Downlink Resource Allocation and Pricing for Wireless Networks

Peter Marbach and Randall Berry

***Abstract*—This paper considers resource allocation and pricing for the downlink of a wireless network. We describe a model that applies to either a time-slotted system (e.g. Qualcomm's HDR proposal) or a CDMA system; the main feature of this model is that the channel quality varies across the users. We study using a pricing scheme for the allocation of radio resources. We show that to maximize revenue in such a system, the base station should allocate resources in a discriminatory manner, where different users are charged different prices based in part on their channel quality. However optimally allocating resources in this way is shown to require knowledge about each user's utility function. We consider a suboptimal scheme which does not require knowledge of the users' utility functions, and show that this scheme is asymptotically optimal, in the limit of large demand. Moreover such a scheme is shown to maximize social welfare. We also consider a heuristic scheme for the case of small demand, which does not require perfect knowledge about the users' utility functions. We provide numerical results that illustrate the performance of this heuristic.**

I. Introduction

In this paper we study the use of pricing for resource allocation in a wireless network. We focus on a downlink communication to a group of untethered users from a single wired network access point, such as a base station in a cellular network or a hub in a wireless LAN. The following is restricted to the downlink for two reasons. First, it is expected that for many data applications, such as down-loading web pages or multimedia, the traffic load on the downlink will be much larger than over the uplink, thus efficiently utilizing this link is of more importance. The second reason is that the resource allocation for the uplink involves several fundamentally different issues than for the downlink, *e.g.* on the downlink all communication in a single cell originates at one point so issues of coordination and interference are of lesser importance than on the uplink.

A basic feature of a wireless network is that channel characteristics will vary across the user population. This variation is due to differences in the proximity of a user to the transmitter as well as multi-path fading and shadowing effects. In this setting, the transmitter can exploit knowledge of the channel quality to send at higher rates to users with better channels, for example by using different modulation and coding choices for different users. Channel knowledge can also be used for allocating "radio resources" such as time-slots, bandwidth or transmission power, among the users. One example of such a system is the Qualcomm High Data Rate (HDR) scheme [1], which is the basis for the IS-856 wireless data standard [3]. In HDR, the base station transmits to a single user during any given time slot; the transmission rate a user can receive in a time slot is determined by estimates of the channel conditions which are fed back from the mobile to the base station. A scheduling algorithm at the base station is used to decide which user is transmitted to during any given time slot, based in part on the channel conditions of the users. In the following, we consider a model which incorporates several features of such a system, and we study the role of pricing on resource allocation in this model.

The use pricing as a means for allocating resources in communication networks has received much attention in recent years. For wire-line networks, a sampling of this literature includes the work in [7], [11], [10], and [12]. A number of authors have also studied pricing in wireless networks, in particular the role of pricing to aid in implementing distributed power control for the uplink in a CDMA setting; examples of this work include [4], [13], [6], and [5]. Regarding related work on the downlink of wireless networks, we mention the work in [9], [17] in which pricing for the downlink of a CDMA network is studied; these papers address both welfare and revenue maximization. The emphasis in [9], [17] is on characterizing optimal prices and the resulting resource allocation given knowledge of the users' utility functions. We also mention work by Borst and Whiting [2] on scheduling for HDR systems; this work shows that a form of "revenue-based" scheduling is optimal for maximizing the normalized long-term expected throughput. Algorithms for adaptively calculating the optimal revenue vector are also given in [2].

Our focus here is on pricing strategies for revenue maximization, *i.e.* how should a service provider price resources to maximize revenue. One key difference between this work and much prior work is that we do not restrict the base station to charge the same price per resource to all users. Indeed we show that to maximize revenue it is often advantageous to employ discriminatory pricing; that is to charge users different prices based on the users' preferences and channel conditions. It is a well known fact in microeconomics that price discrimination can allow a firm with some degree of monopoly power to increase revenue. In the wireless setting we show that the optimal price charged per user depends on both the user's demand for bandwidth as well as the channel conditions.

Another difference with much of the literature is that we consider a pricing framework that is receiver driven, *i.e.* we view the receiver, rather than the sender, as paying for service. For the applications mentioned above, such as down-loading web

P. Marbach is with the Dept. of Computer Science, University of Toronto, 10 King's College Rd., Toronto, Ontario, M5S 1A4, Canada. E-mail: marbach@cs.toronto.edu

R. Berry is with the Dept. of Electrical and Computer Engineering, Northwestern University, 2145 Sheridan Rd., Evanston, IL 60208. E-mail: rberry@ece.nwu.edu. His work was supported in part by the Northwestern-Motorola Center for Communications

pages, this seems to be the natural approach. We note that a receiver driven approach precludes a straightforward extension of pricing techniques as in [11] and [12], in which a price is identified with each packet at the transmitter. Also, since we allow the receiver to charge different prices to various users, this complicates a marking approach as in [8]. Instead of the above approaches, we introduce a pricing scheme, where users submit bids and resources are allocated based on these bids.

The paper is organized as follows. In Section II we describe our model for both a time-slotted system and a CDMA system. Next, we look at some preliminary approaches for resource allocation. In Section IV, we introduce a utility framework to explicitly model a user's preferences. With an appropriate choice of variables, the problem of maximizing social welfare is shown to be mathematically equivalent to the problem studied by Kelly *et al.* in [7]. In Section V we introduce a pricing scheme for allocating radio resources as discussed above. In Section VI an optimal revenue maximizing strategy is given for this pricing scheme; this strategy depends on the base station having perfect knowledge of the user's utilities. In Section VII we study suboptimal strategies that do not require complete knowledge of the user's utilities. In particular, we give an allocation strategy which requires no knowledge about the utilities, and show that this strategy is asymptotically optimal in the regime of large demand. Furthermore such a strategy also maximizes social welfare. Finally in Section VIII we provide some numerical results to illustrate these ideas.

II. Model

In this section we describe a idealized model for the downlink in a single cell of a wireless network; we choose a simple model which highlights the possible disparity between users in such an environment. We first describe this model in the context of a time-slotted system. The resulting model is shown to also apply to a CDMA system similar to that studied in [9].

Time slotted system: Consider a system where the base station transmits to a single user at any any given time. Assume there are M active users, and for $i = 1, \ldots, M$ the channel between the base station and user i is parameterized by a "state," h_i. Assume that this state is known exactly at the base station. Based on this state, let $r(h_i)$ be the rate in packets per second at which the user can receive data in a time slot with acceptable reliability.[1] We assume that over the time scales of interest, the channel state of each user stays fixed; we note this is different than the assumption in work such as [2] and in particular prohibits us from exploiting "multi-user diversity" due to time variations in the users' channels [15]. Suppose that the base station schedules transmissions over a fixed length time frame of T seconds. During a frame, the base station allocates for each user i a time $t_i \in [0, T]$ during which it transmits to that user at rate $r(h_i)$, where

$$\sum_{i=1}^{M} t_i = \sum_{i=1}^{M} \frac{x_i}{r(h_i)} \leq T. \tag{1}$$

[1] In this paper we ignore the possibility of packet losses; one way of incorporating such losses is by assuming that $r(h_i)$ is the long-term average rate at which packets can be successfully sent.

Here x_i is the throughput (in number of packets) of user i. Note we allow the transmission time given to each user to be of an arbitrary length; in practical systems the transmission time would likely be constrained to be an integer multiple of some elementary time-slot, but we ignore this constraint here.

CDMA system: Consider the downlink in a CDMA system, where users are assigned orthogonal spreading codes.[2] In this case, assume that the base station transmits to each user during the entire time frame of T seconds. Let P_i be the transmission power allocated to user i, and assume that the base station has a constraint on its total transmission power so that

$$\sum_{i=1}^{M} P_i \leq P.$$

Let g_i represent the channel gain of user i, so that $g_i P_i$ is the received power at user i. Assume that the base station can send to user i with a rate, r_i, (in packets/sec.) that is proportional[3] to the received power, *i.e.* $r_i = K g_i P_i$, for some constant K. User i will then receive $x_i = r_i T$ packets in the time frame; thus we have that $x_1, \ldots, x_M$ are constrained so that

$$\sum_{i=1}^{M} \frac{x_i}{TKg_i} \leq P. \tag{2}$$

Clearly by identifying $r(h_i)$ with TKg_i and T with P these models are mathematically equivalent. We will describe the following results in the context of the model in (1), but the reader should keep in mind that they apply equally for the second model as well, or any other other situation where a user's transmission rate depends on some resource in a linear fashion and there is a linear constraint on this resource.

In a wire-line network, a similar "link" model can be developed (see *e.g.* [12]) where the throughputs, $\{x_i\}$, are constrained as

$$\sum_{i=1}^{M} x_i \leq CT, \tag{3}$$

and C is the constant transmission rate in packets/sec of the link. Thus the key feature that differentiates the above model from the wire-line case is that in the wireless model, the packets of different users can require a different portion of the "link" resources (depending on the values of $r(h_i)$, and g_i, respectively), while in the wire-line case each user's packets need the same amount of network resources.

III. Maximizing Throughput/Revenue

Suppose that the base station has an infinite supply of data to send to each user, and consider the problem of allocating resources among the users to maximize the total system throughput. For the time-slotted model, this corresponds to choosing a

[2] We ignore any constraints on the available number of spreading codes in the following. Such considerations can be taken into account through a code constraint as in [9].

[3] A linear relationship between rate and power is reasonable for a wide-band system, provided the base station cannot send at too high of a rate. At a high enough rates, channel capacity considerations tell us that this relationship will no longer hold.

transmission time allocation $(t_1, ..., t_M)$ to maximize the total throughput of the system. To achieve this, clearly the base station should always allocate the entire time frame to the user, or users, with the best channel, *i.e.* the user(s) with the maximal transmission rate $r(h_i)$. We note that this is similar in spirit to the conclusion that have been drawn in several other settings, such as [16]. Although the resulting allocation achieves the maximal throughput, this policy is extreme and unfair as most of the users will not be able to receive any data, in particular in the case where the channel states do not change over the time scales of interest. Moreover it does not account for the relative preferences (service requirements) of each user.

Suppose instead that receivers (users) agree to pay the base station (service provider) a price u_i, $i = 1, ..., M$, per received packet. In this scenario, the base station maybe interested in maximizing its revenue, rather than the throughput. The corresponding maximization problem is given as follows

$$\begin{aligned} &\text{maximize} \sum_{i=1}^{M} x_i u_i \\ &\text{subject to:} \sum_{i=1}^{M} \frac{x_i}{r(h_i)} \leq T, \\ &\qquad x_i \geq 0, \qquad i = 1, ..., M. \end{aligned}$$

In order to maximize the revenue, the entire time slot should be allocated to the user, or users, with the highest index $n_i = u_i r(h_i)$. Note that n_i can be interpreted as the revenue per unit time the base station receives when it transmits packets of user i. It is interesting to note that in the case where the base station wants to maximize the revenue, time slots are not necessarily allocated to the users who pay the highest price per packet or the user with the best channel, but to the users who pay the largest amount per unit time. This means that users with a bad channel (small $r(h_i)$) are still able to receive data when they are willing to pay a higher price u_i per packet. In the case where one user has a higher index n_i than all other users, this revenue maximizing policy is again extreme and unfair as the whole time frame gets allocated to only one user, and all others are starved. In the case where the indices of all users are equal, any feasible schedule such that $\sum_{i=1}^{M} t_i = T$ generates the same revenue. Notice that the throughput optimization problem can be thought of as a special case of this revenue maximization problem where all users pay the same amount u per received packet.

IV. MAXIMIZING SOCIAL WELFARE

The above allocation schemes are optimal in the sense that they maximize either the throughput or the revenue. However, as we noted, the resulting allocations are not balanced, *i.e.* most users are starved, except possibly if the indices of all the users are equal. Also, we did not take into account how the prices u_i are generated for the revenue maximization problem. In this section, we introduce utility functions as a means of explicitly characterizing the service preferences of a user. We then explore an approach, where rather than optimizing the throughput, or revenue, the base station maximizes the so called social welfare, or the total utility over all users.

We consider users with elastic traffic [14], *i.e.* users who perceive quality of service solely as a function of the throughput. For this case, we can characterize the service preferences of each user through a quality indicator or utility function that depends on a single variable: the throughput. We will assume that these utility functions are increasing in the throughput. Examples of users with elastic traffic are users sending email, transferring files, and browsing the Web.

Consider a fixed user $i \in \{1, ..., M\}$ and suppose that x_i is the throughput of user i. Then, we associate with user i the utility function $U_i(x_i)$. We make the following assumption.

Assumption 1: For each user $i = 1, ..., M$, the function $U_i : \Re_+ \mapsto \Re_+$ satisfies the following conditions:

a. U_i is increasing, strictly concave and twice differentiable, with $U_i(0) = 0$.
b. There exists constants, K_1, K_2, such that for all i, x, $U_i'(x) \leq K_1$ and $|U_i''(x)| \leq K_2$.

Utility functions for elastic traffic with these characteristics are commonly used in the pricing literature, see *e.g.* [7]. The most restrictive assumption here is condition b, which states that the first and second derivate of the utilities are uniformly bounded. This assumption will be used in Section VII. Notice that the first derivative of U_i is bounded if and only if $U_i'(0) < \infty$. Also notice that Assumption 1 does not require that all users have the same utility function.

We define the demand function $D_i : \Re_+ \mapsto \Re_+$ of user i as follows. For $i = 1, ..., M$, let the function D_i be defined such that $D_i(u)$ is the optimal solution, x^*, to the maximization problem

$$\max_{x \geq 0} \{U_i(x) - xu\}, \quad u \in \Re_+.$$

The goal of this maximization problem is to optimize the user's net benefit given by utility minus cost. Note that in this problem the user faces a trade-off between achieving a high utility (by choosing a large throughput x) and keeping its cost low (by choosing a small throughput x). Under Assumption 1, the above maximization problem has a unique finite solution for all u, so $D_i(u)$ is well defined. $D_i(u)$ can be interpreted as the rate (in packets per second) i would request when the price for receiving one packet is equal to u. Furthermore, from Assumption 1.b it follows that, for each i, there exists some constant $u_{i,max}$ such that

$$D_i(u) = 0, \quad \forall u \geq u_{i,max}.$$

In other words, when user i is charged more than $u_{i,max}$, it will request zero rate.

Assume now that the service provider wants to optimize the total users' utility. This objective is captured by the following optimization problem.

$$\begin{aligned} &\text{maximize} \sum_{i=1}^{M} U_i(x_i) \\ &\text{subject to:} \sum_{i=1}^{M} \frac{x_i}{r(h_i)} \leq T, \\ &\qquad x_i \geq 0, \qquad i = 1, ..., M. \end{aligned} \tag{4}$$

Let us compare this to a single link case of the utility maximization problem studied by Kelly in [7] for a wire-line network. In that case, the constraint in (1) is replaced by (3) resulting in the optimization problem

$$\begin{aligned} \text{maximize} \sum_{i=1}^{M} U_i(x_i) \\ \text{subject to:} \sum_{i=1}^{M} x_i \leq CT, \\ x_i \geq 0, \qquad i = 1, ..., M. \end{aligned} \tag{5}$$

For each i, define $\hat{U}(t_i) = U(t_i r(h_i))$; the quantity $\hat{U}(t_i)$ can be interpreted as an indicator of utility as a function of the amount of time allocated to a user. Note if two users i and j have identical utilities, $U_i(x_i) = U_j(x_j)$, but $r(h_i) \neq r(h_j)$ then $\hat{U}_i(x_i) \neq \hat{U}_j(x_j)$, *i.e.* they will have different utilities as a function of t_i. Using this notation, the optimization in (4) can be re-written as

$$\begin{aligned} \text{maximize} \sum_{i=1}^{M} \hat{U}_i(t_i) \\ \text{subject to:} \sum_{i=1}^{M} t_i \leq T, \\ t_i \geq 0, \qquad i = 1, ..., M. \end{aligned} \tag{6}$$

In this form, this problem can be seen to be mathematically identical to the problem in (5). Thus the results in [7] can be adapted in the current setting as well. In particular, from Theorem 1 in [7] and the above identification, the following proposition directly follows.

Proposition 1: Let Assumption 1 hold. Then the above maximization has an unique solution $(\hat{x}_1, ..., \hat{x}_M)$. In addition, there exists a parameter $\hat{\lambda}$ such that

$$\hat{x}_i = D_i(u_i), \qquad i = 1, ..., M,$$

where

$$u_i = \frac{\hat{\lambda}}{r(h_i)}.$$

The parameter $\hat{\lambda}$ can be interpreted as the (optimal) Lagrange multiplier for the unconstrained optimization problem:

$$\text{maximize} \sum_{i=1}^{M} U_i(x_i) - \hat{\lambda} \left(\sum_{i=1}^{M} \frac{x_i}{r(h_i)} - T \right).$$

The first order conditions for the above optimization problem are then given by

$$U_i'(x_i) - \frac{\hat{\lambda}}{r(h_i)} = U_i'(x_i) - u_i = 0, \qquad i = 1, ..., M.$$

It can be shown that $\hat{\lambda}$ increases when the total demand

$$D(u) = \sum_{i=1}^{M} D_i(u), \qquad u \in \Re_+,$$

increases. The parameter $\hat{\lambda}$ can then also be thought of as a congestion price where each user is charged the same price per time slot in the optimal solution (but different prices, u_i, per packet); this price increases as the demand increases. We also point out that making the equivalent identification for the CDMA model, results in users being charged an equal price per unit of power under the optimal solution.

The above procedure for allocating resources may not be practical for two reasons.

(a) First, it is not clear why the base station should be interested in optimizing the total user utility, rather than the throughput or the revenue.
(b) Even when the base station wants to optimize the total user utility, it may be unrealistic to assume that the base station knows the utility functions of the individual users to carry out the above optimization directly.

In [7], Kelly proposes a pricing mechanism for wire-line packet networks that allows the network to optimize social welfare without requiring the network to have any knowledge about the users' utility functions. In this approach, users bid for resources by indicating a willingness to pay, and the network allocates network resources accordingly. Here, we pursue a similar approach to allocate resources for the downlink of a wireless network. However, rather than considering welfare maximization, we assume that the service provider's objective is revenue maximization.

V. Resource Allocation through Pricing

In this section, we consider a receiver driven pricing mechanism for allocating the transmission times t_i, $i = 1, ..., M$. As discussed in the introduction, this scheme requires users (the receivers) to compete for resources through a bidding mechanism. Specifically, users submit a price bid and the base station allocates transmission times based in part on these bids. In each frame, users pay a price that is equal to their bid, independent of the amount of data they receive, *i.e.* the price is a price per frame, as opposed to a price per packet. For our analysis we assume that users behave in a selfish way (*i.e.* each user is only interested in maximizing its own net benefit) and users act independently (*i.e.* they do not collaborate during the bidding process). User adjust their bids over time, based on the allocation they receive. The goal of the base station is to allocate resources in such way that the resulting bids maximize revenue.

In Section III, we considered optimizing revenue given fixed prices per packet. The mechanism considered here differs from this in several ways. First, the users submit a price per frame (of length T) versus a price per packet, and second, the base station takes into consideration how the users bids change over time. For this framework, which we describe in more detail below, we investigate the following questions,

(a) How will the users bid?
(b) Based on the users' bid, how should the base station allocate the transmission time in order to maximize its revenue?

A. Pricing Mechanism

We consider the following pricing mechanism. In each frame, users bid for resources by submitting price bid $w_i \in \Re_+$. The base station then allocates resources according to a function $f(w) = (f_1(w), ..., f_M(w))$ such that

$$\sum_{i=1}^{M} f_i(w) \le T, \qquad \text{for all } w \in \Re_+^M,$$

where $t_i = f_i(w)$ is the time that the base station allocates in each frame to transmit packets for user i. Notice that we do not require the the entire frame to be allocated. Indeed, some examples can be found where the revenue maximizing allocation does not allocate the entire frame.

To study this pricing scheme, we proceed in a similar manner as in [7] and define a user problem and a base station problem. The user problem addresses the issue of how user i chooses the bid w_i (based on the last price bid and the time $t_i = f_i(w)$ in the last frame). For the base station problem, the goal is to find an allocation strategy f that maximizes the revenue.

B. User Problem

We consider a fixed allocation strategy $f = (f_1, ..., f_M)$. Let $w^{(k)} = (w_1^{(k)}, ..., w_M^{(k)})$ be the bid vector in the kth frame, and $f(w^{(k)}) = (f_1(w^{(k)}), ..., f_M(w^{(k)}))$ the transmission times allocated to users in this frame. The price per packet, $u_i^{(k)}$, that user i pays in the kth frame is then equal to

$$u_i^{(k)} = \frac{w_i^{(k)}}{f_i^{(k)}(w) r(h_i)},$$

and the transmission rate $x_i^{(k)}$ is equal to

$$x_i^{(k)} = \frac{w_i^{(k)}}{u_i^{(k)}}.$$

In the above, if $f_i^{(k)}(w) = 0$, we set $u_i^{(k)} = \infty$ and $x_i^{(k)} = 0$, independent of the other variables. We assume that in the $(k+1)$th frame, user i chooses a bid $w_i^{(k+1)}$ to maximize its net benefit under the price $u_i^{(k)}$, *i.e.* user i solves the following maximization problem,

$$\max_{w_i \ge 0} \left\{ U_i \left(\frac{w_i}{u_i^{(k)}} \right) - w_i \right\}.$$

We note that this problem is equivalent to the user problem from the second decomposition in [7].

Given the allocation strategy f, we then define an equilibrium bid vector as a vector $w^*(f) = (w_1^*(f), ..., w_M^*(f))$ such that for all users $i = 1, ..., M$ we have

$$w_i^*(f) = \arg\max_{w_i \ge 0} \left\{ U_i \left(\frac{w_i}{u_i} \right) - w_i \right\},$$

where

$$u_i = \frac{w_i^*(f)}{f_i(w^*) r(h_i)}.$$

Note that under an equilibrium vector $w^*(f) = (w_1^*(f), ..., w_M^*(f))$, each user i maximizes its own net benefit and therefore has no incentive to deviate from its bid $w_i^*(f)$. The following proposition gives a useful alternative characterization of an equilibrium bid vector.

Proposition 2: A bid vector w^* is an equilibrium bid vector for an allocation strategy f if and only if,

$$f_i(w^*) = \frac{D_i(u_i)}{r(h_i)}, \quad i = 1, \ldots, M,$$

where $u_i = \frac{w_i^*}{f_i(w^*) r(h_i)}$.

The proof of this follows directly from the definitions above.

C. Base Station Problem

Next, we consider the following question. What allocation strategy should the base station choose such that it maximizes the revenue? We will proceed in two steps. First, we assume that the base station has perfect global knowledge (knows the utility function of all users), and derive an allocation f^* whose equilibrium bidding vector maximizes the base station's revenue. In the second step, we derive an allocation strategy $\hat{f}$ for the case where the base station has only imperfect information (does not know the users' utility functions). One would expect that the revenue $P(f^*)$ under the allocation strategy f^* is in general larger than the revenue $P(\hat{f})$ under the allocation strategy $\hat{f}$. However, as we will show in the following, the revenue under the allocation $\hat{f}$ is close to f^* when many users are active.

VI. Optimal Allocation Strategy

In this section, we derive an optimal allocation strategy for the case where the base station knows each user's utility function. Let U_i be the utility function of user i, and let $D_i(u)$ be the corresponding demand (as defined in Section IV). The revenue maximization problem is then given by

$$\begin{aligned} \text{maximize } & \sum_{i=1}^{M} u_i D_i(u_i) \\ \text{subject to: } & \sum_{i=1}^{M} \frac{D_i(u_i)}{r(h_i)} \le T, \\ & u_i \ge 0, \quad i = 1, ...M. \end{aligned} \tag{7}$$

It can be shown that an optimal solution $(u_1^*, ..., u_M^*)$ exists for this maximization problem. If for each i, $u_i D_i(u_i)$ is strictly convex, then this optimum will be unique. Otherwise, there may be multiple optimal solutions, in which case, consider $(u_1^*, ..., u_M^*)$ to be one optimal solution, picked arbitrarily. From the point of view of revenue maximization, which optimal is chosen does not matter.

Let

$$\lambda_i^* = u_i^* r(h_i)$$

be the price per unit time the base station charges user i, and consider the the following allocation strategy. Given the bid vector $w = (w_1, ..., w_M)$, set

$$f_i^*(w) = \frac{w_i}{\phi_i \lambda}, \qquad i = 1, ..., M,$$

where

$$\phi_i = \frac{\lambda_i^*}{\lambda_1^*}, \qquad i = 1, ..., M,$$

and λ is such that

$$\sum_{i=1}^{M} \frac{w_i}{\phi_i \lambda} = \sum_{i=1}^{M} \frac{D_i(u_i^*)}{r(h_i)}.$$

We then have the following result.

Proposition 3: Let Assumption 1 hold. Then there exists a unique equilibrium bidding vector $w^* = (w_1^*, ..., w_M^*)$ for the strategy f^*. In addition, the revenue under the equilibrium bidding vector maximizes the base station's revenue, *i.e.* we have

$$u_i^* = \frac{w_i^*}{f_i^*(w^*) r(h_i)}.$$

Proof: Assume an equilibrium bidding vector w^* exists for the allocation strategy f^*. Using the above definitions, the price per packet paid by user i at this equilibrium is given by

$$\frac{w_i^*}{f_i^*(w^*) r(h_i)} = \frac{\lambda}{\lambda_1^*} u_i^*.$$

From Proposition 2, if w^* is an equilibrium bidding vector it must be that for $i = 1, \ldots, M$,

$$f_i^*(w^*) = \frac{1}{r(h_i)} D_i\left(\frac{\lambda}{\lambda_1^*} u_i^*\right).$$

Thus, from the definition of the allocation strategy,

$$\sum_i \frac{1}{r(h_i)} D_i\left(\frac{\lambda}{\lambda_1^*} u_i^*\right) = \sum_i \frac{D_i(u_i^*)}{r(h_i)}.$$

Since the demand function of each user i is strictly decreasing on $[0, u_{i,max}]$, it follows that there exists a unique λ that satisfies this equation, and $\lambda = \lambda_1^*$. Thus if a equilibrium for f^* exists, it is unique and the equilibrium allocation maximizes revenue.

To see that an equilibrium exists, let $w_i^* = u_i^* D_i(u_i^*)$; under the allocation strategy f^*, this bidding vector can be shown to satisfy Proposition 2. Thus it is an equilibrium bidding vector. ■

The strategy f^* allows the base station to maximize revenue. Under the revenue maximizing strategy, the users may be charged a different price per time slot; this differs from the welfare maximization case discussed above. The difference in price depends on the parameters ϕ_i, $i = 1, \ldots, M$. An example where these parameters vary across the users is given in Sect. VIII. To calculate these parameters requires that the base station knows the utility function of each user. Also note that even if all users had the same utility function they could still be charged a different price per unit time, based on their channel conditions. As we noted above, assuming that the base station knows the utility functions of each user may not be practical. In the next section, we propose a strategy $\hat{f}$ which does not require the base station to know the users' utility function, and show that this strategy is close to optimal for the case where many users are active.

VII. Many User Case

Suppose that the base station uses the allocation strategy $\hat{f}$ given by

$$\hat{f}_i(w) = \frac{w_i}{\lambda}, \qquad i = 1, ..., M,$$

where λ is such that

$$\sum_{i=1}^{M} \frac{w_i}{\lambda} = T.$$

Note this allocation strategy does not depend on the users' utility functions. In the following, we analyze this allocation strategy and show that it has the following properties. In equilibrium, the above strategy, $\hat{f}$,

(a) maximizes the total users' utility, and

(b) maximizes the base station's revenue (in the limit) for the case where many users are active and demand on transmission resources increases (to infinity).

The first property follows immediately from our discussion in Section IV; we summarize this in the following proposition.

Proposition 4: Let Assumption 1 hold. Then there exists a unique equilibrium bidding vector $\hat{w} = (\hat{w}_1, ..., \hat{w}_M)$ for the strategy $\hat{f}$. In addition, the rate vector $(\hat{x}_1, ..., \hat{x}_M)$ under the allocation $(\hat{f}_1(\hat{w}), ..., \hat{f}_M(\hat{w}))$ maximizes the total user's utility.

Next, we address the second property above. We make the following assumption.

Assumption 2: For all $\bar{u}_1, \ldots, \bar{u}_M$, such that

$$\bar{u}_i D_i(\bar{u}_i) = \max_{u_i \geq 0} u_i D_i(u_i), \ \forall i = 1, \ldots, M,$$

we have that

$$\sum_{i=1}^{M} \frac{D_i(\bar{u}_i)}{r(h_i)} \geq T.$$

This assumption implies that when $(u_1^*, ..., u_M^*)$ is an optimal solution to the revenue maximization problem given by Eq. (7), then we have that

$$\sum_{i=1}^{M} \frac{D_i(u_i^*)}{r(h_i)} = T.$$

Intuitively, this states that there are enough users active to ensure that the base station will fully utilize the system (allocate the whole frame duration T) under a strategy that maximizes the revenue.

In the next proposition, we show that in the case where many users are active (and demand on transmission resources increases to infinity), then the $\hat{f}$ is close to optimal.

By Assumption 1.b, there exists a positive constants L such that for all users $i = 1, ..., M$ we have

$$|U_i'(x) - U_i'(x')| \leq L|x - x'|, \quad \text{for all } x, x' \in \Re_+.$$

For each $\lambda \geq 0$, define

$$\bar{D}_T(\lambda) = \max_i \{r(h_i) D_i(\lambda / r(h_i))\}.$$

This can be interpreted as the maximum amount of time any user would demand when charged λ per unit time. From Assumption 1, the function $\bar{D}_T(\lambda)$ is well-defined, decreasing, and there exists λ_{max} such that $\bar{D}_T(\lambda) = 0$, for all $\lambda \geq \lambda_{max}$.

Proposition 5: Let Assumption 1 and 2 hold. Then we have that

$$|P(f^*) - P(\hat{f})| \leq L\bar{D}_T(\hat{\lambda}),$$

where $\hat{\lambda}$ is the price that the base station charges each user in equilibrium per unit time under the strategy $\hat{f}$.

We provide an outline for a proof of this proposition in the appendix.

As the number of active users increases, it can be shown that $\hat{\lambda}$ approaches λ_{max}. Remember that the function $\bar{D}_T(\lambda)$ is decreasing in λ and

$$\lim_{\lambda \to \lambda_{max}} \bar{D}_T(\lambda) = 0.$$

This implies that the the difference $|P(f^*) - P(\hat{f})|$ vanishes as the number of active users increases. Thus Proposition 5 states that the strategy $\hat{f}$ is close to optimal during periods when the demand is high and the base station achieves the highest revenue (under the optimal strategy).

In the following, we provide an informal derivation for Proposition 5. Using Lagrange multipliers, we can rewrite that revenue maximization problem given by Eq. (7) as

$$\begin{aligned} &\text{maximize } \sum_{i=1}^{M} u_i D_i(u_i) - \lambda^* \left(\sum_{i=1}^{M} \frac{D_i(u_i)}{r(h_i)} - T \right) \\ &\text{subject to: } u_i \geq 0, \quad i = 1, ...M. \end{aligned} \tag{8}$$

The optimal prices u_i^*, $i = 1, ..., M$, are then given by

$$u_i^* = \frac{\lambda^*}{r(h_i)} - \frac{D_i(u_i^*)}{D_i'(u_i^*)}$$

and for $\lambda_i^* = u_i^* r(h_i)$ we have

$$\lambda_i^* = \lambda^* - \frac{D_i(u_i^*)}{D_i'(u_i^*) r(h_i)}, \quad i = 1, ..., M. \tag{9}$$

Note from Assumption 1.b, $|D_i'(u_i^*)| \geq \frac{1}{K_2}$ and $D_i(u_i^*) \to 0$ as u_i^* increases. This means that when the number of active users increases (to infinity), then the term $\frac{D_i(u_i^*)}{D_i'(u_i^*)}$ will decrease to zero. This implies that when the number of active users is large, we have (approximately) that

$$\lambda_i^* = \lambda_j^* = \lambda, \qquad i = 1, ..., M,$$

and the strategy $\hat{f}$ becomes (essentially) optimal.

A. A Heuristic Allocation Strategy

We showed in Proposition 5 that in the case where many users are active, the revenue maximization and welfare maximization problem are equivalent (in the limit as the the number of user increases to infinity). However, in the case where a small number of users are active, the strategy $\hat{f}$ may perform significantly worse than f^*.

For this case, the above informal argument can be used to derive a heuristic. In Eq. (9), the term $D_i(u_i^*)/D_i'(u_i^*)$ is the reciprocal of the demand elasticity of user i. Eq. (9) illustrates that under the optimal policy f^* the base station charges users with inelastic demand a higher price to maximize its revenue. Assume the base station does not have the exact knowledge of the utility function of individual users, but some estimate $\hat{\alpha}_i$ about the demand elasticity of user i. The base station could then allocate to user i the transmission time $t_i = f_i(w)$ given by

$$f_i(w) = \frac{w_i}{\lambda + \frac{r(h_i)}{\hat{\alpha}_i}}, \qquad i = 1, ..., M,$$

where λ is such that

$$\sum_{i=1}^{M} \frac{w_i}{\lambda + \frac{r(h_i)}{\hat{\alpha}_i}} = T.$$

Notice that this strategy always fully utilizes the frame (given that there is enough demand), whereas the optimal strategy does not necessarily allocate the full frame duration, T.

The term $D_i(u_i^*)/D_i'(u_i^*)$ is a function of u_i^*; therefore, the base station should (dynamically) change the estimate α_i as the number of active users changes over time. The hope is that when the base station can form a good estimates of the demand elasticity of individual users, then this strategy should perform close to optimal.

VIII. Case Study

We illustrate the above results using a case study. We assume that the user demands are piecewise-linear functions given by

$$D_i(u) = \begin{cases} C_i - a_i u, & 0 \leq u \leq \frac{C_i}{a_i} \\ 0, & \text{otherwise.} \end{cases}$$

We then have that

$$D_i(u)/D_i'(u) = \begin{cases} \frac{C_i}{a_i} - u, & 0 \leq u \leq \frac{C_i}{a_i}, \\ 0, & \text{otherwise.} \end{cases}$$

Letting λ^* be the Lagrange multiplier in (8) and setting

$$u_i^* = \frac{\lambda_i^*}{r(h_i)},$$

the optimal price λ_i^* of (9), as a function of λ^*, is given by

$$\lambda_i^* = \begin{cases} \frac{\lambda^*}{2} + \frac{C_i r(h_i)}{2a_i}, & 0 \leq \lambda^* \leq \frac{C_i r(h_i)}{a_i} \\ \lambda^*, & \text{otherwise.} \end{cases}$$

In the case where the optimal strategy f^* allocates the whole frame duration T, the heuristic that we introduced in Section V is optimal when the base station can exactly estimate parameters C_i and a_i for each user i. However, this means that the base station has exact knowledge of the users' demand functions, or at least perfect knowledge of the ratio $\frac{C_i}{2a_i}$. In practice, the base station might not know the exact ratio $\frac{C_i}{2a_i}$, $i = 1, ..., M$, but only be able to obtain an estimate $\hat{\alpha}_i$.

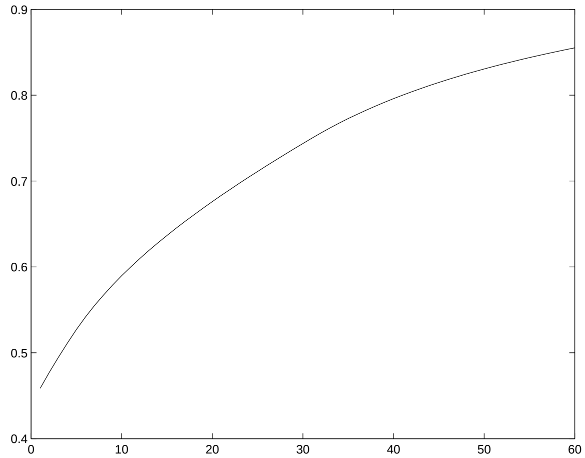

Fig. 1. Revenue under the optimal strategy f^* as a function of of the number of active users of each group given by $N(k) = 10(k-1)+100$, $k = 1, ..., 60$.

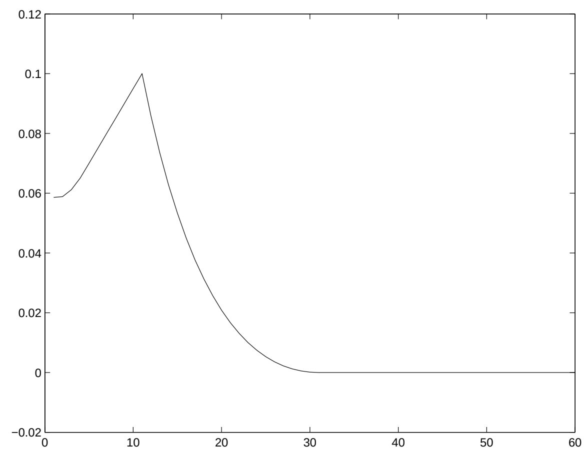

Fig. 2. Revenue difference $P(f^*) - P(\hat{f})$ under f^* and $\hat{f}$ as a function of of the number of active users of each type given by $N(k) = 10(k-1) + 100$, $k = 1, ..., 60$.

We compare, via a numerical study, the performance under the optimal strategy f^*, the strategy $\hat{f}$, and the heuristic strategy f, for the above type of demand function. We set

$$D_i(u) = 0.01(1-u), \qquad i = 1, ..., M.$$

We assume that there are $M = 3N$ active users which we can classify into three different groups corresponding to different channel states; each group with a total of N users. For users of group 1 we set $r(h_i) = 1$, for users of group 2 we set $r(h_i) = 0.5$, and for users of group 3 we set $r(h_i) = 0.3$. We vary the number of active users of each group according to

$$N(k) = 10(k-1) + 100, \qquad k = 1, ..., 60.$$

The frame length is equal to $T = 1$.

Figure 1 shows the revenue under the optimal strategy as a function of the number of active users. As expected, the revenue increases as the number of active users of each group increases.

Figure 2 shows compares the revenue under the optimal strategy f^* and the strategy $\hat{f}$. The strategy f^* always performs better than $\hat{f}$. The difference between the two strategies initially

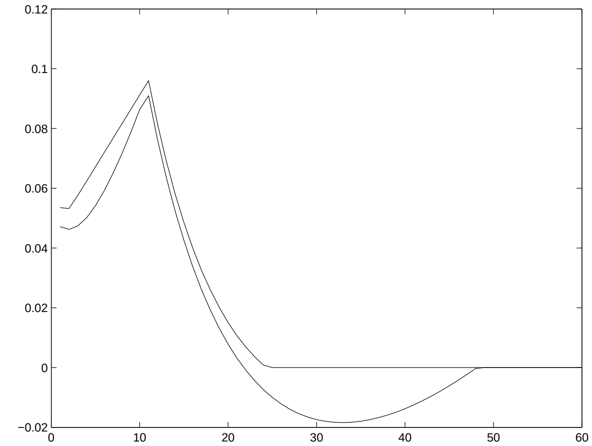

Fig. 3. Revenue difference $P(f) - P(\hat{f})$ under the heuristic f and the strategy $\hat{f}$ as a function of of the number of active users of each type given by $N(k) = 10(k-1) + 100$, $k = 1, ..., 60$. The estimate $\hat{R}$ is equal to $1.3R$ (bottom curve) and $0.8R$ (top curve).

increase as more users are added. However, as the number of active users gets large, the difference decreases to zero, as predicted by Proposition 5.

Figure 3 illustrates how estimation errors of the ratio $R_i = \frac{C_i}{2a_i}$ affect the performance of the heuristic strategy. In our case study, this ratio is given by

$$R_i = R = \frac{0.01}{0.02} = 0.5, \qquad i = 1, ..., M.$$

Let $\hat{R}$ be the estimate by of the ratio R. We consider two scenarios where we assume that the estimate $\hat{R}$ is equal to $1.3R$, and $0.8R$, respectively. Figure 3 that the heuristic outperforms strategy $\hat{f}$ when the number of active users is small. When the number of active users is large, then the performances of the two strategies are identical and equal to the optimal performance. However for $\hat{R} = 1.3$, the heuristic can be worse when the number of active users is moderately large.

A. Conclusions

We have presented a pricing scheme for the downlink in a wireless network where different users can receive data at different rates. In this scheme, users bid in each frame by submitting a price bid and the base station allocates resources to maximize its revenue. We propose a revenue maximizing strategy f^* for the base station. This scheme however requires that the base station know the utility function of each user. We also propose a suboptimal strategy $\hat{f}$ which does not require knowledge of the user's utility functions, and show that this scheme is asymptotically optimal, in the limit of many users and large demand. Moreover, such a scheme is shown to maximize social welfare. We also consider a heuristic strategy f, which does not require perfect knowledge of the users' utility functions. Using a case study, we show that the revenue under $\hat{f}$ is close to f^* for large demand. The case study also illustrates the the heuristic f performs better than $\hat{f}$ for small demand, given that the base station can accurately estimate the elasticities of the individual users.

REFERENCES

[1] P. Bender, *et al.*, "CDMA HDR:A Bandwidth-Efficient High Speed Wireless Data Service for Nomadic Users", *IEEE Comm. Mag.*, July, 2000, pp. 70-77.

[2] S. Borst and P. Whiting, "Dynamic Rate Control Algorithms for HDR Throughput Optimization," *IEEE Infocom 2001*, Anchorage, Alaska, April, 2001.

[3] *TIA/EIA IS-856 CDMA 2000: High rate packet data air interface specification*, Nov., 2000.

[4] D. Famolari, N. Mandayam, and D. Goodman. "A new Framework for power control in wireless data networks: games, utility and pricing," *Allerton Conference on Communication, Control, and Computing*, Monticello, IL, Sept., 1998.

[5] T. Heikkinen, "Optimal Quality of Service and pricing in the wireless Internet," *ITC seminar 2000*, Lillehammer, 2000.

[6] H. Ji, "An Economic Model for Uplink Power Control in Cellular Radio Networks," *Allerton Conference on Communication, Control, and Computing*, Monticello, IL, Sept., 1995.

[7] F. P. Kelly, "Charging and Rate Control for Elastic Traffic," *European Trans on Telecommun* 8 , pp. 33-37, 1997.

[8] F. P. Kelly, A. K. Maulloo, and D. K. H. Tan, "Rate Control for Communication Networks: Shadow Prices, Proportional Fairness and Stability," *Journal of Operation Research Society* 49(1998), 237-252.

[9] P. Liu, M. Honig, and S. Jordan, "Forward-Link CDMA Resource Allocation Based on Pricing," *IEEE Wireless Communications and Networking Conference*, Chicago, IL, Sept., 2000.

[10] S. Low and P. Varaiya, "A new approach to service provisioning in ATM networks," *IEEE/ACM Trans. on Networking*, vol. 1, no. 5, pp. 547-553, 1993.

[11] J. K. MacKie Mason and H. R. Varian, "Pricing Congestible Network Resources", *IEEE Journal of Selected Areas in Communication.* vol. 13, no. 7, pp. 1141-1149, Sept. 1995.

[12] P. Marbach, "Pricing Priority Classes in Differentiated Services Networks," *Allerton Conference on Communication, Control, and Computing*, Monticello, IL, Sept., 1999.

[13] C. Saraydar, N. Mandayam, and D. Goodman, "Pareto efficiency of pricing based power control in wireless data networks", *Wireless Communication and Networking Conference*, 1999.

[14] S. Shenker, "Fundamental design issues for the future Internet," *IEEE J. Selected Areas Comm.*, vol. 13, pp. 1176-1188, 1995.

[15] D. Tse, "Multiuser Diversity in Wireless Networks," Talk at 2001 *IEEE Commun. Theory Workshop*

[16] D. Tse, Optimal Power Allocation over Parallel Gaussian Broadcast Channels, *International Symposium on Information Theory*, Ulm, Germany, June, 1997.

[17] C. Zhou and M. Honig and S. Jordan, "Two-Cell Power Allocation for Wireless Data Based on Pricing", *Allerton Conference on Communication, Control, and Computing*, Monticello, IL, Sept., 2001.

APPENDIX

Let f^* be the allocation strategy from Section VI that maximizes the base station's revenue. Let w^* indicate the equilibrium bid vector for f^* and let

$$\lambda_i^* = \begin{cases} \frac{w_i^*}{f_i^*(w^*)} & \text{if } f_i^*(w^*) > 0, \\ u_{i,max} r(h_i) & \text{if } f_i^*(w^*) = 0. \end{cases} \tag{10}$$

Thus for all active users, λ_i^* is the equilibrium price user i pays per unit time under f^*, and for all inactive users, λ_i^* is the minimum price per unit time such that the demand of that user is zero. Thus, using Proposition 2 and Assumption 2, it follows that

$$\sum_{i=1}^{M} \frac{1}{r(h_i)} D_i \left(\frac{\lambda_i^*}{r(h_i)} \right) = T. \tag{11}$$

Assume that under f^*, for user m, we have,

$$\lambda_m^* \geq \lambda_i^*, \qquad \text{for all } i = 1, ..., M,$$

and for user n we have that $f_n^*(w^*) > 0$ and

$$\lambda_n^* \leq \lambda_i^*, \qquad \text{for all } i \text{ such that } f_i^*(w^*) > 0 .$$

Note that $\lambda_n^* > 0$.

Under the strategy $\hat{f}$ and any given bid vector, notice that all users are charged the same price per unit time; let $\hat{\lambda}$ indicate this price for the equilibrium bid vector under $\hat{f}$. Again using Proposition 2, it follows that

$$\sum_{i=1}^{M} \frac{1}{r(h_i)} D_i \left(\frac{\hat{\lambda}}{r(h_i)} \right) = T. \tag{12}$$

Using (11) and (12), we then have

$$\lambda_m^* \geq \hat{\lambda} \geq \lambda_n^*.$$

Thus,

$$|P(f^*) - P(\hat{f})| \leq T|\lambda_m^* - \hat{\lambda}| \leq T|\lambda_m^* - \lambda_n^*|. \tag{13}$$

Thus if $\lambda_m^* = \lambda_n^*$ the proof is done. Therefore, in the following we assume that $\lambda_n^* \neq \lambda_m^*$.

By the above definitions, we have that $f_n^*(w^*) > 0$ and thus $f_n^*(w^*) < T$. Let Δt be a given small constant such that

$$0 < \Delta t < \min\{f_n^*(w^*), T - f_m^*(w^*)\}.$$

Furthermore, let the constant $\tilde{w}_m$ is chosen so that

$$f_m^*(w^*) + \Delta t = \frac{1}{r(h_m)} D_m \left(\frac{\tilde{w}_m}{(f_m^*(w^*) + \Delta t) r(h_m)} \right)$$

and let the constant $\tilde{w}_n$ is chosen so that

$$f_n^*(w^*) - \Delta t = \frac{1}{r(h_n)} D_n \left(\frac{\tilde{w}_n}{(f_n^*(w^*) - \Delta t) r(h_n)} \right).$$

Consider a new allocation strategy f, defined as follows:

$$f_i(w) = \begin{cases} f_i^*(w^*) & \text{if } i \neq n, m, \text{ and } w_i \geq w_i^*, \\ f_m^*(w^*) + \Delta t & \text{if } i = m \text{ and } w_m \geq \tilde{w}_m, \\ f_n^*(w^*) - \Delta t & \text{if } i = n \text{ and } w_n \geq \tilde{w}_n. \\ 0 & \text{otherwise.} \end{cases}$$

The strategy f can be seen to have the equilibrium bid vector $w^*(f) = (w_1^*, ..., \tilde{w}_m, .., \tilde{w}_n, ..., w_M^*)$, this results in the allocation:

$$f_i(w^*(f)) = \begin{cases} f_i^*(w^*) & \text{if } i \neq m, n, \\ f_m^*(w^*) + \Delta t & \text{if } i = m, \\ f_n^*(w^*) - \Delta t & \text{if } i = n. \end{cases}$$

Let λ_i, $i - 1, \ldots, M$ be defined as in (10) for the equilibrium under f, *i.e.*

$$\lambda_i = \begin{cases} \frac{w_i^*(f)}{f_i(w^*(f))} & \text{if } f_i(w^*(f)) > 0, \\ u_{i,max} r(h_i) & \text{if } f_i(w^*(f)) = 0. \end{cases}$$

We then have that

$$\lambda_i = \begin{cases} \lambda_i^* & \text{if } i \neq m, n, \\ \lambda_m^* - \Delta_m & \text{if } i = m, \\ \lambda_n^* + \Delta_n & \text{if } i = n, \end{cases}$$

for some $\Delta_m > 0$, $\Delta_n > 0$. Let $P(f)$ indicate the revenue under strategy f. Then we have

$$P(f) - P(f^*) = \Delta t\Big(\lambda_m^* - \lambda_n^*\Big) - \Delta_m f_m^*(w^*) \\ + \Delta_n f_n^*(w^*) - \Delta t(\Delta_m + \Delta_n).$$

Using Assumption 1.b, there exist a Lipschitz constant $L > 0$ such that

$$\Delta_m \leq Lr(h_m)^2 \Delta t,$$

and similarly

$$\Delta_n \leq Lr(h_n)^2 \Delta t.$$

Thus

$$P(f) - P(f^*) \geq \Delta t\Big(\lambda_m^* - \lambda_n^* - Lr(h_m)^2 f_m^*(w^*)\Big) \\ + o(\Delta t).$$

¿From this it follows that strategy f would achieve a higher revenue than strategy f^* for some Δt when

$$\lambda_m^* - \lambda_n^* > Lr(h_m)^2 f_m^*(w^*).$$

As f^* is an optimal strategy, it must be that

$$\begin{aligned} \Big|\lambda_m^* - \lambda_n^*\Big| &\leq Lr(h_m)^2 f_m^*(w^*) \\ &= Lr(h_m) D_m \left(\frac{\lambda_m^*}{r(h_m)}\right) \\ &\leq L\bar{D}_T(\lambda_m^*) \\ &\leq L\bar{D}_T(\hat{\lambda}). \end{aligned}$$

Here we have used that (see Proposition 2)

$$f_m^*(w^*) = \frac{1}{r(h_i)} D_m \left(\frac{\lambda_m^*}{r(h_m)}\right)$$

and the definition of $\bar{D}_T(\lambda)$.