{n}}$, let $\boldsymbol{A}_n\boldsymbol{q} \equiv \boldsymbol{q} + \boldsymbol{e}_n$ and $\boldsymbol{D}_n\boldsymbol{q} \equiv (\boldsymbol{q} - \boldsymbol{e}_n)^+$, where $\boldsymbol{q}^+ \equiv (\boldsymbol{y} : y_n = \max\{0, q_n\})$. Let $\gamma \equiv |\boldsymbol{\lambda}| + \max_{\boldsymbol{u}\in\mathcal{C}} |\boldsymbol{u}|$, where $|\cdot|$ denotes L_1 norm. Let τ_k denote the (random) time of the kth transition of $\boldsymbol{Q}(t)$ and $\tau_0 = 0$. Also, let $\boldsymbol{Q}_k = \lim_{t\downarrow\tau_k} \boldsymbol{Q}(t)$. Then, under policy $\boldsymbol{\mu}$, the process $\boldsymbol{Q}(t)$ can be viewed as having a state-independent event rate of γ [i.e., $(\tau_{k+1} - \tau_k) \sim \exp(\gamma)$] and transition probabilities given by, for all $n \in \mathcal{N}$

$$\begin{aligned} P(\boldsymbol{Q}_{k+1} = \boldsymbol{A}_n\boldsymbol{q} | \boldsymbol{Q}_k = \boldsymbol{q}) &= \frac{\lambda_n}{\gamma} \\ P(\boldsymbol{Q}_{k+1} = \boldsymbol{D}_n\boldsymbol{q} | \boldsymbol{Q}_k = \boldsymbol{q}) &= \frac{\mu_n(\boldsymbol{q})}{\gamma} \\ P(\boldsymbol{Q}_{k+1} = \boldsymbol{q} | \boldsymbol{Q}_k = \boldsymbol{q}) &= 1 - \frac{|\boldsymbol{\lambda}| + |\boldsymbol{\mu}(\boldsymbol{q})|}{\gamma}. \end{aligned} \tag{9}$$

Define the cost under policy $\boldsymbol{\mu}$ over $[0, \tau_k)$ when starting in state $\boldsymbol{q}$ as $\mathbb{E}_{\boldsymbol{q}}^{\boldsymbol{\mu}}\left[\int_0^{\tau_k} \langle \boldsymbol{w}, \boldsymbol{Q}(t)\rangle dt\right]$, which, ignoring a constant multiplier γ^{-1}, can be shown to be equal to

$$V_k^{\boldsymbol{\mu}}(\boldsymbol{q}) \equiv \mathbb{E}_{\boldsymbol{q}}^{\boldsymbol{\mu}}\left[\sum_{l=0}^{k-1} \langle \boldsymbol{w}, \boldsymbol{Q}_l\rangle\right]. \tag{10}$$

Likewise, the average cost under policy $\boldsymbol{\mu}$, when starting in state $\boldsymbol{q}$, is given by

$$J_{\boldsymbol{\mu}}(\boldsymbol{q}) = \limsup_{k\to\infty} \frac{1}{k} V_k^{\boldsymbol{\mu}}(\boldsymbol{q}). \tag{11}$$

The optimality criterion given in (3) seeks to minimize this average cost. The problem of finding the minimum average cost and an optimal policy fits the classical dynamic programming framework (e.g., see [22]). Thus, the minimum average cost over all policies (denoted by J^*) is well defined, independent of the starting state, and together with a relative cost function h: $\mathbb{Z}_+^{\bar{n}} \to [0, \infty)$, which is unique up to an additive constant, satisfies Bellman's equation, i.e., for all $\boldsymbol{q} \in \mathbb{Z}_+^{\bar{n}}$

$$\begin{aligned} J^* &= \min_{\boldsymbol{\mu}} \{\langle \boldsymbol{w}, \boldsymbol{q}\rangle + \mathbb{E}^{\boldsymbol{\mu}}[h(\boldsymbol{Q}_{k+1}) - h(\boldsymbol{Q}_k) | \boldsymbol{Q}_k = \boldsymbol{q}]\} \\ &= \min_{\boldsymbol{u}\in\mathcal{C}} \Big\{ \langle \boldsymbol{w}, \boldsymbol{q}\rangle + \sum_{n\in\mathcal{N}} \frac{\lambda_n}{\gamma}(h(\boldsymbol{A}_n\boldsymbol{q}) - h(\boldsymbol{q})) \\ &\qquad + \frac{u_n}{\gamma}(h(\boldsymbol{D}_n\boldsymbol{q}) - h(\boldsymbol{q})) \Big\}. \end{aligned} \tag{12}$$

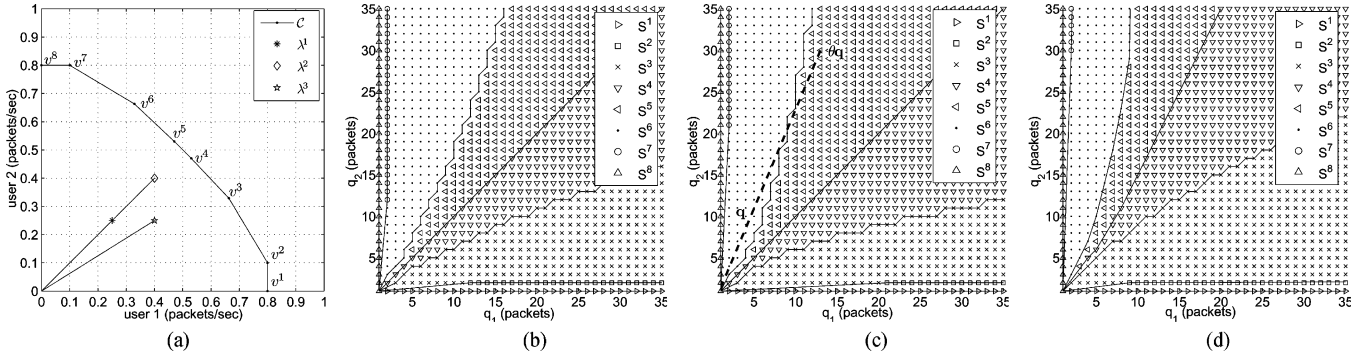

Fig. 1. (a) Partitions under the optimal policy in a two-user capacity region, and partitions corresponding to (b) arrival vector $\lambda^1 = (0.25, 0.25)$, (c) arrival vector $\lambda^2 = (0.4, 0.4)$, and (d) arrival vector $\lambda^3 = (0.4, 0.25)$.

Moreover, let $\boldsymbol{\Delta}h(\boldsymbol{q}) \equiv \big(h(\boldsymbol{q}) - h(\boldsymbol{D}_n\boldsymbol{q}), n \in \mathcal{N}\big)$, and define $\boldsymbol{\mu}^*$ as a policy that achieves the minimum in (12) for every $\boldsymbol{q}$, i.e.,

$$\boldsymbol{\mu}^*(\boldsymbol{q}) \in \arg\max_{\boldsymbol{u}\in\mathcal{C}} \langle \boldsymbol{u}, \boldsymbol{\Delta}h(\boldsymbol{q})\rangle \tag{13}$$

then $\boldsymbol{\mu}^*$ is an optimal policy achieving the minimum average cost J^*. The following lemma characterizes optimal scheduling decisions in *time* [see (1)] and follows from Lemma 1 by interpreting $\boldsymbol{\Delta}h(\boldsymbol{q})$ and $\boldsymbol{\mu}^*(\boldsymbol{q})$ in (13) as $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$, respectively.

Lemma 2: The following policy achieves the minimum average cost [see (3) and (11)]: At any time when the system is in state $(\boldsymbol{Q}, M) = (\boldsymbol{q}, m)$, choose a stochastic vector $\boldsymbol{\sigma}$ that satisfies

$$\sigma_{n^*} = 1 \quad \text{for some } n^* \in \arg\max_{n\in\mathcal{N}} \{\Delta_n h(\boldsymbol{q}) r_n(m)\} \tag{14}$$

where h is a relative cost function satisfying Bellman's (12).

Lemma 2 and (13) relate the tradeoff mentioned in Section I to the geometry of vector field $\boldsymbol{\Delta}h$ associated with the relative cost function. We explore this tradeoff in Section IV, where we use relative value iteration (see, e.g., [23]) to numerically compute h and J^*.

IV. Radial Sum-Rate Monotonicity: Comparing the Optimal Policy With Known Heuristics

In this section, we investigate how delay-optimal schedulers, as well as throughput-optimal policies such as MaxWeight and Exp rule, trade off current service rate versus balancing unequal queues. Specifically, we consider how the service rate vector chosen by each policy changes as the queues grow proportionally from a state $\boldsymbol{q} \in \mathbb{Z}_+^{\bar{n}}$ to a state $\theta\boldsymbol{q} \in \mathbb{Z}_+^{\bar{n}}$ for θ>1. Note that $\boldsymbol{q}$ and $\theta\boldsymbol{q}$ lie on a *radial* line in $\mathbb{R}_+^{\bar{n}}$ that passes through origin. For any $\boldsymbol{q}$, let $\mathcal{N}_{\boldsymbol{q}} \equiv \{n \in \mathcal{N}: q_n \neq 0\}$, i.e., the set of nonempty queues. We begin by defining an interesting property that we refer to as *radial sum-rate monotonicity*, as well as a weaker version of this property.

Definition 1: Given a weight vector $\boldsymbol{w}>\boldsymbol{0}$, we say a scheduling policy $\boldsymbol{\mu}$ is *radial sum-rate monotone* with respect to vector $\boldsymbol{w}$ if it satisfies two conditions. For any $\boldsymbol{q}$ and scalar θ such that $\theta\boldsymbol{q} \in \mathbb{Z}_+^{\bar{n}}$:

1) the total weighted service rate $\langle \boldsymbol{w}, \boldsymbol{\mu}(\theta\boldsymbol{q})\rangle$ is an increasing function of θ;
2) $\lim_{\theta\to\infty}\langle \boldsymbol{w}, \boldsymbol{\mu}(\theta\boldsymbol{q})\rangle = \max_{\boldsymbol{u}}(\langle \boldsymbol{w}, \boldsymbol{u}\rangle : \boldsymbol{u} \in \mathcal{C}(\mathcal{N}_{\boldsymbol{q}}))$.

Moreover, we say that $\boldsymbol{\mu}$ is *weakly* RSM if it satisfies 1).

Hence, as the queue grows proportionally, an RSM policy allocates service rates in a manner that *deemphasizes queue balancing* in favor of *increasing the total weighted service rate* (with respect to weight vector $\boldsymbol{w}$). Another useful and natural property, called *transition monotonicity* [24], describes the behavior of a policy along an *axial* line, i.e., as a queue grows from state $\boldsymbol{q}$ to a state $\boldsymbol{q} + \theta\boldsymbol{e}_n$ for any integer θ>0 and $n \in \mathcal{N}$.

Definition 2: A scheduling policy $\boldsymbol{\mu}$ is *transition monotone* if for all $n \in \mathcal{N}$ and $\boldsymbol{q} \in \mathbb{Z}_+^{\bar{n}}$, we have[2] $\mu_n(\boldsymbol{q}) \leq \mu_n(\boldsymbol{q} + \boldsymbol{e}_n)$, and $\lim_{\theta\to\infty} \mu_n(\boldsymbol{q} + \theta\boldsymbol{e}_n) = \max_{\boldsymbol{u}\in\mathcal{C}} u_n$.

Hence, as a single queue grows while others remain unchanged, a transition monotone policy allocates more service to the growing queue, and in the limit, only the longest queue is scheduled (whenever it sees a nonzero channel). As a result, asymptotically along an axial line, the total weighted service rate decreases.

Remark 1: Radial sum-rate and transition monotonicities both describe the above-mentioned tradeoff as $\boldsymbol{q}$ is taken to ∞, however along *different* paths. Indeed, a policy can be both radial sum-rate *and* transition monotone, in which case the total weighted service rate will increase along radial lines, however (asymptotically) decrease along axial lines. Moreover, it is simple to show that a throughput-optimal policy must be transition monotone except possibly on a compact subset of queue state space; see, e.g., [25].

A. Tradeoff Under Delay-Optimal Schedulers

Since the capacity region $\mathcal{C}$ is a polyhedron, instead of searching over the entire region for the maximum in (13), it suffices to only consider the vertices of $\mathcal{C}$

$$\boldsymbol{\mu}^*(\boldsymbol{q}) \in \arg\max_{\boldsymbol{u}\in\mathcal{C}^v} \langle \boldsymbol{u}, \boldsymbol{\Delta}h(\boldsymbol{q})\rangle. \tag{15}$$

Hence, the optimal policy partitions the state space $\mathbb{Z}_+^{\bar{n}}$ into at most $\bar{l}$ nonempty decision regions $\mathcal{S}^1, \mathcal{S}^2, \ldots, \mathcal{S}^{\bar{l}}$, each of which is associated with a distinct vertex, i.e., $\mathcal{S}^l \equiv \{\boldsymbol{q} : \boldsymbol{\mu}^*(\boldsymbol{q}) = \boldsymbol{v}^{(l)}\}$. In each region $\mathcal{S}^l$, the scheduler tries to *push* the queue process $\boldsymbol{Q}_k$ along vector $\boldsymbol{\lambda} - \boldsymbol{v}^{(l)}$. Fig. 1 shows the optimal policy's partitioning of $\mathbb{Z}_+^{\bar{n}}$ for a two-user system with weight vector $\boldsymbol{w} = (1, 1)$ under three different arrival vectors. Fig. 1(a) shows the hypothetical two-user capacity region and arrival vectors considered. Fig. 1(b) depicts the partition for $\boldsymbol{\lambda} = (0.25, 0.25)$ packets/s. Fig. 1(c) exhibits

[2]By coordinate-convexity of $\mathcal{C}$, it follows that for all $n' \neq n$, we have $\mu_{n'}(\boldsymbol{q}) \geq \mu_{n'}(\boldsymbol{A}_n\boldsymbol{q})$).

a more pronounced radial sum-rate monotonicity when arrival rate is increased to $\boldsymbol{\lambda} = (0.4, 0.4)$ packets/s, and the last plot is intended to exhibit the warping effect on the partition resulting from asymmetric arrival rates $\boldsymbol{\lambda} = (0.4, 0.25)$ packets/s. The boundaries between decision region are referred to as the *switching curves*.

For the optimal policy to be RSM, we must have that as $\theta \to \infty$ such that $\theta\boldsymbol{q} \in \mathbb{Z}_+^{\bar{n}}$ and $\boldsymbol{q} > 0$, the ratio $\frac{\Delta_{n'}h(\theta\boldsymbol{q})}{\Delta_n h(\theta\boldsymbol{q})}$ monotonically converges to $\frac{w_{n'}}{w_n}$, and therefore

$$\lim_{\theta\to\infty} \boldsymbol{\Delta} h(\theta\boldsymbol{q}) \propto (w_n \mathbb{1}_{\{q_n>0\}}, n \in \mathcal{N}) \tag{16}$$

whereas for the optimal policy to be *weakly* RSM, we only need that as θ increases, the ratio $\frac{\Delta_{n'}h(\theta\boldsymbol{q})}{\Delta_n h(\theta\boldsymbol{q})}$ monotonically gets closer to $\frac{w_{n'}}{w_n}$ (but is not required to converge to $\frac{w_{n'}}{w_n}$).

By computing the relative cost function h and the optimal scheduling policy for various arrival rate vectors and capacity regions, we observe that the optimal policies exhibit weak radial sum-rate monotonicity. An intuitive explanation of why the delay-optimal policies exhibit weak radial sum-rate monotonicity can be based on the following two observations.

1) The cost incurred per unit time in state $\theta\boldsymbol{q}$ for $\theta>1$ is more than the cost incurred per unit time in state $\boldsymbol{q}$.
2) The state $\theta\boldsymbol{q}$ is *farther* from any axis than the state $\boldsymbol{q}$.

Both observations suggest that in state $\theta\boldsymbol{q}$, an optimal policy would indeed deemphasize queue balancing in favor of increasing the current total weighted service rate.

To verify if the optimal policy is RSM, we need the asymptotics of the relative cost function $h(\theta\boldsymbol{q})$ for large θ [see (16)], which we cannot *compute* through relative value iteration. One can however use deterministic *fluid models* to obtain[3] $\lim_{\theta\to\infty} \frac{h(\theta\boldsymbol{q})}{\theta^2}$; see [26, Theorem 10.0.5]. Such limits can be used to determine the asymptotic slope of the switching curves on the state space of the *fluid-scaled* queue process. For details, see Appendix A, where we solve the fluid models of some nontrivial systems and show that, in general, the optimal policy may not be RSM. However, for a symmetric system subject to sufficient load, RSM policies are *fluid-scale asymptotic optimal*, and therefore RSM policies and optimal policies have similar switching curves on the state space of the fluid-scaled queue process.

Weighted Max-Rate Horn: Consider the decision regions $\mathcal{S}^4$ and $\mathcal{S}^5$, i.e., regions corresponding to those vertices of $\mathcal{C}$ that have the largest projection along vector $\boldsymbol{w}$. Under the optimal policy, union of these decision regions is shaped like a French horn (referred to as weighted max-rate horn). As we shall see next, under the Exp rule (with appropriately chosen constants), the union of the same partitions is shaped like a cylinder with gradually increasing diameter, whereas under MaxWeight, all partitions are simply cones.

B. Tradeoffs Under MaxWeight and Exp Rule

The MaxWeight and the Exp rule policies can also be expressed in a similar form as (13). These policies replace $\boldsymbol{\Delta}h$ with a suitable vector field on $\mathbb{Z}_+^{\bar{n}}$ such that the system is stable for any stabilizable $\boldsymbol{\lambda}$. Hence, the tradeoff under each policy can be investigated by considering how the vector fields change direction as queues grow proportionally.

[3]Function h exhibits quadratic growth [26, Theorem 9.0.5], therefore the limit is meaningful.

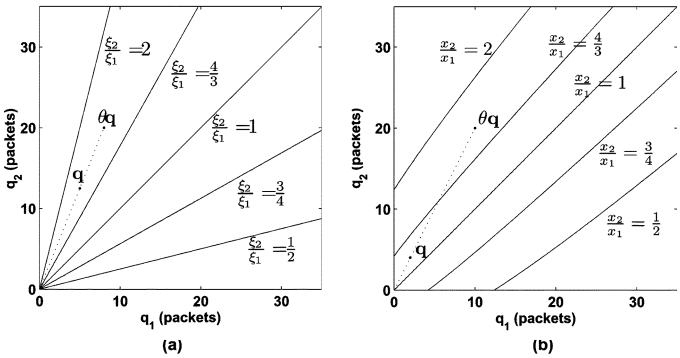

Fig. 2. Curves along which the direction is held constant by the vector field (a) $\boldsymbol{\xi}(\cdot)$ with $b_n = 1\ \forall\ n \in \mathcal{N}$, $\alpha = 0.5$, and (b) $\boldsymbol{x}(\cdot)$ with $a_n = 0.1$, $b_n = 1\ \forall\ n \in \mathcal{N}$, $c = 1$, $\eta = 0.5$.

MaxWeight policies [6] can be defined as follows: When the system is in state $(\boldsymbol{Q}, M) = (\boldsymbol{q}, m)$, choose a stochastic vector $\boldsymbol{\sigma}$ that satisfies

$$\sigma_{n^*} = 1 \quad \text{for some } n^* \in \arg\max_{n\in\mathcal{N}} \{\xi_n(\boldsymbol{q})\, r_n(m)\}$$

where $\xi_n(\boldsymbol{q})$ is the nth component of $\boldsymbol{\xi}(\boldsymbol{q}) \equiv (b_n q_n^\alpha, n \in \mathcal{N})$, for any fixed positive b_n's and α. Equivalently, when the queue state is $\boldsymbol{q}$, the policy uses a service rate vector $\boldsymbol{\mu}^W(\boldsymbol{q})$ given by

$$\boldsymbol{\mu}^W(\boldsymbol{q}) \in \arg\max_{\boldsymbol{u}\in\mathcal{C}^v} \langle \boldsymbol{u}, \boldsymbol{\xi}(\boldsymbol{q})\rangle. \tag{17}$$

Similarly, the Exp rule [7] is given by

$$\boldsymbol{\mu}^X(\boldsymbol{q}) \in \arg\max_{\boldsymbol{u}\in\mathcal{C}^v} \langle \boldsymbol{u}, \boldsymbol{x}(\boldsymbol{q})\rangle \tag{18}$$

where

$$\boldsymbol{x}(\boldsymbol{q}) \equiv \left(b_n \exp\left(\frac{a_n q_n}{c + (\bar{n}^{-1} \sum_{j\in\mathcal{N}} a_j q_j)^\eta} \right), n \in \mathcal{N} \right)$$

for any fixed positive a_n's, b_n's, c, and $0<\eta<1$.

While both the MaxWeight and the Exp rule are transition monotone, neither is radially sum-rate monotone. For $n = 2$ and extending the domain of $\boldsymbol{\xi}$ and $\boldsymbol{x}$ to $\mathbb{R}_+^{\bar{n}}$, Fig. 2 shows the curves in $\mathbb{R}_+^2$ along which the vector fields $\boldsymbol{\xi}$ and $\boldsymbol{x}$ hold their direction, i.e.,

$$\left\{\boldsymbol{q} : \frac{\xi_2(\boldsymbol{q})}{\xi_1(\boldsymbol{q})} = \text{``constant''}\right\} \quad \left\{\boldsymbol{q} : \frac{x_2(\boldsymbol{q})}{x_1(\boldsymbol{q})} = \text{``constant''}\right\}$$

for various values of "constant." Curves like these form the boundaries of the decision regions, i.e., the switching curves. The vector field $\boldsymbol{\xi}$ is homogeneous, hence the service rate allocation under MaxWeight is invariant as the queues grow from state $\boldsymbol{q}$ to state $\theta\boldsymbol{q}$. By contrast, in the case of the Exp rule (with $\boldsymbol{b}$ set to $\boldsymbol{w}$), the total weighted service rate $\langle \boldsymbol{w}, \boldsymbol{\mu}^X(\theta\boldsymbol{q})\rangle$ decreases with θ, and the emphasis shifts to queue -balancing, so much so that as $\theta \to \infty$, only the longest weighted queue(s) receives service.

V. Improved Throughput-Optimal Policies

We begin this section with a sufficiency theorem regarding throughput-optimal policies.

Theorem 1: Let $\boldsymbol{g}$: $\mathbb{R}_+^{\bar{n}} \to \mathbb{R}_+^{\bar{n}}$ be a gradient field (i.e., $\boldsymbol{g} = \boldsymbol{\nabla} G$ for some G: $\mathbb{R}_+^{\bar{n}} \to \mathbb{R}$). Moreover, suppose $\boldsymbol{g}$ is differentiable on $\mathbb{R}_+^{\bar{n}}$ and for all $n \in \mathcal{N}$ and satisfies

$$\lim_{\boldsymbol{y}\to\infty:y_n=0} \frac{g_n(\boldsymbol{y})}{|\boldsymbol{g}(\boldsymbol{y})|} = 0 \tag{19}$$

$$\lim_{\boldsymbol{y}\to\infty} \frac{\partial g_n(\boldsymbol{y})/\partial y_n}{|\boldsymbol{g}(\boldsymbol{y})|} = 0 \tag{20}$$

and for some $\epsilon>0$, $|\boldsymbol{g}(\boldsymbol{y})|>\epsilon$ for all $\boldsymbol{y}$ outside a compact subset of $\mathbb{R}_+^{\bar{n}}$, then any policy $\hat{\boldsymbol{\mu}}$ satisfying the following, for all $\boldsymbol{q} \in \mathbb{Z}_+^{\bar{n}}$, is throughput-optimal:

$$\hat{\boldsymbol{\mu}}(\boldsymbol{q}) \in \arg\max_{\boldsymbol{u}\in\mathcal{C}} \langle \boldsymbol{u}, \boldsymbol{g}(\boldsymbol{q})\rangle. \tag{21}$$

Remark 2: The condition that $\boldsymbol{g}$ be a gradient field and $|\boldsymbol{g}(\boldsymbol{y})|>\epsilon$ outside a compact set is used to establish the existence of a potential (Lyapunov) function G such that $\boldsymbol{\nabla} G = \boldsymbol{g}$. Condition (19) is needed to ensure that when queue state vector $\boldsymbol{q}$ is large, the policy given by $\boldsymbol{g}$ is work-conserving, i.e., it does not allocate any service rate to an empty queue at the cost of nonempty queues. Condition (20) is used to ensure that the Hessian of G can be dominated by its gradient in the Taylor expansion (see proof). See Appendix B for proof. See [25] for a similar result with a slightly different system model; it improves upon the above result by not requiring $\boldsymbol{g}$ to be a gradient field.

Examples: Examples of functions that satisfy the conditions of Theorem 1 are $\boldsymbol{\xi}(\cdot)$ with its domain extended to $\mathbb{R}_+^{\bar{n}}$, $\boldsymbol{g}(\boldsymbol{y}) = (\exp(y_n^{\alpha}), n \in \mathcal{N})$ for $\alpha \in (0,1)$, and $\boldsymbol{g}(\boldsymbol{y}) = (\log(1+\log(1+y_n)), n \in \mathcal{N})$, indicating that a throughput-optimal policy can exhibit anywhere from sublogarithmical to almost exponential sensitivity to changes in queue lengths. Another interesting example is $\boldsymbol{g}(\boldsymbol{y}) = (y_n^{\alpha_1}(y_n + c)^{\alpha_2}, n \in \mathcal{N})$ for $\alpha_1 \geq \alpha_2 > 0$ and $c\geq 0$, which behave as MaxWeight with exponent α_1 near the origin in $\mathbb{Z}_+^{\bar{n}}$ and as MaxWeight with exponent α_2 radially far from the origin.

A. Log Rule

In this section, we consider a class of schedulers satisfying Theorem 1, which we refer to as the Log rule.

Definition 3: Arbitrarily fix $\boldsymbol{a} = (a_n, n \in \mathcal{N})>0$, $\boldsymbol{b} = (b_n, n \in \mathcal{N})>0$, and $c\geq 1$. For all $\boldsymbol{y} \in \mathbb{R}_+^{\bar{n}}$, let $\boldsymbol{g}^{\mathrm{L}}(\boldsymbol{y}) = \left(g_n^{\mathrm{L}}(\boldsymbol{y}), n \in \mathcal{N}\right)$, where $g_n^{\mathrm{L}}(\boldsymbol{y}) = b_n \log(c + a_n y_n)$. When the system is in state $(\boldsymbol{Q}, M) = (\boldsymbol{q}, m)$ [see (1)], choose a stochastic vector $\boldsymbol{\sigma}$ that satisfies

$$\sigma_{n^*} = 1 \quad \text{for some } n^* \in \arg\max_{n\in\mathcal{N}} \left\{ g_n^{\mathrm{L}}(\boldsymbol{q}) r_n(m) \right\}. \tag{22}$$

Theorem 2: The Log rule is radial sum-rate monotone w.r.t. weight vector b (with $c = 1$) and throughput-optimal.

Proof: First, for the throughput optimality, let $\boldsymbol{\mu}^{\mathrm{L}}(\boldsymbol{q}) = \left(\mu_n^{\mathrm{L}}(\boldsymbol{q}), n \in \mathcal{N}\right)$ denote the vector of service rates (under the Log rule) seen by the queues when $\boldsymbol{Q}(t) = \boldsymbol{q}$, i.e.,

$$\mu_n^{\mathrm{L}}(\boldsymbol{q}) = \mathbb{E}\left[r_n(M) f_n^{\mathrm{L}}(\boldsymbol{q}, M) \right] \tag{23}$$

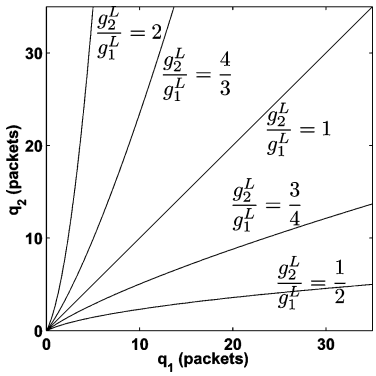

Fig. 3. Partitions under the Log rule with $a_n = 1$, $b_n = 1\ \forall\ n \in \mathcal{N}$, $c = 1$.

where $\boldsymbol{f}^{\mathrm{L}}(\boldsymbol{q}, m)$ denotes the stochastic vector chosen by the Log rule [i.e., satisfying (22)] in state $(\boldsymbol{q}, m)$. By (22) and Lemma 1, we have $\langle \boldsymbol{g}^{\mathrm{L}}(\boldsymbol{q}), \boldsymbol{\mu}^{\mathrm{L}}(\boldsymbol{q})\rangle = \max_{\boldsymbol{u}\in\mathcal{C}} \langle \boldsymbol{g}^{\mathrm{L}}(\boldsymbol{q}), \boldsymbol{u}\rangle$. Noting that the functions $\boldsymbol{g}^{\mathrm{L}}$ and $\boldsymbol{\mu}^{\mathrm{L}}$ satisfy the conditions of Theorem 1, the throughput-optimality of Log rule follows. To verify the radial sum-rate monotonicity of the Log rule, we note that for any $\boldsymbol{q} \in \mathbb{Z}_+^{\bar{n}}$ such that $0 < a_n q_n < a_{n'} q_{n'}$, we have $\frac{g_n^{\mathrm{L}}(\theta \boldsymbol{q})}{g_{nX'}^{\mathrm{L}}(\theta \boldsymbol{q})} \uparrow b_n/b_{n'}$ as $\theta \to \infty$. ∎

With $c = 1$ and extending the domain of $\boldsymbol{g}^{\mathrm{L}}$ to $\mathbb{R}_+^{\bar{n}}$, we note that as $\theta \to 0$, we have $g_n^{\mathrm{L}}(\theta \boldsymbol{q})/g_{n'}^{\mathrm{L}}(\theta \boldsymbol{q}) \to a_n q_n / a_{n'} q_{n'}$, i.e., close to origin in $\mathbb{Z}_+^{\bar{n}}$, the Log rule behaves similar to the MaxWeight with $\alpha = 1$, whereas radially far away from origin (as $\theta \to \infty$), $\boldsymbol{g}^{\mathrm{L}}(\theta \boldsymbol{q})$ becomes parallel to the vector $(b_n \mathbb{1}_{\{q_n>0\}}, n \in \mathcal{N})$, and thus the Log rule ignores queue balancing in favor of maximizing the total weighted service rate, $\left\langle \boldsymbol{b}, \boldsymbol{\mu}^{\mathrm{L}}(\theta \boldsymbol{q}) \right\rangle$.

Fig. 3 shows the curves along which the direction of the gradient field $\boldsymbol{g}^{\mathrm{L}}$ is constant; curves like these form the switching curves and define partition of the queue state-space into decision regions. A good choice for w_n (hence b_n) is $1/\mathbb{E}[R_n]$, as suggested for the Exp rule in [27]. The line $\left\{ \boldsymbol{q} \in \mathbb{Z}_+^{\bar{n}} : a_n q_n = a_{n'} q_{n'}\ \forall\ n, n' \in \mathcal{N} \right\}$ defines the axis of the weighted max-rate horn, whereas the magnitude of the vector $\boldsymbol{a}$ controls the width of the horn. Increasing the magnitude of $\boldsymbol{a}$ widens the horn and reduces the emphasis of Log rule on balancing user queues (this is opposite to the role this parameter plays in the Exp rule). By choosing $c>1$, the Log rule can be made to behave similar to the Exp rule, instead of MaxWeight with $\alpha = 1$, near the origin in $\mathbb{Z}_+^{\bar{n}}$.

Asymptotic Probability of Sum-Queue Overflow Under the Log Rule: Due to space limitations, we have relegated the proof of the Log rule's asymptotic optimality to a separate paper [17]. By leveraging the refined sample path large deviations principle, recently introduced in [15] to study nonhomogenous schedulers such as the Exp and the Log rules, we are able to show that for a $\bar{n} = 2$ user system, a Log rule-like radial sum-rate monotone policy [w.r.t. a given weight vector (w_1, w_2)] indeed minimizes the asymptotic probability of weighted-sum-queue overflow, i.e.,

$$\limsup_{k\to\infty} \frac{1}{k} \log \mathbb{P}\left(\sum_{n\in\mathcal{N}} w_n Q_n(0) > k \right)$$

where $\mathbb{P}(\cdot)$ denotes the stationary distribution of the Markov chain $\boldsymbol{Q}$ under a stable scheduling policy.

The *most-likely mode of queue overflow* under an RSM policy like the Log rule is in general quite different from the mode under the Exp rule. Recall that the latter minimizes the asymptotic probability of *max*-queue overflow [15], whereas the former minimizes the asymptotic probability of (weighted) *sum*-queue overflow.

This leads to basic questions as to which design objective is appropriate in designing opportunistic schedulers, and whether the asymptotic results are sufficiently accurate to dictate which class of scheduler should be used. We consider this in Section VI.

VI. Evaluating Opportunistic Scheduler Design Objectives—Simulations

In this section, we discuss a simulation-based evaluation of opportunistic schedulers from various perspectives:

1) performance, including mean packet delays and 99th-percentile delays of individual users as well the overall system;
2) sensitivity to both scheduler parameters and channel characteristics;
3) graceful degradation, in terms of the fraction of users that meet QoS objectives under overloads.

Note we consider a system as overloaded if it can no longer meet users' QoS requirements; this might be due to a change in the channel characteristics, mobility, etc. These perspectives are clearly interrelated, yet for clarity we discuss them separately.

A. Simulation Model and Operational Scenarios

We choose an HDR-like wireless downlink [18] to compare various scheduling rules, namely Log rule, MaxWeight, and Exp rule. Performance comparisons for an HDR downlink under Proportional Fair scheduling, MaxWeight, and Exp rule were presented in [27] and showed the Exp rule to be superior to the others. Note that the HDR downlink model differs from the system model presented in Section II, however we choose this as our simulation model to demonstrate the practical significance of our proposed scheduling rule and allow comparison to other simulations and theoretical work in the literature. Thus, instead of i.i.d. channel, continuous time scheduling, and Poisson arrivals with exponentially distributed packets sizes, here we assume that channels are *correlated* over time, scheduling decisions are made once in each time slot of duration 1.67 ms, and each user's packets are 1 kb and arrive as i.i.d. Bernoulli processes.

We consider $\bar{n} = 12$ heterogenous users connected to a single access point. The locations of the $\bar{n}$ users are taken to be uniformly distributed in a circular cell. The wireless link between the access point and each user is taken as an independent Rayleigh fading channel with a Doppler frequency of 18 Hz. Specifically, in any time slot $t \in \mathbb{Z}$, the channel state (rate supported by the channel) of the nth user is given by

$$R_n(t) \equiv \mathrm{BW} \times \log_2\left(1 + \mathrm{SINR}_n(t)\right) \text{ bits/s}$$

where $\mathrm{BW} = 800$ b/s and signal-to-interference-plus-noise ratio (SINR) is assumed to hold its value over the duration of the time slot. During each time slot, data is transmitted to a single user who is selected according to the scheduling policy. If user n is selected in time slot t, then (at most) $1.67\ \mathrm{ms} \times R_n(t)$ bits are transmitted from its queue. Table I gives the mean data rate $\mathbb{E}[R_n]$ in bits/second that the wireless channel of each user can support.

TABLE I
Mean Data Rate Supported by Wireless Channel of Each User

User i	1	2	3	4	5	6
$\mathbb{E}[R_n]$ kbps	572.8	392.1	304.6	250.1	215.1	187.9
User i	7	8	9	10	11	12
$\mathbb{E}[R_n]$ kbps	167.6	151.3	138.0	127.2	117.1	109.6

Let $\mathcal{C}^{\mathrm{no}} \in \mathbb{R}_+^{\bar{n}}$ denote the simplex obtained as the convex hull of origin and the following $\bar{n}$ points:

$$\mathbb{E}[R_1]\boldsymbol{e}_1, \ldots, \mathbb{E}[R_{\bar{n}}]\boldsymbol{e}_{\bar{n}}.$$

That is, $\mathcal{C}^{\mathrm{no}} \subset \mathcal{C}$ is the capacity region achievable by *nonopportunistic* (but possibly channel-aware) schedulers. Let $\boldsymbol{r}^* \equiv (r_n^*, n \in \mathcal{N}) \in \mathcal{C}^{\mathrm{no}}$ be the maximal point of $\mathcal{C}^{\mathrm{no}}$ that satisfies $r_1^* = r_2^* = \cdots = r_{\bar{n}}^*$, then we have [18]

$$r_n^* = \left(\sum_{j=1}^{\bar{n}} \frac{1}{\mathbb{E}[R_j]}\right)^{-1} \text{bits/s}, \quad n \in \mathcal{N}.$$

We present simulation results for five operational scenarios. In the first three scenarios, users see heterogenous channels, but have homogenous traffic with *low* $\boldsymbol{\lambda}^{(\mathrm{s,l})}$, *medium* $\boldsymbol{\lambda}^{(\mathrm{s,m})}$, or *high* $\boldsymbol{\lambda}^{(\mathrm{s,h})}$ packet arrival rates given by

$$\begin{aligned}\boldsymbol{\lambda}^{(\mathrm{s,m})} &= 2.3\boldsymbol{r}^* \times \frac{1}{1024 \text{ bits/packet}} \text{ packets/s}\\ \boldsymbol{\lambda}^{(\mathrm{s,l})} &= 0.98\boldsymbol{\lambda}^{(\mathrm{s,m})}\\ \boldsymbol{\lambda}^{(\mathrm{s,h})} &= 1.02\boldsymbol{\lambda}^{(\mathrm{s,m})}.\end{aligned}$$

In words, for the medium case, a user's arrival rate is 2.3 times higher than that stabilizable by a nonopportunistic scheduler; the low and the high arrival rates are respectively 2% lower and higher than the medium. Fig. 4(a) and (b) shows performance results under these three homogenous load scenarios; see the caption for a detailed explanation. In the fourth scenario, the arrival rate is kept *low*, but User 7 (see Table I) is moved to the edge of cell, which increases the system load. Fig. 4(c) exhibits the results for this case. For the fifth scenario, users have heterogenous arrival rates given by

$$\lambda_n = 2.35 \times \frac{\mathbb{E}[R_n]}{n} \times \frac{1}{1024} \text{ packets/s}$$

i.e., arrival rate vector $\boldsymbol{\lambda}$ is proportional to the mean channel rate vector $\mathbb{E}[\boldsymbol{R}]$ and 2.35 times higher than that stabilizable by a nonopportunistic scheduler. Fig. 4(d) exhibits the performance results for this case.

B. Discussion of Results and Insights

1) Performance Comparisons: As seen in Fig. 4(a) and (b) under *low* traffic, Log and Exp rules are comparable and outperform MaxWeight. Although users see heterogenous channels, the performance they see is very similar, verifying that we have a good choice for the scheduling policy parameters; see Table II. However, as the traffic rate increases, there are clear trends: The users' and overall means are better under the Log rule (up to 20% reduction), while the variability or spread of

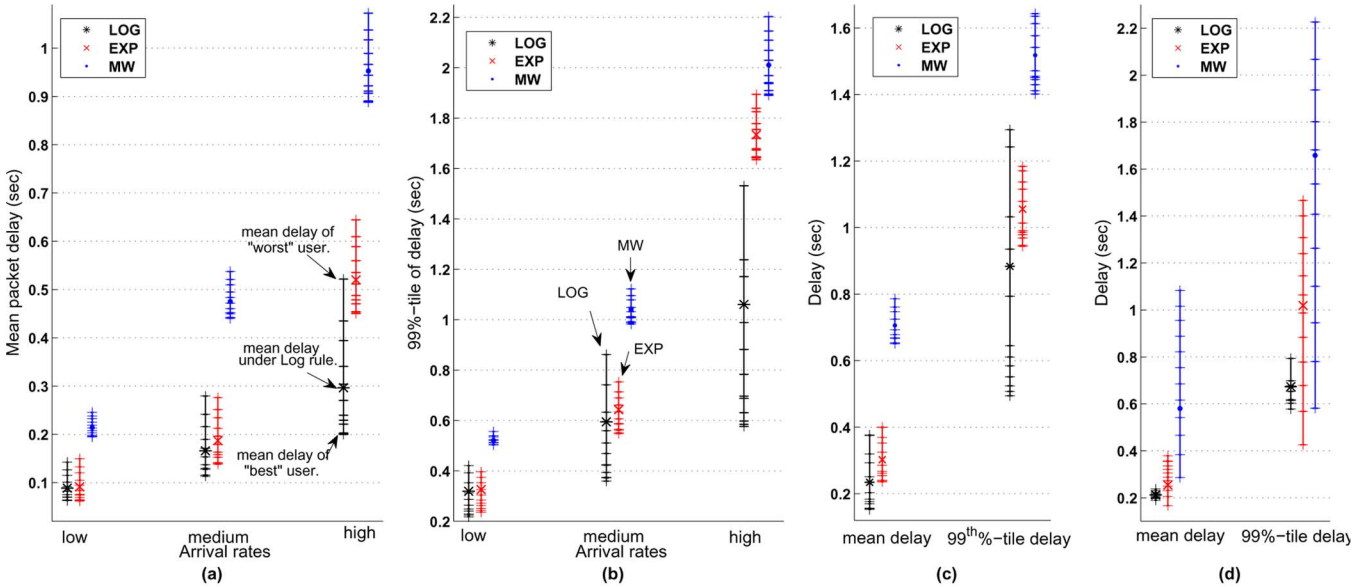

Fig. 4. Simulation-based performance comparisons for three opportunistic scheduling policies: Log rule, Exp rule, and Max Weight. (a) Mean delay and (b) 99th-percentile delay for each user and overall system, under *low*, *medium*, and *high* symmetric loads. (c) Mean and 99th-percentile delay for each user and overall system under *low* symmetric traffic, but when User 7 is moved to the cell edge. (d) Mean and 99th-percentile delay for users and overall system for the asymmetric traffic. Each cross-tick on a vertical line marks a user's performance.

TABLE II
PARAMETERS USED FOR EACH SCHEDULING POLICY

LOG	EXP	MW
$b_n = \frac{1}{\mathbb{E}[R_n]}$ $a_n = 10$ $c = 10$	$b_n = \frac{1}{\mathbb{E}[R_n]}$ $a_n = 0.05$, $c = 1, \eta = 0.5$	$b_n = \frac{1}{\mathbb{E}[R_n]}$ $\alpha = 1$

the 99th-percentile delay across users is lower under the Exp rule (the 99th-percentile delay spread is halved). Note, however, that all but two users have 5%–70% better 99th-percentile delay under the Log rule versus the Exp rule. The situation is even more favorable to the Log rule at higher loads, where all users experience 20%–80% lower mean and 99th-percentile delays versus the Exp rule (which still maintains a lower delay spread than the Log rule). Clearly for heterogenous channels, the Exp Rule's strong bias toward balancing queues is excessively compromising the realized throughput, and eventually the mean delays and tails for almost all users. Although asymptotically Exp rule should be optimal, the pre-exponent must also be playing a role in determining the systems performance.

2) Sensitivity: Another way to view this is that the actual performance (not the theoretical asymptotic tail) achieved by the Exp rule is more sensitive to the absolute values of $\boldsymbol{a}$.Fig. 4(a) and (b) exhibits the degeneration in the relative performance of Exp versus Log rule for a set of fixed parameters as the load is scaled up. The RSM property of the Log rule naturally calibrates the scheduler to increased load. Similarly, comparing the *low* and the *medium* results in Fig. 4(a) and (b) to those in Fig. 4(c) and (d), we see the performance sensitivity to changes in the channel or load characteristics. In both cases, for most users, the mean and 99th-percentile delays are better under the Log rule, and in the case of heterogenous loads, i.e., Fig. 4(d), the delay spreads are also improved. So unless parameters can be carefully tuned to possibly changing loads and unpredictable channel capacities, the Log rule appears to be a more robust scheduling policy. Intuitively, this is what one would expect from optimizing for the overall average versus worst-case asymptotic tail.

3) Graceful Degradation: Suppose the user flows correspond to buffered streaming audio sessions with a QoS requirement of 99th-percentile delay below 1 s; see, e.g., [27]. Under medium traffic [Fig. 4(b)], all users comfortably meet the QoS requirement for both the Log and the Exp rule. However, if User 7 moves to the cell edge [Fig. 4(c)], then under the Log rule, 9 out of 12 users versus 6 out of 12 for Exp rule meet the QoS requirement. If instead, the traffic loads associated with the users were to change, then, as shown in Fig. 4(d), all users meet the QoS requirement under the Log rule versus only 6 out of 12 under the Exp rule. Unless system resource is provisioned extremely conservatively, i.e., for worst case, we can expect such scenarios to arise, and this work suggests Log rule would provide a more graceful degradation of service.

VII. SCHEDULING IN MULTICHANNEL SYSTEMS

This section focuses on implementation of the Log rule for scheduling and resource allocation in OFDMA-based multichannel systems, e.g., WiMax, LTE. We begin by appropriately modifying the single-channel HDR-like system model used in the previous section to now capture an OFDMA-based multichannel system where power and bandwidth can be shared across multiple users over a scheduling/transmission time interval.

So far, the implied meaning of a channel state $m \in \mathcal{M}$ has been "a collection of quantized SNRs measured and reported by each user." Since we were considering a TDMA system where the scheduled user was allocated all the resources (power and bandwidth), we implicitly converted the SNR reported by the nth user into the supported transmission rate $r_n(m)$. Therefore, the region $\mathcal{C}_m$—available service rate region conditional on channel being in state m—was given by a simplex defined

by rates $(r_n(m), n \in \mathcal{N})$. All scheduling policies considered so far [see (15), (17), (18), and (22)] picked a vertex of the simplex $\mathcal{C}_m$ or, equivalently, scheduled a single user.

However, in wideband/multichannel systems, it is undesirable (and oftentimes even infeasible) to allocate all the resources to one user over a scheduling time interval [28]. The main reasons are as follows and will be addressed by our proposed implementation.

- Allocating all resources to one user in the presence of typically hundreds of active users will result in bursty service with long delays between successive allocations to a user.
- The service rate region available under all possible power and bandwidth allocations to *multiple* users over a scheduling time interval is larger than the simplex $\mathcal{C}_m$ defined above.

A. Modifications to the System Model

As in the previous section, we consider a time-slotted system. We can capture a multichannel system by appropriately redefining the meaning of channel state m and associating with it a suitable service region $\mathcal{C}_m$.

- The channel state is now defined as the collection of quantized SNRs reported by each user and measured at a reference power level for each *resource block* (RB)[4] group (collection of a few consecutive resource blocks). We continue to denote by $\mathcal{M}$ the set of all possible channel states.
- For each $m \in \mathcal{M}$, the service region $\mathcal{C}_m$ is the convex hull of the service rates (in bits/time slot) that can be jointly offered to the $\bar{n}$ users under all *feasible* resource (power and bandwidth) allocations, conditional on the channel being in state m.

Feasibility is determined by the system specifications and computational complexity afforded, e.g., limits on the minimum and maximum bandwidth that can be allocated to a user, and limits on power per user and per RB; see, e.g., [29] for a formal description of the rate region $\mathcal{C}_m$ in terms of various feasibility constraints. Also, we let $\boldsymbol{q}$ denote the queue length vector in number of *bits* rather than packets.

B. Scheduling and Resource Allocation Polices

The general channel-aware (but queue-oblivious) rate-adaptive scheduling and resource allocation problem for an OFDMA system is typically defined as follows (see [28], [29], and references therein): When the channel is in state m, allocate resources to users so that the long-run average offered service rate conditional on the channel being in state m solves the following program:

$$\begin{aligned} \text{maximize} \quad & \mathcal{U}(\boldsymbol{u}) \\ \text{subject to} \quad & \boldsymbol{u} \in \mathcal{C}_m \end{aligned} \tag{24}$$

where $\mathcal{U}$: $\mathbb{R}_+^{\bar{n}} \to \mathbb{R}$ is a given *utility* function satisfying concavity, smoothness, and separability properties. An optimal rate $\boldsymbol{u}^*$ [i.e., maximizer of (24)] corresponds to an allocation of RBs and transmit power across users. However, much like the single-channel case, such a queue-oblivious scheduler will not be throughput-optimal.

[4]Resource blocks are the smallest chunks of bandwidth that can be allocated to a user over a scheduling time interval.

Using a convex program formulation like that in (24), we shall define the Log rule for multichannel systems as follows.

Definition 4: When the system is in state $(\boldsymbol{q}, m)$, allocate resources so that the long-run average offered service rate conditional on the system being in state $(\boldsymbol{q}, m)$ is given by the solution to the following program:

$$\begin{aligned} \text{minimize} \quad & h^{\mathrm{L}}(\boldsymbol{q} - \boldsymbol{u}) \\ \text{subject to} \quad & \boldsymbol{u} \in \mathcal{C}_m \end{aligned} \tag{25}$$

where h^{L}: $\mathbb{R}_+^{\bar{n}} \to \mathbb{R}$ is given by

$$h^{\mathrm{L}}(\boldsymbol{y}) \equiv \sum_{n \in \mathcal{N}} b_n \left(\left(\frac{c}{a_n} + y_n \right) \log(c + a_n y_n) - y_n \right)$$

and the constants $(a_n, n \in \mathcal{N})$, $(b_n, n \in \mathcal{N})$, and c are as in the definition of the single-channel Log rule (see Definition 3).

The function h^{L} is convex-increasing and can be viewed as an approximation for the relative *cost function* on the queue state-space, satisfying the Bellman's equation. Also, note that

$$\begin{aligned} \nabla h^{\mathrm{L}}(\boldsymbol{q} - \boldsymbol{0}) &\equiv \left(\frac{\partial h^{\mathrm{L}}(\boldsymbol{q} - \boldsymbol{u})}{\partial u_n}, n \in \mathcal{N} \right) \Bigg|_{\boldsymbol{u}=\boldsymbol{0}} \\ &= -\boldsymbol{g}^{\mathrm{L}}(\boldsymbol{q}). \end{aligned}$$

Therefore, for a single-channel system where $\mathcal{C}_m$ was a *small* simplex, the single-channel Log rule [see (22)] can be viewed as linearizing the convex program (25) using the first-order Taylor expansion of $h^{\mathrm{L}}(\boldsymbol{q} - \boldsymbol{u})$ at $\boldsymbol{u} = \boldsymbol{0}$, i.e.,

$$\begin{aligned} \text{maximize} \quad & \langle \boldsymbol{g}^{\mathrm{L}}(\boldsymbol{q}), \boldsymbol{u} \rangle \\ \text{subject to} \quad & \boldsymbol{u} \in \mathcal{C}_m. \end{aligned}$$

Remark 3: The linearized version stated above may not be suitable in a multichannel system because if the region $\mathcal{C}_m$ is a simplex (or *close* to a simplex), the linearized program reduces to picking a vertex and thus a single user (or allocating most of the resources to a single user) even if all weighted queues $a_n q_n$ are equal. As mentioned earlier, this is undesirable in a multichannel system. The region $\mathcal{C}_m$ in a multichannel system can still be a simplex if, for instance, the users only report one effective SNR over the entire bandwidth (e.g., wideband CQI in LTE [30]) and the power allocated per resource block is fixed (i.e., the mapping from the reported CQI to the chosen modulation and coding scheme is fixed). Moreover, the analysis in [31] shows that for certain symmetric ON–OFF multichannel systems, any resource allocation policy given by a linear program will have a zero large-deviation rate function associated with the max-queue (asymptotically in the number of users and channels and in the *small buffer* regime.)

Fast computation algorithms to solve program (24) [or (25)] for a general concave-increasing separable utility function $\mathcal{U}(\cdot)$ are given in [28] and [29] (and references therein). The complexity of the algorithm depends on the feasibility constraints that define the region $\mathcal{C}_m$. For example, assuming that service rate at any SNR is equal to the Shannon's capacity (with a possible gap), a total power constraint, and that each resource block group can be shared by an arbitrary number of users, the algorithm obtained in [28] has a complexity of $O(\bar{n}\bar{b})$ per *iteration*, where $\bar{b}$ is the number of resource block groups over which the users report their measured SNR/CQI. It is reported

that typically about 25 iterations are needed for the algorithm to converge. For a similar system where users only report a single effective SNR measured across the entire bandwidth (however, the power and/or bandwidth can still be shared by multiple users), the complexity of the algorithm reduces to $O(\bar{n})$ per iteration, which is the complexity of single-channel scheduling algorithms [see (17), (18), and (22)] that pick only a single user.

See [32] for an implementation of a complete queue-and-channel-aware scheduler for an LTE downlink and a performance comparison through simulation of the Log and Exp rules. The simulation results presented in [32] for a multi-channel system agree with ones presented in this paper for a simple single-channel system and reinforce the observations made in Section VI-B.

VIII. Conclusion

This paper has made the case not only for a new class of opportunistic scheduling policies, but also for new metrics to design and evaluate such schedulers. Our conclusion is simple and, in retrospect, intuitive. A scheduler "optimized" for the overall system performance is likely to be more robust to changes in the traffic and channel statistics than the one optimized for the worst case. The numerical results presented in this paper show that mean-delay optimal schedulers are weakly RSM and, in some cases, even RSM. Further asymptotic results in a companion paper show that an RSM policy minimizes the tail of sum-queue distribution. The proposed Log rule policy is RSM and, although not necessarily mean-delay-optimal for a given scenario, exhibits the promised robustness versus the Exp and MaxWeight rules. The set of presented simulations (and others not included) lends support to the practical benefits of this new class of policies.

Appendix A
Fluid-Scale Asymptotic Optimality and RSM Policies

In Section IV-A, we observed that optimal policies $\boldsymbol{\mu}^*$ for the MDP, computed through relative value iteration for different capacity regions and arrival vectors, satisfied condition 1) in the definition of RSM (see Definition 1). In this Appendix, we will consider if and when RSM polices are *fluid-scale asymptotic optimal* (FSAO), which is formally defined in the second section. Roughly speaking, the asymptotic slopes of the switching curves under $\boldsymbol{\mu}^*$ and FSAO policies are identical.

For the MDP, we have already defined an optimal policy $\boldsymbol{\mu}^*$ by (12) and (13), and a representative RSM policy, namely the Log rule $\boldsymbol{\mu}^{\mathrm{L}}$, by (23). Paralleling this, next, for a deterministic fluid model, we will introduce an optimal fluid policy $\boldsymbol{\mu}^{\mathrm{F}*}$ and a greedy fluid policy $\boldsymbol{\mu}^{\mathrm{Fg}}$. The two policies for the MDP are related to the two policies for the fluid model, as shown in Fig. 5. Using these relationships, one can show the following.

1) If $\boldsymbol{\mu}^{\mathrm{Fg}}$ is not an optimal policy for the fluid model, then $\boldsymbol{\mu}^*$ does not satisfy condition 2) in the definition of RSM.
2) Otherwise, RSM policies like the Log rule are FSAO. Furthermore, if $\boldsymbol{\mu}^{\mathrm{Fg}}$ is the *unique* optimal policy for the fluid model, then it follows that $\boldsymbol{\mu}^*$ and $\boldsymbol{\mu}^{\mathrm{L}}$ have similar switching curves on the state-space of appropriately scaled queue process.

The formal description follows. Subsequently, we will also consider examples of fluid models where the greedy policy may or may not be optimal.

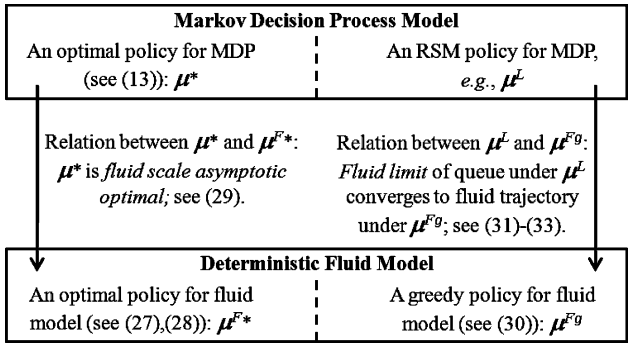

Fig. 5. Relation between policy $\boldsymbol{\mu}^*$ for MDP and policy $\boldsymbol{\mu}^{\mathrm{F}*}$ for fluid model, and between policy $\boldsymbol{\mu}^{\mathrm{L}}$ for MDP and policy $\boldsymbol{\mu}^{\mathrm{Fg}}$ for fluid model.

A. Deterministic Fluid Model

One can associate a fluid model (see, e.g., [26] and [33]) with the MDP defined in Section III, as follows. Let $(\boldsymbol{x}(t), t \geq 0)$ be a deterministic continuous trajectory starting at point $\boldsymbol{x}(0) = \boldsymbol{y} \in \mathbb{R}_+^{\bar{n}}$ and evolving as

$$\boldsymbol{x}(t) = \boldsymbol{y} + \boldsymbol{\lambda} t - \int_0^t \boldsymbol{u}(\tau) d\tau$$

where the *control* $\boldsymbol{u}(\cdot) \in \mathcal{C}$ and the control policy $(\boldsymbol{u}(t), t \geq 0)$ is measurable and keeps $\boldsymbol{x}(\cdot)$ in $\mathbb{R}_+^{\bar{n}}$. To indicate the dependence of the trajectory on the control policy and initial condition, subsequently we will write $\boldsymbol{x}^{\boldsymbol{u}}(t; \boldsymbol{y})$ instead of $\boldsymbol{x}(t)$. The cost of trajectory $(\boldsymbol{x}^{\boldsymbol{u}}(t; \boldsymbol{y}), t \geq 0)$ in turn is given by

$$J_{\boldsymbol{u}}^{\mathrm{F}}(\boldsymbol{y}) \equiv \int_0^\infty \langle \boldsymbol{w}, \boldsymbol{x}^{\boldsymbol{u}}(t; \boldsymbol{y}) \rangle dt. \tag{26}$$

Since $\boldsymbol{\lambda}$ lies in the interior of $\mathcal{C}$ [see (7)], there exists at least one control policy $(\boldsymbol{u}(t), t \geq 0)$ under which the cost is finite for all finite starting points $\boldsymbol{x}(0)$.

If $\boldsymbol{u}(t) = \boldsymbol{\mu}^{\mathrm{F}}(\boldsymbol{x}(t))$, then we will refer to $\boldsymbol{\mu}^{\mathrm{F}}$: $\mathbb{R}_+^{\bar{n}} \to \mathcal{C}$ as a state-feedback control policy for the fluid model. It is shown in [33] that there exists a state-feedback policy $\boldsymbol{\mu}^{\mathrm{F}*}$ with the following properties.

- For any starting point $\boldsymbol{y} \in \mathbb{R}_+^{\bar{n}}$, the trajectory $(\boldsymbol{x}^{\boldsymbol{\mu}^{\mathrm{F}*}}(t; \boldsymbol{y}), t \geq 0)$ is absolutely continuous (and thus differentiable a.e.) and satisfies for all regular t

$$\frac{d}{dt}\boldsymbol{x}(t) = \boldsymbol{\lambda} - \boldsymbol{\mu}^{\mathrm{F}*}(\boldsymbol{x}(t)).$$

- The policy $\boldsymbol{\mu}^{\mathrm{F}*}$ is optimal for the fluid model. That is, for any admissible policy $(\boldsymbol{u}(t), t \geq 0)$, we have

$$J_{\boldsymbol{\mu}^{\mathrm{F}*}}^{\mathrm{F}}(\boldsymbol{y}) = \int_0^\infty \left\langle \boldsymbol{w}, \boldsymbol{x}^{\boldsymbol{\mu}^{\mathrm{F}*}}(t; \boldsymbol{y}) \right\rangle dt \leq J_{\boldsymbol{u}}^{\mathrm{F}}(\boldsymbol{y}). \tag{27}$$

Moreover, *any* state-feedback optimal fluid policy $\boldsymbol{\mu}^{\mathrm{F}*}$ must satisfy the Hamilton–Jacobi–Bellman equation (see [26, Proposition 4.3.2]), i.e., for all $\boldsymbol{x} \in \mathbb{R}_+^{\bar{n}}$

$$\boldsymbol{\mu}^{\mathrm{F}*}(\boldsymbol{x}) \in \arg\max_{\boldsymbol{u} \in \mathcal{C}_{\boldsymbol{x}}} \left\langle \nabla J_{\boldsymbol{\mu}^{\mathrm{F}*}}^{\mathrm{F}}(\boldsymbol{x}), \boldsymbol{u} \right\rangle \tag{28}$$

where $\mathcal{C}_{\boldsymbol{x}} = \{\boldsymbol{u} \in \mathcal{C} : \forall\, n \in \mathcal{N}, \text{ if } x_n = 0 \Rightarrow u_n \leq \lambda_n\}$ is the set of admissible controls in state $\boldsymbol{x}$. Subsequently we will use (28) to test a candidate fluid policy for optimality.

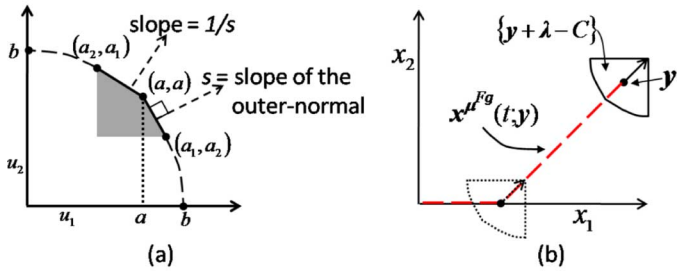

Fig. 6. Fluid model: (a) capacity region $\mathcal{C}$, (b) fluid trajectory $\boldsymbol{x}^{\boldsymbol{\mu}^{Fg}}(\cdot;\boldsymbol{y})$ starting at point $\boldsymbol{y}$ with $y_1 > y_2$.

B. Fluid-Scale Asymptotic Optimality

For each integer $\theta > 0$, consider an independent Markov chain, $\left(\boldsymbol{Q}_k^{(\theta)}, k\geq 0\right)$, starting in state $\boldsymbol{q}^{(\theta)}(0) = (\lfloor \theta y_n \rfloor, n \in \mathcal{N}) \in \mathbb{Z}_+^{\bar{n}}$ and evolving under a scheduling policy $\boldsymbol{\mu}$. Let $\bar{\boldsymbol{Q}}^{(\theta)}(t)$ denote the *fluid-scaled* version of the Markov chain, i.e.,

$$\bar{\boldsymbol{Q}}^{(\theta)}(t) \equiv \frac{1}{\theta}\boldsymbol{Q}^{(\theta)}_{\lfloor \theta t \rfloor}, \quad t \in \mathbb{R}_+.$$

Then, the policy $\boldsymbol{\mu}$ is said to be FSAO if it satisfies the following [33]:

$$\lim_{t\to\infty}\lim_{\theta\to\infty} \mathbb{E}^{\boldsymbol{\mu}}_{\bar{\boldsymbol{q}}^{(\theta)}(0)}\left[\int_0^t \left\langle \boldsymbol{w}, \bar{\boldsymbol{Q}}^{(\theta)}(\tau)\right\rangle d\tau\right] = J^{\mathrm{F}}_{\boldsymbol{\mu}^{\mathrm{F}*}}(\boldsymbol{y}). \quad (29)$$

C. Relation Between $\boldsymbol{\mu}^*$ and $\boldsymbol{\mu}^{\mathrm{F}*}$

The optimal policies $\boldsymbol{\mu}^*$ for the MDP are FSAO; see [26, Theorem 10.0.5 and Sec. 10.6.2].

D. Relation Between RSM Policy for MDP and Greedy Policy for Fluid Model

For ease of exposition, we assume that the capacity region $\mathcal{C}$ and vectors $\boldsymbol{\lambda}$ and $\boldsymbol{w}$ are such that the $\arg\max_{\boldsymbol{u}\in\mathcal{C}_{\boldsymbol{x}}}\langle \boldsymbol{w}, \boldsymbol{u}\rangle$ is unique for all $\boldsymbol{x} \in \mathbb{R}_+^{\bar{n}}$. Let us define a policy $\boldsymbol{\mu}^{\mathrm{Fg}}$ that is *greedy* with respect to the weight vector $\boldsymbol{w}$. More precisely, $\boldsymbol{\mu}^{\mathrm{Fg}}$ is given as follows: For any $\boldsymbol{x} \in \mathbb{R}_+^{\bar{n}}$, we have

$$\boldsymbol{\mu}^{\mathrm{Fg}}(\boldsymbol{x}) \equiv \arg\max_{\boldsymbol{u}\in\mathcal{C}_{\boldsymbol{x}}} \langle \boldsymbol{w}, \boldsymbol{u}\rangle. \quad (30)$$

For example, for any $\boldsymbol{x}>\boldsymbol{0}$, the control $\boldsymbol{\mu}^{\mathrm{Fg}}(\boldsymbol{x})$ is equal to the (weighted) max-sum-rate vertex $\arg\max_{\boldsymbol{u}\in\mathcal{C}}\langle \boldsymbol{w}, \boldsymbol{u}\rangle$.

Next, we describe the relation between the RSM policy and the greedy policy. For each integer $\theta>0$, consider an independent Markov chain, $\left(\boldsymbol{Q}_k^{(\theta)}, k \geq 0\right)$, starting in state $\boldsymbol{q}^{(\theta)}(0) = (\lfloor \theta y_n \rfloor, n \in \mathcal{N}) \in \mathbb{Z}_+^{\bar{n}}$ and evolving under an RSM policy. The analysis in [17, Lemmas 1-3] and [34] shows that under an RSM policy like the Log rule $\boldsymbol{\mu}^{\mathrm{L}}$, we have the following uniform convergence over compact sets along some subsequence $\{\theta_j\}$:

$$\left(\bar{\boldsymbol{Q}}^{(\theta_j)}(t), t \geq 0\right) \to (\bar{\boldsymbol{q}}(t), t \geq 0)$$

where the *fluid-limit* $(\bar{\boldsymbol{q}}(t),\ t \geq 0)$ with $\bar{\boldsymbol{q}}(0) = \boldsymbol{y}$ satisfies

$$\frac{d}{dt}\bar{\boldsymbol{q}}(t) = \boldsymbol{\lambda} - \arg\max_{\boldsymbol{u}\in\mathcal{C}_{\bar{\boldsymbol{q}}(t)}} \langle \boldsymbol{w}, \boldsymbol{u}\rangle. \quad (31)$$

Subsequently, we will write $(\bar{\boldsymbol{q}}(t;\boldsymbol{y}), t \geq 0)$ to explicitly indicate the starting point of the fluid-limit trajectory. By comparing (30) and (31), we have that the fluid limit under the RSM policy and the fluid trajectory under the greedy policy are identical, i.e.,

$$\boldsymbol{x}^{\boldsymbol{\mu}^{\mathrm{Fg}}}(t;\boldsymbol{y}) = \bar{\boldsymbol{q}}(t;\boldsymbol{y}), \quad t \geq 0. \quad (32)$$

While the proof in [17] and [34] is rather lengthy, the following observation captures the intuition: Recall the weighted max-rate horn described in Section IV-A; for any point of the fluid-limit trajectory in the interior of $\mathbb{R}_+^{\bar{n}}$, i.e., $\bar{\boldsymbol{q}}(t)>\boldsymbol{0}$, the unscaled queue $\theta_j\bar{\boldsymbol{Q}}^{(\theta_j)}(t)$ for large enough θ_j lies in the weighted max-rate horn.

Lastly, we have the following convergence in the mean:

$$\begin{aligned}\lim_{t\to\infty}\lim_{\theta\to\infty} &\mathbb{E}^{\boldsymbol{\mu}^{\mathrm{L}}}_{\bar{\boldsymbol{q}}^{(\theta)}(0)}\left[\int_0^t \left\langle \boldsymbol{w}, \bar{\boldsymbol{Q}}^{(\theta)}(\tau)\right\rangle d\tau\right] \\ &= \int_0^\infty \langle \boldsymbol{w}, \bar{\boldsymbol{q}}(\tau;\boldsymbol{y})\rangle\, d\tau \\ &= J^{\mathrm{F}}_{\boldsymbol{\mu}^{\mathrm{Fg}}}(\boldsymbol{y}). \end{aligned} \quad (33)$$

Main Result: See Fig. 5 and recall the relation between $\boldsymbol{\mu}^*$ and $\boldsymbol{\mu}^{\mathrm{F}*}$ and between $\boldsymbol{\mu}^{\mathrm{L}}$ and $\boldsymbol{\mu}^{\mathrm{Fg}}$. Suppose that for a given fluid model, the greedy policy $\boldsymbol{\mu}^{\mathrm{Fg}}$ is optimal, i.e., $J^{\mathrm{F}}_{\boldsymbol{\mu}^{\mathrm{F}*}}(\cdot) = J^{\mathrm{F}}_{\boldsymbol{\mu}^{\mathrm{Fg}}}(\cdot)$. Then, it follows from (29) and (33) that $\boldsymbol{\mu}^{\mathrm{L}}$ is FSAO. Furthermore, suppose $\boldsymbol{\mu}^{\mathrm{Fg}}$ is the *unique* (a.e.) optimal policy. Then, since $\boldsymbol{\mu}^*$ is also FSAO, the fluid limit $(\boldsymbol{q}(t;\cdot), t \geq 0)$ of the queue process under $\boldsymbol{\mu}^*$ (if the limit exists) and under $\boldsymbol{\mu}^{\mathrm{L}}$ will be identical and, in turn, identical to the deterministic fluid trajectory $\left(\boldsymbol{x}^{\boldsymbol{\mu}^{\mathrm{Fg}}}(t;\cdot), t \geq 0\right)$. In other words, the switching curves under $\boldsymbol{\mu}^*$ and $\boldsymbol{\mu}^{\mathrm{L}}$ on the state-space of fluid-scaled queue will be identical and, in turn, identical to the switching curves under the greedy policy $\boldsymbol{\mu}^{\mathrm{Fg}}$. Next, we show through a representative example that if $\boldsymbol{\lambda}$ is *not too small*, then $\boldsymbol{\mu}^{\mathrm{Fg}}$ is indeed the unique optimal policy for the fluid model.

To simplify the exposition, we restrict the system to $\bar{n} = 2$ users, weight vector $\boldsymbol{w} = (1,1)$, and a capacity region $\mathcal{C} \in \mathbb{R}_+^{\bar{n}}$ depicted in Fig. 6(a). That is, $\mathcal{C}$ has a unique max-sum-rate vertex (a,a), with the two adjacent vertices (a_1,a_2) and (a_2,a_1) satisfying $a_1+a_2<2a$. Let the point $(0,\hat{b})$ be such that the line segment joining points (a,a) and $(0,\hat{b})$ passes through vertex (a_1,a_2). Then, the region $\mathcal{C}$ intercepts with the two axes at points $(b,0)$ and $(0,b)$ for some $b \in (a,\hat{b})$. The remaining shape of $\mathcal{C}$ is unspecified and can be anything (as long as $\mathcal{C}$ remains convex).

Consider any symmetric vector $\boldsymbol{\lambda}$ in the shaded region of $\mathcal{C}$ in Fig. 6, i.e., $\lambda_1 = \lambda_2 \in [a_2, a)$. Fig. 6 depicts the trajectory $\left(\boldsymbol{x}^{\boldsymbol{\mu}^{\mathrm{Fg}}}(t;\boldsymbol{y}), t \geq 0\right)$ starting from some point $\boldsymbol{y}>\boldsymbol{0}$ such that $y_1>y_2$. The trajectory evolves as follows: If $\boldsymbol{x}^{\boldsymbol{\mu}^{\mathrm{Fg}}}(t;\boldsymbol{y})>\boldsymbol{0}$, then

$$\frac{d}{dt}\boldsymbol{x}^{\boldsymbol{\mu}^{\mathrm{Fg}}}(t;\boldsymbol{y}) = \boldsymbol{\lambda} - (a,a)$$

and if $x_1^{\boldsymbol{\mu}^{\mathrm{Fg}}}(t;\boldsymbol{y}) > 0, x_2^{\boldsymbol{\mu}^{\mathrm{Fg}}}(t;\boldsymbol{y}) = 0$, then

$$\frac{d}{dt}\boldsymbol{x}^{\boldsymbol{\mu}^{\mathrm{Fg}}}(t;\boldsymbol{y}) = \boldsymbol{\lambda} - (c(\lambda_2), \lambda_2)$$

where $c(\lambda_2)$ is the first coordinate of the point $(c(\lambda_2), \lambda_2)$ on the boundary of $\mathcal{C}$. Then, for any starting point $\boldsymbol{y}$ such that $y_1 \geq y_2$, it is easy to show that

$$J^{\mathrm{F}}_{\boldsymbol{\mu}^{\mathrm{Fg}}}(\boldsymbol{y}) = \frac{(y_1-y_2)^2}{2(c(\lambda_2)-\lambda_1)} + \frac{y_1 y_2}{a-\lambda_2}.$$

Note that $J^{\mathrm{F}}_{\boldsymbol{\mu}^{\mathrm{Fg}}}(\cdot)$ is homogenous, i.e., $J^{\mathrm{F}}_{\boldsymbol{\mu}^{\mathrm{Fg}}}(\theta \boldsymbol{y}) = \theta^2 J^{\mathrm{F}}_{\boldsymbol{\mu}^{\mathrm{Fg}}}(\boldsymbol{y})$, and so $\nabla J^{\mathrm{F}}_{\boldsymbol{\mu}^{\mathrm{Fg}}}(\boldsymbol{y}) \propto \nabla J^{\mathrm{F}}_{\boldsymbol{\mu}^{\mathrm{Fg}}}(\theta \boldsymbol{y})$. Now we are ready to test the policy $\boldsymbol{\mu}^{\mathrm{Fg}}$ for optimality using (28). Let s denote the slope of the outer-normal to the facet joining the vertex (a, a) and (a_1, a_2) of $\mathcal{C}$, i.e., $s = \frac{a_1 - a}{a - a_2}$; see Fig. 6. It can be shown that for any $\boldsymbol{y}>0$ such that $y_1 \geq y_2$, the slope of gradient $\nabla J^{\mathrm{F}}_{\boldsymbol{\mu}^{\mathrm{Fg}}}(\boldsymbol{y})$ lies in $(s, 1]$, i.e.,

$$\frac{\nabla_2 J^{\mathrm{F}}_{\boldsymbol{\mu}^{\mathrm{Fg}}}(\boldsymbol{y})}{\nabla_1 J^{\mathrm{F}}_{\boldsymbol{\mu}^{\mathrm{Fg}}}(\boldsymbol{y})} \in (s, 1]$$

whereas, on the x_1-axis, the slope is equal to s. Then, by (28), we have that $\boldsymbol{\mu}^{\mathrm{Fg}}$ is the unique optimal policy. In fact, along the lines of [35], one can show that $\boldsymbol{\mu}^{\mathrm{Fg}}$ is the optimal policy for any $\boldsymbol{\lambda}$, not necessarily symmetric, that lies in the shaded region of $\mathcal{C}$. Therefore, for the corresponding MDP with $\boldsymbol{\lambda}$, $\mathcal{C}$, and $\boldsymbol{w}$ as described above, $\boldsymbol{\mu}^{\mathrm{L}}$ is FSAO and has same switching curves on the state-space of fluid-scaled queue as those of $\boldsymbol{\mu}^*$.

However, it can also be shown that for any $\boldsymbol{\lambda}<(a_2, a_2)$, the policy $\boldsymbol{\mu}^{\mathrm{Fg}}$ is *not* optimal for the fluid model. That is, there exist states $\boldsymbol{x} \in \mathbb{R}^{\bar{n}}_+$ for which $J^{\mathrm{F}}_{\boldsymbol{\mu}^{\mathrm{Fg}}}(\boldsymbol{x}) > J^{\mathrm{F}}_{\boldsymbol{\mu}^{\mathrm{F}*}}(\boldsymbol{x})$. It follows from (29) and (33) that RSM policies like $\boldsymbol{\mu}^{\mathrm{L}}$ will *not* be FSAO. Since the optimal policies $\boldsymbol{\mu}^*$ are FSAO, they cannot be RSM.

Appendix B
Proof of Theorem 1

We will use Foster's Criterion to show that $(\boldsymbol{Q}_k, k \geq 0)$ is positive-recurrent for any stabilizable $\boldsymbol{\lambda}$ [see (7)]. Specifically, take a (Lyapunov) function G: $\mathbb{R}^{\bar{n}}_+ \to \mathbb{R}_+$ such that $\nabla G = \boldsymbol{g}$ and $G(\boldsymbol{0}) = 0$. Then

$$\begin{aligned}\mathbb{E}^{\hat{\boldsymbol{\mu}}}[G(\boldsymbol{Q}_{k+1}) - G(\boldsymbol{Q}_k) | \boldsymbol{Q}_k = \boldsymbol{q}] &= \gamma^{-1} \sum_{n \in \mathcal{N}} \lambda_n (G(\boldsymbol{A}_n \boldsymbol{q}) - G(\boldsymbol{q})) \\ &\quad + \gamma^{-1} \sum_{n \in \mathcal{N}_{\boldsymbol{q}}} \hat{\mu}_n(\boldsymbol{q}) (G(\boldsymbol{D}_n \boldsymbol{q}) - G(\boldsymbol{q})) \qquad (34)\end{aligned}$$

where, as before, $\mathcal{N}_{\boldsymbol{q}} = \{n: q_n \neq 0\}$. Since $\boldsymbol{g}$ is differentiable, define $\dot{g}_n(\cdot) = \partial g_n(\cdot)/\partial x_n$, $\forall\, n \in \mathcal{N}$. Then, G has the following Taylor expansion for any $\boldsymbol{q} \in \mathbb{Z}^{\bar{n}}_+$:

$$G(\boldsymbol{A}_n \boldsymbol{q}) - G(\boldsymbol{q}) = g_n(\boldsymbol{q}) + \frac{1}{2} \dot{g}_n(\boldsymbol{q} + \alpha_n \boldsymbol{e}_n) \qquad \forall\, n \in \mathcal{N}$$

$$G(\boldsymbol{D}_n \boldsymbol{q}) - G(\boldsymbol{q}) = -\, g_n(\boldsymbol{q}) + \frac{1}{2} \dot{g}_n(\boldsymbol{q} - \beta_n \boldsymbol{e}_n) \qquad \forall\, n \in \mathcal{N}_{\boldsymbol{q}}$$

for some $\alpha_n \in [0, 1]$ and $\beta_n \in [0, 1]$ that depend on $\boldsymbol{q}$. One can rewrite (34) as follows:

$$\begin{aligned}\mathbb{E}^{\hat{\boldsymbol{\mu}}}[G(\boldsymbol{Q}_{k+1}) - G(\boldsymbol{Q}_k) | \boldsymbol{Q}_k = \boldsymbol{q}] &= \gamma^{-1} \Big(\sum_{n \in \mathcal{N}} \lambda_n g_n(\boldsymbol{q}) + \frac{1}{2} \sum_{n \in \mathcal{N}} \lambda_n \dot{g}_n(\boldsymbol{q} + \alpha_n \boldsymbol{e}_n) \\ &\quad - \sum_{n \in \mathcal{N}_{\boldsymbol{q}}} \hat{\mu}_n(\boldsymbol{q}) g_n(\boldsymbol{q}) \\ &\quad + \frac{1}{2} \sum_{n \in \mathcal{N}_{\boldsymbol{q}}} \hat{\mu}_n(\boldsymbol{q}) \dot{g}_n(\boldsymbol{q} - \beta_n \boldsymbol{e}_n) \Big).\end{aligned}$$

Adding and subtracting $\gamma^{-1} \sum_{n \in \mathcal{N} \setminus \mathcal{N}_{\boldsymbol{q}}} \hat{\mu}_n(\boldsymbol{q}) g_n(\boldsymbol{q})$ from the left side of the above yields

$$\begin{aligned}&\mathbb{E}^{\hat{\boldsymbol{\mu}}}[G(\boldsymbol{Q}_{k+1}) - G(\boldsymbol{Q}_k) | \boldsymbol{Q}_k = \boldsymbol{q}] \\ &\quad = \gamma^{-1} \langle \boldsymbol{\lambda} - \hat{\boldsymbol{\mu}}(\boldsymbol{q}), \boldsymbol{g}(\boldsymbol{q}) \rangle + \gamma^{-1} \sum_{n \in \mathcal{N} \setminus \mathcal{N}_{\boldsymbol{q}}} \hat{\mu}_n(\boldsymbol{q}) g_n(\boldsymbol{q}) \\ &\qquad + \frac{\gamma^{-1}}{2} \sum_{n \in \mathcal{N}} \lambda_n \dot{g}_n(\boldsymbol{q} + \alpha_n \boldsymbol{e}_n) \\ &\qquad + \frac{\gamma^{-1}}{2} \sum_{n \in \mathcal{N}_{\boldsymbol{q}}} \hat{\mu}_n(\boldsymbol{q}) \dot{g}_n(\boldsymbol{q} - \beta_n \boldsymbol{e}_n) \\ &\quad \leq \gamma^{-1} \langle \boldsymbol{\lambda} - \hat{\boldsymbol{\mu}}(\boldsymbol{q}), \boldsymbol{g}(\boldsymbol{q}) \rangle + \gamma^{-1} \sum_{n \in \mathcal{N} \setminus \mathcal{N}_{\boldsymbol{q}}} \hat{\mu}_n(\boldsymbol{q}) g_n(\boldsymbol{q}) \\ &\qquad + \max \Big\{ \max_n \{\dot{g}_n(\boldsymbol{q} + \alpha_n \boldsymbol{e}_n) : n \in \mathcal{N}\}, \\ &\qquad\quad \max_n \{\dot{g}_n(\boldsymbol{q} - \beta_n \boldsymbol{e}_n) : n \in \mathcal{N}_{\boldsymbol{q}}\} \Big\}. \qquad (35)\end{aligned}$$

Let $\boldsymbol{u} \in \mathcal{C}$ be a service rate vector satisfying $\lambda_n < u_n$ for all $n \in \mathcal{N}$. Define $\epsilon_1 = \gamma^{-1} \min_{n \in \mathcal{N}} \{(u_n - \lambda_n)\}$, then clearly $\epsilon_1 > 0$. Moreover, by Lemma 1, for all $\boldsymbol{q} \in \mathbb{Z}^{\bar{n}}_+$

$$\langle \boldsymbol{\lambda} - \hat{\boldsymbol{\mu}}(\boldsymbol{q}), \boldsymbol{g}(\boldsymbol{q}) \rangle \leq \langle \boldsymbol{\lambda} - \boldsymbol{u}, \boldsymbol{g}(\boldsymbol{q}) \rangle \leq -\epsilon_1 \gamma |\boldsymbol{g}(\boldsymbol{q})|.$$

Substituting in (35)

$$\begin{aligned}&\mathbb{E}^{\hat{\boldsymbol{\mu}}}[G(\boldsymbol{Q}_{k+1}) - G(\boldsymbol{Q}_k) | \boldsymbol{Q}_k = \boldsymbol{q}] \\ &\quad \leq -\epsilon_1 |\boldsymbol{g}(\boldsymbol{q})| + \gamma^{-1} \sum_{n \in \mathcal{N} \setminus I_{\boldsymbol{q}}} \hat{\mu}_n(\boldsymbol{q}) g_n(\boldsymbol{q}) \\ &\qquad + \max \Big\{ \max_n \{\dot{g}_n(\boldsymbol{q} + \alpha_n \boldsymbol{e}_n) : n \in \mathcal{N}\}, \\ &\qquad\quad \max_n \{\dot{g}_n(\boldsymbol{q} - \beta_n \boldsymbol{e}_n) : n \in \mathcal{N}_{\boldsymbol{q}}\} \Big\}. \qquad (36)\end{aligned}$$

By using (19), when $\boldsymbol{q}$ is suitably large, $g_n(\boldsymbol{q})$ for each $n \in \mathcal{N} \setminus \mathcal{N}_{\boldsymbol{q}}$ in the second term of the above summation can be bounded above by $\frac{\epsilon_1}{4} |\boldsymbol{g}(\boldsymbol{q})|$. Similarly, using (20), the third term of the above summation can be bounded above by $\frac{\epsilon_1}{4} |\boldsymbol{g}(\boldsymbol{q})|$. Hence, for $\boldsymbol{q}$ large enough, (36) becomes

$$\mathbb{E}^{\hat{\boldsymbol{\mu}}}[G(\boldsymbol{Q}_{k+1}) - G(\boldsymbol{Q}_k) | \boldsymbol{Q}_k = \boldsymbol{q}] \leq -\frac{\epsilon_1}{2} |\boldsymbol{g}(\boldsymbol{q})|.$$

Since $0 < \epsilon < |\boldsymbol{g}(\boldsymbol{q})|$ for all large $\boldsymbol{q}$, the proof is complete.

Acknowledgment

The authors would like to thank an anonymous reviewer and J. Hasenbein for invaluable suggestions and discussions on fluid models.

References

[1] R. Knopp and P. Humblet, "Information capacity and power control in single-cell multiuser communications," in *Proc. IEEE ICC*, Jun. 1995, vol. 1, pp. 331–335.

[2] P. Viswanath, D. Tse, and R. Laroia, "Opportunistic beamforming using dumb antennas," *IEEE Trans. Inf. Theory*, vol. 48, no. 6, pp. 1277–1294, Jun. 2002.

[3] X. Liu, E. K. P. Chong, and N. B. Shroff, "A framework for opportunistic scheduling in wireless networks," *Comput. Netw.*, vol. 41, no. 4, pp. 451–474, 2003.

[4] M. J. Neely, "Order optimal delay for opportunistic scheduling in multi-user wireless uplinks and downlinks," *IEEE/ACM Trans. Netw.*, vol. 16, no. 5, pp. 1188–1199, 2008.
[5] M. Andrews, "Instability of the proportional fair scheduling algorithm for HDR," *IEEE Trans. Wireless Commun.*, vol. 3, no. 5, pp. 1422–1426, Sep. 2004.
[6] M. Andrews, K. Kumaran, K. Ramanan, A. Stolyar, R. Vijayakumar, and P. Whiting, "Scheduling in a queuing system with asynchronously varying service rates," *Probab. Eng. Inf. Sci.*, vol. 18, no. 2, pp. 191–217, 2004.
[7] S. Shakkottai and A. Stolyar, "Scheduling for multiple flows sharing a time-varying channel: The exponential rule," *Amer. Math. Soc. Transl.*, ser. 2, vol. 207, pp. 185–202, 2002.
[8] L. Tassiulas and A. Ephremides, "Dynamic server allocation to parallel queues with randomly varying connectivity," *IEEE Trans. Inf. Theory*, vol. 39, no. 2, pp. 466–478, Mar. 1993.
[9] E. M. Yeh and A. S. Cohen, "Delay optimal rate allocation in multiaccess fading communications," in *Proc. Allerton Conf. Commun., Control Comput.*, 2004, pp. 140–149.
[10] B. Hajek, "Optimal control of two interacting service stations," *IEEE Trans. Autom. Control*, vol. 29, no. 6, pp. 491–499, Jun. 1984.
[11] V. Subramanian, "Large deviations of max-weight scheduling policies on convex rate regions," in *Proc. IEEE Inf. Theory Workshop*, Feb. 2008, pp. 535–544.
[12] V. Subramanian, T. Javidi, and S. Kittipiyakul, "Many sources large deviations of max-weight scheduling," in *Proc. Allerton Conf.*, 2008, pp. 1495–1502.
[13] D. Bertsimas, I. Paschalidis, and J. Tsitsiklis, "Asymptotic buffer overflow probabilities in multiclass multiplexers: An optimal control approach," *IEEE Trans. Autom. Control*, vol. 43, no. 3, pp. 315–335, Mar. 1998.
[14] V. J. Venkataramanan and X. Lin, "On wireless scheduling algorithms for minimizing the queue-overflow probability," *IEEE/ACM Trans. Netw.*, vol. 18, no. 3, pp. 788–801, Jun. 2010.
[15] A. L. Stolyar, "Large deviations of queues sharing a randomly time-varying server," *Queue. Syst. Theory Appl.*, vol. 59, no. 1, pp. 1–35, 2008.
[16] D. Shah and D. Wischik, "Optimal scheduling algorithms for input-queued switches," in *Proc. IEEE INFOCOM*, Apr. 2006, pp. 1–11.
[17] B. Sadiq and G. de Veciana, "Large deviations sum-queue optimality of a radial sum-rate monotone opportunistic scheduler," *IEEE Trans. Inf. Theory*, vol. 56, no. 7, pp. 3395–3412, Jul. 2010.
[18] P. Bender, P. Black, M. Grob, R. Padovani, N. Sindhushyana, and S. Viterbi, "CDMA/HDR: A bandwidth-efficient high-speed wireless data service for nomadic users," *IEEE Commun. Mag.*, vol. 38, no. 7, pp. 70–77, Jul. 2000.
[19] S. Borst, "User-level performance of channel-aware scheduling algorithms in wireless data networks," *IEEE/ACM Trans. Netw.*, vol. 13, no. 3, pp. 636–647, Jun. 2005.
[20] R. Prakash and V. V. Veeravalli, "Centralized wireless data networks with user arrivals and departures," *IEEE Trans. Inf. Theory*, vol. 53, no. 2, pp. 695–713, Feb. 2007.
[21] T. Bonald, S. C. Borst, and A. Proutiére, "How mobility impacts the flow-level performance of wireless data systems," in *Proc. IEEE INFOCOM*, 2004, vol. 3, pp. 1872–1881.
[22] V. S. Borkar, "Controlled Markov chains and stochastic networks," *SIAM J. Control Optim.*, vol. 21, no. 4, pp. 652–666, 1983.
[23] D. P. Bertsekas, *Dynamic Programming and Optimal Control*. Belmont, MA: Athena Scientific, 1995, vol. 2.
[24] S. Stidham and R. Weber, "A survey of Markov decision models for control of networks of queues," *Queue. Syst.*, vol. 13, no. 1, pp. 291–314, Mar. 1993.
[25] C. Zhou and G. Wunder, "General stability conditions in wireless broadcast channels," in *Proc. Allerton Conf. Commun., Control Comput.*, Sep. 2008, pp. 675–682.
[26] S. Meyn, *Control Techniques for Complex Networks*. Cambridge, U.K.: Cambridge Univ. Press, 2008.
[27] S. Shakkottai and A. Stolyar, "Scheduling algorithms for a mixture of real-time and non-real-time data in HDR," in *Proc. ITC*, Sep. 2001, pp. 793–804.
[28] R. Madan, S. P. Boyd, and S. Lall, "Fast algorithms for resource allocation in cellular networks," *IEEE/ACM Trans. Netw.*, vol. 18, no. 3, pp. 973–984, Jun. 2010.
[29] J. Huang, V. G. Subramanian, R. Berry, and R. Agrawal, "Scheduling and resource allocation in OFDMA wireless systems," in *Orthogonal Frequency Division Multiple Access*. New York: Auerbach, CRC Press, 2010.
[30] *LTE, the UMTS Long Term Evolution: From Theory to Practice*, S. Sesia, I. Toufik, and M. E. Baker, Eds. Hoboken, NJ: Wiley, 2009.
[31] S. Bodas, S. Shakkottai, L. Ying, and R. Srikant, "Low-complexity scheduling algorithms for multi-channel downlink wireless networks," in *Proc. IEEE INFOCOM*, 2010, pp. 1–9.
[32] B. Sadiq, R. Madan, and A. Sampath, "Downlink scheduling for multiclass traffic in LTE," *EURASIP J. Wireless Commun. Netw.*, vol. 2009, pp. 1–18, Aug. 2009.
[33] C. Maglaras, "Discrete-review policies for scheduling stochastic networks: Trajectory tracking and fluid-scale asymptotic optimality," *Ann. Appl. Probab.*, vol. 10, no. 3, pp. 897–929, Aug. 2000.
[34] V. J. Venkataramanan and X. Lin, "On the large-deviations optimality of scheduling policies minimizing the drift of a Lyapunov function," in *Proc. Allerton Conf. Commun., Control Comput.*, Oct. 2009, pp. 919–926.
[35] I. M. Verloop and R. Núñez-Queija, "Asymptotically optimal parallel resource assignment with interference," *Queue. Syst.*, vol. 65, no. 1, pp. 43–92, 2010.

Bilal Sadiq received the B.S. degree in electronics engineering from Ghulam Ishaq Khan Institute, Topi, Pakistan, in 2000, and is currently Ph.D. candidate with the Department of Electrical and Computer Engineering, The University of Texas at Austin.

From 2000 to 2003, he was with Schlumberger Oilfield Services in the UAE and Scotland. From July to December 2008, he was with Qualcomm Flarion Technologies, Bridgewater, NJ. His research interests include network asymptotics, queueing theory, control, algorithms, signal processing, and engineering design.

Mr. Sadiq was the recipient of the Ghulam Ishaq Khan Gold Medal for Best Academic Performance in 2000.

Seung Jun Baek (S'04–M'09) received the B.S. degree from Seoul National University in 1998, and the M.S. and Ph.D. degrees from the University of Texas at Austin in 2002 and 2007, respectively, all in electrical and computer engineering.

From 2007 to 2009, he was a Member of Technical Staff with the Communications Systems Lab, DSPS R&D Center, Texas Instruments, Dallas, TX. Since 2009, he has been an Assistant Professor with the College of Information and Communications, Korea University, Seoul, Korea. His research interests include modeling and analysis of broadband access networks, wireless ad hoc networks, and sensor networks.

Gustavo de Veciana (S'88–M'94–SM'01–F'09) received the B.S., M.S, and Ph.D. degrees in electrical engineering from the University of California, Berkeley, in 1987, 1990, and 1993, respectively.

He is currently a Professor with the Department of Electrical and Computer Engineering, The University of Texas at Austin, and the Director of the Wireless Networking and Communications Group (WNCG). His research focuses on the design, analysis, and control of telecommunication networks. Current interests include measurement, modeling, and performance evaluation; wireless and sensor networks; and architectures and algorithms to design reliable computing and network systems.

Dr. de Veciana is the recipient of the General Motors Foundation Centennial Fellowship in Electrical Engineering and a 1996 National Science Foundation CAREER Award, co-recipient of the IEEE William McCalla Best ICCAD Paper Award for 2000, and co-recipient of the Best Paper in the ACM *Transactions on Design Automation of Electronic Systems*, January 2002–2004. He has been an Editor for the IEEE/ACM TRANSACTIONS ON NETWORKING.