The Optimality of Two Prices: Maximizing Revenue in a Stochastic Communication System

Longbo Huang, Michael J. Neely

Abstract—**This paper considers the problem of pricing and transmission scheduling for an Access Point (AP) in a wireless network, where the AP provides service to a set of mobile users. The goal of the AP is to maximize its own time-average profit. We first obtain the optimum time-average profit of the AP and prove the "Optimality of Two Prices" theorem. We then develop an online scheme that jointly solves the pricing and transmission scheduling problem in a dynamic environment. The scheme uses an admission price and a business decision as tools to regulate the incoming traffic and to maximize revenue. We show the scheme can achieve any average profit that is arbitrarily close to the optimum, with a tradeoff in average delay. This holds for general Markovian dynamics for channel and user state variation, and does not require a-priori knowledge of the Markov model. The model and methodology developed in this paper are general and apply to other stochastic settings where a single party tries to maximize its time-average profit.**

Index Terms—**Wireless Mesh Network, Pricing, Queueing, Dynamic Control, Lyapunov analysis, Optimization**

I. Introduction

In this paper, we consider the profit maximization problem of an access point (AP) in a wireless mesh network. Mobile users connect to the mesh network via the AP. The AP receives the user data and transmits it to the larger network via a wireless link. Time is slotted with integral slot boundaries $t \in \{0, 1, 2, \ldots\}$, and every timeslot the AP chooses an admission price $p(t)$ (cost per unit packet) and announces this price to all present mobile users. The users react to the current price by sending data, which is queued at the AP. While the AP gains revenue by accepting this data, it in turn has to deliver all the admitted packets by transmitting them over its wireless link. Therefore, it incurs a transmission cost for providing this service (for example, the cost might be proportional to the power consumed due to transmission). The mission of the AP is to find strategies for both packet admission and packet transmission so as to maximize its time average profit while ensuring queue stability.

We assume that the expected number of new packets sent to the AP is determined every timeslot by a *demand state variable* $M(t)$ and a *user demand function* $F(M(t), p(t))$. Specifically, the state variable $M(t)$ represents the current condition of the user population that affects its aggregate spending ability. For example, $M(t)$ can represent the integer number of users present at time t, or can be a rough estimate of the aggregate "willingness-to-pay" (such as "Low", "Medium", and "High"). The demand function $F(M(t), p(t))$ is equal to the expected number of packets that arrive on slot t under a given user condition $M(t)$ and a given price $p(t)$. We assume the AP knows the current demand state $M(t)$ and the demand function $F(M(t), p(t))$ for each slot t. However, $M(t)$ is assumed to vary according to a general finite state ergodic Markov chain, and the transition and steady state probabilities of $M(t)$ may be unknown. Similarly, the condition of the wireless channel from AP to the mesh network is potentially time varying and is determined by a Markov modulated channel state process $S(t)$. The AP is assumed to know the current channel state $S(t)$ on each timeslot t, although the transition and steady state probabilities of $S(t)$ are potentially unknown.

We develop a pricing and transmission scheduling algorithm (PTSA) for the AP. The PTSA algorithm has low complexity and can be viewed as making greedy decisions every timeslot. Despite its simplicity, PTSA is able to dynamically react to the time varying network conditions, and yields an average net profit that can be pushed arbitrarily close to the optimum, with a corresponding tradeoff in average queueing delay.

Many existing works on revenue maximization can be found. Work in [1] [2] models the problem of maximizing revenue as a dynamic program. Work in [3] and [4] model revenue maximization as static optimization problems. A game theoretic perspective is considered in [5], where equilibrium results are obtained. Works [6], [7], and [8] also use game theoretic approaches with the goal of obtaining efficient strategies for both the AP and the users. The paper [9] looks at the problem from a mechanism design perspective, and [10], [11] consider profit maximization with Qos guarantees. Early work on network pricing in [12], [13], and [14] consider throughput-utility maximization rather than revenue maximization. There, prices play the role of Lagrange multipliers, and are used mainly to facilitate better utilization of the shared network resource. This is very different from the revenue maximization problem, where the service provider is only interested in its own profit. Indeed, the revenue maximization problem can be much more complex due to non-convexity issues.

The above prior work does not directly solve the profit maximization problem for APs in a wireless network for one or more of the following reasons: (1) Most works consider time-invariant systems, i.e., the network condition does not change with time. (2) Works that model the problem as an

Longbo Huang (email: longbohu@usc.edu) and Michael J. Neely (web: http://www-rcf.usc.edu/~mjneely) are with the Department of Electrical Engineering, University of Southern California, Los Angeles, CA 90089, USA.

This work was presented in part as an invited paper at the 45th Annual Allerton Conference on Communication, Control, and Computing (Monticello, IL), Sept. 2007.

This material is supported in part by one or more of the following: the DARPA IT-MANET program grant W911NF-07-0028, the NSF grant OCE 0520324, the NSF Career grant CCF-0747525.

optimization problem rely heavily on the assumption that the user utility function or the demand function is concave. (3) Many of the prior works adopt the flow rate allocation model, where a single fixed operating point is obtained and used for all time. However, in a wireless network, the network condition can easily change due to channel fading and/or node mobility, so that a fixed resource allocation decision may not be efficient. Also, although the utility functions can generally be assumed to be concave, it is easy to construct examples where the demand function is non-concave/non-convex even if users have concave utility functions. Indeed, profit maximization problems are often non-convex in nature. Hence, they are generally hard to solve, even in the static case where the channel condition, user condition, and demand function is fixed for all time. It is also common to look for single-price solutions in these static network problems. Our results show that single-price solutions are not always optimal, and that even for static problems the AP can only maximize time average profit by providing a "regular" price some fraction of the time, and a "reduced price" at other times. Moreover, most network pricing work considers flow allocation that neglects the packet-based nature of the traffic, and neglects issues of queueing delay. An exception is the recent work in [15] that considers a packet-based model for a free market wireless network. However, [15] focuses on network-wide efficiency and on guarantees of non-negative profit to all participants, and does not consider the very different problem of maximizing revenue for a single AP.

In order to enable the AP to better react to the varying network condition and to overcome the difficulty of solving non-convex/non-concave optimization problems, we propose a novel joint pricing and transmission scheduling algorithm (PTSA). PTSA has the same nature as the schemes proposed in [15], which are "state-dependent" [12], although it solves a very different problem. PTSA bypasses the non-concavity/non-convexity difficulty by turning the static optimization problem into a stochastic optimization problem. Our analysis of the performance of PTSA uses the Lyapunov techniques and general utility-optimization framework developed in [16] [17] [18]. In particular, we note that our resulting pricing algorithm can be viewed as imposing a flow control mechanism that is similar to [16]. However, the algorithm in [16] is designed to provide fairness and social optimality, whereas the algorithm in this paper maximizes revenue and therefore often exhibits "two-price" behavior that is different from [16]. Our analysis also considers general Markovian demand states and channels, which change the structure of the performance-delay tradeoff by a logarithmic factor when compared to the tradeoff derived in [16] for arrivals and channels that are i.i.d. over slots.

We first consider a single channel AP system. In Section II we describe the network model. In Section III, we discuss some practical issues of our model. In Section IV we characterize the optimal time average profit and prove the "Optimality of Two Prices" theorem. The PTSA algorithm is presented in Section V, where performance optimality is proven. We extend our results to a multi-channel AP system in Section VI. Simulation results are provided in Section VII.

II. Network Model

We consider the network as shown in Fig 1. The network is assumed to operate in slotted time, i.e. $t \in \{0, 1, 2, ...\}$.

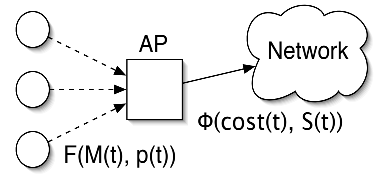

Fig. 1. An Access Point (AP) that connects mobile users to a larger network.

A. Arrival Model: The Demand Function

We first describe the packet arrival model. Let $M(t)$ be the demand state at time t. $M(t)$ might be the number of present mobile users, or could represent the current demand situation, such as the demand being "High", "Medium" or "Low". We assume that $M(t)$ evolves according to a finite state ergodic Markov chain with state space $\mathcal{M}$. Let π_m represent the steady state probability that $M(t) = m$. The value of $M(t)$ is assumed known at the beginning of each slot t, although the transition and steady state probabilities are potentially unknown.

Every timeslot, the AP first makes a *business decision* by deciding whether or not to allow new data (this decision can be based on knowledge of the current $M(t)$ state). Let $Z(t)$ be a $0/1$ variable for this decision, defined as:[1]

$$Z(t) = \begin{cases} 1 & \text{if the AP allows new data on slot } t, \\ 0 & \text{else.} \end{cases} \tag{1}$$

If the AP chooses $Z(t) = 1$, it then chooses a per-unit price $p(t)$ for incoming data and advertises this price to the mobile users. We assume that price is restricted to a set of price options $\mathcal{P}$, so that $p(t) \in \mathcal{P}$ for all t. We assume the set $\mathcal{P}$ includes the constraint that prices are non-negative and bounded by some finite maximum price p_{max}. Let $R(t)$ be the total number of packets that are sent by the mobile users in reaction to this price. The income earned by the AP on slot t is thus $Z(t)R(t)p(t)$.

The arrival $R(t)$ is a random variable that depends on the demand state $M(t)$ and the current price $p(t)$ via a demand function $F(M(t), p(t))$:

$$F : (M(t), p(t)) \mapsto \mathbb{E}\left\{R(t)\right\}. \tag{2}$$

Specifically, the demand function maps $M(t)$ and $p(t)$ into the *expected value* of arrivals $\mathbb{E}\left\{R(t)\right\}$. We further assume that there is a maximum value R_{max}, so that $R(t) \leq R_{max}$ for all t, regardless of $M(t)$ and $p(t)$. The higher order statistics for $R(t)$ (beyond its expectation and its maximum value) are arbitrary. The random variable $R(t)$ is assumed to be conditionally independent of past history given the current

[1]The $Z(t)$ decisions are introduced to allow stability even in the possible situation where user demand is so high that incoming traffic would exceed transmission capabilities, even if price were set to its maximum value p_{max}.

$M(t)$ and $p(t)$. The demand function $F(m, p)$ is only assumed to satisfy $0 \leq F(m, p) \leq R_{max}$ for all $m \in \mathcal{M}$ and all $p \in \mathcal{P}$.

Example: In the case when $M(t)$ represents the number of mobile users in range of the AP at time t, a useful example model for $F(M(t), p(t))$ is:

$$F(M(t), p(t)) = A(M(t))\hat{F}(p(t)),$$

where $\hat{F}(p)$ is the expected number of packets sent by a single user in reaction to price p, a curve that is possibly obtained via empirical data; and $A(M(t))$ is a non-negative function of $M(t)$, e.g. $A(M(t)) = \theta M(t)$, $\theta \geq 0$, which represents the "effective number of participating users" generated by the $M(t)$ present users. In this case, we assume that the $A(M(t))$ is bounded by some value A_{max} and the maximum number of packets sent by any single user is bounded by some value R_{max}^{single}, so that $R_{max} = A_{max} R_{max}^{single}$.

In Section V, we show that this type of demand function (i.e., $F(m, p) = A(m)\hat{F}(p)$) leads to an interesting situation where the AP can make "*demand state blind*" pricing decisions, where prices are chosen without knowledge of $M(t)$.

B. Transmission Model: The Rate-Cost Function

Let $S(t)$ represent the channel condition of the wireless link from AP to the mesh network on slot t. We assume that the channel state process $S(t)$ is a finite state ergodic Markov chain with state space $\mathcal{S}$. Let π_s represent the steady state probability that $S(t) = s$. The transition and steady state probabilities of $S(t)$ are potentially unknown to the AP, although we assume the AP knows the current $S(t)$ value at the beginning of each slot t.

Every slot t, the AP decides how much resource to allocate for transmission. We model this decision completely by its *cost* to the AP, denoted as $cost(t)$. We assume that $cost(t)$ is chosen within some set of costs $\mathcal{C}$, and that $\mathcal{C}$ includes the constraint $0 \leq cost \leq C_{max}$ for some finite maximum cost C_{max}. The transmission rate is then determined by $cost(t)$ and the channel state $S(t)$ according to the *rate-cost*[2] function $\mu(t) = \Phi(cost(t), S(t))$. In our problem, we assume $\Phi(0, S(t)) = 0$ for all $S(t)$. Further, we assume there is a finite maximum transmission rate, so that:

$$\Phi(cost(t), S(t)) \leq \mu_{max} \quad \text{for all } cost(t), S(t), t. \tag{3}$$

We assume that packets can be continuously split, so that $\mu(t) = \Phi(cost(t), S(t))$ determines the portion of packets that can be sent over the link from AP to the network at slot t (for this reason, the rate function can also be viewed as taking units of *bits*). Of course, the set $\mathcal{C}$ can be restricted to a finite set of costs that correspond to integral units for $\Phi(cost(t), S(t))$ in systems where packets cannot be split.

C. Queueing Dynamics and other Notations

Let $U(t)$ be the queue backlog of the AP at time t, in units of packets.[3] Note that this is a *single commodity* problem as we do not distinguish packets from different users.[4] We assume the following queueing dynamics for $U(t)$:

$$U(t+1) = \max[U(t) - \mu(t), 0] + Z(t)R(t), \tag{4}$$

where $\mu(t) = \Phi(cost(t), S(t))$. Throughout the paper, we adopt the following notion of queue stability:

$$\mathbb{E}\{U\} \triangleq \limsup_{t\to\infty} \frac{1}{t}\sum_{\tau=0}^{t-1} \mathbb{E}\{U(\tau)\} < \infty. \tag{5}$$

III. Discussion of Model

A. Application Examples

Our model for pricing of access point services is general and is useful in several different contexts. Below we outline three important scenarios that have different assumptions on the size of a timeslot and the nature of the transmission costs.

1) The "Packet Dropbox" Model: The simplest scenario is one where there are many users, each with a single file to send. Timeslots are large, on the order of tens of seconds to tens of minutes. Each user is in range of the AP only for one slot, and makes a single accept/reject decision based on the advertised price. Such a scenario can be envisioned, for example, in shopping areas where there might be a densely deployed wireless mesh network with APs that act as "drop boxes" for delay-tolerant data.

2) The Extended Hotspot Uplink Model: The second example is where the AP represents a wireless hotspot, such as at an airport or a coffee shop, where users spend a relatively long period of time (perhaps one hour), while the timeslots are very small (less than one second) and the delay requirements are small (e.g., 1 to 50 slots). It extends the original model to also allow multiple service channels. Specifically, the AP maintains a "schedule queue" for each user. At every time slot, the users accepting the price first send request "tokens" (which are small compared to data packets) to the AP indicating the amount of data to send. The AP then places the same amount of virtual bits in the corresponding "schedule queues". The AP then measures the channel conditions that reflect the costs to receive data from users at the current slot, and polls the chosen users to receive packets from them. The sizes of the "schedule queues" kept at the AP actually reflect the data in all users that must be eventually received by the AP over uplink channels, and the costs incurred by the AP are reception costs rather than transmission costs. As we will see in Section VI, this extended model can be used to allow for opportunistic scheduling in cases where different users have different channel conditions.

3) The Extended Hotspot Downlink Model: The third example is in a setting similar to the second example, but where users are interested in obtaining data from the network, such as downloading files. At each slot, users accepting the price will send their requests to the AP, indicating the content and amount of data they want to purchase. The AP will then obtain the data from the network (assuming this is done quickly), and sends them to the users over downlinks. In this case, the AP aggregates all user data for transmission, polls users in a

[2]This is essentially the same as the rate-power curve in [17].

[3]The packet units can be fractional. Alternatively, the backlog could be expressed in units of *bits*.

[4]Our analysis can be extended to treat multi-commodity models, although that is omitted for brevity.

manner similar as in the second example, and incurs downlink transmission costs when delivering the data. In these cases, there might be a limit to the amount of data that can be sent or received over one slot, so that a typical user makes repeated transactions (over multiple slots) based on current prices.

B. Discussion of the Demand Model

Our traffic arrival model assumes the demand at slot t depends only on the current demand state $M(t)$ and price $p(t)$ through a general (possibly non-linear and non-convex) function $F(M(t), p(t))$. This model yields a revenue-maximizing strategy where the AP has a time-varying price. The resulting dynamic prices are qualitatively similar to the time-varying prices seen in many real world economic markets. Such behavior cannot be captured by mathematical models that a-priori assume fixed price strategies.

While our model is quite general, we note the following limitation: The demand state $M(t)$ is assumed to be independent of the pricing decisions made by the AP. This does not capture situations where a user can anticipate a lower future price and hence changes its demand. This problem is most relevant in the second and third application examples of section III-A, where users are in range of the AP for many slots. While it is still profitable to charge a dynamic price in such a scenario (particularly to reap profit from users who cannot afford to wait for a reduced "sale price"), the dynamics of this anticipatory behavior are complex and would require demand to depend not only on the current demand state and advertised price, but also on the evolution of the pricing process itself. However, we note that our resulting pricing algorithm generates prices where the long term averages and peaks can be learned, but where it would be difficult to "guess" the short term price dynamics (see simulations in Section VII). Furthermore, users with much data to send over multiple slots cannot afford to always wait until a lower price state is offered, as the fraction of time such lower prices are offered can limit their overall throughput (See Section V-E for an example). Therefore, it is reasonable to assume that users send data when they find the currently advertised price to be acceptable.

IV. CHARACTERIZING THE MAXIMUM PROFIT

In this section, we characterize the optimal average profit that is achievable over the class of all possible control polices that stabilize the queue at the AP. We show that it suffices for the AP to use only *two prices* for every demand state $M(t)$ to maximize its profit.

A. The Maximum Profit

To describe the maximum average profit, we use an analysis that is similar to the analysis of the capacity region in [19], [20] and the minimum average power for stability problem in [17]. Note that in [19], [20] and [17], the arrival rate is taken as a given parameter, while in our case, the AP needs to balance between the profit from data admission and the cost for packet transmission. The following theorem shows that optimality over all possible policies can be characterized by a simpler class of stationary randomized strategies with the following structure: Every slot, the AP observes $M(t) = m$, and makes a business decision $Z(t)$ by independently and randomly choosing $Z(t) = 1$ with probability $\phi^{(m)}$ (for some $\phi^{(m)}$ values defined for each $m \in \mathcal{M}$). If $Z(t) = 1$, then the AP allocates a price randomly from a countable *collection of prices* $\{p_1^{(m)}, p_2^{(m)}, p_3^{(m)}, \ldots\}$, with probabilities $\{\alpha_k^{(m)}\}_{k=1}^{\infty}$. Similarly, the AP observes $S(t) = s$ and makes a transmission decision by choosing $cost(t)$ randomly from a set of costs $\{cost_k^{(s)}\}_{k=1}^{\infty}$ with probabilities $\{\beta_k^{(s)}\}_{k=1}^{\infty}$.

Theorem 1: (Maximum Profit with Stability) The optimal average profit for the AP, with its queue being stable, is given by $Profit_{av}^{opt}$, where $Profit_{av}^{opt}$ is defined as the following:

$$Profit_{av}^{opt} = sup\{Income_{av} - Cost_{av}\} \tag{6}$$

$$\text{s.t. } Income_{av} = \mathbb{E}_m\left\{\phi^{(m)}\sum_{k=1}^{\infty}\alpha_k^{(m)}F(m, p_k^{(m)})p_k^{(m)}\right\} \tag{7}$$

$$Cost_{av} = \mathbb{E}_s\left\{\sum_{k=1}^{\infty}\beta_k^{(s)}cost_k^{(s)}\right\} \tag{8}$$

$$\lambda_{av} = \mathbb{E}_m\left\{\phi^{(m)}\sum_{k=1}^{\infty}\alpha_k^{(m)}F(m, p_k^{(m)})\right\} \tag{9}$$

$$\mu_{av} = \mathbb{E}_s\left\{\sum_{k=1}^{\infty}\beta_k^{(s)}\Phi\left(cost_k^{(s)}, s\right)\right\} \tag{10}$$

$$\mu_{av} \geq \lambda_{av} \tag{11}$$

$$0 \leq \phi^{(m)} \leq 1 \qquad \forall m \in \mathcal{M} \tag{12}$$

$$p_k^{(m)} \in \mathcal{P} \qquad \forall k, \forall m \in \mathcal{M} \tag{13}$$

$$cost_k^{(s)} \in \mathcal{C}, \qquad \forall k, \forall s \in \mathcal{S} \tag{14}$$

$$\sum_{k=1}^{\infty}\alpha_k^{(m)} = 1 \qquad \forall m \in \mathcal{M} \tag{15}$$

$$\sum_{k=1}^{\infty}\beta_k^{(s)} = 1 \qquad \forall s \in \mathcal{S} \tag{16}$$

where $sup\{\}$ denotes the *supremum*, $\mathbb{E}_s$ and $\mathbb{E}_m$ denote the expectations over the steady state distribution for $S(t)$ and $M(t)$, respectively, and $\phi^{(m)}$, $\alpha_k^{(m)}$, $p_k^{(m)}$, $\beta_k^{(s)}$, and $cost_k^{(s)}$ are auxiliary variables with the interpretation given in the text preceding Theorem 1.

The proof of Theorem 1 contains two parts. Part I shows that no algorithm that stabilizes the AP can achieve an average profit that is larger than $Profit_{av}^{opt}$. Part II shows that we can achieve a profit of at least $\rho Profit_{av}^{opt}$ (for any ρ such that $0 < \rho < 1$) with a particular stationary randomized algorithm that also yields average arrival and transmission rates λ_{av} and μ_{av} that satisfy $\lambda_{av} < \mu_{av}$. The formal proof is similar to the proof in [21] and is omitted for brevity. The following important corollary to Theorem 1 is simpler and is useful for analysis of the online algorithm described in Section V.

Corollary 1: For any $Profit_{av}^{nopt} = Profit_{av}^{opt} - \epsilon^* > 0$,[5] where $\epsilon^* > 0$, there exists a control algorithm $STAT^*$ that makes stationary and randomized business and pricing decisions $Z^*(t)$ and $p^*(t)$ depending only on the current demand state $M(t)$ (and independent of queue backlog), and makes stationary randomized transmission decisions $cost^*(t)$

[5]The case when $Profit_{av}^{nopt} = 0$ can trivially be satisfied and thus is not considered here.

depending only on the current channel state $S(t)$ (and independent of queue backlog) such that:

$$\mathbb{E}\{Z^*(t)R^*(t)\} \leq \mathbb{E}\{\mu^*(t)\}, \tag{17}$$

$$\mathbb{E}\{Z^*(t)p^*(t)F(M(t),p^*(t))\} - \mathbb{E}\{cost^*(t)\} = Profit_{av}^{nopt}, \tag{18}$$

where $\mu^*(t) = \Phi(cost^*(t), S(t))$. The above expectations are taken with respect to the steady state distributions for $M(t)$ and $S(t)$. Specifically:

$$\begin{aligned} \mathbb{E}\{Z^*(t)R^*(t)\} &= \mathbb{E}_m\{Z^*(t)F(m,p^*(t))\}, \\ \mathbb{E}\{\mu^*(t)\} &= \mathbb{E}_s\{\Phi(cost^*(t),s)\}. \quad \square \end{aligned}$$

B. The Optimality of Two Prices

The following two theorems show that instead of considering a countably infinite collection of prices $\{p_1^{(m)}, p_2^{(m)}, \ldots\}$ for the stationary algorithm of Corollary 1, it suffices to consider only *two* price options for each distinct demand state $M(t) \in \mathcal{M}$.

***Theorem** 2:* Let $(\lambda^{(m)*}, Income^{(m)*})$ represent any rate-income tuple formed by a stationary randomized algorithm that chooses $Z(t) \in \{0,1\}$ and $p(t) \in \mathcal{P}$, so that:

$$\begin{aligned} \mathbb{E}\{Z(t)F(M(t),p(t)) \mid M(t)=m\} &= \lambda^{(m)*}, \\ \mathbb{E}\{Z(t)p(t)F(M(t),p(t)) \mid M(t)=m\} &= Income^{(m)*}, \end{aligned}$$

then:

a) $(\lambda^{(m)*}, Income^{(m)*})$ can be expressed as a convex combination of at most three points in the set $\Omega^{(m)}$, defined:

$$\Omega^{(m)} \triangleq \{(ZF(m,p), ZpF(m,p)) \mid Z \in \{0,1\}, p \in \mathcal{P}\}.$$

b) If $(\lambda^{(m)*}, Income^{(m)*})$ is on the boundary of the convex hull of $\Omega^{(m)}$, then it can be expressed as a convex combination of at most two elements of $\Omega^{(m)}$, corresponding to at most two business-price tuples (Z_1, p_1), (Z_2, p_2).

c) If the demand function $F(m,p)$ is continuous in p for each $m \in \mathcal{M}$, and if the set of price options $\mathcal{P}$ is *connected*, then any $(\lambda^{(m)*}, Income^{(m)*})$ point (possibly not on the boundary of the convex hull of $\Omega^{(m)}$) can be expressed as a convex combination of at most two elements of $\Omega^{(m)}$.

Proof: Part (a): It is known that for any vector random variable $\vec{X}$ that takes values within a set Ω, the expected value $\mathbb{E}\{\vec{X}\}$ is in the convex hull of Ω (see, for example, Appendix 4.B in [20]). Therefore, the 2-dimensional point $(\lambda^{(m)*}, Income^{(m)*})$ is in the convex hull of the set $\Omega^{(m)}$. By Caratheodory's theorem (see, for example, [22]), any point in the convex hull of the 2-dimensional set $\Omega^{(m)}$ can be achieved by a convex combination of at most three elements of $\Omega^{(m)}$.

Part (b): We know from part (a) that $(\lambda^{(m)*}, Income^{(m)*})$ can be expressed as a convex combination of at most three elements of $\Omega^{(m)}$ (say, $\boldsymbol{\omega}_1$, $\boldsymbol{\omega}_2$, and $\boldsymbol{\omega}_3$). Suppose these elements are distinct. Because $(\lambda^{(m)*}, Income^{(m)*})$ is on the boundary of the convex hull of $\Omega^{(m)}$, it cannot be in the interior of the triangle formed by $\boldsymbol{\omega}_1$, $\boldsymbol{\omega}_2$, and $\boldsymbol{\omega}_3$. Hence, it must be on an edge of the triangle, so that it can be reduced to a convex combination of two or fewer of the $\boldsymbol{\omega}_i$ points.

Part (c): We know from part (a) that $(\lambda^{(m)*}, Income^{(m)*})$ is in the convex hull of the 2-dimensional set $\Omega^{(m)}$. An extension to Caratheodory's theorem in [23] shows that any such point can be expressed as a convex combination of at most *two* points in $\Omega^{(m)}$ if $\Omega^{(m)}$ is the union of at most two connected components. The set $\Omega^{(m)}$ can be written as:

$$\Omega^{(m)} = \{(0,0)\} \cup \{(F(m,p), pF(m,p)) \mid p \in \mathcal{P}\},$$

which corresponds to the cases $Z = 0$ and $Z = 1$. Let $\hat{\Omega}^{(m)}$ represent the set on the right hand side of the above union, so that $\Omega^{(m)} = \{(0,0)\} \cup \hat{\Omega}^{(m)}$. Because the $F(m,p)$ function is continuous in p for each $m \in \mathcal{M}$, the set $\hat{\Omega}^{(m)}$ is the image of the connected set $\mathcal{P}$ through the continuous function $(F(m,p), pF(m,p))$, and hence is itself connected [24]. Thus, $\Omega^{(m)}$ is the union of at most two connected components. It follows that $(\lambda^{(m)*}; Income^{(m)*})$ can be achieved via a convex combination of at most two elements in $\Omega^{(m)}$. ■

***Theorem** 3:* (Optimality of Two Prices) Let $(\lambda^*, Income^*)$ represent the rate-income tuple corresponding to any stationary randomized policy $Z^*(t)$, $p^*(t)$, $cost^*(t)$, possibly being the policies of Corollary 1 that achieve any near optimal profit $Profit_{av}^{nopt}$. Specifically, assume the algorithm yields an average profit $Profit_{av}^*$ (defined by the left hand side of (18)), and that:

$$\begin{aligned} \lambda^* &= \mathbb{E}_m\{Z^*(t)F(m,p^*(t))\}, \\ Income^* &= \mathbb{E}_m\{Z^*(t)p^*(t)F(m,p^*(t))\}. \end{aligned}$$

Then for each $m \in \mathcal{M}$, there exists two business-price tuples $(Z_1^{(m)}, p_1^{(m)})$ and $(Z_2^{(m)}, p_2^{(m)})$ and two probabilities $q_1^{(m)}, q_2^{(m)}$ (where $q_1^{(m)} + q_2^{(m)} = 1$) such that:

$$\begin{aligned} \lambda^* &= \sum_{m\in\mathcal{M}} \pi_m \sum_{i=1}^{2} \left[q_i^{(m)} Z_i^{(m)} F(m, p_i^{(m)})\right], \\ Income^* &\leq \sum_{m\in\mathcal{M}} \pi_m \sum_{i=1}^{2} \left[q_i^{(m)} Z_i^{(m)} p_i^{(m)} F(m, p_i^{(m)})\right]. \end{aligned}$$

That is, a new stationary randomized pricing policy can be constructed that yields the same average arrival rate λ^* and an average income that is greater than or equal to $Income^*$, but uses at most two prices for each state $m \in \mathcal{M}$.[6]

Proof: For the stationary randomized policy $Z^*(t)$ and $p^*(t)$, define:

$$\begin{aligned} \lambda^{(m)*} &\triangleq \mathbb{E}\{Z^*(t)F(m,p^*(t)) \mid M(t)=m\}, \\ Income^{(m)*} &\triangleq \mathbb{E}\{Z^*(t)p^*(t)F(m,p^*(t)) \mid M(t)=m\}. \end{aligned}$$

Note that the point $(\lambda^{(m)*}, Income^{(m)*})$ can be expressed as a convex combination of at most three points $\boldsymbol{\omega}_1^{(m)}$, $\boldsymbol{\omega}_2^{(m)}$, $\boldsymbol{\omega}_3^{(m)}$ in $\Omega^{(m)}$ (from Theorem 2 part (a)). Then $(\lambda^{(m)*}, Income^{(m)*})$ is inside (or on an edge of) the triangle formed by $\boldsymbol{\omega}_1^{(m)}$, $\boldsymbol{\omega}_2^{(m)}$, $\boldsymbol{\omega}_3^{(m)}$. Thus, for some value $\delta \geq 0$ the point $(\lambda^{(m)*}, Income^{(m)*} + \delta)$ is on an *edge* of the triangle. Hence, the point $(\lambda^{(m)*}, Income^{(m)*} + \delta)$ can be achieved by a convex combination of at most two of the $\boldsymbol{\omega}_i^{(m)}$ values. Hence, for each $m \in \mathcal{M}$, we can find a convex combination of two elements of $\Omega^{(m)}$, defining a stationary randomized

[6]Because the new average income is greater than or equal to $Income^*$, the new average profit is greater than or equal to $Profit_{av}^*$ when this new pricing policy is used together with the old $cost^*(t)$ scheduling policy.

pricing policy with two business-price choices $(Z_1^{(m)}, p_1^{(m)})$, $(Z_2^{(m)}, p_2^{(m)})$ and two probabilities $q_1^{(m)}$, $q_2^{(m)}$. This new policy yields exactly the same average arrival rate λ^*, and has an average income that is greater than or equal to $Income^*$. ■

Most work in network pricing has focused on achieving optimality over the class of single-price solutions, and indeed in some cases it can be shown that optimality can be achieved over this class (so that two prices are not needed). However, such optimality requires special properties of the demand function (Section IV-D provides a sufficient condition for the existence of a single optimal price). Instead, Theorem 3 shows that for *any* demand function $F(m,p)$, the AP can optimize its average profit by using only two prices for every demand state $m \in \mathcal{M}$. We note that there are similar logical arguments about using finite price options to achieve good performance in the economic literature. For example, [25] shows that under certain conditions, the social value of using two price classes is at least half of the optimal value. However, we note that problems there typically consider selling a certain amount of goods in a given time interval, e.g., [26], or assume excessive demand will be lost, e.g., [27], thus are different from our problem, which can be viewed as queueing the excessive demand and serve them later.

Theorem 3 is also related to a classical result of Markov decision theory that bounds the number of required modes for constrained optimization over the class of stationary policies [28]. Indeed, using a more detailed argument as in [28] together with the stationarity and separability of pricing and transmission scheduling that arise from Theorems 1 and 2, our two-price result can likely be extended to show that there exists a policy that achieves maximum revenue (or arbitrarily close to it) where most demand states $m \in \mathcal{M}$ use only *one* price, while at most one demand state requires two prices. We omit this extended analysis for brevity. In fact, the following example shows that the number two is *tight*, in that a single fixed price does not always suffice to achieve optimality.

C. Example Demonstrating Necessity of Two Prices

For simplicity, we consider a static situation where the transmission rate is equal to $\mu = 1$ with zero cost for all t (so that $\Phi(cost(t), S(t)) = 1$ for all $S(t)$ and all $cost(t)$, including $cost(t) = 0$). The demand state $M(t)$ is also assumed to be fixed for all time, so that $F(m,p)$ can be simply written as $F(p)$. Let $\mathcal{P}$ represent the interval $0 \leq p \leq p_{max}$, with $p_{max} = 9$. We consider the following $F(p)$ function:

$$F(p) = \begin{cases} 10 - \frac{9}{2}p & 0 \leq p \leq 2, \\ \frac{9}{7} - \frac{1}{7}p & 2 < p \leq 9. \end{cases} \quad (19)$$

Note that the demand curve (19) is convex and monotone. Indeed, it can represent a market demand generated by two groups of customers having demands $F(p) = \frac{61}{7} - \frac{61p}{14}$, $0 \leq p \leq 2$ and $F(p) = \frac{9}{7} - \frac{p}{7}$, $0 \leq p \leq 9$. Such demand functions are common in the microeconomic literature, e.g., [29], for modeling real world problems. The $F(p)$ and $pF(p)$ functions corresponding to (19) are plotted in Fig. 2. Now consider the situation when the AP only uses one price. First we consider the case when $Z(t) = 1$ for all time. Since $\mu = 1$, in order to stabilize the queue, the AP has to choose a price p such that

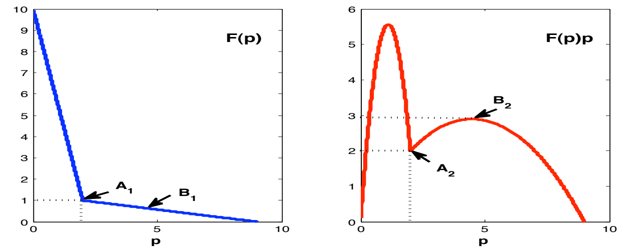

Fig. 2. $A_1 = (2,1)$, $B_1 = (\frac{9}{2}, \frac{9}{14})$, $A_2 = (2,2)$ and $B_2 = (\frac{9}{2}, \frac{81}{28})$.

$\lambda = F(p) \leq 1$. Thus we obtain that p has to be no less than 2 (points A_1 and A_2 in Fig. 2 show $F(p)$ and $F(p)p$ for $p = 2$). It is easy to show that in this case the best single-price is $p = \frac{9}{2}$ (points B_1 and B_2 in Fig. 2 show its $F(p)$ and $F(p)p$), which yields an average profit of $Profit_{single} = 81/28 \approx 2.8929$.[7] However, we see that in this case the average arrival rate $F(p)$ is only $9/14 \approx 0.6429$, which is only 65% of μ. Now consider an alternative scheme that uses two prices $p_1 = \frac{13}{9}$ and $p_2 = \frac{9}{2}$, with probabilities of $\frac{1}{10}$ and $\frac{9}{10}$, respectively. Then the resulting profit is:

$$\begin{aligned} Profit_{Two} &= \frac{1}{10}F(p_1)p_1 + \frac{9}{10}F(p_2)p_2 \\ &= \frac{1}{10} \cdot \frac{7}{2} \cdot \frac{13}{9} + \frac{9}{10} \cdot \frac{9}{14} \cdot \frac{9}{2} \approx 3.1091 \end{aligned}$$

which is strictly larger than $Profit_{single}$. Further, the resulting arrival rate is only:

$$\lambda_{Two} = \frac{1}{10}F(p_1) + \frac{9}{10}F(p_2) = \frac{1}{10} \cdot \frac{7}{2} + \frac{9}{10} \cdot \frac{9}{14} \approx 0.9286$$

which is strictly less than $\mu = 1$. Therefore the queue is stable under this scheme [18].

Now consider the case when the AP uses a varying $Z(t)$ and a single fixed price. From Theorem 1 we see that this is equivalent to using a probability $0 < \phi < 1$ to decide whether or not to allow new data for all time.[8] In order to stabilize the queue, the AP has to choose a price p such that $F(p)\phi < \mu$. Thus the average profit in this case would be $F(p)p\phi < p\mu$. If $p \leq 2$, then $F(p)p\phi < 2 \cdot 1 = 2$ (note that this is indeed just an upper bound); else if $2 < p \leq 9$, $F(p)p\phi < F(\frac{9}{2}) \cdot \frac{9}{2} = 81/28$. Both are less than $Profit_{Two}$ obtained above. Therefore, no single price policy is optimal.

It is interesting to note that the demand curve (19) actually has two unit-elasticity points (which are usually profit maximization points in the economic literature) [29]: $p = \frac{10}{9}$ and $p = \frac{9}{2}$. However, none of them alone achieves the optimal profit under the capacity constraint. Furthermore, it can be verified that the optimal revenue is not achieved by any time sharing between them. This indeed highlights the importance of Theorem 3 and the need of an efficient algorithm.

D. Existence of a Single Optimal Price

While two prices are sufficient and necessary to achieve optimal profit, here we provide a sufficient condition for

[7] Throughout the paper, numbers of this type are numerical results and are accurate enough for our arguments.

[8] The case when ϕ=0 is trivial and thus is excluded.

the existence of a single optimal price for a demand state. We first state the condition in the following corollary from Theorem 2 and 3. Then we show that the class of non-increasing linear demand functions satisfy this condition. Let $(\lambda^{(m)*}, Income^{(m)*})$ represent any rate-income tuple formed by a stationary randomized algorithm as in Theorem 2.

***Corollary** 2:* Let $m \in \mathcal{M}$. Suppose $\mathcal{P} = [a, b]$ for some $0 \leq a \leq b \leq p_{max}$, and $F(m, p)$ is continuous in p. Also suppose for any $p_1, p_2, p_3 \in \mathcal{P}$ and $0 \leq \theta \leq 1$ that satisfy:

$$F(m, p_3) = \theta F(m, p_1) + (1-\theta) F(m, p_2), \tag{20}$$

we have:

$$p_3 F(m, p_3) \geq \theta p_1 F(m, p_1) + (1-\theta) p_2 F(m, p_2), \tag{21}$$

Then there exist a single price $p^{(m)}$ and a probability $q^{(m)}$ for m such that:

$$\begin{aligned} \lambda^{(m)*} &= q^{(m)} F(m, p^{(m)}), \\ Income^{(m)*} &\leq q^{(m)} p^{(m)} F(m, p^{(m)}). \end{aligned}$$

Specifically, $q^{(m)}$ is the probability of choosing $Z = 1$ and the price $p^{(m)}$. Further, if there exists a price p such that $F(m, p) = 0$, then $q^{(m)} = 1$.

Intuitively, equations (20) and (21) together imply that for Z=1, the 2-dimensional curve $(F(m, p), pF(m, p))$ traced out by p is concave in the (x, y) space (See Figure 3). We now use this property to prove Corollary 2.

Proof: (Corollary 2) From Theorem 2 and 3 we see that for any $(\lambda^{(m)*}, Income^{(m)*})$ tuple, there exist two points $\boldsymbol{\omega}_1^{(m)}$ and $\boldsymbol{\omega}_2^{(m)}$ in $\Omega^{(m)}$ so that one convex combination of the two points yields $(\lambda^{(m)*}, Income^{(m)*} + \delta)$ for some $\delta \geq 0$. Now we consider three different cases: If (a) $\boldsymbol{\omega}_1^{(m)} = \boldsymbol{\omega}_2^{(m)}$, or (b) one of them is the point $(0, 0)$ according to choosing $Z = 0$, the corollary follows with $q^{(m)}$ being the probability of *not* choosing $Z = 0$. Now consider (c) $\boldsymbol{\omega}_1^{(m)} \neq \boldsymbol{\omega}_2^{(m)}$ and are both in (Figure 3 shows an example of such situation)

$$\hat{\Omega}^{(m)} = \{(F(m, p), pF(m, p)) \,|\, p \in \mathcal{P}\}.$$

We know that the $(\lambda^{(m)*}, Income^{(m)*} + \delta)$ tuple can be

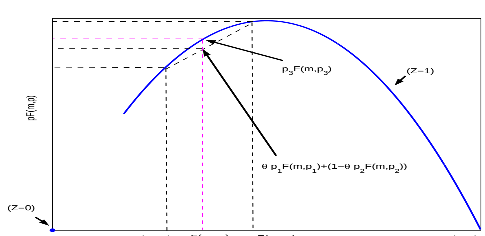

Fig. 3. An example of (1) the $(ZF(m, p), ZpF(m, p))$ curve and (2) the case when $\boldsymbol{\omega}_1^{(m)} \neq \boldsymbol{\omega}_2^{(m)}$ are in $\hat{\Omega}^{(m)}$.

written as: $(\lambda^{(m)*}, Income^{(m)*} + \delta) = q_1^{(m)} \boldsymbol{\omega}_1^{(m)} + q_2^{(m)} \boldsymbol{\omega}_2^{(m)}$, where $\boldsymbol{\omega}_i^{(m)} = (F(m, p_i^{(m)}), p_i^{(m)} F(m, p_i^{(m)}))$, and $q_i^{(m)}$ are probabilities such that $q_1^{(m)} + q_2^{(m)} = 1$. Now since $\mathcal{P} = [a, b]$ and $F(m, p)$ is continuous in p, by the intermediate value theorem, there exists $p_3^{(m)}$ such that:

$$F(m, p_3^{(m)}) = q_1^{(m)} F(m, p_1^{(m)}) + q_2^{(m)} F(m, p_2^{(m)}).$$

Thus by (21), we have:

$$p_3^{(m)} F(m, p_3^{(m)}) \geq q_1^{(m)} p_1^{(m)} F(m, p_1^{(m)}) + q_2^{(m)} p_2^{(m)} F(m, p_2^{(m)}).$$

Therefore using the single price $p_3^{(m)}$ with $q^{(m)} = 1$ yields an income no less than $Income^{(m)*} + \delta$ and yields a same arrival rate.

When there exists p such that $F(m, p) = 0$, the point $(0, 0)$ is also in $\hat{\Omega}^{(m)}$. In this case (a) and (b) are included in (c). Thus we always have $q^{(m)} = 1$. ■

***Lemma** 1:* Let $m \in \mathcal{M}$. If $\mathcal{P} = [a, b]$ with $0 \leq a \leq b \leq p_{max}$, and $F(m, p) = -Q(m)p + W(m)$, for some functions $Q(m), W(m) \geq 0$ such that $F(m, p) \geq 0$, then there exists a single optimal price for m.

Proof: We only need to show that $F(m, p)$ satisfies the conditions in Corollary 2, the details are omitted for brevity. ■

V. Achieving the Maximum Profit

Even though Theorem 2 and 3 show the possibility of achieving the optimum average profit by using only two prices for each demand state, in practice, we still need to solve the problem in Theorem 1. This often involves a very large number of variables and would require the exact demand state and channel state distributions, which are usually hard to obtain. To overcome these difficulties, here we develop the dynamic Pricing and Transmission Scheduling Algorithm (PTSA). The algorithm offers a control parameter $V > 0$ that determines the tradeoff between the queue backlog and the proximity to the optimal average profit. For simplicity, we assume $\mathcal{P}$ is compact and $F(m, p)$ is continuous in $p \in \mathcal{P}$. Likewise, we assume $\mathcal{C}$ is compact and $\Phi(cost, s)$ is continuous in $cost \in \mathcal{C}$.[9]

Admission Control: Every slot t, the AP observes the current backlog $U(t)$ and the user demand $M(t)$ and chooses the price $p(t)$ to be the solution of the following problem:

$$\begin{aligned} Max: \quad & VF(M(t), p)p - 2U(t)F(M(t), p) \\ s.t. \quad & p \in \mathcal{P}. \end{aligned} \tag{22}$$

If for all $p \in \mathcal{P}$ the resulting maximum is less than or equal to zero, the AP sends the "CLOSED" signal ($Z(t) = 0$) and does not accept new data. Else, it sets $Z(t) = 1$ and announces the chosen $p(t)$.

Cost/Transmission: Every slot t, the AP observes the current channel state $S(t)$ and backlog $U(t)$ and chooses $cost(t)$ to be the solution of the following problem:

$$\begin{aligned} Max: \quad & 2U(t)\Phi(cost, S(t)) - V cost \\ s.t. \quad & cost \in \mathcal{C}. \end{aligned} \tag{23}$$

The AP then sends out packets according to $\mu(t) = \Phi(cost(t), S(t))$.

The control policy is thus decoupled into separate algorithms for pricing and transmission scheduling. Note from (22) that a larger $U(t)$ increases the negative term $-2U(t)F(M(t), p)$ in the optimization metric, and hence tends to lead to a higher price $p(t)$. Intuitively, such a slow

[9] These assumptions are only made to ensure the existence of a well defined *max* in equations (22) and (23). Without these assumptions, the algorithm and the analysis can similarly be described and obtained, but are more involved.

down of the packet arrival helps alleviate the congestion in the AP. Note that the metric in (22) can be written as $F(M(t),p)\big(Vp - 2U(t)\big)$. This is positive only if p is larger than $2U(t)/V$. Thus, we have the following simple fact:

***Lemma** 2:* Under the PTSA algorithm, if $2U(t)/V > p_{max}$, then $Z(t) = 0$. □

Notice that PTSA only requires the AP to solve the problems (22) and (23), which use current $M(t)$ and $S(t)$ states but do not require statistical knowledge of how these states evolve. While these problems may be non-convex, we note that they are both optimizing a function of *one* variable over an interval, and hence can easily be solved to obtain highly accurate solutions. For instance, if the pricing set $\mathcal{P}$ contains 100 pricing options, the pricing decision is made just by comparing the metric in (22) over each option. Alternatively, for continuous price options, the function typically has a small number of sub-intervals over which it is piecewise linear or convex, so that the solution can be obtained by comparing the optimums over each sub-interval.

A. Performance Results

In this section we evaluate the performance of PTSA. The following theorem summarizes the performance results:

***Theorem** 4:* PTSA stabilizes the AP and achieves the following bounds (assuming $U(0) = 0$):

$$U(t) \leq U_{max} \triangleq V p_{max}/2 + R_{max}, \quad \forall\, t \quad (24)$$

$$Profit_{av} \geq Profit_{av}^{opt} - \frac{\tilde{B}}{V}, \quad (25)$$

where:

$$Profit_{av} \triangleq \liminf_{t\to\infty} \frac{1}{t}\sum_{\tau=0}^{t-1} \mathbb{E}\left\{Z(\tau)P(\tau)R(\tau) - cost(\tau)\right\},$$

and where $Profit_{av}^{opt}$ is the optimal profit characterized by (6) in Theorem 1, and $\tilde{B}$ is defined in equation (41) of the proof, and $\tilde{B} = O(\log(V))$.

Because $\tilde{B}/V = O(log(V)/V)$, the V parameter can be increased to push the profit arbitrarily close to the optimum value, while the worst case backlog bound grows linearly with V. In fact, we can see from (22) and (23) that these results are quite intuitive: when using a larger V, the AP is more inclined to admit packets (setting $p(t)$ to a smaller value and only requiring $p(t) \geq 2U(t)/V$). Also, a larger V implies that the AP is more careful in choosing the transmission opportunities (indeed, $\Phi\left(cost(t), S(t)\right)$ must be more cost effective, i.e. larger than $Vcost(t)/2U(t)$). Therefore a larger V would yield a better profit, at the cost of larger backlog. The proof of Theorem 4 is given in Section V-C.

B. Discussion of Worst case Delay

Note that in the special case of a fixed $\mu(t) = \mu$ for all t, the worst case delay of any packet is upper bounded by $(\frac{1}{2}Vp_{max} + R_{max})/\mu$. This is a very useful result. For instance, if the users also require the worst case delay to be no more than some constant D, the AP can choose V to be such that $D \geq (\frac{1}{2}Vp_{max} + R_{max})/\mu$ (provided this inequality is achievable). Then the delay requirement is met and the revenue lost is less than $\tilde{B}/V = O(\log V/V)$. This is due to the fact that the delay constrained optimal revenue is no more than $Profit_{av}^{opt}$, while PTSA gets within $\tilde{B}/V$ of $Profit_{av}^{opt}$. This is a unique feature of our algorithm, previous results on QoS pricing are usually obtained based on queueing approximations, e.g., [10], [11].

C. Proof of Performance

We first prove (24) in Theorem 4:

Proof: ((24) in Theorem 4) We prove this by induction. It is easy to see that (24) is satisfied at time 0. Now assume $U(t) \leq Vp_{max}/2 + R_{max}$ for some integer slot $t \geq 0$. We will prove that $U(t+1) \leq Vp_{max}/2 + R_{max}$. We have the following two cases:

(a) $U(t) \leq Vp_{max}/2$: In this case, $U(t+1) \leq Vp_{max}/2 + R_{max}$ by the definition of R_{max}.

(b) $U(t) > Vp_{max}/2$: In this case, $2U(t)/V > p_{max}$. By Lemma 2 the AP will decide not to admit any new data. Therefore $U(t+1) \leq U(t) \leq Vp_{max}/2 + R_{max}$. □

In the following we prove (25) in Theorem 4 via a Lyapunov analysis, using the framework of [18]. First define the Lyaponov function $L(U(t))$ to be: $L(U(t)) \triangleq U^2(t)$. Define the one-step *unconditional Lyapunov drift* as $\Delta(t) \triangleq \mathbb{E}\{L(U(t+1)) - L(U(t))\}$. Squaring both sides of (4) and rearranging the terms, we see that the drift satisfies:

$$\Delta(t) \leq B - \mathbb{E}\big\{2U(t)\big[\Phi(cost(t), S(t)) - Z(t)R(t)\big]\big\}, \quad (26)$$

where $B = R_{max}^2 + \mu_{max}^2$. For a given number $V > 0$, we subtract from both sides the instantaneous profit (scaled by V) and rearrange terms to get:

$$\begin{aligned}\Delta(t) - V\mathbb{E}\big\{Z(t)p(t)R(t) - cost(t)\big\} \\ \leq B - \mathbb{E}\big\{2U(t)\Phi(cost(t), S(t)) - Vcost(t)\big\} \\ -\mathbb{E}\big\{Z(t)\big[Vp(t)R(t) - 2U(t)R(t)\big]\big\}. \quad (27)\end{aligned}$$

Now we see that the PTSA algorithm is designed to *minimize the right hand side of the drift expression (27) over all alternative control decisions that could be chosen on slot t.* Thus, we have that the drift of PTSA satisfies:

$$\begin{aligned}\Delta^P(t) - V\mathbb{E}\big\{Z^P(t)p^P(t)R^P(t) - cost^P(t)\big\} \\ \leq B - \mathbb{E}\big\{2U^P(t)\Phi(cost^*(t), S(t)) - Vcost^*(t)\big\} \\ -\mathbb{E}\big\{Z^*(t)\big[Vp^*(t)R^*(t) - 2U^P(t)R^*(t)\big]\big\}, \quad (28)\end{aligned}$$

where the decisions $Z^*(t)$, $p^*(t)$, and $cost^*(t)$ (and the resulting random arrival $R^*(t)$) correspond to any other feasible control action that can be implemented on slot t (subject to the same constraints $p^*(t) \in \mathcal{P}$ and $cost^*(t) \in \mathcal{C}$). Note that we have used notations $\Delta^P(t)$, $Z^P(t)$, $p^P(t)$, $R^P(t)$, and $cost^P(t)$ on the left hand side of the above inequality to emphasize that this left hand side corresponds to the variables associated with the PTSA policy. Note also that, because the PTSA policy has been implemented up to slot t, the queue backlog on the right hand side at time t is the backlog associated with the PTSA algorithm and hence is also denoted $U^P(t)$. We emphasize that the right hand side of the drift inequality (28) has been modified *only* in those control variables that can be chosen on slot t. Note further that $R^*(t)$ is a random variable that is conditionally independent of the past given the $p^*(t)$ price and the current value of $M(t)$.

Now consider the alternative control policy $STAT^*$ described in Corollary 1, which chooses decisions $Z^*(t), p^*(t)$ and $cost^*(t)$ on slot t as a pure function of the observed $M(t)$ and $S(t)$ states and yields:

$$\begin{aligned} Profit_{av}^{nopt} &= \mathbb{E}_m\{Z^*(t)R^*(t)p^*(t)\} - \mathbb{E}_s\{cost^*(t)\}, \quad (29) \\ \lambda_{av}^* &\triangleq \mathbb{E}_m\{Z^*(t)R^*(t)\} \\ &\leq \mu_{av}^* \triangleq \mathbb{E}_s\{\mu^*(t)\}, \quad (30) \end{aligned}$$

where $Profit_{av}^{nopt} = Profit_{av}^{opt} - \epsilon^*$ and $Profit_{av}^{opt}$ is the optimal average profit defined in Theorem 1, $\mu^*(t) = \Phi(cost^*(t), S(t))$, and $R^*(t)$ is the random arrival for a given $p^*(t)$ and $M(t)$. Recall that $\mathbb{E}_m\{\}$ and $\mathbb{E}_s\{\}$ denote expectations over the steady state distributions for $M(t)$ and $S(t)$, respectively. Of course, the expectations in (29) and (30) cannot be directly used in the right hand side of (28) because the $M(t)$ and $S(t)$ distributions at time t may not be the same as their steady state distributions. However, regardless of the initial condition of $M(0)$ and $S(0)$ we have:

$$\lim_{t\to\infty} \frac{1}{t}\sum_{\tau=0}^{t-1} \mathbb{E}\left\{Z^*(\tau)p^*(\tau)R^*(\tau) - cost^*(\tau)\right\} = Profit_{av}^{nopt}. \quad (31)$$

Let $f^P(t)$ represent a short-hand notation for the left-hand side of (28), and define $g^*(t)$ as the right hand side of (28), so that:

$$\begin{aligned} g^*(t) \triangleq B &- \mathbb{E}\{2U^P(t)[\mu^*(t) - Z^*(t)R^*(t)]\} \\ &-V\mathbb{E}\{Z^*(t)p^*(t)R^*(t) - cost^*(t)\}, \quad (32) \end{aligned}$$

where we have rearranged terms and have used $\mu^*(t)$ to represent $\Phi(cost^*(t), S(t))$. Thus, the inequality (28) is equivalent to $f^P(t) \leq g^*(t)$. To compute a simple upper bound on $g^*(t)$, note that for any integer $d \geq 0$, we have:

$$U^P(t-d) - d\mu_{max} \leq U^P(t) \leq U^P(t-d) + dR_{max}.$$

These inequalities hold since the backlog at time t is no smaller than the backlog at time $t-d$ minus the maximum departures during the interval from $t-d$ to t, and is no larger than the backlog at time $t-d$ plus the largest possible arrivals during this interval. Plugging these two inequalities directly into the definition of $g^*(t)$ in (32) yields:

$$\begin{aligned} g^*(t) \leq\ & B + 2d(\mu_{max}^2 + R_{max}^2) \\ &-\mathbb{E}\left\{2U^P(t-d)[\mu^*(t) - Z^*(t)R^*(t)]\right\} \\ &-V\mathbb{E}\left\{Z^*(t)p^*(t)R^*(t) - cost^*(t)\right\}. \quad (33) \end{aligned}$$

Also note that (by the law of iterated expectations):

$$\begin{aligned} &\mathbb{E}\{U^P(t-d)[\mu^*(t) - Z^*(t)R^*(t)]\} \\ &= \mathbb{E}\Big\{U^P(t-d)\mathbb{E}\{[\mu^*(t) - Z^*(t)R^*(t)] \mid \chi(t-d)\}\Big\}, \quad (34) \end{aligned}$$

where $\chi(t) \triangleq [M(t), S(t), U(t)]$ is the joint demand state, channel state, and queue state of the system. Since $M(t)$ and $S(t)$ are Markovian and both have well defined steady state distributions, and the $STAT^*$ policy makes $p^*(t)$ and $cost^*(t)$ decisions as a stationary and random function of the observed $M(t)$ and $S(t)$ states (and independent of queue backlog), we see that the resulting processes $\mu^*(t)$ and $Z^*(t)R^*(t)$ are Markovian and have well defined steady state averages. Further, they converge exponentially fast to their steady state values [30]. Of course, we know the steady state averages are given by μ_{av}^* and λ_{av}^*, respectively. Therefore there exist positive constants θ_1, θ_2, and $0 < \gamma_1, \gamma_2 < 1$, such that:

$$\begin{aligned} \mathbb{E}\{\mu^*(t) \mid \chi(t-d)\} &\geq \mu_{av}^* - \theta_1\gamma_1^d, \quad (35) \\ \mathbb{E}\{Z^*(t)R^*(t) \mid \chi(t-d)\} &\leq \lambda_{av}^* + \theta_2\gamma_2^d. \quad (36) \end{aligned}$$

Plugging (35) and (36) into (34) yields:

$$\begin{aligned} &\mathbb{E}\{U^P(t-d)[\mu^*(t) - Z^*(t)R^*(t)]\} \\ &\geq -\mathbb{E}\Big\{U^P(t-d)\left[\theta_1\gamma_1^d + \theta_2\gamma_2^d\right]\Big\}, \quad (37) \end{aligned}$$

where we have used the fact that $\lambda_{av}^* \leq \mu_{av}^*$ (from (30)). Plugging (37) directly into (33) yields:

$$\begin{aligned} g^*(t) \leq\ & B_1 + 2\mathbb{E}\{U^P(t-d)(\theta_1\gamma_1^d + \theta_2\gamma_2^d)\} \\ &-V\mathbb{E}\left\{Z^*(t)p^*(t)R^*(t) - cost^*(t)\right\}, \quad (38) \end{aligned}$$

where $B_1 \triangleq B + 2d(\mu_{max}^2 + R_{max}^2)$. However, the queue backlog under PTSA is always bounded by U_{max} (by (24) in Theorem 4). We now choose d large enough so that $\theta_i\gamma_i^d \leq 1/(2U_{max})$ for $i \in \{1, 2\}$. Specifically, by choosing:

$$d \triangleq \left\lceil \max_{i=1,2}\left\{\frac{\log(2\theta_i U_{max})}{\log(1/\gamma_i)}\right\}\right\rceil, \quad (39)$$

we have $2U_{max}[\theta_1\gamma_1^d + \theta_2\gamma_2^d] \leq 2$. Inequality (38) becomes:

$$g^*(t) \leq B_1 + 2 - V\mathbb{E}\left\{Z^*(t)p^*(t)R^*(t) - cost^*(t)\right\}. \quad (40)$$

Now define $\tilde{B}$ as follows:

$$\tilde{B} \triangleq B_1 + 2 = (2d+1)(R_{max}^2 + \mu_{max}^2) + 2, \quad (41)$$

where d is defined in (39). Because $U_{max} = Vp_{max}/2 + R_{max}$ (by (24) in Theorem 4), the value of d is $O(\log(V))$, and hence $\tilde{B} = O(\log(V))$. Recalling that $f^P(t) \leq g^*(t)$, where $f^P(t)$ is the left hand side of (28), we have:

$$\begin{aligned} &\Delta^P(t) - V\mathbb{E}\{Z^P(t)p^P(t)R^P(t) - cost^P(t)\} \\ &\leq \tilde{B} - V\mathbb{E}\left\{Z^*(t)p^*(t)R^*(t) - cost^*(t)\right\}. \end{aligned}$$

The above inequality holds for all t. Summing both sides over $\tau \in \{0, 1, \ldots, t-1\}$ and using $\Delta^P(t) = \mathbb{E}\left\{L(U^P(t+1)) - L(U^P(t))\right\}$, we get:

$$\begin{aligned} &\mathbb{E}\left\{L(U^P(t))\right\} - \mathbb{E}\left\{L(U^P(0))\right\} \\ &-V\sum_{\tau=0}^{t-1}\mathbb{E}\left\{Z^P(\tau)p^P(\tau)R^P(\tau) - cost^P(\tau)\right\} \\ &\leq \tilde{B}t - V\sum_{\tau=0}^{t-1}\mathbb{E}\left\{Z^*(\tau)p^*(\tau)R^*(\tau) - cost^*(\tau)\right\}. \end{aligned}$$

Dividing by Vt, using the fact that $L(U^P(t)) \geq 0$, $L(U(0)) = 0$, and taking limits yields:

$$\begin{aligned} &\liminf_{t\to\infty} \frac{1}{t}\sum_{\tau=0}^{t-1}\mathbb{E}\left\{Z^P(\tau)p^P(\tau)R^P(\tau) - cost^P(\tau)\right\} \geq \\ &Profit_{av}^{nopt} - \tilde{B}/V = Profit_{av}^{opt} - \epsilon^* - \tilde{B}/V, \quad (42) \end{aligned}$$

where we have used (31). The LHS of (42) is the liminf time-average profit of the PTSA algorithm. Now let $\epsilon^* \to 0$ completes the proof of Theorem 4.

D. Demand Blind Pricing

In the special case when the demand function $F(m,p)$ takes the form of $F(m,p) = A(m)\hat{F}(p)$ for $A(m) \geq 0$, PTSA can in fact choose the current price *without* looking at the current demand state $M(t)$. To see this, note in this case that (22) can be written as:

$$\begin{aligned} Max: \quad & A(M(t))\big[V\hat{F}(p)p - 2U(t)\hat{F}(p)\big] \\ s.t. \quad & p \in \mathcal{P}. \end{aligned} \tag{43}$$

Thus we see that the price set by the AP under PTSA is *independent* of $M(t)$. So in this case, PTSA can make decisions just by looking at the queue backlog value $U(t)$. This will be very useful if acquiring the demand state incurs some cost to the AP.

E. The Effect of Users Anticipating Prices

Here we provide an example where some users try to anticipate the AP's price and only send packets when the price is low. As we will see, anticipating the price can lead to a significant throughput loss of the anticipating users. When each user's transmission is associated with some utility, we see that anticipating the price can also lead to a significant payoff (i.e., utility minus cost) loss of the anticipating users.

We consider an AP with a constant service rate μ serving three homogeneous users. Each user will send 0, 1 or 2 packets if the AP sets the price to be 1, $1/2$ or $1/3$. This can happen, for instance, if each user has a utility function $Utility(x) = \log(1+x)$ and chooses $x = 0$, 1 or 2 to maximize its payoff $po(x) = Utility(x) - px$ when it sees a price p. We assume that each user can send no more than 2 packets at a time, due to some system constraint, such as a peak power constraint. We thus obtain the demand function as:

$$F(p) = \begin{cases} 6 & p = 1/3, \\ 3 & p = 1/2, \\ 0 & p = 1. \end{cases} \tag{44}$$

An AP implementing the PTSA algorithm will choose the price p to maximize the following at every time slot :

$$VF(p)p - 2U(t)F(p) = \begin{cases} 2V - 12U(t) & p = 1/3, \\ 3V/2 - 6U(t) & p = 1/2, \\ 0 & p = 1. \end{cases}$$

Assume $V = 100$. It is easy to see that whenever $U(t) < V/12 \approx 8.33$, $p = 1/3$; else if $U(t) < V/4 = 25$, $p = 1/2$; else $p = 1$. Recall that the queueing dynamic is given by:

$$U(t+1) = \max[U(t) - \mu, 0] + R(t).$$

We consider $\mu = 1.5$ and $\mu = 4$. First we look at the throughput of each user when users react to the price "normally" and do not try to wait for a low price. We summarize the results in Fig. 4. For $\mu = 1.5$, each user gets a throughput of $1/2$, i.e., sends a packet every other slot. The average packet price for each user is $1/2$ per packet. For $\mu = 4$, each user gets a throughput of $4/3$ and pays an average of $5/12$ per packet.

Now suppose user 1 algorithm changes so that it only sends packets only when $p(t) = 1/3$. We similarly obtain Fig. 5. For $\mu = 1.5$, we see that after time slot 2, the price $1/3$ will not appear again. Thus user 1 will not be able to send again after time 2, resulting a throughput loss of 100%. The other two users each get a throughput of $3/4$ and pay $1/2$ per packet on average. For $\mu = 4$, user 1 will get a throughput of 1, resulting a throughput loss of 25%, while the other two users each get a throughput of $3/2$ and pay $7/18$ per packet on average.

$\mu = 1.5$			$\mu = 4$		
Time t	$U(t)$	Price $p(t)$	Time t	$U(t)$	Price $p(t)$
0	0	1/3	0	0	1/3
1	6	1/3	1	6	1/3
2	10.5	1/2	2	8	1/3
3	12	1/2	3	10	1/2
.	.	.	4	9	1/2
.	.	.	5	8	1/3
11	24	1/2	6	10	1/2
12	25.5	1	7	9	1/2
13	24	1/2	8	8	1/3

Fig. 4. The case when all users follow the AP's prices. For $\mu = 1.5$, we see that $(U(13), p(13)) = (U(11), p(11))$. Thus the $(U(t), p(t))$ process will repeat the pattern $(24, 1/2), (25.5, 1)$ starting at time 11. For $\mu = 4$, we see that $(U(5), p(5)) = (U(2), p(2))$. The $(U(t), p(t))$ process repeats the pattern $(8, 1/3), (10, 1/2), (9, 1/2)$ starting at time 2.

$\mu = 1.5$			$\mu = 4$		
Time t	$U(t)$	Price $p(t)$	Time t	$U(t)$	Price $p(t)$
0	0	1/3	0	0	1/3
1	6	1/3	1	6	1/3
2	10.5	1/2	2	8	1/3
3	11	1/2	3	10	1/2
.	.	.	4	8	1/3
.	.	.	5	10	1/2
28	23.5	1/2	6	8	1/3
29	24	1/2	7	10	1/2
30	24.5	1/2	8	8	1/3
31	25	1	9	10	1/2
32	23.5	1/2	10	8	1/3

Fig. 5. The case when one user anticipates the AP's prices. For $\mu = 1.5$, we have $(U(32), p(32)) = (U(28), p(28))$. Thus $(U(t), p(t))$ repeats the pattern $(23.5, 1/2), (24, 1/2), (24.5, 1/2), (25, 1)$ starting at time 28. For $\mu = 4$, we see that $(U(4), p(4)) = (U(2), p(2))$. Thus $(U(t), p(t))$ repeats the pattern $(8, 1/3), (10, 1/2)$ starting at time 2.

The cases when two users only send packets at price $1/3$ and when all three users are anticipating the price can be similarly analyzed. To also study the possible effect of price anticipation on the user's average payoff, we consider the case where each user has the utility function $Utility(x) = \log(1+x)$. We compute each user's average payoff, i.e., the average of $po(x) = Utility(x) - px$, under the four different cases. The results are shown in Fig. 6.

		$\mu = 1.5$			$\mu = 4$		
Case	User	$\overline{x}$	$\overline{p}$	$\overline{po}$	$\overline{x}$	$\overline{p}$	$\overline{po}$
3 NU	NU	1/2	1/2	0.097	4/3	5/12	0.273
2 NU	NU	3/4	1/2	0.145	3/2	7/18	0.313
1 AU	AU	0	0	0	1	1/3	0.216
1 NU	NU	11/10	31/66	0.217	8/5	3/8	0.336
2 AU	AU	1/5	1/3	0.043	6/5	1/3	0.259
3 AU	AU	1/2	1/3	0.108	4/3	1/3	0.288

Fig. 6. $\overline{x}, \overline{p}, \overline{po}$ denote the average throughput, cost per packet and payoff. NU and AU refer to the "normal" user and the price-anticipating user.

We see that when all users anticipate the price, the users can achieve higher payoffs and pay lower prices. For example, when $\mu = 4$ and all users anticipate the price, each user pays the lowest price $1/3$ per packet and gets a throughput-payoff pair of $(\frac{4}{3}, 0.288)$, which is better than $(\frac{4}{3}, 0.273)$, obtained if they all follow the price. However, in all other cases, the anticipating users will lose in both throughput and payoff. For instance, when $\mu = 1.5$ and two users anticipate the price, the two anticipating users each get a throughput-payoff pair of $(\frac{1}{5}, 0.043)$. This is more than 50% less than $(\frac{1}{2}, 0.097)$, which is what they could achieve if they simply follow the price. When $\mu = 1.5$ and only one user tries to anticipate the price, the anticipating user even gets zero throughput and zero payoff (100% loss). Moreover, the situation when all users anticipate the price is not likely going to happen as it requires cooperation among all the users, which are non-cooperative in general. Even if the users do cooperate, we see that such a situation is "unstable", as each user has the motivation to break the cooperation since it can get a better performance by not cooperating. For example, when $\mu = 4$ and all users are anticipating the price, if user 1 switches to follow the price, it will get a throughput-payoff pair of $(\frac{8}{5}, 0.336)$, which is strictly better than $(\frac{4}{3}, 0.288)$, obtained if it continues to anticipate the price.

We note that in the above example, if the user's average payoff is instead defined as $Utility(\bar{x}) - \overline{px}$, where $\bar{x}$ and $\overline{px}$ are the average throughput and the average payment respectively, then in some cases a user can improve its payoff performance by anticipating the price. Understanding the effects of general user anticipating strategies is an important topic for future research.

VI. Multi-channel Pricing and Transmission Scheduling

In this section, we consider the extended hotspot uplink model in Section III. For simplicity, we consider a case where there is a fixed number of users accessing a network via an AP, so the demand state $M(t)$ does not change with time. Such a situation can arise, e.g., when the number of users changes in a time scale much larger than the time slot size.

A. System Model

1) Arrival Model: We consider an AP provides service to a set of K users, denoted as $\mathcal{K} = \{1, 2, ..., K\}$. Each user is assumed to generate traffic that is going into the network, such as uploading files to network servers. Each user pays for every packet it wants to send to the network. At every time slot, the AP first makes a business decision on whether or not to allow new data. We use $Z(t)$ as in (1) to denote this decision. If the AP decides to accept new packets ($Z(t) = 1$), it chooses a price $p(t) \in \mathcal{P}$ and announces it to all the K users. $\mathcal{P}$ is assumed to include the constraint $p_{max} \geq p(t) \geq 0$. Let $R_i(t)$ be the total number of packets that user i decides to pay for at time t in reaction to $p(t)$. We assume $R_i(t)$ to be a random variable depending on the current price $p(t)$ via a demand function $F_i(p(t))$: $F_i : p(t) \mapsto \mathbb{E}\{R_i(t)\}$ for all i.

We assume there is a maximum value R_{max} so that $R_i(t) \leq R_{max}$ for all i and all t. The random variable $R_i(t)$ in this case is assumed to be conditionally independent of the past history given $p(t)$. The demand functions are assumed to satisfy $0 \leq F_i(p) \leq R_{max}$ for all $p \in \mathcal{P}$ and all i. We also assume there exists a constant number r_m such that $F_i(p) \leq r_m F_j(p)$ for all i, j and for all $p \in \mathcal{P}$. It is easy to see that $r_m \geq 1$. The last assumption indicates that for a given price p, the demand of one user is "not too different" from those of other users.

2) Transmission Model: Store and Poll: The packets that user i pays for at time t are first stored in a buffer at user i. The users then send their request tokens to the AP, indicating the numbers of packets that they want to send under the current price. The AP maintains a "schedule queue" for each user recording how many packets are waiting to be sent at the user. Every time slot, the AP first collects the tokens and updates its record of each user's buffer size. The AP then looks at the schedule queues and the channel conditions and decides which users to serve as well as the rates at which they are served. The chosen users will then be polled and send packets with the corresponding rates.

Each user connects to the AP with a wireless link. Let $S_i(t)$ represent the channel condition of the link from user i to the AP at t. We assume the $S_i(t)$ process is a finite state ergodic Markov chain with state space $\mathcal{S}_i$ for all i. Every slot, the AP decides how much resource to allocate for receiving packets from each user. This decision is modeled by its cost to the AP, denoted as $cost(t) = (cost_1(t), cost_2(t), ..., cost_K(t))$, where $cost_i(t)$ is the cost spent to receive packets from user i. We assume $cost(t)$ is chosen within some set of cost vectors $\mathcal{C}$ and $\mathcal{C}$ includes the constraint $0 \leq cost_i(t) \leq C_{max}$ for all i and t. The transmission rate of user i is then determined by $cost(t)$ and $S_i(t)$ according to the cost-rate function $\mu_i(t) = \Phi_i(cost(t), S_i(t))$. We assume $\Phi_i(cost(t), S_i(t)) = 0$ if $cost_i(t) = 0$ and $\Phi_i(cost(t), S_i(t)) \leq \mu_{max}$ for all i, $S_i(t)$ and for all time.

3) Queueing Dynamics: Let $U_i(t)$ be the queue backlog at user i at time t.[10] We assume the following queueing dynamics: $U_i(t+1) = \max[U_i(t) - \mu_i(t), 0] + Z(t)R_i(t), \forall i$. Notice that if channels are the same for all users and the AP can poll multiple users in a single slot, then the problem is exactly the same as the one in Section II.

B. Multi-channel PTSA (MPTSA)

In this section we develop the multi-channel PTSA algorithm (MPTSA). Define the Lyapunov function to be $L(t) = \sum_{i \in \mathcal{K}} (U_i(t))^2$. We compute the *unconditional* drift as follows:

$$\Delta(t) \leq B - \mathbb{E}\big\{2 \sum_{i \in \mathcal{K}} U_i(t)[\mu_i(t) - Z(t)R_i(t)]\big\},$$

where $B = K(\mu_{max}^2 + R_{max}^2)$. Now subtract from both sides the optimization metric: $h(t) = V\mathbb{E}\big\{\sum_{i \in \mathcal{K}} Z(t)p(t)R_i(t) - \sum_{i \in \mathcal{K}} cost_i(t)\big\}$. Rearranging the terms, we obtain:

$$\begin{aligned} &\Delta(t) - V\mathbb{E}\big\{\sum_{i \in \mathcal{K}} Z(t)p(t)R_i(t) - \sum_{i \in \mathcal{K}} cost_i(t)\big\} \\ &\leq B - \mathbb{E}\big\{V \sum_{i \in \mathcal{K}} Z(t)p(t)R_i(t) - 2\sum_{i \in \mathcal{K}} U_i(t)Z(t)R_i(t)\big\} \quad (45) \\ &\quad - \mathbb{E}\{2 \sum_{i \in \mathcal{K}} U_i(t)\mu_i(t) - V\sum_{i \in \mathcal{K}} cost_i(t)\}. \end{aligned}$$

This drift is similar as in the single channel case. Now choose the corresponding prices and costs to minimize the right hand

[10] This is also the size of the schedule queue at the AP for user i.

side of (45), we obtain the following multi-channel PTSA algorithm (MPTSA):

Admission Control: Every slot t, the AP observes all $U_i(t)$ and chooses the price $p(t)$ to be the solution of the following problem:

$$Max : \sum_{i \in \mathcal{K}} \big[VF_i(p)p - 2U_i(t)F_i(p)\big], \quad s.t. : p \in \mathcal{P}. \quad (46)$$

If for all $p \in \mathcal{P}$, the maximum is less than or equal to zero, the AP sends the "CLOSED" signal ($Z(t) = 0$) and does not admit any new packet. Else it sets $Z(t) = 1$ and announces the chosen price $p(t)$.

Cost/Reception: Every slot, the AP observes all the channel states $S_i(t)$ and backlogs $U_i(t)$ and chooses $cost(t)$ to be the vector that solves the following problem:

$$\begin{aligned} Max : \quad & \sum_{i \in \mathcal{K}} \big[2U_i(t)\Phi_i(cost, S_i(t)) - V cost_i\big] \\ s.t. : \quad & cost \in \mathcal{C}. \end{aligned} \quad (47)$$

Then the AP polls all the users with $cost_i > 0$ and coordinates them to send their data according to the chosen rates $\Phi_i(cost, S_i(t))$. Idle fill is used if needed.

Notice even though the multi-channel problem is more complicated than the single channel one, (46) retains the simplicity of (22): maximize a function of one variable over an interval. The difference is that now the objective function is a sum of several functions. The complexity of solving (47) depends largely on the structure of $\mathcal{C}$. But in the special case when only one channel can be active in a slot, (47) can be simply solved by computing the optimization metric for each user assuming it is the only user.

C. MPTSA Performance

The performance results of MPTSA are summarized in the following theorem.

***Theorem** 5:* The MPTSA algorithm stabilizes all queues and achieves the following (Assuming $U_i(0) = 0$ for all i):

$$\begin{aligned} \sum_{i \in \mathcal{K}} U_i(t) \quad &\leq \quad U_{max} = V p_{max} r_m^2 K/2 + K R_{max}, \quad (48) \\ Profit_{av} \quad &\geq \quad Profit_{av}^{opt} - O(\log(V)/V). \quad (49) \end{aligned}$$

Proof: We first prove (48) by induction:

(a) Suppose at time t, we have $\sum_{i \in \mathcal{K}} U_i(t) \leq V p_{max} r_m^2 K/2$. Then by definition, we have $\sum_{i \in \mathcal{K}} U_i(t+1) \leq V p_{max} r_m^2 K/2 + K R_{max}$.

(b) Now suppose we have $\sum_{i \in \mathcal{K}} U_i(t) > V p_{max} r_m^2 K/2$. For any feasible price p, denote the set of users having $2U_i(t) - Vp > 0$ as $\mathcal{I}_p$. We first have the following for all $0 \leq p \leq p_{max}$ from the fact that $r_m \geq 1$:

$$\begin{aligned} V p r_m^2 (K - |\mathcal{I}_p|) + V p |\mathcal{I}_p| \quad &< \quad 2 \sum_{i \in \mathcal{K}} U_i(t) \\ &\leq \quad \sum_{j \in \mathcal{I}_p} 2U_j(t) + 2r_m^2 \sum_{i \notin \mathcal{I}_p} U_i(t), \end{aligned}$$

here $|\mathcal{I}_p|$ is the cardinality of $\mathcal{I}_p$. Rearrange terms, we have

$$r_m^2 \sum_{i \notin \mathcal{I}_p} [Vp - 2U_i(t)] < \sum_{j \in \mathcal{I}_p} [2U_j(t) - Vp]. \quad (50)$$

Now choose an $F_i(p)$ to be such that $F_i(p) > 0$. If there is no such i, we see that $F_i(t) = 0$ for all i, thus no new packets will be admitted. Without loss of generality, let $F_1(p) > 0$. From (50) we have:

$$r_m \sum_{i \notin \mathcal{I}_p} F_1(p)[Vp - 2U_i(t)] < \frac{1}{r_m} \sum_{j \in \mathcal{I}_p} F_1(p)[2U_j(t) - Vp].$$

This means:

$$\sum_{i \notin \mathcal{I}_p} F_i(p)[Vp - 2U_i(t)] < \sum_{j \in \mathcal{I}_p} F_j(p)[2U_j(t) - Vp], \quad (51)$$

and thus $\sum_{i \in \mathcal{K}} F_i(p)[Vp - 2U_i(t)] < 0$. Therefore no new data will be accepted.

Now we prove (49). Similar as in the proof of Theorem 4, we define $g^*(t)$ to be the right hand side of (45) when an alternate policy is used:

$$\begin{aligned} g^*(t) \triangleq B - \mathbb{E}\big\{2 \sum_{i \in \mathcal{K}} U_i^{MP}(t)[\mu_i^*(t) - Z^*(t)R_i^*(t)]\big\} \\ -V\mathbb{E}\big\{\sum_{i \in \mathcal{K}} Z^*(t)p^*(t)R_i^*(t) - \sum_{i \in \mathcal{K}} cost_i^*(t)\big\}, \end{aligned}$$

where the decisions $Z^*(t)$, $p^*(t)$ and $cost_i^*(t)$ correspond to the feasible control action that can be implemented on slot t of the alternate policy and $U_i^{MP}(t)$ denotes the backlog of user i when the MPTSA algorithm is implemented up to time t. Now plug in the optimal stationary randomized policy and use the same approach as in the proof of Theorem 4, we obtain:

$$\begin{aligned} g^*(t) \quad &\leq \quad B + 2Kd(\mu_{max}^2 + R_{max}^2) \\ &\quad -2\mathbb{E}\big\{\sum_{i \in \mathcal{K}} U_i^{MP}(t-d)[\mu_i^*(t) - Z^*(t)R_i^*(t)]\big\} \\ &\quad -V\mathbb{E}\big\{\sum_{i \in \mathcal{K}} Z^*(t)p^*(t)R_i^*(t) - \sum_{i \in \mathcal{K}} cost_i^*(t)\big\} \\ &\leq \quad B_1 + 2\mathbb{E}\big\{\sum_{i \in \mathcal{K}} U_i^{MP}(t-d)\theta_i \gamma_i^d\big\} \\ &\quad -V\mathbb{E}\big\{\sum_{i \in \mathcal{K}} Z^*(t)p^*(t)R_i^*(t) - \sum_{i \in \mathcal{K}} cost_i^*(t)\big\}, \end{aligned}$$

where $B_1 = B + 2Kd(\mu_{max}^2 + R_{max}^2)$ and $\theta_i > 0$, $0 < \gamma_i < 1$ are constants depend on the Markov Chain processes $S_i(t)$. Since $U_i(t) \leq U_{max}$ for all i and t, by choosing $d \triangleq \left\lceil \max_{i \in \mathcal{K}} \left\{ \frac{\log(2\theta_i U_{max})}{\log(1/\gamma_i)} \right\} \right\rceil = O(\log(V))$, we obtain:

$$g^*(t) \leq B_1 + K - V\mathbb{E}\big\{\sum_{i \in \mathcal{K}} Z^*(t)p^*(t)R_i^*(t) - \sum_{i \in \mathcal{K}} cost_i^*(t)\big\}.$$

The rest follows exactly as in the proof of Theorem 4. ∎

VII. Simulation

In this section, we provide simulation results for the PTSA algorithm. For simplicity, we only simulate the single channel system. We compare two types of arrival processes. In the first case, the arrival $R(t)$ is deterministic and is exactly equal to $F(M(t), p(t))$. In the other case, we assume that $R(t)$ is a Bernoulli random variable, i.e., $R(t) = 2F(M(t), p(t))$ or $R(t) = 0$ with equal probabilities. [11]

Now we provide our simulation results. We assume $\mathcal{M} = \{Low, High\}$, $\mathcal{S} = \{Good, Bad\}$. The demand curve for $M(t) = Low$ is given by:

$$F(Low, p) = \begin{cases} 4 & 0 \leq p \leq 1, \\ -6p + 10 & 1 < p \leq \frac{3}{2}, \\ -\frac{2}{17}p + \frac{20}{17} & \frac{3}{2} < p \leq 10. \end{cases} \tag{52}$$

The demand curve for $M(t) = High$ is given by:

$$F(High, p) = \begin{cases} 10 - p & 0 \leq p \leq 2, \\ -6p + 20 & 2 < p \leq 3, \\ -\frac{1}{7}p + \frac{17}{7} & 3 < p \leq 10. \end{cases} \tag{53}$$

The rate-cost curve is given by :

$$\Phi(cost(t), S(t)) = log(1 + \gamma_{S(t)} cost(t)), \tag{54}$$

where $0 \leq cost \leq 10$, $\gamma_{S(t)} = 2$ if $S(t) = Good$ and $\gamma_{S(t)} = 1$ else. Both the demand state and the channel are assumed to vary according to a two-state Markov Chain with transition probabilities to the other state being both 0.4. Fig. 7 shows the backlog and profit performance of PTSA under this dynamic setting. We see that the profit converges quickly to the optimum value and the backlog is no larger than the worst case bound. We now also look at the prices chosen by PTSA: We see in Fig. 8 that in fact PTSA quickly determines the optimum prices for each state, and consequently determines the optimum traffic share of the two different demand states.

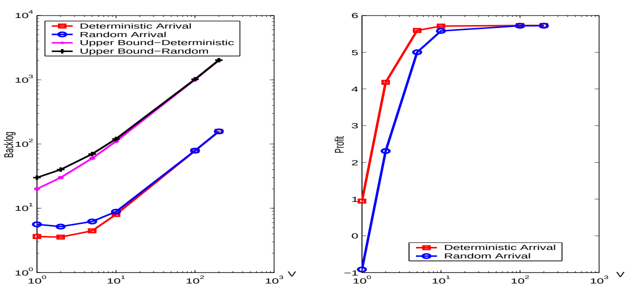

Fig. 7. Average Backlog and Average Profit v.s. V

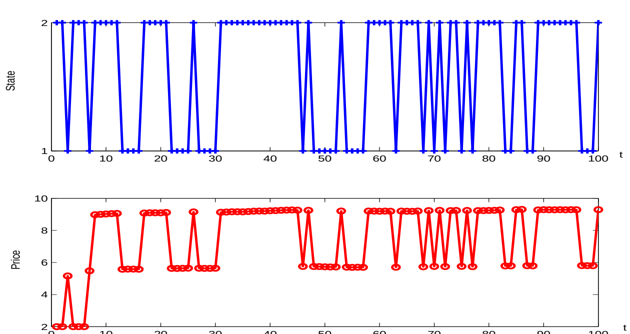

Fig. 8. Prices chosen according to demand state $M(t)$ for V=100

In this case, we see that for each demand state, only one price is chosen. This is different from the simulation results in [21]. In [21], we simulated an example similar to the one of Fig. 2 and indeed observed a two-price phenomenon.

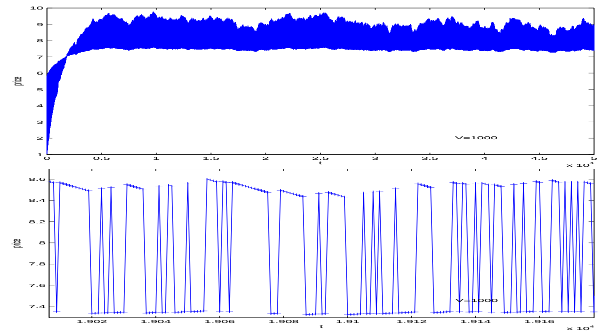

Fig. 9. Upper: prices chosen in the first 5×10^4 slots; Lower: prices chosen over an interval of 200 slots (V=1000)

However, it is not always the case that prices chosen by PTSA converge to a two-price process. Fig. 9 shows the result for the following setting: $\mathcal{M} = \{Low, High\}$, state transition probabilities being both 0.4, service rate $\mu = 2.28$ with no cost, and demand functions are $F(High, p) = 10 - \frac{1}{10}p^2$ and $F(Low, p) = \frac{10}{1+p^2}$, both with $0 \leq p \leq 10$. We see that prices fluctuate within some interval over the whole period of time.

[11] For simplicity here, we assume $R(t)$ can take fractional values. Alternatively, we could restrict packet sizes to integral units and make the probabilities be such that $\mathbb{E}\{R(t) \mid p(t), M(t)\} = F(M(t), p(t))$.

REFERENCES

[1] I. Ch. Paschalidis and J. N. Tsitsiklis. Congestion-dependent pricing of network services. *IEEE/ACM Trans. Netw.*, 8(2):171–184, 2000.

[2] X. Lin and N. B. Shroff. Simplification of network dynamics in large systems. *IEEE/ACM Transactions on Networking*, vol. 13, no. 4, pp. 813-826, August 2005.

[3] D. Acemoglu, A. Ozdaglar, and R. Srikant. The marginal user principle for resource allocation in wireless networks. *Proc. of IEEE Conf. on Decision and Control, Dec. 2004.*

[4] P. Marbach and R. Berry. Downlink resource allocation and pricing for wireless networks. *Proceedings of IEEE INFOCOM*, 2002.

[5] T. Basar and R. Srikant. Revenue-maximizing pricing and capacity expansion in a many-users regime. *Proceedings of IEEE INFOCOM*, 2002.

[6] J. Musacchio and J. Walrand. Wifi access point pricing as a dynamic game. *IEEE/ACM Trans. Netw.*, 14(2):289–301, 2006.

[7] R. K. Lam, D. Chiu, and J. C. S. Lui. On the access pricing and network scaling issues of wireless mesh networks. *IEEE Transactions on Computers : Accepted for future publication,Digital Object Identifier: 10.1109/TC.2007.70753*, 2007.

[8] S. Shakkottai and R. Srikant. Economics of network pricing with multiple isps. *IEEE/ACM Trans. Netw.*, 14(6):1233–1245, 2006.

[9] E. J. Friedman and D. C. Parkes. Pricing wifi at starbucks– issues in online mechanism design. *In Proc. Fourth ACM Conf. on Elec. Commerce (EC'03), pp.240-241*, 2003.

[10] N. J. Keon and G. Anandalingam. Optimal pricing for multiple services in telecommunications networks offering quality-of-service guarantees. *IEEE/ACM Trans. Netw.*, 11(1):66–80, 2003.

[11] Z. Zhang, D. Dey, and Y. Tan. Pricing communication services with delay guarantee. *INFORMS Journal on Computing, Vol.19, No.2, pp248-260*, 2007.

[12] J. K. MacKie-Mason and H. R. Varian. Pricing congestible network resources. *IEEE Journal on Selected Areas in Communications*, vol. 13, no. 7, pp. 1141–1149, 1995.

[13] F. Kelly. Charging and rate control for elastic traffic. *European Transactions on Telecommunications*, vol. 8, pp. 33-37, 1997.

[14] S. H. Low and D. E. Lapsley. Optimization flow control, i: Basic algorithm and convergence. *IEEE/ACM Transactions on Networking*, vol. 7(6): 861-75, Dec. 1999.

[15] M. J. Neely. Optimal pricing in a free market wireless network. *Proc. of IEEE INFOCOM, May 2007.*

[16] M. J. Neely, E. Modiano, and C. Li. Fairness and optimal stochastic control for heterogeneous networks. *IEEE INFOCOM Proceedings*, March 2005.
[17] M. J. Neely. Energy optimal control for time-varying wireless networks. *IEEE Transactions on Information Theory 52(7): 2915-2934*, July 2006.
[18] L. Georgiadis, M. J. Neely, and L. Tassiulas. *Resource Allocation and Cross-Layer Control in Wireless Networks*. Foundations and Trends in Networking Vol. 1, no. 1, pp. 1-144, 2006.
[19] M. J. Neely, E. Modiano, and C. E. Rohrs. Dynamic power allocation and routing for time-varying wireless networks. *IEEE Journal on Selected Areas in Communications, Vol 23, NO.1*, January 2005.
[20] M. J. Neely. *Dynamic Power Allocation and Routing for Satellite and Wireless Networks with Time Varying Channels*. PhD thesis, Massachusetts Institute of Technology, Laboratory for Information and Decision Systems (LIDS), 2003.
[21] L. Huang and M. J. Neely. The optimality of two prices: Maximizing revenue in a stochastic network. *Proc. of 45th Annual Allerton Conference on Communication, Control, and Computing (invited paper)*, Sept. 2007.
[22] D. P. Bertsekas, A. Nedic, and A. E. Ozdaglar. *Convex Analysis and Optimization*. Boston: Athena Scientific, 2003.
[23] O. Hanner and H. Rådström. A generalization of a theorem of Fenchel. *Proc. American Mathematical Society*, vol. 2, no. 4, pp. 589-593, Aug. 1951.
[24] J. R. Munkres. *Topology*. NJ: Prentice Hall, Inc., 2000.
[25] R. Preston McAfee. Coarse matching. *Econometrica, Vol. 70, Number 5, 2025-2034*, September 2002.
[26] G. R. Bitran and R. Caldentey. An overview of pricing models for revenue management. *Manufacturing and Service Operations Management*, Vol. 5, No. 3, pp. 203-209, Summer 2003.
[27] M. A. Crew, C. S. Fernando, and P. R. Kleindorfer. The theory of peak-load pricing: A survey. *Journal of Regulatory Economics*, 8(3):215–48, November 1995.
[28] E. Altman. *Constrained Markov Decision Processes*. Chapman and Hall/CRC, 1999.
[29] E. K. Browning and M. A. Zupan. *Microeconomic theory and applications*. Addison-Wesley Educational Publishers, Inc., 1999.
[30] S. Ross. *Stochastic Processes*. John Wiley and Sons Inc., New York, 1996.