Energy-Aware Wireless Scheduling with Near Optimal Backlog and Convergence Time Tradeoffs

Michael J. Neely
University of Southern California

***Abstract*— This paper considers a wireless link with randomly arriving data that is queued and served over a time-varying channel. It is known that any algorithm that comes within ϵ of the minimum average power required for queue stability must incur average queue size at least $\Omega(\log(1/\epsilon))$. However, the optimal convergence time is unknown. This paper develops a scheduling algorithm that, for any $\epsilon > 0$, achieves the optimal $O(\log(1/\epsilon))$ average queue size tradeoff with a convergence time of $O(\log(1/\epsilon)/\epsilon)$. An example system is presented for which all algorithms require convergence time at least $\Omega(1/\epsilon)$, and so the proposed algorithm is within a logarithmic factor of the optimal convergence time. The method uses the simple drift-plus-penalty technique with an improved convergence time analysis.**

I. Introduction

This paper considers power-aware scheduling in a wireless link with a time-varying channel and randomly arriving data. Arriving data is queued for eventual transmission. The transmission rate out of the queue is determined by the current channel state and the current power allocation decision. Specifically, the controller can make an *opportunistic scheduling* decision by observing the channel before allocating power. For a given $\epsilon > 0$, the goal is to push average power to within ϵ of the minimum possible average power required for queue stability while ensuring optimal queue size and convergence time tradeoffs.

A major difficulty is that the data arrival rate and the channel probabilities are unknown. Hence, the convergence time of an algorithm includes the *learning time* associated with estimating probability distributions or "sufficient statistics" of these distributions. The optimal learning time required to achieve the average power and backlog objectives, as well as the appropriate sufficient statistics to learn, are unknown. This open question is important because it determines how fast an algorithm can adapt to its environment. A contribution of the current paper is the development of an algorithm that, under suitable assumptions, provides an optimal power-backlog tradeoff while provably coming within a logarithmic factor of the optimal convergence time. This is done via the existing drift-plus-penalty algorithm but with an improved convergence time analysis.

Work on opportunistic scheduling was pioneered by Tassiulas and Ephremides in [2], where the Lyapunov method and the *max-weight* algorithms were introduced for queue stability. Related opportunistic scheduling work that focuses on utility optimization is given in [3][4][5][6][7][8][9][10] using dual, primal-dual, and stochastic gradient methods, and in [11] using index policies. The basic drift-plus-penalty algorithm of Lyapunov optimization can be viewed as a *dual method*, and is known to provide, for any $\epsilon > 0$, an ϵ-approximation to minimum average power with a corresponding $O(1/\epsilon)$ tradeoff in average queue size [10][12]. This tradeoff is not optimal. Work by Berry and Gallager in [13] shows that, for queues with *strictly concave* rate-power curves, any algorithm that achieves an ϵ-approximation must incur average backlog of $\Omega(\sqrt{1/\epsilon})$, even if that algorithm knows all system probabilities. Work in [14] shows this tradeoff is achievable (to within a logarithmic factor) using an algorithm that does *not* know the system probabilities. The work [14] further considers the exceptional case when rate-power curves are *piecewise linear*. In that case, an improved tradeoff of $O(\log(1/\epsilon))$ is both achievable and optimal. This is done using an exponential Lyapunov function together with a drift-steering argument. Work in [15][16] shows that similar logarithmic tradeoffs are possible via the basic drift-plus-penalty algorithm with Last-in-First-Out scheduling.

Now consider the question of *convergence time*, being the time required for the average queue size and power guarantees to kick in. This convergence time question is unique to problems of *stochastic scheduling when system probabilities are unknown*. If probabilities were known, the optimal fractions of time for making certain decisions could be computed offline (possibly via a very complex optimization), so that system averages would "kick in" immediately at time 0. Thus, convergence time in the context of this paper should not be confused with *algorithmic complexity* for non-stochastic optimization problems.

Unfortunately, prior work that treats stochastic scheduling with unknown probabilities, including the basic drift-plus-penalty algorithm as well as extensions that achieve square root and logarithmic tradeoffs, give only $O(1/\epsilon^2)$ convergence time guarantees. Recent work in [17] treats convergence time for a related problem of flow rate allocation and concludes that constraint violations decay as $c(\epsilon)/t$, where $c(\epsilon)$ is a constant that depends on ϵ and t is the total time the algorithm has been in operation. While [17] does not specify the size of $c(\epsilon)$, it can be shown that $c(\epsilon) = O(1/\epsilon)$. Intuitively, this is because $c(\epsilon)$ is related to an average queue size, which is $O(1/\epsilon)$. The time t needed to ensure constraint violations are at most ϵ is found by solving $c(\epsilon)/t = \epsilon$. The simple answer is $t = O(1/\epsilon^2)$, again exhibiting $O(1/\epsilon^2)$ convergence time! This leads one to suspect that $O(1/\epsilon^2)$ is *optimal*. However, a recent result in [18] shows that a technique for Lagrange multiplier estimation can, in some cases, reduce queue backlog to $O(1/\epsilon^{2/3})$ and

The author is with the Electrical Engineering department at the University of Southern California, Los Angeles, CA.

This material was presented in part at the IEEE INFOCOM conference, Hong Kong, 2015 [1]. This work was supported in part by the NSF Career grant CCF-0747525.

convergence time to $O(1/\epsilon^{1+2/3})$. [1]

This paper pushes the tradeoff further to achieve a convergence time result that is near optimal. Specifically, under the same piecewise linear assumption in [14], and for the special case of a system with just one queue, it is shown that the existing drift-plus-penalty algorithm yields an ϵ-approximation with both $O(\log(1/\epsilon))$ average queue size and $O(\log(1/\epsilon)/\epsilon)$ convergence time. This is an encouraging result that shows learning times for power-aware scheduling can be pushed much smaller than expected. Further, an example system is demonstrated for which every algorithm incurs convergence time at least $\Omega(1/\epsilon)$, so that the proposed algorithm is within a logarithmic factor of optimality. The techniques in this paper can likely be extended to more general multi-queue situations (see recent results in this direction in [19]).

The next section specifies the problem formulation. Section III shows a lower bound on convergence time of $\Omega(1/\epsilon)$. Section IV develops an algorithm that achieves this bound to within a logarithmic factor.

II. System model

Consider a wireless link with randomly arriving traffic. The system operates in slotted time with slots $t \in \{0, 1, 2, \ldots\}$. Data arrives every slot and is queued for transmission. Define:

$$\begin{aligned} Q(t) &= \text{queue backlog on slot } t \\ a(t) &= \text{new arrivals on slot } t \\ \mu(t) &= \text{service offered on slot } t \end{aligned}$$

The values of $Q(t), a(t), \mu(t)$ are nonnegative and their units depend on the system of interest. For example, they can take integer units of *packets* (assuming packets have fixed size), or real units of *bits*. Assume the queue is initially empty, so that $Q(0) = 0$. The queue dynamics are:

$$Q(t+1) = \max[Q(t) + a(t) - \mu(t), 0] \tag{1}$$

Assume that $\{a(t)\}_{t=0}^{\infty}$ is an independent and identically distributed (i.i.d.) sequence with mean $\lambda = \mathbb{E}[a(t)]$. For simplicity, assume the amount of arrivals in one slot is bounded by a constant a_{max}, so that $0 \leq a(t) \leq a_{max}$ for all slots t.

If the controller decides to transmit data on slot t, it uses one unit of power. Let $p(t) \in \{0, 1\}$ be the power used on slot t. The amount of data that can be transmitted depends on the current channel state. Let $\omega(t)$ be the amount of data that can be transmitted on slot t if power is allocated, so that:

$$\mu(t) = p(t)\omega(t)$$

Assume that $\omega(t)$ is i.i.d. over slots and takes values in a finite set $\Omega = \{\omega_0, \omega_1, \omega_2, \ldots, \omega_M\}$, where $\omega_0 = 0$ and ω_i is a positive real number for all $i \in \{1, \ldots, M\}$. Assume these values are ordered so that:

$$0 = \omega_0 < \omega_1 < \omega_2 < \cdots < \omega_M$$

For each $\omega_k \in \Omega$, define $\pi(\omega_k) = Pr[\omega(t) = \omega_k]$.

[1]The work [18] shows the *transient time* for backlog to come close to a Lagrange multiplier vector is $O(1/\epsilon^{2/3})$. For transients to be amortized, the total time for averages to be within ϵ of optimality is $O(1/\epsilon^{1+2/3})$.

Every slot t the system controller observes $\omega(t)$ and then chooses $p(t) \in \{0, 1\}$. The choice $p(t) = 1$ activates the link for transmission of $\omega(t)$ units of data. Fewer than $\omega(t)$ units are transmitted if $Q(t) < \mu(t)$ (see the queue equation (1)). The largest possible average transmission rate is $\mathbb{E}[\omega(t)]$, which is achieved by using $p(t) = 1$ for all t. It is assumed throughout that $0 \leq \lambda \leq \mathbb{E}[\omega(t)]$.

A. Optimization goal

For a real-valued random process $b(\tau)$ that evolves over slots $\tau \in \{0, 1, 2, \ldots\}$, define its time average expectation over $t > 0$ slots as:

$$\overline{b}(t) \triangleq \frac{1}{t}\sum_{\tau=0}^{t-1} \mathbb{E}[b(\tau)] \tag{2}$$

where "$\triangleq$" represents "defined to be equal to." With this notation, $\overline{\mu}(t)$, $\overline{p}(t)$, $\overline{Q}(t)$ respectively denote the time average expected transmission rate, power, and queue size over the first t slots.

The basic stochastic optimization problem of interest is:

$$\begin{aligned} \text{Minimize:} \quad & \limsup_{t\to\infty} \overline{p}(t) && (3) \\ \text{Subject to:} \quad & \liminf_{t\to\infty} \overline{\mu}(t) \geq \lambda && (4) \\ & p(t) \in \{0, 1\} \quad \forall t \in \{0, 1, 2, \ldots\} && (5) \end{aligned}$$

The assumption $\lambda \leq \mathbb{E}[\omega(t)]$ ensures the above problem is always *feasible*, so that it is possible to satisfy constraints (4)-(5) using $p(t) = 1$ for all t. Define p^* as the infimum average power for the above problem. An algorithm is said to produce an *ϵ-approximation* at time t if, for a given $\epsilon \geq 0$:

$$\begin{aligned} \overline{p}(t) &\leq p^* + \epsilon && (6) \\ \lambda - \overline{\mu}(t) &\leq \epsilon && (7) \end{aligned}$$

Given $\epsilon > 0$, an algorithm is said to have a *convergence time of T_ϵ* if it ensures (6)-(7) hold for all $t \geq T_\epsilon$.

An algorithm is said to produce an *$O(\epsilon)$-approximation* if the ϵ symbols on the right-hand-side of (6)-(7) are replaced by some constant multiples of ϵ. Convergence times to an $O(\epsilon)$ approximation shall also be denoted T_ϵ to emphasize dependence on ϵ. Fix $\epsilon > 0$. This paper shows that a simple drift-plus-penalty algorithm that takes ϵ as an input parameter (and that has no knowledge of the arrival rate or channel probabilities) can ensure there is a time T_ϵ for which:

- The algorithm produces an $O(\epsilon)$-approximation for all $t \geq T_\epsilon$.
- The algorithm ensures the following for all $t \in \{0, 1, 2, \ldots\}$:
$$\overline{Q}(t) \leq O(\log(1/\epsilon)) \tag{8}$$
- $T_\epsilon = O(\log(1/\epsilon)/\epsilon)$.

The average queue size bound (8) is known to be optimal, in the sense that no algorithm can provide a sub-logarithmic guarantee [14]. The next section shows that the convergence time $O(\log(1/\epsilon)/\epsilon)$ is within a logarithmic factor of the optimal convergence time.

B. Discussion of the time average expectation

Can convergence be defined using the sample path time average, rather than its expectation? This does not work because, over a finite horizon, it is impossible to ensure sample paths of transmission rate and power are both close to their ergodic optimums (consider the example when all channels are bad for the first t slots). Hence, this paper uses a time average expectation. This represents expected behavior over the first t slots. Fortunately, the sample path "tends" to stay close to its expectation, as demonstrated via simulation in Section VII and via a deviation probability calculation in Section V-H.

III. A LOWER BOUND ON CONVERGENCE TIME

A. Expectations and randomization

One type of power allocation policy is an *ω-only policy* that, every slot t, observes $\omega(t)$ and independently chooses $p(t) \in \{0, 1\}$ according to some stationary conditional probabilities $Pr[p(t) = 1|\omega(t) = \omega]$ that are specified for all $\omega \in \Omega$. The resulting average power and transmission rate is:

$$\begin{aligned}\mathbb{E}[p(t)] &= \textstyle\sum_{k=0}^{M} \pi(\omega_k) Pr[p(t) = 1|\omega(t) = \omega_k] \\ \mathbb{E}[\mu(t)] &= \textstyle\sum_{k=0}^{M} \pi(\omega_k)\omega_k Pr[p(t) = 1|\omega(t) = \omega_k]\end{aligned}$$

It is known that the problem (3)-(5) is solvable over the class of ω-only policies [10]. Specifically, if the arrival rate λ and the channel probabilities $\pi(\omega_k)$ were known in advance, one could offline compute an ω-only policy to satisfy:

$$\mathbb{E}[p(t)] = p^* \tag{9}$$
$$\mathbb{E}[\mu(t)] = \lambda \tag{10}$$

This is a 0-approximation for all $t \geq 0$. In particular, such a randomized algorithm achieves a convergence time of 0. However, such an algorithm would typically incur infinite average queue size (since the service rate equals the arrival rate). Further, it is not possible to implement this algorithm without perfect knowledge of λ and $\pi(\omega_k)$ for all $\omega_k \in \Omega$.

This section develops a convergence time lower bound in the case when system probabilities are unknown. A different type of lower bound is given in [17]. It shows the sample path time average of an integer sequence that approaches a non-integer real number has error magnitude that decays like $\Omega(1/t)$ (for example, the error might be $1/t$ on odd slots and $-1/t$ on even slots). This is a non-probabilistic result that holds regardless of whether or not probabilities are known. Of course, if probabilities are known, one can design a randomized algorithm to have optimal expectations every slot, so the *expected* error is zero. This section shows that, if probabilities are *unknown*, even the *expected* error is necessarily $\Omega(1/t)$. The proof is necessarily different from [17] and has nothing to do with averages of integer sequences.

B. Intuition

Suppose one temporarily allows for infinite average queue size. Consider the following thought experiment. Consider an algorithm that does not know the system probabilities and hence makes a *single mistake* at time 0, so that:

$$\mathbb{E}[p(0)] = p^* + c$$

where $c > 0$ is some constant gap away from the optimal average power p^*. However, suppose a genie gives the controller perfect knowledge of the system probabilities at time 1, and then for slots $t \geq 1$ the network makes decisions to achieve the ideal averages (9)-(10). The resulting time average expected power over the first $t > 1$ slots is:

$$\overline{p}(t) = \frac{p^* + c}{t} + \frac{(t-1)p^*}{t} = p^* + \frac{c}{t}$$

Thus, to reach an ϵ-approximation, this genie-aided algorithm requires a convergence time $t = c/\epsilon = \Theta(1/\epsilon)$.

C. An example with $\Omega(1/\epsilon)$ convergence time

The above thought experiment does not prove an $\Omega(1/\epsilon)$ bound on convergence time because it assumes the algorithm makes decisions according to (9)-(10) for all slots $t \geq 1$, which may not be the optimal way to compensate for the mistake on slot 0. This section defines a simple system for which convergence time is at least $\Omega(1/\epsilon)$ under any algorithm.

Consider a system with deterministic arrivals of 1 packet every slot (so $\lambda = 1$). There are three possible channel states $\omega(t) \in \{1, 2, 3\}$, with probabilities:

$$\pi(3) = y, \pi(2) = z, \pi(1) = 1 - y - z$$

For each slot $t > 0$, define the *system history* $\mathcal{H}(t) = \{(a(0), \omega(0), p(0)), \ldots, (a(t-1), \omega(t-1), p(t-1))\}$. Define $\mathcal{H}(0) = 0$. For each slot t, a general algorithm has conditional probabilities $\theta_i(t)$ defined for $i \in \{1, 2, 3\}$ by:

$$\theta_i(t) = Pr[p(t) = 1|\omega(t) = i, a(t), \mathcal{H}(t)]$$

On a single slot, it is not difficult to show that the minimum average power $\mathbb{E}[p(t)]$ required to achieve a given average service rate $\mu = \mathbb{E}[\mu(t)]$ is characterized by the following function $h(\mu)$:

$$h(\mu) \triangleq \begin{cases} \mu/3 & \text{if } 0 \leq \mu \leq 3y \\ y + (\mu - 3y)/2 & \text{if } 3y \leq \mu \leq 3y + 2z \\ \mu - 2y - z & \text{if } 3y + 2z \leq \mu \leq 2y + z + 1 \end{cases}$$

There are two significant vertex points $(\mu, h(\mu))$ for this function. The first is $(3y, y)$, achieved by allocating power if and only if $\omega(t) = 3$. The second is $(3y + 2z, y + z)$, achieved by allocating power if and only if $\omega(t) \in \{2, 3\}$.

Define $\mathcal{R}$ as the set of points (μ, p) that lie on or above the curve $h(\mu)$:

$$\mathcal{R} = \{(\mu, p) \in \mathbb{R}^2 | 0 \leq \mu \leq 2y + z + 1, h(\mu) \leq p \leq 1\}$$

The set $\mathcal{R}$ is convex. Under any algorithm one has:

$$(\mathbb{E}[\mu(\tau)], \mathbb{E}[p(\tau)]) \in \mathcal{R} \quad \forall \tau \in \{0, 1, 2, \ldots\}$$

For a given $t > 1$, the following two vectors must be in $\mathcal{R}$:

$$\begin{aligned}(\mu_0, p_0) &\triangleq (\mathbb{E}[\mu(0)], \mathbb{E}[p(0)]) \\ (\mu_1, p_1) &\triangleq \frac{1}{t-1}\sum_{\tau=1}^{t-1}(\mathbb{E}[\mu(\tau)], \mathbb{E}[p(\tau)])\end{aligned}$$

That (μ_1, p_1) is in $\mathcal{R}$ follows because it is the average of points in $\mathcal{R}$, and $\mathcal{R}$ is convex. By definition of $(\overline{\mu}(t), \overline{p}(t))$:

$$(\overline{\mu}(t), \overline{p}(t)) = \frac{1}{t}(\mu_0, p_0) + \frac{t-1}{t}(\mu_1, p_1) \tag{11}$$

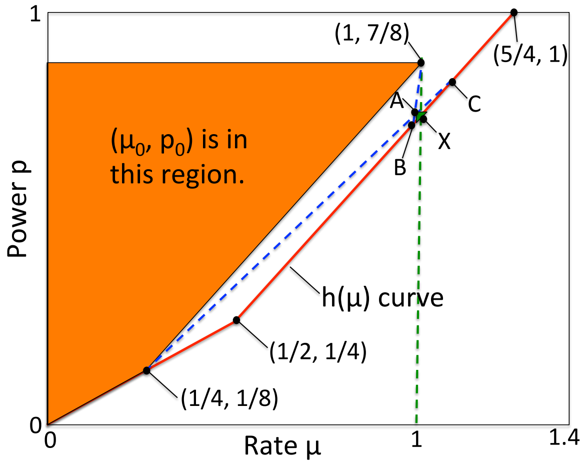

Fig. 1. The performance region for case 1. The line segments between $(1/4, 1/8)$ and C and between $(1, 7/8)$ and B intersect at point A.

Fix a small value $\epsilon > 0$. The algorithm must ensure $(\overline{\mu}(t), \overline{p}(t))$ is an ϵ-approximation to the target point $(1, h(1))$, so that:

$$\overline{\mu}(t) \geq 1 - \epsilon \ , \ \overline{p}(t) \leq h(1) + \epsilon$$

The algorithm has no knowledge of the probabilities y and z at time 0, so $\theta_1(0)$, $\theta_2(0)$, $\theta_3(0)$ are arbitrary. Suppose a genie reveals y and z on slot 1, and the network makes decisions on slots $\{1, \ldots, t-1\}$ that result in a (μ_1, p_1) vector that *optimally compensates for any mistake on slot* 0. Thus, (μ_1, p_1) is assumed to be the vector in $\mathcal{R}$ that ensures (11) produces an ϵ-approximation in the smallest time t.

The following proof considers two cases: The first case assumes $\theta_2(0) \leq 1/2$, but considers probabilities y and z for which minimizing average power requires *always transmitting when* $\omega(t) = 2$. The second case assumes $\theta_2(0) > 1/2$, but then considers probabilities y and z for which minimizing average power requires *never transmitting when* $\omega(t) = 2$. In both cases, the nonlinear structure of the $h(\mu)$ curve prevents a fast recovery from the initial mistake.

- Case 1: Suppose $\theta_2(0) \leq 1/2$. Consider $y = 0, z = 1/4$. Then $\omega(t) \in \{1, 2\}$ for all t, $\pi(1) = 3/4, \pi(2) = 1/4$, and $\omega(t) = 2$ is the most efficient state. The $h(\mu)$ curve is shown in Fig. 1. The minimum average power to support $\lambda = 1$ is $h(1) = 3/4$, and so the target point is $X = (1, 3/4)$. The point $(\mu_0, p_0) = (\mathbb{E}[\mu(0)], \mathbb{E}[p(0)])$ is:

$$(\mu_0, p_0) = \left(\frac{\theta_2(0)}{2} + \frac{3\theta_1(0)}{4}, \frac{\theta_2(0)}{4} + \frac{3\theta_1(0)}{4}\right)$$

  The set of possible (μ_0, p_0) is formed by considering all $\theta_2(0) \in [0, 1/2]$, $\theta_1(0) \in [0, 1]$. This set lies inside the left (orange) shaded region of Fig. 1. To see this, note that if $\theta_2(0)$ is fixed at a certain value, the resulting (μ_0, p_0) point lies on a line segment of slope 1 that is formed by sweeping $\theta_1(0)$ through the interval $[0, 1]$. If $\theta_2(0) = 1/2$, that line segment is between points $(1/4, 1/8)$ and $(1, 7/8)$ in Fig. 1. If $\theta_2(0) < 1/2$ then the line segment is shifted to the left.
  The small triangular (green) shaded region in Fig. 1, with one vertex at point A, is the *target region*. The vector $(\overline{\mu}(t), \overline{p}(t))$ must be in this region to be an ϵ-approximation. The point A is defined:

$$A = X + (-\epsilon, \epsilon) = (1 - \epsilon, 3/4 + \epsilon)$$

  It suffices to search for an optimal compensation vector (μ_1, p_1) on the curve $(\mu, h(\mu))$. This is because the average power p_1 from a point (μ_1, p_1) *above* the curve $(\mu, h(\mu))$ can be reduced, without affecting μ_1, by choosing a point *on* the curve. By geometry, (μ_1, p_1) must lie on the line segment between points B and C in Fig. 1. Specifically, the point C is defined as the point where the line between $(1/4, 1/8)$ and point A pierces the curve $(\mu, h(\mu))$, and the point B is defined similarly. To understand this, note that if (μ_1, p_1) were on the $(\mu, h(\mu))$ curve but *not* in between points B and C, it would be impossible for a convex combination of (μ_1, p_1) and (μ_0, p_0) to be in the target region (which is required by (11)). For example, suppose (μ_1, p_1) is on the line segment between points C and $(5/4, 1)$ (not including point C itself). Define L as the infinite line between (μ_1, p_1) and the point $(1/4, 1/8)$. The green target triangle is strictly below line L, while the orange region containing (μ_0, p_0) lies on or above line L, as does the line segment between (μ_0, p_0) and (μ_1, p_1).
  By basic algebra for intersecting lines we have:

$$\begin{aligned} B &= X - \left(\frac{\epsilon}{1 - 16\epsilon}, \frac{\epsilon}{1 - 16\epsilon}\right) \\ C &= X + \left(\frac{11\epsilon}{1 - 16\epsilon}, \frac{11\epsilon}{1 - 16\epsilon}\right) \end{aligned}$$

  Observe that:

$$||(\overline{\mu}(t), \overline{p}(t)) - X|| \leq \epsilon\sqrt{2} \quad (12)$$

$$||(\mu_1, p_1) - X|| \leq O(\epsilon) \quad (13)$$

$$||(\mu_0, p_0) - (\mu_1, p_1)|| \geq \sqrt{2}/16 \quad (14)$$

  where (12) follows by considering the maximum distance between X and any point in the target region, (13) holds because any vector on the line segment between B and C is $O(\epsilon)$ distance away from X, and (14) holds because the distance between any point on the line segment between B and C and a point in the left (orange) shaded region is at least $\sqrt{2}/16$ (being the distance between the two parallel lines of slope 1). Starting from (12) one has:

$$\begin{aligned} \epsilon\sqrt{2} &\geq ||(\overline{\mu}(t), \overline{p}(t)) - X|| \\ &= ||(1/t)(\mu_0, p_0) + (1 - 1/t)(\mu_1, p_1) - X|| \\ &= ||(1/t)[(\mu_0, p_0) - (\mu_1, p_1)] - [X - (\mu_1, p_1)]|| \\ &\geq (1/t)||(\mu_0, p_0) - (\mu_1, p_1)|| - ||X - (\mu_1, p_1)|| \\ &\geq \sqrt{2}/(16t) - O(\epsilon) \end{aligned}$$

  where the first equality holds by (11), the second-to-last inequality uses the triangle inequality $||W - Z|| \geq ||W|| - ||Z||$ for any vectors W, Z, and the final inequality uses (13) and (14). So $\sqrt{2}/(16t) \leq O(\epsilon)$. It follows that $t \geq \Omega(1/\epsilon)$.
- Case 2: Suppose $\theta_2(0) > 1/2$. However, suppose $y = z = 1/2$. So $\omega(t) \in \{2, 3\}$, $\pi(2) = \pi(3) = 1/2$, and

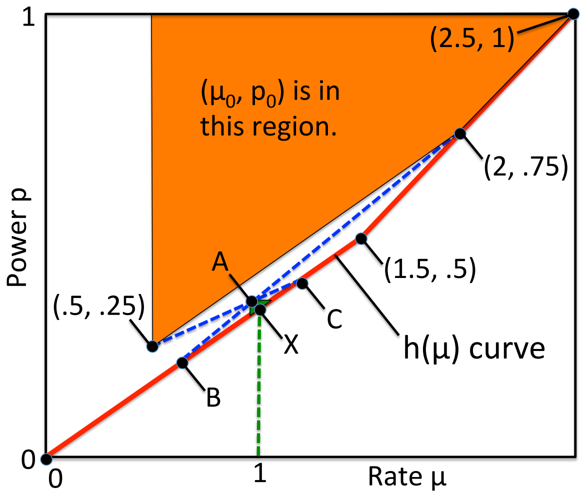

Fig. 2. The performance region for case 2.

$\omega(t) = 2$ is the *least efficient* state. The $h(\mu)$ curve is shown in Fig. 2. Note that $h(1) = 1/3$, and so the target point is $X = (1, 1/3)$. The point $A = (1 - \epsilon, 1/3 + \epsilon)$ is shown in Fig. 2. Point A is one vertex of the small triangular (green) target region that defines all points $(\overline{\mu}(t), \overline{p}(t))$ that are ϵ-approximations.
Because $\theta_2(0) \geq 1/2$, the point (μ_0, p_0) lies somewhere in the (orange) shaded region in Fig. 2. Indeed, if $\theta_2(0) = 1/2$, then (μ_0, p_0) is on the line segment between points $(0.5, 0.25)$ and $(2, 0.75)$. It is above this line segment if $\theta_2(0) > 1/2$. As before, the geometry of the problem ensures an optimal compensation vector (μ_1, p_1) lies somewhere on the line segment of the $h(\mu)$ curve between points B and C of Fig. 2. As before, it holds that:

$$\begin{aligned} B &= X - (O(\epsilon), O(\epsilon)) \\ C &= X + (O(\epsilon), O(\epsilon)) \end{aligned}$$

and:

$$\begin{aligned} ||(\overline{\mu}(t), \overline{p}(t)) - X|| &\leq O(\epsilon) \\ ||(\mu_1, p_1) - X|| &\leq O(\epsilon) \\ ||(\mu_0, p_0) - (\mu_1, p_1)|| &\geq \Theta(1) \end{aligned}$$

As before, it follows that $t \geq \Omega(1/\epsilon)$.

D. Discussion

The above $\Omega(1/\epsilon)$ bound can be interpreted as a Cramer-Rao type result for controlled queues. Classical Cramer-Rao theory lower-bounds the error of any algorithm for estimating a mean in a system with unknown probabilities. Similarly, the above result lower-bounds the convergence time for controlling a queue with unknown probabilities. There are many ways of controlling a queueing system. Some ways might rely on estimation of certain quantities of interest. This approach is taken in the recent work [18], where a Lagrange multiplier estimate is used to improve convergence time from $O(1/\epsilon^2)$ to $O(1/\epsilon^{1+2/3})$. The result in [18] holds for more general multi-queue systems. The next section shows that, in the special case of one link, convergence time can be improved further to within a logarithmic factor of the $\Omega(1/\epsilon)$ bound.

IV. The Dynamic Algorithm

This section shows that a simple drift-plus-penalty algorithm achieves $O(\log(1/\epsilon)/\epsilon)$ convergence time and $O(\log(1/\epsilon))$ average queue size.

A. Problem structure

Without loss of generality, assume that $\pi(\omega_k) > 0$ for all $k \in \{1, \ldots, M\}$ (else, remove ω_k from the set Ω). The value of $\pi(\omega_0)$ is possibly zero. For each real number μ in the interval $[0, \mathbb{E}[\omega(t)]]$, define $h(\mu)$ as the minimum average power required to achieve an average transmission rate of μ. It is known that $p^* = h(\lambda)$. Further, it is not difficult to show that $h(\mu)$ is non-decreasing, convex, and *piecewise linear* with $h(0) = 0$ and $h(\mathbb{E}[\omega(t)]) = 1 - \pi(\omega_0)$. The point $(0, 0)$ is a vertex point of the piecewise linear curve $h(\mu)$. There are M other vertex points, achieved by the ω-only policies of the form:

$$p(t) = \begin{cases} 1 & \text{if } \omega(t) \geq \omega_k \\ 0 & \text{otherwise} \end{cases} \tag{15}$$

for $k \in \{1, \ldots, M\}$. This means that a vertex point is achieved by only using channel states $\omega(t)$ that are on or above a certain threshold ω_k. Lowering the threshold value by selecting a smaller ω_k allows for a larger $\mathbb{E}[\mu(t)]$ at the expense of sometimes using less efficient channel states. The proof that this class of policies achieves the vertex points follows by a simple interchange argument that is omitted for brevity.

For ease of notation, define $\omega_{M+1} \triangleq \infty$ and $\mu_{M+1} \triangleq 0$. Let $\{\mu_1, \mu_2, \ldots, \mu_M, \mu_{M+1}\}$ be the set of transmission rates at which there are vertex points. Specifically, for $k \in \{1, \ldots, M\}$, μ_k corresponds to the threshold ω_k in the policy (15). That is:

$$\mu_k \triangleq \sum_{i=k}^{M} \omega_i \pi(\omega_i) \tag{16}$$

Note that:

$$0 = \mu_{M+1} < \mu_M < \mu_{M-1} < \cdots < \mu_1 = \mathbb{E}[\omega(t)]$$

It follows that $h(\mu_k)$ is the corresponding average power for vertex k, so that $(\mu_k, h(\mu_k))$ is a vertex point of the curve $h(\mu)$:

$$h(\mu_k) = \sum_{i=k}^{M} \pi(\omega_i) \tag{17}$$

The numbers $\{\mu_1, \mu_2, \ldots, \mu_M, \mu_{M+1}\}$ represent a set of measure 0 in the interval $[0, \mathbb{E}[\omega(t)]]$. It is assumed that the arrival rate λ is a number in $[0, \mathbb{E}[\omega(t)]]$ that lies strictly between two points μ_{b+1} and μ_b for some index $b \in \{1, \ldots, M\}$. That is:

$$\mu_{b+1} < \lambda < \mu_b$$

Thus, the point $(\lambda, h(\lambda))$ can be achieved by timesharing between the vertex points $(\mu_{b+1}, h(\mu_{b+1}))$ and $(\mu_b, h(\mu_b))$:

$$\begin{aligned} \lambda &= \theta\mu_{b+1} + (1-\theta)\mu_b & (18) \\ p^* = h(\lambda) &= \theta h(\mu_{b+1}) + (1-\theta)h(\mu_b) & (19) \end{aligned}$$

for some probability θ that satisfies $0 < \theta < 1$. In particular:

$$\theta = \frac{\mu_b - \lambda}{\mu_b - \mu_{b+1}}$$

B. The drift-plus-penalty algorithm

For each slot $t \in \{0, 1, 2, \ldots\}$, define $L(t) = \frac{1}{2}Q(t)^2$ and $\Delta(t) = L(t+1) - L(t)$. Let V be a nonnegative real number. The *drift-plus-penalty* algorithm from [10][12] makes a power allocation decision that, every slot t, minimizes a bound on $\Delta(t) + Vp(t)$. The value V can be chosen as desired and affects a performance tradeoff. This technique is known to yield average queue size of $O(V)$ with deviation from optimal average power no more than $O(1/V)$ [10][12]. This holds for general multi-queue networks. By defining $\epsilon = 1/V$, this produces an $O(\epsilon)$ approximation with average queue size $O(1/\epsilon)$. Further, it can be shown that convergence time is $O(1/\epsilon^2)$ (see Appendix D in [20]).

In the context of the simple one-queue system of the current paper, the drift-plus-penalty algorithm reduces to the following: Every slot t, observe $Q(t)$ and $\omega(t)$ and choose $p(t) \in \{0, 1\}$ to minimize:

$$Vp(t) - Q(t)\omega(t)p(t)$$

That is, choose $p(t)$ according to the following rule:

$$p(t) = \begin{cases} 1 & \text{if } Q(t)\omega(t) \geq V \\ 0 & \text{otherwise} \end{cases} \quad (20)$$

The current paper shows that, for this special case of a system with only one queue, the above algorithm leads to an improved queue size and convergence time tradeoff.

C. The induced Markov chain

The drift-plus-penalty algorithm induces a Markov structure on the system. The system state is $Q(t)$ and the state space is the set of nonnegative real numbers. Observe from (20) that the drift-plus-penalty algorithm has the following behavior:

- $Q(t) \in [V/\omega_{b+1}, V/\omega_b) \implies p(t) = 1$ if and only if $\omega(t) \geq \omega_{b+1}$. In this case one has (from (16) and (17)):

$$\mathbb{E}[\mu(t)|Q(t)] = \mu_{b+1}, \quad \text{if } Q(t) \in [V/\omega_{b+1}, V/\omega_b)$$
$$\mathbb{E}[p(t)|Q(t)] = h(\mu_{b+1}), \quad \text{if } Q(t) \in [V/\omega_{b+1}, V/\omega_b)$$

- $Q(t) \in [V/\omega_b, V/\omega_{b-1}) \implies p(t) = 1$ if and only if $\omega(t) \geq \omega_b$. In this case one has:

$$\mathbb{E}[\mu(t)|Q(t)] = \mu_b, \quad \text{if } Q(t) \in [V/\omega_b, V/\omega_{b-1})$$
$$\mathbb{E}[p(t)|Q(t)] = h(\mu_b), \quad \text{if } Q(t) \in [V/\omega_b, V/\omega_{b-1})$$

where $V/0$ is defined as ∞ (in the case $\omega_{b-1} = \omega_0 = 0$), and $\omega_{M+1} = \infty$ so that $V/\omega_{M+1} = 0$.

Now define intervals $\mathcal{I}^{(1)}, \mathcal{I}^{(2)}, \mathcal{I}^{(3)}, \mathcal{I}^{(4)}$ (see Fig. 3):

$$\begin{aligned} \mathcal{I}^{(1)} &\triangleq [0, V/\omega_{b+1}) \\ \mathcal{I}^{(2)} &\triangleq [V/\omega_{b+1}, V/\omega_b) \\ \mathcal{I}^{(3)} &\triangleq [V/\omega_b, V/\omega_{b-1}) \\ \mathcal{I}^{(4)} &\triangleq [V/\omega_{b-1}, \infty) \end{aligned}$$

If $V/\omega_{b+1} = 0$ then $\mathcal{I}^{(1)}$ is defined as the empty set, and if $V/\omega_{b-1} = \infty$ then $\mathcal{I}^{(4)}$ is defined as the empty set. The above

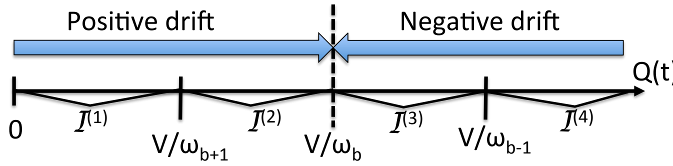

Fig. 3. An illustration of the four intervals $\mathcal{I}_i$ for $i \in \{1, 2, 3, 4\}$.

equalities can be rewritten as:

$$\mathbb{E}[\mu(t)|Q(t)] = \mu_{b+1}, \quad \text{if } Q(t) \in \mathcal{I}^{(2)} \quad (21)$$
$$\mathbb{E}[p(t)|Q(t)] = h(\mu_{b+1}), \quad \text{if } Q(t) \in \mathcal{I}^{(2)} \quad (22)$$
$$\mathbb{E}[\mu(t)|Q(t)] = \mu_b, \quad \text{if } Q(t) \in \mathcal{I}^{(3)} \quad (23)$$
$$\mathbb{E}[p(t)|Q(t)] = h(\mu_b), \quad \text{if } Q(t) \in \mathcal{I}^{(3)} \quad (24)$$

Recall that under the drift-plus-penalty algorithm (20), if $Q(t) \in \mathcal{I}^{(2)}$ then the set of all $\omega(t)$ that lead to a transmission is equal to $\{\omega \in \Omega | \omega \geq \omega_{b+1}\}$. If $Q(t) \in \mathcal{I}^{(1)}$, then the set of all $\omega(t)$ that lead to a transmission depends on the particular value of $Q(t)$. However, since interval $\mathcal{I}^{(1)}$ is to the left of interval $\mathcal{I}^{(2)}$, the set of all $\omega(t)$ that lead to a transmission when $Q(t) \in \mathcal{I}^{(1)}$ is always a subset of $\{\omega \in \Omega | \omega \geq \omega_{b+1}\}$. Similarly, since $\mathcal{I}^{(4)}$ is to the right of $\mathcal{I}^{(3)}$, the set of all $\omega(t)$ that lead to a transmission when $Q(t) \in \mathcal{I}^{(4)}$ is a superset of the set of all $\omega(t)$ that lead to a transmission when $Q(t) \in \mathcal{I}^{(3)}$. Therefore, under the drift-plus-penalty algorithm one has:

$$\mathbb{E}[\mu(t)|Q(t)] \leq \mu_{b+1}, \quad \text{if } Q(t) \in \mathcal{I}^{(1)} \quad (25)$$
$$\mathbb{E}[p(t)|Q(t)] \leq h(\mu_{b+1}), \quad \text{if } Q(t) \in \mathcal{I}^{(1)} \quad (26)$$
$$\mathbb{E}[\mu(t)|Q(t)] \geq \mu_b, \quad \text{if } Q(t) \in \mathcal{I}^{(4)} \quad (27)$$
$$\mathbb{E}[p(t)|Q(t)] \geq h(\mu_b), \quad \text{if } Q(t) \in \mathcal{I}^{(4)} \quad (28)$$

For each $i \in \{1, 2, 3, 4\}$ define the indicator function:

$$1\{Q(t) \in \mathcal{I}^{(i)}\} = \begin{cases} 1 & \text{if } Q(t) \in \mathcal{I}^{(i)} \\ 0 & \text{otherwise} \end{cases}$$

For each slot $t > 0$ and each $i \in \{1, 2, 3, 4\}$, define $\overline{1}^{(i)}(t)$ as the expected fraction of time that $Q(t) \in \mathcal{I}^{(i)}$:

$$\overline{1}^{(i)}(t) \triangleq \frac{1}{t}\sum_{\tau=0}^{t-1} \mathbb{E}\left[1\{Q(t) \in \mathcal{I}^{(i)}\}\right]$$

It follows that (using (22), (24), (26)):

$$\begin{aligned} \overline{p}(t) \quad &\leq \quad \overline{1}^{(2)}(t)h(\mu_{b+1}) + \overline{1}^{(3)}(t)h(\mu_b) \\ &\quad + \overline{1}^{(1)}(t)h(\mu_{b+1}) + \overline{1}^{(4)}(t) \end{aligned} \quad (29)$$

where the final term follows because $p(t) \leq 1$ for all slots t. Similarly (using (21), (23), (25)):

$$\begin{aligned} \overline{\mu}(t) \quad &\leq \quad \overline{1}^{(2)}(t)\mu_{b+1} + \overline{1}^{(3)}(t)\mu_b \\ &\quad + \overline{1}^{(1)}(t)\mu_{b+1} + \overline{1}^{(4)}(t)\mathbb{E}[\omega(t)] \end{aligned} \quad (30)$$

where the final term follows because $\mathbb{E}\left[\mu(t)|Q(t) \in \mathcal{I}^{(4)}\right] \leq \mathbb{E}[\omega(t)]$. Likewise (using (21), (23), (27)):

$$\overline{\mu}(t) \quad \geq \quad \overline{1}^{(2)}(t)\mu_{b+1} + \overline{1}^{(3)}(t)\mu_b + \overline{1}^{(4)}(t)\mu_b \quad (31)$$

which holds because $\mathbb{E}\left[\mu(t)|Q(t) \in \mathcal{I}^{(1)}\right] \geq 0$.

In the next section it is shown that:

- $\overline{\mu}(t)$ is close to λ when t is sufficiently large.
- $\overline{1}^{(1)}(t)$ and $\overline{1}^{(4)}(t)$ are close to 0 when t and V are sufficiently large.
- $\overline{1}^{(2)}(t)$ and $\overline{1}^{(3)}(t)$ are close to θ and $1-\theta$, respectively, when t and V are sufficiently large.
- $\overline{p}(t)$ is close to p^* when t and V are sufficiently large.

Furthermore, to address the issue of convergence time, the notion of "sufficiently large" must be made precise. A key step is establishing bounds on the average queue size.

V. Analysis

A. The distance between $\overline{\mu}(t)$ and λ

Recall that ω_M is the largest possible value of $\omega(t)$. Assume that $V \geq \omega_M^2$.

Lemma 1: If $V \geq \omega_M^2$, then under the drift-plus-penalty algorithm:

a) One has $p(t) = \mu(t) = 0$ whenever $Q(t) < \omega_M$.

b) The queueing equation (1) can be replaced by the following for all slots $t \in \{0, 1, 2, \ldots\}$:

$$Q(t+1) = Q(t) + a(t) - \mu(t)$$

Proof: Suppose $V \geq \omega_M^2$. To prove (a), suppose that $Q(t) < \omega_M$. Since $\omega(t) \leq \omega_M$ for all t, one has:

$$\begin{aligned} Q(t)\omega(t) &\leq Q(t)\omega_M \\ &< \omega_M^2 \\ &\leq V \end{aligned}$$

and so the algorithm (20) chooses $p(t) = 0$, so that $\mu(t)$ is also 0. This proves part (a).

To prove (b), note that part (a) implies $Q(t) \geq \mu(t)$ for all slots t. Indeed, this holds in the case $Q(t) < \omega_M$ (since part (a) ensures $\mu(t) = 0$ in this case), and also holds in the case $Q(t) \geq \omega_M$ (since $\omega_M \geq \mu(t)$ always). Thus:

$$\begin{aligned} Q(t+1) &= \max[Q(t) + a(t) - \mu(t), 0] \\ &= Q(t) + a(t) - \mu(t) \end{aligned}$$

□

Lemma 2: If $V \geq \omega_M^2$ and $Q(0) = q_0$ with probability 1 (for some constant $q_0 \geq 0$), then for every slot $t > 0$:

$$\overline{\mu}(t) = \lambda - \mathbb{E}\left[Q(t) - q_0\right]/t$$

Proof: By Lemma 1 one has for all slots $\tau \in \{0, 1, 2, \ldots\}$:

$$Q(\tau + 1) - Q(\tau) = a(\tau) - \mu(\tau)$$

Summing the above over $\tau \in \{0, 1, 2, \ldots, t-1\}$ and dividing by t gives:

$$\frac{Q(t) - q_0}{t} = \frac{1}{t}\sum_{\tau=0}^{t-1} a(\tau) - \frac{1}{t}\sum_{\tau=0}^{t-1}\mu(\tau) \tag{32}$$

Taking expectations proves the result. □

The above lemma implies that if $V \geq \omega_M^2$, then $\overline{\mu}(t)$ converges to λ whenever $\mathbb{E}\left[Q(t) - q_0\right]/t$ converges to 0.

B. The distance between $\overline{1}^{(2)}(t)$ and θ

The following lemma shows that if $\overline{1}^{(1)}(t)$, $\overline{1}^{(4)}(t)$, and $\mathbb{E}\left[Q(t) - q_0\right]/t$ are close to 0, then $\overline{1}^{(2)}(t)$ is close to θ.

Lemma 3: If $V \geq \omega_M^2$ and $Q(0) = q_0$ with probability 1 (for some constant $q_0 \geq 0$), then for all slots $t > 0$:

$$\begin{aligned} \theta - \frac{[\mu_b \overline{1}^{(1)}(t) - \psi(t)]}{\mu_b - \mu_{b+1}} &\leq \overline{1}^{(2)}(t) \\ &\leq \theta + \frac{\overline{1}^{(4)}(t)\mathbb{E}\left[\omega(t)\right] + \psi(t)}{\mu_b - \mu_{b+1}} \end{aligned}$$

where $\psi(t)$ is defined:

$$\psi(t) \triangleq \mathbb{E}\left[Q(t) - q_0\right]/t$$

Proof: Fix $t > 0$. Lemma 2 implies:

$$\begin{aligned} \lambda &= \overline{\mu}(t) + \psi(t) \\ &\geq \overline{1}^{(2)}(t)\mu_{b+1} + \overline{1}^{(3)}(t)\mu_b + \overline{1}^{(4)}(t)\mu_b + \psi(t) \\ &= \overline{1}^{(2)}(t)\mu_{b+1} + (1 - \overline{1}^{(2)}(t))\mu_b - \overline{1}^{(1)}(t)\mu_b + \psi(t) \end{aligned}$$

where the first inequality holds by (31). Substituting the identity for λ given in (18) into the above inequality gives:

$$\begin{aligned} &\theta\mu_{b+1} + (1-\theta)\mu_b \\ &\geq \overline{1}^{(2)}(t)\mu_{b+1} + (1 - \overline{1}^{(2)}(t))\mu_b - \overline{1}^{(1)}(t)\mu_b + \psi(t) \end{aligned}$$

Rearranging terms proves that:

$$\theta - \frac{[\mu_b \overline{1}^{(1)}(t) - \psi(t)]}{\mu_b - \mu_{b+1}} \leq \overline{1}^{(2)}(t)$$

To prove the second inequality, note that:

$$\begin{aligned} \lambda &= \overline{\mu}(t) + \psi(t) && (33) \\ &\leq \overline{1}^{(2)}(t)\mu_{b+1} + \overline{1}^{(3)}(t)\mu_b + \overline{1}^{(1)}(t)\mu_{b+1} \\ &\quad + \overline{1}^{(4)}(t)\mathbb{E}\left[\omega(t)\right] + \psi(t) && (34) \\ &= \overline{1}^{(2)}(t)\mu_{b+1} + (1 - \overline{1}^{(2)}(t))\mu_b \\ &\quad - \overline{1}^{(1)}(t)\mu_b - \overline{1}^{(4)}(t)\mu_b + \overline{1}^{(1)}(t)\mu_{b+1} \\ &\quad + \overline{1}^{(4)}(t)\mathbb{E}\left[\omega(t)\right] + \psi(t) \\ &\leq \overline{1}^{(2)}(t)\mu_{b+1} + (1 - \overline{1}^{(2)}(t))\mu_b \\ &\quad + \overline{1}^{(4)}(t)\mathbb{E}\left[\omega(t)\right] + \psi(t) && (35) \end{aligned}$$

where (33) holds by Lemma 2, (34) holds by (30), and (35) holds because $\mu_{b+1} < \mu_b$. Substituting the identity for λ given in (18) gives:

$$\begin{aligned} \theta\mu_{b+1} + (1-\theta)\mu_b &\leq \overline{1}^{(2)}(t)\mu_{b+1} + (1 - \overline{1}^{(2)}(t))\mu_b \\ &\quad + \overline{1}^{(4)}(t)\mathbb{E}\left[\omega(t)\right] + \psi(t) \end{aligned}$$

Rearranging terms proves the result. □

C. Positive and negative drift

Define $\mathbb{E}\left[Q(t+1) - Q(t) | Q(t)\right]$ as the *conditional drift*. Assume that $V \geq \omega_M^2$, so that Lemma 1 implies $Q(t+1) - Q(t) = a(t) - \mu(t)$ for all slots t. Thus:

$$\begin{aligned} \mathbb{E}\left[Q(t+1) - Q(t) | Q(t)\right] &= \mathbb{E}\left[a(t) - \mu(t) | Q(t)\right] \\ &= \lambda - \mathbb{E}\left[\mu(t) | Q(t)\right] \end{aligned}$$

where the final equality follows because $a(t)$ is independent of $Q(t)$. From (21) and (25) one has for all slots t:

$$\mathbb{E}\left[\mu(t)|Q(t)\right] \leq \mu_{b+1} \text{ , if } Q(t) < V/\omega_b$$

Likewise, from (23) and (27) one has:

$$\mathbb{E}\left[\mu(t)|Q(t)\right] \geq \mu_b \text{ , } \quad \text{if } Q(t) \geq V/\omega_b$$

Define positive constants β_L and β_R (associated with drift when $Q(t)$ is to the *Left* and *Right* of the threshold V/ω_b) by:

$$\beta_L \triangleq \lambda - \mu_{b+1} \quad , \quad \beta_R \triangleq \mu_b - \lambda$$

It follows that:

$$\mathbb{E}\left[Q(t+1) - Q(t)|Q(t)\right] \geq \beta_L \quad \text{ if } Q(t) < V/\omega_b \quad (36)$$

$$\mathbb{E}\left[Q(t+1) - Q(t)|Q(t)\right] \leq -\beta_R \quad \text{ if } Q(t) \geq V/\omega_b \quad (37)$$

Thus, the system has *positive drift* if $Q(t) < V/\omega_b$, and *negative drift* otherwise (see Fig. 3). It is remarkable that the probability-unaware drift-plus-penalty algorithm yields a drift picture of Fig. 3 that is qualitatively similar to the algorithm of [13] that is designed offline using system probabilities.

D. A basic drift lemma

Consider a real-valued random process $Z(t)$ over slots $t \in \{0, 1, 2, \ldots\}$. The following drift lemma is similar in spirit to results in [21][16], but focuses on a finite time horizon with an arbitrary initial condition $Z(0) = z_0$ (rather than on steady state), and on expectations at a given time (rather than time averages). These distinctions are crucial to convergence time analysis. The lemma will be applied using $Z(t) = Q(t)$ for bounds on average queue size and on $\overline{1}^{(4)}(t)$. It will then be applied using $Z(t) = V/\omega_b - Q(t)$ to bound $\overline{1}^{(1)}(t)$. Assume there is a constant $\delta_{max} > 0$ such that with probability 1:

$$|Z(t+1) - Z(t)| \leq \delta_{max} \ \forall t \in \{0, 1, 2, \ldots\} \quad (38)$$

Suppose there are constants $\theta \in \mathbb{R}$ and $\beta > 0$ such that:

$$\mathbb{E}\left[Z(t+1) - Z(t)|Z(t)\right] \leq \begin{cases} \delta_{max} & \text{if } Z(t) < \theta \\ -\beta & \text{if } Z(t) \geq \theta \end{cases} \quad (39)$$

Note that if (38) holds then (39) automatically holds for the special case $Z(t) < \theta$. Thus, the negative drift case $Z(t) \geq \theta$ is the important case for condition (39). Further, if (38)-(39) both hold, then the constant β necessarily satisfies:

$$0 < \beta \leq \delta_{max}$$

Lemma 4: Suppose $Z(t)$ is a random process that satisfies (38)-(39) for given constants θ, δ_{max}, β (with $\theta \in \mathbb{R}$ and $0 < \beta \leq \delta_{max}$). Suppose $Z(0) = z_0$ for some $z_0 \in \mathbb{R}$. Then for every slot $t \geq 0$ the following holds:

$$\mathbb{E}\left[e^{rZ(t)}\right] \leq D + \left(e^{rz_0} - D\right)\rho^t \quad (40)$$

where constants r, ρ, D are defined:

$$r \triangleq \frac{\beta}{\delta_{max}^2 + \delta_{max}\beta/3} \quad (41)$$

$$\rho \triangleq 1 - r\beta/2 \quad (42)$$

$$D \triangleq \frac{(e^{r\delta_{max}} - \rho)e^{r\theta}}{1-\rho} \quad (43)$$

Note that the property $0 < \beta \leq \delta_{max}$ can be used to show that $0 < \rho < 1$.

Proof: (Lemma 4) The proof is by induction. The inequality (40) trivially holds for $t = 0$. Suppose (40) holds at some slot $t \geq 0$. The goal is to show that it also holds on slot $t + 1$. Let r be a positive number that satisfies $0 < r\delta_{max} < 3$. It is known from results in [21] that for any real number x that satisfies $|x| \leq \delta_{max}$:

$$e^{rx} \leq 1 + rx + \frac{(r\delta_{max})^2}{2(1 - r\delta_{max}/3)} \quad (44)$$

Define $\delta(t) = Z(t+1) - Z(t)$ and note that $|\delta(t)| \leq \delta_{max}$ for all t. Then:

$$\begin{aligned} e^{rZ(t+1)} &= e^{rZ(t)}e^{r\delta(t)} \\ &\leq e^{rZ(t)}\left[1 + r\delta(t) + \frac{(r\delta_{max})^2}{2(1 - r\delta_{max}/3)}\right] \quad (45) \end{aligned}$$

where the final inequality holds by (44). Choose r such that:

$$\frac{(r\delta_{max})^2}{2(1 - r\delta_{max}/3)} = \frac{r\beta}{2} \quad (46)$$

It is not difficult to show that the value of r given in (41) simultaneously satisfies (46) and $0 < r\delta_{max} < 3$. For this value of r, substituting (46) into (45) gives:

$$e^{rZ(t+1)} \leq e^{rZ(t)}\left[1 + r\delta(t) + \frac{r\beta}{2}\right] \quad (47)$$

Now consider the following two cases:

- Case 1: Suppose $Z(t) \geq \theta$. Taking conditional expectations of (47) gives:

$$\begin{aligned} \mathbb{E}\left[e^{rZ(t+1)}|Z(t)\right] &\leq \mathbb{E}\left[e^{rZ(t)}[1 + r\delta(t) + \frac{r\beta}{2}]|Z(t)\right] \\ &\leq e^{rZ(t)}[1 - r\beta + \frac{r\beta}{2}] \quad (48) \\ &= e^{rZ(t)}\rho \end{aligned}$$

where (48) follows by (39), and the final equality holds by definition of ρ in (42).

- Case 2: Suppose $Z(t) < \theta$. Then:

$$\begin{aligned} \mathbb{E}\left[e^{rZ(t+1)}|Z(t)\right] &= \mathbb{E}\left[e^{rZ(t)}e^{r\delta(t)}|Z(t)\right] \\ &\leq e^{rZ(t)}e^{r\delta_{max}} \end{aligned}$$

Putting these two cases together gives:

$$\begin{aligned} &\mathbb{E}\left[e^{rZ(t+1)}\right] \\ &\leq \rho\mathbb{E}\left[e^{rZ(t)}|Z(t) \geq \theta\right] Pr[Z(t) \geq \theta] \\ &\quad + e^{r\delta_{max}}\mathbb{E}\left[e^{rZ(t)}|Z(t) < \theta\right] Pr[Z(t) < \theta] \\ &= \rho\mathbb{E}\left[e^{rZ(t)}\right] \\ &\quad + (e^{r\delta_{max}} - \rho)\mathbb{E}\left[e^{rZ(t)}|Z(t) < \theta\right] Pr[Z(t) < \theta] \\ &\leq \rho\mathbb{E}\left[e^{rZ(t)}\right] + (e^{r\delta_{max}} - \rho)e^{r\theta} \end{aligned}$$

where the final inequality uses the fact that $e^{r\delta_{max}} > 1 > \rho$. By the induction assumption it is known that (40) holds on slot t. Substituting (40) into the right-hand-side of the above inequality gives:

$$\begin{aligned}\mathbb{E}\left[e^{rZ(t+1)}\right] &\leq \rho\left[D + (e^{rz_0} - D)\,\rho^t\right] \\ &\quad +(e^{r\delta_{max}} - \rho)e^{r\theta} \\ &= D + (e^{rz_0} - D)\,\rho^{t+1}\end{aligned}$$

where the final equality holds by the definition of D in (43). This completes the induction step. $\square$

Let $1\{Z(\tau) \geq \theta + c\}$ be an indicator function that is 1 if $Z(\tau) \geq \theta + c$, and 0 else. The next corollary shows that the expected fraction of time that this indicator is 1 decays exponentially in c.

Corollary 1: If the assumptions of Lemma 4 hold, then for any $c > 0$ and any slots T and t that satisfy $0 \leq T < t$:

$$\begin{aligned}&\frac{1}{t}\sum_{\tau=0}^{t-1}\mathbb{E}\left[1\{Z(\tau) \geq \theta + c\}\right] \\ &\leq \frac{(e^{r\delta_{max}} - \rho)e^{-rc}}{1-\rho} + \left[\frac{T}{t} + \frac{e^{r(z_0-c-\theta)}\rho^T}{t(1-\rho)}\right]\end{aligned} \tag{49}$$

where r and ρ are defined in (41)-(42). Further, if $z_0 \leq \theta$ then for any $t > 0$:

$$\frac{1}{t}\sum_{\tau=0}^{t-1}\mathbb{E}\left[1\{Z(\tau) \geq \theta + c\}\right] \leq \frac{e^{-rc}(e^{r\delta_{max}} - \rho + 1/t)}{(1-\rho)} \tag{50}$$

The intuition behind the right-hand-side of (49) is that the first term represents a "steady state" bound as $t \to \infty$, which decays like e^{-rc}. The last two terms (in brackets) are due to the transient effect of the initial condition z_0. This transient can be significant when $z_0 > \theta$. In that case, $e^{r(z_0-c-\theta)}$ might be large, and a time T is required to shrink this term by multiplication with the factor ρ^T.

Proof: (Corollary 1) One has for $t > T$:

$$\sum_{\tau=0}^{t-1}\mathbb{E}\left[1\{Z(\tau) \geq \theta + c\}\right] \leq T + \sum_{\tau=T}^{t-1}\mathbb{E}\left[1\{Z(\tau) \geq \theta + c\}\right] \tag{51}$$

However, for every slot $\tau \geq 0$ one has:

$$e^{rZ(\tau)} \geq e^{r(\theta+c)}1\{Z(\tau) \geq \theta + c\}$$

Taking expectations of both sides gives:

$$\mathbb{E}\left[e^{rZ(\tau)}\right] \geq e^{r(\theta+c)}\mathbb{E}\left[1\{Z(\tau) \geq \theta + c\}\right]$$

Rearranging the above shows that for every slot $\tau \geq 0$:

$$\begin{aligned}\mathbb{E}\left[1\{Z(\tau) \geq \theta + c\}\right] &\leq e^{-r(\theta+c)}\mathbb{E}\left[e^{rZ(\tau)}\right] \\ &\leq e^{-r(\theta+c)}[D + (e^{rz_0} - D)\rho^\tau)]\end{aligned}$$

where the final inequality uses (40). Substituting the above inequality into the right-hand-side of (51) gives:

$$\begin{aligned}&\sum_{\tau=0}^{t-1}\mathbb{E}\left[1\{Z(\tau) \geq \theta + c\}\right] \\ &\leq T + e^{-r(\theta+c)}\sum_{\tau=T}^{t-1}\left[D + (e^{rz_0} - D)\rho^\tau\right] \\ &= T + e^{-r(\theta+c)}\left[(t-T)D + (e^{rz_0} - D)\rho^T\frac{(1-\rho^{t-T})}{(1-\rho)}\right] \\ &\leq T + e^{-r(\theta+c)}\left[tD + \frac{e^{rz_0}\rho^T}{(1-\rho)}\right]\end{aligned}$$

Dividing by t and substituting the definition of D proves (49). Inequality (50) follows immediately from (49) by choosing $T = 0$. $\square$

E. Bounding $\mathbb{E}\left[Q(t)\right]$ *and* $\overline{1}^{(4)}(t)$

Let $Q(t)$ be the backlog process under the drift-plus-penalty algorithm. Assume that $V \geq \omega_M^2$ and the initial condition is $Q(0) = q_0$ for some constant q_0. Define $\delta_{max} \triangleq \max[\omega_M, a_{max}]$ as the largest possible change in $Q(t)$ over one slot, so that:

$$|Q(t+1) - Q(t)| \leq \delta_{max} \;\; \forall t \in \{0, 1, 2, \ldots\}$$

From (37) it holds that:

$$\mathbb{E}\left[Q(t+1) - Q(t)|Q(t)\right] \leq \begin{cases} \delta_{max} & \text{if } Q(t) < V/\omega_b \\ -\beta_R & \text{if } Q(t) \geq V/\omega_b \end{cases}$$

It follows that the process $Q(t)$ satisfies the conditions (38)-(39) required for Lemma 4. Specifically, define $Z(t) = Q(t)$, $z_0 = q_0$, $\theta = V/\omega_b$, $\beta = \beta_R$.

Lemma 5: If $0 \leq q_0 \leq V/\omega_b$ and $V \geq \omega_M^2$, then for all slots $t \geq 0$ one has:

$$\mathbb{E}\left[Q(t)\right] \leq \frac{V}{\omega_b} + \frac{1}{r_R}\log\left(1 + \frac{e^{r_R\delta_{max}} - \rho_R}{1 - \rho_R}\right) = O(V)$$

where constants r_R and ρ_R are defined:

$$r_R \triangleq \frac{\beta_R}{\delta_{max}^2 + \delta_{max}\beta_R/3} \tag{52}$$

$$\rho_R \triangleq 1 - r_R\beta_R/2 \tag{53}$$

The lemma provides a bound on $\mathbb{E}\left[Q(t)\right]$ that does not depend on t. The bound holds whenever the initial condition satisfies $0 \leq q_0 \leq V/\omega_b$. Typically, the initial condition is $q_0 = 0$. However, a *place-holder* technique in Section VI requires a nonzero initial condition that still satisfies the desired inequality $0 \leq q_0 \leq V/\omega_b$.

Proof: For ease of notation, let "r" and "ρ" respectively denote "r_R" and "ρ_R" given in (52) and (53). Define $\theta = V/\omega_b$ and $\beta = \beta_R$. By (40) one has for all $t \geq 0$ (using $Z(0) = Q(0) = q_0$):

$$\begin{aligned}\mathbb{E}\left[e^{rQ(t)}\right] &\leq D + (e^{rq_0} - D)\rho^t \\ &\leq D + e^{rV/\omega_b}\end{aligned} \tag{54}$$

where D is given in (43), and where the final inequality uses $D\rho^t \geq 0$ and $q_0 \leq V/\omega_b$. Using Jensen's inequality gives:

$$e^{r\mathbb{E}[Q(t)]} \leq D + e^{rV/\omega_b}$$

Taking a log of both sides and dividing by r gives:

$$\begin{aligned}\mathbb{E}\left[Q(t)\right] &\leq \frac{\log(D+e^{rV/\omega_b})}{r} \\ &= \frac{1}{r}\log\left(e^{rV/\omega_b} + \frac{(e^{r\delta_{max}}-\rho)e^{rV/\omega_b}}{1-\rho}\right) \\ &= \frac{V}{\omega_b} + \frac{1}{r}\log\left(1+\frac{e^{r\delta_{max}}-\rho}{1-\rho}\right)\end{aligned}$$

□

Lemma 6: If $0 \leq q_0 \leq V/\omega_b$ and $V \geq \omega_M^2$, then for all slots $t > 0$:

$$\overline{1}^{(4)}(t) \leq O(e^{-r_R V(\frac{1}{\omega_{b-1}} - \frac{1}{\omega_b})})$$

where r_R is given by (52).

Proof: For ease of notation, this proof uses "r" to denote "r_R." If the interval $\mathcal{I}^{(4)}$ does not exist then $\overline{1}^{(4)}(t) = 0$ and the result is trivial. Now suppose interval $\mathcal{I}^{(4)}$ exists (so that the interval $\mathcal{I}^{(3)}$ is not the final interval in Fig. 3). Define $\theta = V/\omega_b$, $c = V(1/\omega_{b-1} - 1/\omega_b)$, $\beta = \beta_R$, $\rho = 1 - r\beta_R/2$. Then $1\{Q(\tau) \geq \theta + c\} = 1$ if and only if $Q(\tau) \geq V/\omega_{b-1}$, which holds if and only if $Q(\tau) \in \mathcal{I}_4$. Thus, for all slots $t > 0$:

$$\begin{aligned}\overline{1}^{(4)}(t) &= \frac{1}{t}\sum_{\tau=0}^{t-1}\mathbb{E}\left[1\{Q(\tau) \geq \theta + c\}\right] \\ &\leq \frac{e^{-rc}(e^{r\delta_{max}} - \rho + 1/t)}{1-\rho} \quad (55) \\ &= \frac{e^{-rV(\frac{1}{\omega_{b-1}}-\frac{1}{\omega_b})}(e^{r\delta_{max}} - \rho + 1/t)}{1-\rho}\end{aligned}$$

where (55) holds by (50) (which applies since $z_0 = q_0 \leq \theta$). The right-hand-side of the above inequality is indeed of the form $O(e^{-rV(\frac{1}{\omega_{b-1}}-\frac{1}{\omega_b})})$. □

F. Bounding $\overline{1}^{(1)}(t)$

One can similarly prove a bound on $\overline{1}^{(1)}(t)$. The intuition is that the positive drift in region $\mathcal{I}^{(2)}$ of Fig. 3, together with the fact that the size of interval $\mathcal{I}^{(2)}$ is $\Theta(V)$, makes the fraction of time the queue is to the left of V/ω_b decay exponentially as we move further left. The result is given below. Recall that $Q(0) = q_0$ for some constant $q_0 \geq 0$.

Lemma 7: If $q_0 \geq 0$ and $V \geq \omega_M^2$, then for all slots $t > 0$ one has:

$$\overline{1}^{(1)}(t) \leq O(V)/t + O(e^{-r_L V(\frac{1}{\omega_b} - \frac{1}{\omega_{b+1}})})$$

where r_L is defined:

$$r_L \triangleq \frac{\beta_L}{\delta_{max}^2 + \delta_{max}\beta_L/3}$$

Intuitively, the first term in the above lemma (that is, the $O(V)/t$ term) bounds the contribution from the *transient time* starting from the initial state $Q(0) = q_0$ and ending when the threshold V/ω_b is crossed. The second term represents a "steady state" probability assuming an initial condition V/ω_b. The proof defines a new process $Z(t) = V/\omega_b - Q(t)$. It then applies inequality (49) of Corollary 1, with a suitably large time $T > 0$, to handle the initial condition $z_0 = V/\omega_b - q_0$.

Proof: (Lemma 7) Define $Z(t) = V/\omega_b - Q(t)$ and note that $|Z(t+1) - Z(t)| \leq \delta_{max}$ still holds. Further, from (36) it holds:

$$\mathbb{E}\left[Z(t+1) - Z(t)|Z(t)\right] \leq \begin{cases} \delta_{max} & \text{if } Z(t) \leq 0 \\ -\beta_L & \text{if } Z(t) > 0 \end{cases}$$

Now define θ as any positive value. It follows that:

$$\mathbb{E}\left[Z(t+1) - Z(t)|Z(t)\right] \leq \begin{cases} \delta_{max} & \text{if } Z(t) < \theta \\ -\beta_L & \text{if } Z(t) \geq \theta \end{cases}$$

Thus, the conditions (38)-(39) hold for this $Z(t)$ process, with initial condition $z_0 = V/\omega_b - q_0$. Therefore, Corollary 1 can be applied.

For ease of notation let "r" represent "r_L," let "β" represent "β_L," and let "ρ" represent "ρ_L," where $\rho_L \triangleq 1 - r_L\beta_L/2$. Define $c = V/\omega_b - V/\omega_{b+1}$. From (49) of Corollary 1, the following holds for all slots T, t such that $0 \leq T < t$:

$$\begin{aligned}&\frac{1}{t}\sum_{\tau=0}^{t-1}\mathbb{E}\left[1\{Z(\tau) \geq \theta + c\}\right] \\ &\leq \frac{(e^{r\delta_{max}} - \rho)e^{-rc}}{1-\rho} + \left[\frac{T}{t} + \frac{e^{r(z_0 - c - \theta)}\rho^T}{t(1-\rho)}\right]\end{aligned}$$

This holds for all $\theta > 0$. Taking a limit as $\theta \to 0^+$ gives:

$$\begin{aligned}&\frac{1}{t}\sum_{\tau=0}^{t-1}\mathbb{E}\left[1\{Z(\tau) > c\}\right] \\ &\leq \frac{(e^{r\delta_{max}} - \rho)e^{-rc}}{1-\rho} + \left[\frac{T}{t} + \frac{e^{r(z_0 - c)}\rho^T}{t(1-\rho)}\right]\end{aligned}$$

Notice that the event $1\{Z(\tau) > c\}$ is equivalent to the event $\{Q(\tau) < V/\omega_{b+1}\}$, which is the same as the event $Q(\tau) \in \mathcal{I}_1$ (see Fig. 3). Thus, the left-hand-side of the above inequality is the same as $\overline{1}^{(1)}(t)$. Hence:

$$\begin{aligned}\overline{1}^{(1)}(t) &\leq \frac{(e^{r\delta_{max}} - \rho)e^{-rc}}{1-\rho} + \left[\frac{T}{t} + \frac{e^{r(z_0 - c)}\rho^T}{t(1-\rho)}\right] \\ &\leq \frac{(e^{r\delta_{max}} - \rho)e^{-rc}}{1-\rho} + \left[\frac{T}{t} + \frac{e^{rV/\omega_{b+1}}\rho^T}{t(1-\rho)}\right]\end{aligned}$$

where the final inequality uses the fact that $z_0 \leq V/\omega_b$. By definition of c, the first term on the right-hand-side is $O(e^{-rV(\frac{1}{\omega_b} - \frac{1}{\omega_{b+1}})})$. It remains to choose a value $T > 0$ for which the remaining two terms (in brackets) are $O(V)/t$. To this end, define $x \triangleq r/(\omega_{b+1}\log(1/\rho))$. Choose T as the smallest integer that is greater than or equal to xV. Then $T = O(V)$ and:

$$\begin{aligned}\frac{T}{t} &\leq O(V)/t \\ \frac{e^{rV/\omega_{b+1}}\rho^T}{t(1-\rho)} &\leq \frac{e^{rV/\omega_{b+1}}\rho^{xV}}{t(1-\rho)} \\ &= \frac{1}{t(1-\rho)} \\ &\leq O(V)/t\end{aligned}$$

□

G. Optimal backlog and near-optimal convergence time

Define:

$$\gamma \triangleq \min\left[r_R\left(\frac{1}{\omega_{b-1}} - \frac{1}{\omega_b}\right), r_L\left(\frac{1}{\omega_b} - \frac{1}{\omega_{b+1}}\right)\right]$$

Results of Lemmas 5-7 imply that if the drift-plus-penalty algorithm (20) is used with $V \geq \omega_M^2$, and if the initial queue state satisfies $0 \leq q_0 \leq V/\omega_b$, then for all $t > 0$:

$$\begin{aligned} \overline{Q}(t) &\leq O(V) &(56)\\ \mathbb{E}\left[Q(t)\right]/t &\leq O(V)/t &(57)\\ \overline{1}^{(4)}(t) &\leq O(e^{-\gamma V}) &(58)\\ \overline{1}^{(1)}(t) &\leq O(e^{-\gamma V}) + O(V)/t &(59)\end{aligned}$$

Indeed, (56)-(57) follow from Lemma 5, while (58) and (59) follow from Lemmas 6 and 7, respectively.

Fix $\epsilon > 0$ and define:

$$\begin{aligned} V &= \max[(1/\gamma)\log(1/\epsilon), \omega_M^2] &(60)\\ T_\epsilon &= \log(1/\epsilon)/\epsilon \end{aligned}$$

Inequalities (56)-(59) can be used to easily derive the following facts:[2]

- Fact 1: For all slots $t > 0$ one has $\overline{Q}(t) \leq O(\log(1/\epsilon))$.
- Fact 2: For all slots $t > T_\epsilon$ one has $\mathbb{E}\left[Q(t)\right]/t \leq O(\epsilon)$.
- Fact 3: For all slots $t > 0$ one has $\overline{1}^{(4)}(t) \leq O(\epsilon)$.
- Fact 4: For all slots $t > T_\epsilon$ one has $\overline{1}^{(1)}(t) \leq O(\epsilon)$.

Fact 2 and Lemma 2 ensure that for $t > T_\epsilon$:

$$\overline{\mu}(t) \geq \lambda - O(\epsilon) \quad (61)$$

Facts 2, 3, 4 and Lemma 3 ensure that for $t > T_\epsilon$:

$$|\overline{1}^{(2)}(t) - \theta| \leq O(\epsilon) \quad , \quad |\overline{1}^{(3)}(t) - (1-\theta)| \leq O(\epsilon)$$

Substituting the above into (29) proves that for $t > T_\epsilon$:

$$\begin{aligned} \overline{p}(t) &\leq \theta h(\mu_{b+1}) + (1-\theta)h(\mu_b) + O(\epsilon)\\ &= p^* + O(\epsilon) &(62)\end{aligned}$$

The guarantees (61) and (62) show that the drift-plus-penalty algorithm gives an $O(\epsilon)$-approximation with convergence time $T_\epsilon = O(\log(1/\epsilon)/\epsilon)$. This is within a factor $\log(1/\epsilon)$ of the convergence time lower bound given in Section III. Hence, the algorithm has near-optimal convergence time.

In [14] it is shown that, under mild system assumptions, any algorithm that yields an $O(\epsilon)$-approximation must have average queue size of $\overline{Q}(t) \geq \Omega(\log(1/\epsilon))$. Fact 1 shows the drift-plus-penalty algorithm meets this bound with equality. Hence, not only does it provide near optimal convergence time, it provides an optimal average queue size tradeoff.

[2]For example, Fact 1 follows from (56) and the fact that $V = O(\log(1/\epsilon))$, Fact 2 follows from (57) and the fact that $V/T_\epsilon \leq O(\epsilon)$, and so on.

H. Sample path constraint violation probability

Does the sample path transmission rate usually have behavior similar to its expectation? Recall from (32) that if $Q(0) = 0$, the gap between $\frac{1}{t}\sum_{\tau=0}^{t-1} a(\tau)$ and $\frac{1}{t}\sum_{\tau=0}^{t-1}\mu(\tau)$ is $Q(t)/t$. This gap is probabilistically bounded as follows:

$$\begin{aligned} Pr[Q(t)/t > \epsilon] &= Pr[e^{rQ(t)} > e^{r\epsilon t}]\\ &\leq \mathbb{E}\left[e^{rQ(t)}\right]e^{-r\epsilon t} &(63)\\ &\leq (D + e^{rV/\omega_b})e^{-r\epsilon t} &(64)\end{aligned}$$

where (63) holds by the Markov inequality, and (64) holds from the moment generating function bound (54). Recall that $D + e^{rV/\omega_b}$ is polynomial in the $1/\epsilon$ value.[3] Thus, there is a constant $c > 0$ such that the right-hand-side of (64) is $O(\epsilon)$ whenever $t \geq c\log(1/\epsilon)/\epsilon$.

VI. Practical improvements

A. Place-holders

The structure of this problem admits a practical improvement in queue size via the *place-holder technique* of [10]. This does not change the $O(\log(1/\epsilon))$ average queue size tradeoff with ϵ, but can reduce the *coefficient* that multiplies the $\log(1/\epsilon)$ term. Assume that $V \geq 0$ and define the following nonnegative parameter:

$$q_{place} \triangleq \max\left[\frac{V}{\omega_M} - \omega_M, 0\right] \quad (65)$$

The technique uses a nonzero initial condition $Q(0) = q_{place}$, where the initial backlog q_{place} is *fake data*, also called *place-holder backlog*. Note that $q_{place} > 0$ if and only if $V > \omega_M^2$.

The following lemma refines Lemma 1 and shows that this place-holder backlog is never transmitted. Hence, it acts only to shift the queue size up to a value required to make desirable power allocation decisions via (20).

Lemma 8: If $V \geq \omega_M^2$ and $Q(0) = q_{place}$, then the drift-plus-penalty algorithm (20) chooses $p(t) = \mu(t) = 0$ whenever $Q(t) < V/\omega_M$. Thus, $Q(t) \geq q_{place}$ for all t.

Proof: The proof is similar to that of Lemma 1 and is omitted for brevity. □

Consequently, at every slot t the queue can be decomposed as $Q(t) = q_{place} + Q^{real}(t)$, where $Q^{real}(t)$ is the *real* queue backlog from actual arrivals. The sample path of $Q(t)$ and all power decisions $p(t)$ are the same as when the drift-plus-penalty algorithm is implemented with the nonzero initial condition q_{place}. Of course, every transmission $\mu(t)$ sends real data from the queue, rather than fake data. The resulting algorithm is:

- Initialize $Q^{real}(0) = 0$.
- Every slot t, observe $Q^{real}(t)$ and $\omega(t)$ and choose:

$$p(t) = \begin{cases} 1 & \text{if } (q_{place} + Q^{real}(t))\omega(t) \geq V\\ 0 & \text{otherwise}\end{cases}$$

- Update $Q^{real}(t)$ by:

$$Q^{real}(t+1) = \max[Q^{real}(t) + a(t) - p(t)\omega(t), 0] \quad (66)$$

[3]Indeed, recall from (60) and the proof of Lemma 5 that $V = O(\log(1/\epsilon))$, D is given in (43), and $\theta = V/\omega_b$.

If $q_{place} > 0$ then $q_{place} = V/\omega_M - \omega_M \leq V/\omega_b$. Thus, $0 \leq q_{place} \leq V/\omega_b$, and so the initial condition $Q(0) = q_{place}$ still meets the requirements of the lemmas of the previous section. Therefore, the same performance bounds hold for the power process $p(t)$ and the queue size process $Q(t)$. However, at every instant of time, the *real* queue size $Q^{real}(t)$ is reduced by exactly q_{place} in comparison to $Q(t)$.

B. LIFO scheduling

The queue update equations (66) and (1) allow for any work-conserving scheduling mechanism. The default mechanism is First-In-First-Out (FIFO). However, Last-In-First-Out (LIFO) scheduling can provide significant delay improvements for 98% of the packets [22][15]. Intuitively, this is because the backlog $Q(t)$ is almost always to the right of the V/ω_{b+1} point in Fig. 3. Packets that arrive when $Q(t) \geq V/\omega_{b+1}$ must wait for at least V/ω_{b+1} units of data to be served under FIFO, but are transmitted more quickly under LIFO. Work in [15] mathematically formalizes this observation. Roughly speaking, most packets have average delay reduced by at least $V/(\omega_{b+1}\lambda)$ under LIFO (and without the place-holder technique). With the place-holder technique, this reduction is changed to $(V/\omega_{b+1} - q_{place})/\lambda$ (since the place-holder technique already reduces average delay of *all packets* by q_{place}/λ). One caveat is that, under LIFO, a finite amount of arriving data *might never be transmitted*. Of course, using LIFO as opposed to FIFO does not change the total queue size or the fundamental tradeoff between total average queue size and average power. These issues are explored via simulation in the next section.

VII. Simulation

A. Two channel states

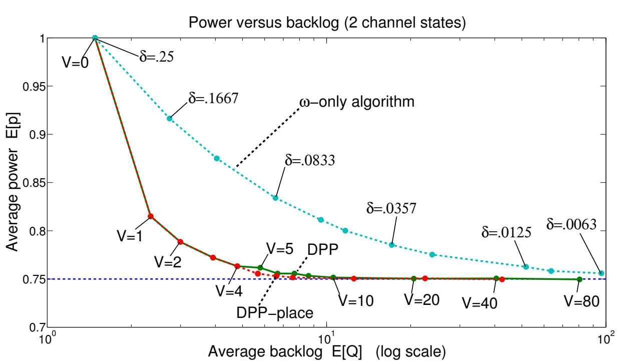

Fig. 4. Average power versus average backlog for the case of 2 channel states. All data points are averages obtained after simulation over 1 million slots. Three algorithms are shown. The drift-plus-penalty (DPP) algorithms use various values of V. The V values are labeled for select points on the DPP curve (green). The ω-only algorithm uses various values of δ.

Consider the scenario of Case 1 in Section III. There are two channel states $\omega(t) \in \{1, 2\}$ with $\pi(1) = 3/4, \pi(2) = 1/4$. The $h(\mu)$ curve is shown in Fig. 1. Assume the arrival process $a(t)$ is i.i.d. over slots with:

$$Pr[a(t) = 0] = \frac{2}{5}\ ,\ Pr[a(t) = 1] = \frac{1}{5}\ ,\ Pr[a(t) = 2] = \frac{2}{5}$$

The arrival rate is $\lambda = \mathbb{E}[a(t)] = 1$, and the minimum average power required for stability is $p^* = h(1) = 3/4$. Three different algorithms are considered below:

- Drift-plus-penalty (DPP) with $Q(0) = 0$.
- DPP with place-holder (DPP-place) with $q_{place} = \max[V/2 - 2, 0]$ (from (65)) and $Q^{real}(0) = 0$.
- An ω-only policy designed to satisfy $\mathbb{E}[\mu(t)] = \lambda + \delta$ and $\mathbb{E}[p(t)] = h(\lambda + \delta)$.

The DPP algorithms operate online without knowledge of λ, $\pi(1)$, $\pi(2)$, while the ω-only policy is designed offline with knowledge of these values. Results are plotted in Fig. 4 for various values of $V \geq 0$ and $\delta \geq 0$. *The DPP algorithms significantly outperform the ω-only algorithm even though they do not have knowledge of the system probabilities*. The theoretical tradeoffs of the previous section were derived under the assumption that $V \geq \omega_M^2$ (in this case, $\omega_M^2 = 2^2 = 4$). However, the DPP algorithms can be implemented for any value $V \geq 0$. Observe from the figure that average power starts approaching optimality even for values $V < 4$, and converges to the optimal $p^* = 3/4$ as V is increased beyond 4. It can be shown that the ω-only algorithm achieves an $O(\epsilon)$-approximation with average queue size $\Theta(1/\epsilon)$, whereas results in the previous section prove the DPP algorithms achieve an $O(\epsilon)$-approximation with average queue size $\Theta(\log(1/\epsilon))$. The simulations verify these theoretical results.

In this example, the DPP place-holder algorithm gives performance very close to standard DPP, with only a modest gain in the range $V \in [4, 10]$. For values $V \leq 4$ the DPP and DPP-place algorithms are identical.

Convergence time to the desired constraint $\overline{\mu}(t) \geq \lambda$ is illustrated in Fig. 5 by plotting the empirical value of $\mathbb{E}[\mu(t)]$ versus time. The ω-only policy is not plotted because it achieves the constraint immediately by its offline design. The DPP-place algorithm shows a slight convergence time improvement over DPP. Both DPP algorithms demonstrate that $|\overline{\mu}(t) - \lambda|$ decays like V/t. This is consistent with the theoretical guarantees derived in the previous section. Indeed, for an $O(\epsilon)$-approximation, one sets $V = \Theta(\log(1/\epsilon))$, so after time $t \geq \Theta(\log(1/\epsilon)/\epsilon)$ the deviation from the constraint is at most $O(V/t) \leq O(\epsilon)$. The corresponding average power $\mathbb{E}[p(t)]$ is plotted in Fig. 6.

B. Nine channel states

Now consider a process $\omega(t)$ with 9 possible rates $\{\omega_0, \ldots, \omega_9\}$:

$$\Omega = \{0, 3, 7, 11, 18, 22, 24, 36, 46\}$$

The probabilities are:

$$\pi(\omega_i) \;=\; \begin{cases} 1/15 & \text{if } i \in \{0, 1, 2\} \\ 2/9 & \text{if } i \in \{3, 4, 5\} \\ 2/45 & \text{if } i \in \{6, 7, 8\} \end{cases}$$

The arrival process $a(t)$ has probabilities:

$$Pr[a(t) = 0] = 0.42, Pr[a(t) = 20] = 0.58$$

with arrival rate $\lambda = 11.6$ packets/slot. The DPP-place algorithm uses $q_{place} = \max[V/46 - 46, 0]$ (as in (65)), and

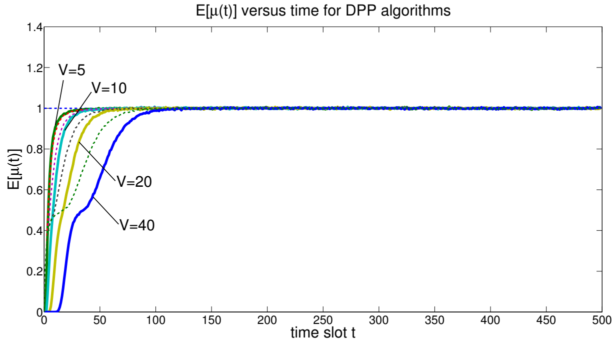

Fig. 5. Average transmission rate $\mathbb{E}\left[\mu(t)\right]$ versus time, obtained from 10^5 independent simulation runs over the first 500 slots. The DPP curves are thick, solid, and labeled with $V \in \{5, 10, 20, 40\}$. The DPP-place curves are thin dashed curves where $V \in \{5, 10, 20, 40\}$ corresponds to red, purple, grey, green, respectively.

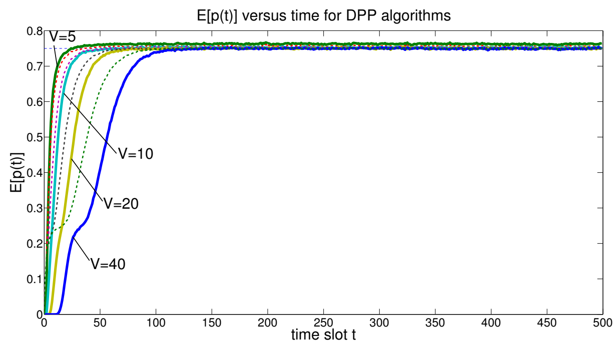

Fig. 6. Average power $\mathbb{E}\left[p(t)\right]$ versus time for the same experiments, V parameters, and color scheme as Fig. 5.

$q_{place} > 0$ if and only if $V > 46^2 = 2116$. It can be shown that $p^* = h(\lambda) = 7/15$ for this system. Simulations for DPP and DPP-place are in Fig. 7. As before, the DPP algorithms outperform the ω-only policy, although the improvements are not as dramatic as they are in Fig. 4. This is because the arrival rate vector in this case is close to a vertex point of the $h(\mu)$ curve. As before, the DPP-place algorithm performance is similar to that of DPP with a shifted V parameter.

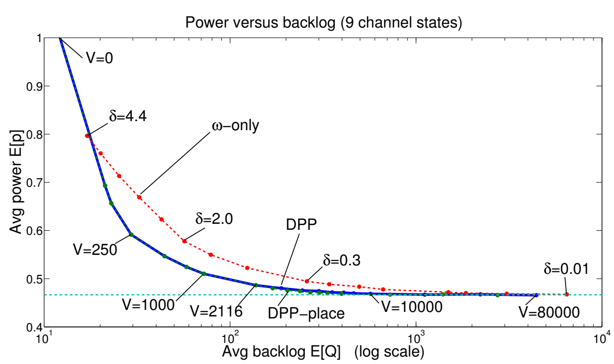

Fig. 7. Simulation for the ergodic system with 9 channel states.

C. Robustness to non-ergodic changes

This subsection illustrates how the algorithm reacts to nonergodic changes. The system with 9 possible channel states from the previous section is considered. The simulation is run over 6000 slots, broken into three phases of 2000 slots each. The system probabilities are changed at the beginning of each phase. The algorithm is not aware of the changes and must adapt. Specifically:

1) First phase: The same parameters of the previous subsection are used (so $\lambda = 11.6$).
2) Second phase: Channel probabilities are the same as phase 1. The arrival rate is increased to $\lambda = 13$ by using $Pr[a(t) = 20] = .65$, $Pr[a(t) = 0] = 0.35$.
3) Third phase: The same arrival rate $\lambda = 13$ of phase 2 is used. However, channel probabilities are changed to:

$$\pi(\omega_i) \quad = \quad \begin{cases} 1/15 & \text{if } i \in \{0, 1, 2\} \\ 1/9 & \text{if } i \in \{3, 4, 5\} \\ 7/45 & \text{if } i \in \{6, 7, 8\} \end{cases}$$

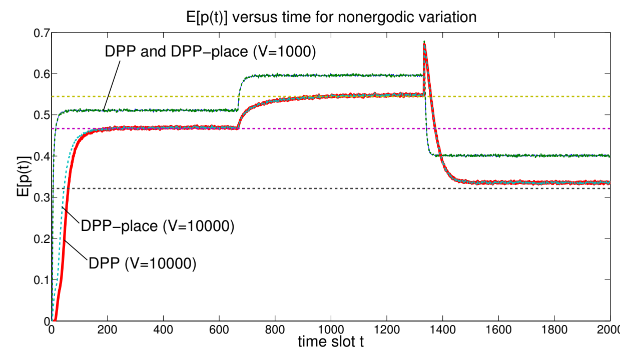

Fig. 8. Average power (obtained from 10000 independent simulation runs) versus time for a system that changes nonergodically over 3 phases.

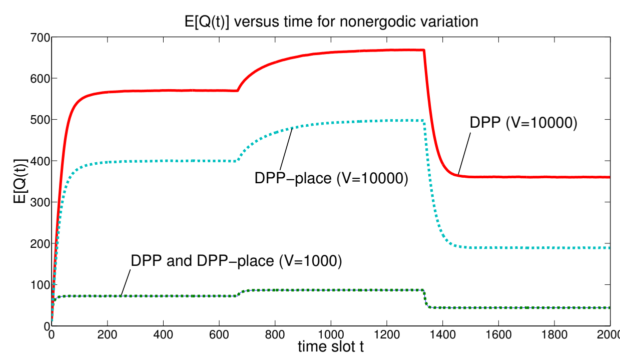

Fig. 9. Average queue size (obtained from 10000 independent simulation runs) versus time for a system that changes nonergodically over 3 phases.

The resulting power and queue size averages are plotted in Figs. 8 and 9. The data is obtained by averaging sample paths over 10000 independent runs. Fig. 8 shows that for large V, average power converges to a value close to the long-term optimum associated with each phase. Thus, the DPP algorithms adapt to changing environments. For each V, average power of DPP-place is roughly the same as DPP (Fig. 8). Average queue size of DPP-place is smaller than that of DPP when V is large (Fig. 9).

D. Delay improvements under LIFO

Fig. 10 illustrates the gains of Last-in-First-Out (LIFO) scheduling (as in [22][15]) for the 9-channel state system with parameters described in Section VII-B (the system is the same as that of Fig. 7). Average power is plotted versus average delay (in slots) for DPP-place with and without LIFO. The LIFO data considers only the 98% of all packets with the smallest delay (so that 2% of the packets are ignored in the delay computation). LIFO scheduling significantly reduces delay for these packets. For example, when $V = 80000$, average delay is 236.3 slots without LIFO, and only 20.0 slots with LIFO (average power is the same for both algorithms).

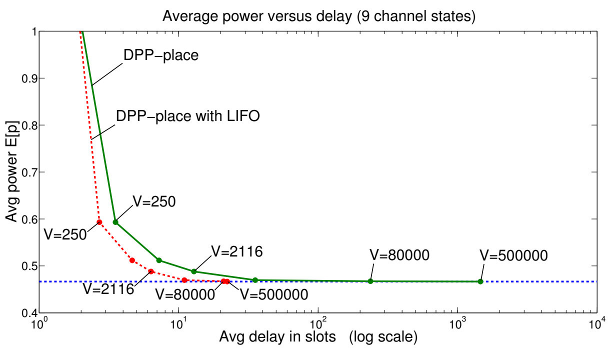

Fig. 10. A comparison of the power-delay tradeoff for DPP-place with and without LIFO. The LIFO data for average delay considers the best 98% of all traffic. Each data point represents a simulation over 10^6 slots. Average power is the same for both algorithms whenever V is the same.

VIII. Conclusions

This paper considers convergence time for minimizing average power in a wireless transmission link with time varying channels and random traffic. It shows that no algorithm can get convergence time better than $O(1/\epsilon)$. It then shows that this ideal convergence time can be approached to within a logarithmic factor. Furthermore, the resulting average queue size is at most $O(\log(1/\epsilon))$, which is known to be an optimal tradeoff. This establishes fundamental convergence time, queue size, and power characteristics of wireless links. It shows that learning times in an unknown environment can be pushed much faster than expected.

References

[1] M. J. Neely. Energy-aware wireless scheduling with near optimal backlog and convergence time tradeoffs. *Proc. IEEE INFOCOM*, Hong Kong, April 2015.

[2] L. Tassiulas and A. Ephremides. Dynamic server allocation to parallel queues with randomly varying connectivity. *IEEE Transactions on Information Theory*, vol. 39, no. 2, pp. 466-478, March 1993.

[3] A. Eryilmaz and R. Srikant. Fair resource allocation in wireless networks using queue-length-based scheduling and congestion control. *IEEE/ACM Transactions on Networking*, vol. 15, no. 6, pp. 1333-1344, Dec. 2007.

[4] A. Eryilmaz and R. Srikant. Joint congestion control, routing, and MAC for stability and fairness in wireless networks. *IEEE Journal on Selected Areas in Communications, Special Issue on Nonlinear Optimization of Communication Systems*, vol. 14, pp. 1514-1524, Aug. 2006.

[5] J. W. Lee, R. R. Mazumdar, and N. B. Shroff. Opportunistic power scheduling for dynamic multiserver wireless systems. *IEEE Transactions on Wireless Communications*, vol. 5, no.6, pp. 1506-1515, June 2006.

[6] H. Kushner and P. Whiting. Asymptotic properties of proportional-fair sharing algorithms. *Proc. 40th Annual Allerton Conf. on Communication, Control, and Computing, Monticello, IL*, Oct. 2002.

[7] R. Agrawal and V. Subramanian. Optimality of certain channel aware scheduling policies. *Proc. 40th Annual Allerton Conf. on Communication, Control, and Computing, Monticello, IL*, Oct. 2002.

[8] A. Stolyar. Maximizing queueing network utility subject to stability: Greedy primal-dual algorithm. *Queueing Systems*, vol. 50, no. 4, pp. 401-457, 2005.

[9] M. J. Neely, E. Modiano, and C. Li. Fairness and optimal stochastic control for heterogeneous networks. *IEEE/ACM Transactions on Networking*, vol. 16, no. 2, pp. 396-409, April 2008.

[10] M. J. Neely. *Stochastic Network Optimization with Application to Communication and Queueing Systems*. Morgan & Claypool, 2010.

[11] X. Liu, E. K. P. Chong, and N. B. Shroff. A framework for opportunistic scheduling in wireless networks. *Computer Networks*, vol. 41, no. 4, pp. 451-474, March 2003.

[12] M. J. Neely. Energy optimal control for time varying wireless networks. *IEEE Transactions on Information Theory*, vol. 52, no. 7, pp. 2915-2934, July 2006.

[13] R. Berry and R. Gallager. Communication over fading channels with delay constraints. *IEEE Transactions on Information Theory*, vol. 48, no. 5, pp. 1135-1149, May 2002.

[14] M. J. Neely. Optimal energy and delay tradeoffs for multi-user wireless downlinks. *IEEE Transactions on Information Theory*, vol. 53, no. 9, pp. 3095-3113, Sept. 2007.

[15] L. Huang, S. Moeller, M. J. Neely, and B. Krishnamachari. LIFO-backpressure achieves near optimal utility-delay tradeoff. *IEEE/ACM Transactions on Networking*, vol. 21, no. 3, pp. 831-844, June 2013.

[16] L. Huang and M. J. Neely. Delay reduction via Lagrange multipliers in stochastic network optimization. *IEEE Transactions on Automatic Control*, vol. 56, no. 4, pp. 842-857, April 2011.

[17] B. Li, R. Li, and A. Eryilmaz. On the optimal convergence speed of wireless scheduling for fair resource allocation. *IEEE Transactions on Networking*, vol. 23, no. 2:pp. 631–643, April 2015.

[18] L. Huang, X. Liu, and X. Hao. The power of online learning in stochastic network optimization. *Proc. SIGMETRICS*, 2014.

[19] S. Supittayapornpong and M. J. Neely. Time-average stochastic optimization with non-convex decision set and its convergence. In *Proc. 13th Intl. Symp. on Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks (WiOpt)*, May 2015.

[20] M. J. Neely. Distributed stochastic optimization via correlated scheduling. *ArXiv technical report, arXiv:1304.7727v2*, May 2013.

[21] F. Chung and L. Lu. Concentration inequalities and martingale inequalities–a survey. *Internet Mathematics*, vol. 3, pp. 79-127, 2006.

[22] S. Moeller, A. Sridharan, B. Krishnamachari, and O. Gnawali. Routing without routes: The backpressure collection protocol. *Proc. 9th ACM/IEEE Intl. Conf. on Information Processing in Sensor Networks (IPSN)*, April 2010.