Opportunistic scheduling in cellular systems in the presence of noncooperative mobiles

Veeraruna Kavitha, Eitan Altman, Rachid El-Azouzi and Rajesh Sundaresan

Kavitha.Voleti_Veeraruna@sophia.inria.fr, Eitan.Altman@sophia.inria.fr,

rachid.elazouzi@univ-avignon.fr, rajeshs@ece.iisc.ernet.in

Abstract

A central scheduling problem in wireless communications is that of allocating resources to one of many mobile stations that have a common radio channel. Much attention has been given to the design of efficient and fair scheduling schemes that are centrally controlled by a base station (BS) whose decisions depend on the channel conditions of each mobile. The BS is the only entity taking decisions in this framework based on truthful information from the mobiles on their radio channel conditions. In this paper, we study the scheduling problem from a game-theoretic perspective in which some of the mobiles may be noncooperative. We model this as a signaling game and study its equilibria. We then propose various approaches to enforce truthful signaling of the radio channel conditions: a pricing approach, an approach based on some knowledge of the mobiles' policies, and an approach that replaces this knowledge by a stochastic approximations approach that combines estimation and control. We further identify other equilibria that involve non-truthful signaling.

I. Introduction

Short-term fading arises in a mobile wireless radio communication system in the presence of scatterers, resulting in time-varying channel gains. Various cellular networks have downlink shared data channels that use scheduling mechanisms to exploit fluctuations in radio conditions (e.g. 3GPP HSDPA [2] and CDMA/HDR [7] or 1xEV-DO [1]). The scheduler design and the obtained gain are predicated on the mobiles sending information concerning the downlink channel gains in a truthful fashion. In a frequency-division duplex system, the base station (BS) has no direct information on the channel gains, but transmits downlink pilots, and relies on the mobiles' reported values of gains on these pilots for scheduling. A cooperative mobile will truthfully report this information to the BS. A noncooperative mobile will however send a signal that is likely to induce the scheduler to behave in a manner beneficial to the mobile.

For some examples of nonstandard, noncooperative and aggressive transmission behavior in WLANs, we refer the reader to Mare et al. [3], where nodes attempt more frequently than the specifications in the IEEE 802.11 standard. See also Bianchi et al. [4]. This is presumably because the particular equipment provider wants to make its devices more competitive. Such behavior may occur in cellular phones with respect to channel reports for similar reasons of competitiveness since compliance testing for cooperation is over only limited, published, and standardized scenarios. Both HSDPA and EV-DO use an opportunistic scheduler in the downlink to profit from multi-user diversity. For instance, a non-cooperative mobile can modify their 3G mobile devices or laptops 3G PC cards, either by using Software Development Kit (SDK) (see [20]) or the device firmware [27], in order to usurp time slots at the expense of cooperative mobiles, hence denying them network access.

This paper deals with game-theoretic analysis of downlink scheduling in the presence of noncooperative mobiles. We initially assume that the identity of players that do not cooperate is common knowledge. In the later parts of the paper, while discussing the stochastic approximation based approach, the BS can detect the non cooperative mobiles and hence will not require this knowledge. We model this non cooperative downlink scheduling initially as a *signaling game*. Mobiles send signals that correspond to reported channel states and play the role of *leaders* in the signaling game. The BS allocates the channel resource and plays the role of a *follower* that reacts to the signals. Mobile utilities (throughputs) are determined by BS's allocation. For efficient scheduling, BS optimizes the sum of the utilities of all the mobiles and hence naturally the sum utility forms its utility in the game formulation. We initially focus on the study of equilibria of this game and later on concentrate on robustification of the policies of the BS against noncooperation.

Contribution of the paper: We begin with the case in which BS does not use any extra intelligence to deal with noncooperative mobiles (BS makes scheduling decisions based only the signals from the mobiles). The only Perfect Bayesian Equilibrium (PBE) of the resulting signaling game are of the *babbling* type: the noncooperative mobiles send signals independent of their channel states, and the BS simply ignores them to allocate channels based only on prior channel statistics (Section III). Fortunately, the BS can use more intelligent strategies to achieve a truth revealing equilibrium henceforth called as TRE. We present three ways to obtain these as equilibria of an appropriate form of the game (Section IV). The first relies on capabilities of the BS to estimate the real downlink channel quality (perhaps obtained at a later stage based on the rate at which the actual transmission took place), combined with a pricing mechanism that creates incentives for truthful signaling. In the second approach, the BS learns mobiles' signaling statistics, correlates them with the true channel statistics, and punishes the deceivers. We next come up with a practical strategy to achieve a TRE in the form of a variant of the proportional fair sharing algorithm (PFA) which elicits truthful signals from mobiles (Section V). Further, in Section VI, we establish the existence of other equilibria at which the BS improves its utility in comparison to that obtained at a babbling equilibrium; the noncooperative mobiles also improve their utilities over their cooperative shares (their utilities at a TRE).

Prior work: PFA and related algorithms were intensely analyzed as applied to CDMA/HDR and 3GPP HSDPA systems ([10], [7], [6], [25], [5], [9], [19]). Kushner & Whiting [17] showed using stochastic approximation techniques that the asymptotic averaged throughput can be driven to optimize a certain system utility function (sum of logarithms of offset-rates). All the above works assume that the centralized scheduler has true information of channel states.

However, as seen in the simulations of Kong et al. [15], noncooperative mobiles can gain in throughput (by 5%) but cause a decrease in the overall system throughput (by 20%). Nuggehalli et al. [21] considered noncooperation by low-priority latency-tolerant mobiles in an 802.11e LAN setting capable of providing differentiated quality of service. Price & Javidi [24] consider an uplink version of the problem with mobiles being the informed parties on valuations of uplinks (queue state information is available only at the respective mobiles).

Our problem is closely related to signaling games with *cheap talk*, i.e., signals incur no costs [16, Sec. 7], for which it is well-known that babbling equilibria exist. While the above works [21] and [24] use *mechanism design* techniques [12] to induce truth revelation, with pricing implemented via "carrots" on the opposite link, our problem differs from mechanism design not only in not having pricing but also in considering the BS as a player.

We now begin with a precise formulation of the problem before proceeding to solutions.

II. PROBLEM FORMULATION

Consider the downlink of a wireless network with one base station (BS). M mobiles compete for the downlink data channel. Time is divided into small intervals or slots. In each slot one of the M mobiles is allocated the channel. Mobile m can be in one of the channel states $h_m \in \mathcal{H}_m$, where $|\mathcal{H}_m| < \infty$. Fading characteristics are independent across mobiles. Let $\mathbf{h} := [h_1, h_2, \cdots, h_M]^t$ be the vector of channel gains in a particular slot. Its distribution is $p_{\mathbf{H}}(\mathbf{h}) = \prod_m p_{H_m}(h_m)$, where p_{H_m} is that of the random variable h_m. We assume mobile m can estimate h_m perfectly using pilots transmitted by the BS. Mobile m sends signal s_m to the BS to indicate its channel gain. Some mobiles (say those with indices $1 \leq m \leq M_1$ where $M_1 \leq M$) are noncooperative and may signal a different (say good) channel condition other than their true channel (say bad) in order to be allocated the channel. Channel statistics and noncooperative mobile identities are common knowledge to all players. Signal values are chosen from the channel space itself, i.e., $s_m \in \mathcal{H}_m$. BS makes a scheduling decision based on signals $\mathbf{s} := [s_1, s_2, \cdots, s_M]^t$.

1) Utilities: Let A denote the mobile to which channel is allocated. If $A = m$, mobile m gets a utility $U_m(s_m, h_m, A)$ given by $f(h_m)$ which depends only on its own channel state and the allocation, but not on the signal. Thus $U_m(s_m, h_m, A) = 1_{\{A=m\}} f(h_m)$[1]. An example f function is $f(h_m) = \log(1 + h_m^2 \mathsf{SNR})$ where SNR is the average received signal-to-noise ratio. BS utility

[1]This is the case if the BS allocates based on the mobile's signal when the signaled channel gain is more than or equal to the true channel gain. In the later sections that develop robust BS policies, we will come across situations when the BS allocates to provide a utility (say $\tilde{u}$) different from the one requested. In these cases, $U_m(s_m, h_m, A) = 1_{\{A=m\}} \min\{\tilde{u}, f(h_m)\}$.

is the sum of mobile utilities:

$$U_{BS}(\mathbf{s}, \mathbf{h}, A) = \sum_m U_m(s_m, h_m, A).$$

Optimizing the BS's sum utility results in an *efficient* solution, our main object of study. Fairness may be incorporated appropriately; see our extensions [13] where utilities are concave functions of long-term average throughputs.

How is $f(h_m)$ achievable when the transmitting BS does not know the true h_m? Even if a mobile signals more than its true value and the BS attempts to transmit at that higher transmitted rate, the actual rate at which the transmission takes place will still be $f(h_m)$. This is reasonable given the following observations. The reported channel is usually subject to estimation errors and delays, an aspect that we do not consider explicitly in this paper. To address this issue, the BS employs a *rate-less* code, i.e., starts at an aggressive modulation and coding rate, gets feedback from the mobile after each transmission, and stops as soon as sufficient number of redundant bits are received to meet the decoding requirements. This incremental redundancy technique supported by hybrid ARQ is already implemented in the aforementioned standards (3GPP HSDPA and 1xEV-DO). Then a rate close to the true utility may be achieved.

2) A Motivating Example: To illustrate the main concepts, consider two mobiles in the toy one-shot game with channel states, probabilities, and achieved throughputs as given in Table I. The fifth column shows utilities when allocation $A \equiv m$ (BS always allocates mobile m). The sixth column shows utilities when mobiles signal truthfully and allocation $A^*(s_1, s_2) = \arg\max_m\{s_m\}$ is to mobile with the best channel, yielding the best total utility of 6 + 2.50 = 8.50 for the BS. If mobile 2 is strategic, noncooperative, and therefore always signals 10, $s_m \equiv 10$, an ignorant BS always allocates to mobile 2 and attains a utility of $3.75 < 8.50$. Since the mobile 2 noncooperative utility is 3.75 which is greater than the 2.50 attained under cooperative signaling, mobile 2 will not cooperate. If the BS is aware of such noncooperative behavior, we will soon see that it will always allocate the channel to mobile 1 (based only on priors) yielding utilities of 8 to mobile 1, 0 to mobile 2, and 8 to BS; the last quantity is less than 8.50 under cooperative signaling. We will also see that 8 and 8.50 are two extremes of what the BS can achieve.

TABLE I
PARAMETERS AND UTILITIES FOR THE MOTIVATING EXAMPLE

Player	$\mathcal{H}_m$	p_{H_m}	$f(\mathcal{H}_m)$	$\mathbb{E}[U_m(\cdot, h_m, m)]$ $(A \equiv m)$	$\mathbb{E}[U_m(h_m, h_m, A^*)]$ $A^* = \arg\max_m\{s_m\}$
Mobile 1	$\{h_1^0\}$	$\{1\}$	$\{8\}$	8	6
Mobile 2	$\{h_2^1, h_2^2, h_2^3\}$	$\{\frac{1}{4}, \frac{1}{2}, \frac{1}{4}\}$	$\{10, 2, 1\}$	3.75	2.50

3) Terminology: We define $[M] := \{1, 2, \cdots, M\}$. For a set $\mathcal{C}$ let $\mathcal{P}(\mathcal{C})$ denote the set of probability measures on $\mathcal{C}$. As is usual in games, all players employ randomized strategies. Hence, in the one-shot model, a policy of mobile m is a mapping $h_m \mapsto \mu_m(\cdot|h_m) \in \mathcal{P}(\mathcal{H}_m)$, i.e., a random signal is generated as per the mapped distribution given the channel state h_m. A policy of the BS is a mapping $\mathbf{s} \mapsto \beta(\cdot|\mathbf{s}) \in \mathcal{P}([M])$, i.e., a random allocation is made based on the signal vector. Let

$$\mu(\mathbf{s}|\mathbf{h}) := \prod_{m \leq M_1} \mu_m(s_m|h_m) \prod_{j > M_1} \delta(h_j - s_j).$$

As is usual, to exclude mobile m, define $\mathbf{h}_{-m} := [h_1, \cdots, h_{m-1}, h_{m+1}, \cdots, h_M]$, $p_{H_{-m}}(\mathbf{h}_{-m}) := \prod_{j \neq m} p_{H_j}(h_j)$ and

$$\mu_{-m}(\mathbf{s}_{-m}|\mathbf{h}_{-m}) := \prod_{j \neq m; j \leq M_1} \mu_j(s_j \mid h_j) \prod_{j \neq m; j > M_1} \delta(h_j - s_j).$$

We reuse U_m to denote the instantaneous utility of mobile m when its channel condition is h_m, when the mobiles use strategies μ, and when the BS uses strategy β. More precisely,

$$\begin{aligned} U_m(\mu, h_m, \beta) &:= \mathbb{E}_{\mathbf{h}_{-m}}\left[\sum_{\mathbf{s}} U_m(s_m, h_m, m)\beta(m \mid \mathbf{s})\mu(\mathbf{s} \mid \mathbf{h})\right] \\ &= f(h_m)\mathbb{E}_{\mathbf{h}_{-m}}\left[\sum_{\mathbf{s}} \beta(m \mid \mathbf{s})\ \mu(\mathbf{s} \mid \mathbf{h})\right]. \end{aligned}$$

Similarly, $U_{BS}(\mu, \mathbf{h}, \beta) := \sum_m U_m(\mu, h_m, \beta)$.

4) Strategic Form Game : This noncooperative downlink game results in a strategic form game with $M_1 + 1$ players $1, 2, \cdots, M_1$, and the BS. The strategy set of the players are $\mu_1, \mu_2, \cdots, \mu_{M_1}$, and β, respectively. The payoffs are $\mathbb{E}[U_1], \mathbb{E}[U_2], \cdots, \mathbb{E}[U_{M_1}]$, and $\mathbb{E}[U_{BS}]$, respectively, where the expectations are with respect to $\mathbf{h}$. A *Nash Equilibrium (NE)* for this game is a strategy-tuple $(\mu_1^*, \cdots, \mu_{M_1}^*; \beta^*)$ that satisfies

$$\begin{aligned} \mu_m^* &\in \arg\max_{\mu_m} \mathbb{E}_{h_m}[U_m((\mu_m, \mu_{-m}^*), h_m, \beta^*)] \quad (1 \leq m \leq M_1) \\ \beta^* &\in \arg\max_{\beta} \mathbb{E}_{\mathbf{h}} \left[\sum_m U_m(\mu^*, h_m, \beta) \right]. \end{aligned}$$

An ϵ*-Nash Equilibrium (*ϵ*-NE)* for this game is a strategy-tuple $(\mu_1^*, \cdots, \mu_{M_1}^*; \beta^*)$ that satisfies

$$\begin{aligned} \mu_m^* &> \max_{\mu_m} \mathbb{E}_{h_m}[U_m((\mu_m, \mu_{-m}^*), h_m, \beta^*)] - \epsilon \quad (1 \leq m \leq M_1) \\ \beta^* &> \max_{\beta} \mathbb{E}_{\mathbf{h}} \left[\sum_m U_m(\mu^*, h_m, \beta) \right] - \epsilon. \end{aligned}$$

5) Signaling Game : For the problem under study, the BS has to act based on the signals sent by the mobiles and hence this is better modeled by the two stage signaling game. For such signaling games, a refinement of NE based on the rationale of credible posterior beliefs (Kreps & Sobel [16, Sec. 5], Sobel [26]) is the Perfect Bayesian Equilibrium defined below.

Definition *2.1 (Posterior beliefs):* $\pi := \{\pi_m; m \leq M_1\}$ is the set of posterior beliefs where $\pi_m(h_m \mid s_m)$ is the BS's belief of the posterior probability that the mobile's true channel is h_m given that its signal is s_m. □

Definition *2.2 (Perfect Bayesian Equilibrium (PBE)):* A strategy profile $(\mu_1^*, \cdots, \mu_{M_1}^*; \ \beta^*)$ and a posterior belief π^* constitute a PBE if (a) for each signal vector s, we have

$$\beta^*(\cdot \mid \mathbf{s}) \in \arg \max_{\gamma \in \mathcal{P}([M])} \sum_{m > M_1} \gamma(m) f(s_m) + \sum_{m \leq M_1} \gamma(m) \sum_{h_m} \pi_m^*(h_m \mid s_m) f(h_m); \tag{1}$$

(b) for each mobile $m \leq M_1$ and $h_m \in \mathcal{H}_m$, we have

$$\mu_m^*(\cdot \mid h_m) \in \arg \max_{\alpha \in \mathcal{P}(\mathcal{H}_m)} \sum_{s_m \in \mathcal{H}_m} \alpha(s_m) \left[\sum_{\mathbf{s}_{-m}} \beta^*(m \mid \mathbf{s}) \sum_{\mathbf{h}_{-m}} p_{H_{-m}}(\mathbf{h}_{-m}) \mu_{-m}^*(\mathbf{s}_{-m} | \mathbf{h}_{-m}) f(h_m) \right]; \tag{2}$$

(c) for each $m \leq M_1$ and $s_m \in \mathcal{H}_i$, the BS updates

$$\pi_m^*(h_m \mid s_m) = \frac{p_{H_m}(h_m) \mu_m^*(s_m \mid h_m)}{\sum_{h'_m \in \mathcal{H}_m} p_{H_m}(h'_m) \mu_m^*(s_m \mid h'_m)}, \tag{3}$$

if the denominator in (3) is nonzero, and $\pi_m^*(\cdot \mid s_m)$ is any element in $\mathcal{P}(\mathcal{H}_m)$ otherwise. □

In Definition 2.2, equation (2) ensures that $(\mu_1^*, \cdots, \mu_{M_1}^*)$ is a NE of the subgame of noncooperative mobiles, (1) ensures that β^* is the Bayes-Nash equilibrium of the subgame with the BS, and (3) determines a consistent Bayesian approach to determining posterior beliefs.

In the sequel, we will come across two types of PBE ([26]). The first is the *babbling equilibrium* where the *sender's* (mobile's) strategy is independent of the channel state, and the *receiver's* (BS's) strategy is independent of signals. The second is the desirable *separating equilibrium* where sender sends signals from disjoint subsets of the set of available signals for each channel state. Clearly then, the receiver gets complete information about the true channel states of the leaders (mobiles). If this equilibrium is achieved, the BS can design a scheduling algorithm as in a fully cooperative environment. Hence a separating PBE is a Truth Revealing Equilibrium (TRE).

The question then is what kind of equilibria do we encounter in the above signaling game. Refinements to handle the more realistic repeated game over multiple slots and availability of more information at the BS are handled subsequently.

III. Babbling Equilibrium

The following theorem characterizes all the possible PBE of the signaling game as babbling equilibria.

Theorem 1: The $M_1 + 1$ player signaling game has a PBE of the following type:

$$\begin{aligned}\pi_m^*(\cdot|s_m) &= p_{H_m} && \text{(for all } m,\ s_m),\\ \mu_m^*(\cdot|h_m) &\text{ equals any fixed } \mu_m \in \mathcal{P}(\mathcal{H}_m) && \text{(for all } m,\ h_m),\\ \beta^*(\cdot|\mathbf{s}) &\text{ equals any fixed } \gamma_{\mathbf{s}} \in \mathcal{P}(M^*(\mathbf{s})) && \text{(for all } \mathbf{s}),\end{aligned}$$

where $M^*(\mathbf{s})$ is the set of mobiles with the best expected throughput among noncooperative mobiles and the conditional throughput given s among cooperative mobiles, i.e.,

$$\begin{aligned}M^*(\mathbf{s}) &:= \arg \max_{m \in M^*_{NC} \cup M^*_C(\mathbf{s})} \mathbb{E}_{h_m}\left[f\left(h_m\right) \mid s_m\right] \text{ with}\\ M^*_{NC} &:= \arg\max_{m \le M_1} \mathbb{E}_{h_m}\left[f(h_m)\right] \text{ and } M^*_C(\mathbf{s}) := \arg\max_{m > M_1} f(s_m).\end{aligned}$$

Further, any PBE for this game is of the above type. □

Proof: See Appendix A. ■

The above theorem shows that if the BS makes scheduling decisions based only on the signals from the (noncooperative) mobiles: at any equilibrium $\pi_m^*(\cdot|s_m) = p_{H_m}$ for all $m \le M_1$, i.e., mobile signals do not improve BS's knowledge of current channel states. BS allocates based only on prior statistics and signals of cooperative mobiles. As a consequence, multiuser diversity gain cannot be exploited and the best possible BS utility under this situation is

$$U^*_{cop} := \mathbb{E}_{h_m; m > M_1}\left[\max\left\{\max_{m > M_1} f(h_m), \max_{j \le M_1} \mathbb{E}_{h_j}[f(h_j)]\right\}\right] \tag{4}$$

To do better, we exploit the fact that typical connections last several slots enabling the BS to learn more about mobiles' strategies, for example the statistics of their signals. We first study two punitive strategies to elicit truthful signals from mobiles and then go on to study other equilibria.

IV. Separating Equilibrium

We showed in the previous section that there exist only babbling equilibria in the presence of noncooperative mobiles. In this section we obtain the desired TRE using two different approaches.

A. Penalty for deviant reporting

The BS does not have access to the true channel state of the mobile. But based on the actual throughput seen on the allocated link, the BS may extract the squared error $\{(h_m - s_m)^2; m \leq M_1\}$ after the transmission is over. This error can used to punish mobiles for deviant reporting.

More precisely, let mobile m report s_m when its channel is h_m and suppose it succeeds in getting the channel. For a $\Delta \in (0, \infty)$, let us impose a penalty proportional to the squared error if it exceeds a threshold c_m, as follows:

$$U_m(h_m, s_m, A) = 1_{\{A=m\}} \left(f(h_m) - \Delta 1_{\{(h_m - s_m)^2 > c_m\}} (h_m - s_m)^2 \right), \tag{5}$$

where $c_m > 0$ is chosen small enough such that for all $h_m \in \mathcal{H}_m$,

$$\left\{ s_m : (s_m - h_m)^2 \leq c_m \right\} \cap \mathcal{H}_m = \{h_m\}.$$

If we now choose Δ such that

$$\Delta > \max_{\{m \leq M_1,\, (h_m, s_m) \in \mathcal{H}_m \times \mathcal{H}_m : (h_m - s_m)^2 > c_m\}} \left\{ \frac{f(h_m)}{(h_m - s_m)^2} \right\}, \tag{6}$$

then it is clear that $U_m(h_m, s_m, A)$ is negative whenever the action A is m, i.e., any deviant signaling results in negative utility to the mobile. This new utility function is closely related to a pricing mechanism, a powerful tool for achieving a more socially desirable result. Typically, pricing is used to encourage the mobile to use system resource more efficiency and generate revenue for the system. Usage-based pricing is an approach commonly encountered in the literature. In usage-based pricing, the price a mobile pays for using resource is proportional to the amount of resource consumed by the mobile. In our case, the price corresponds to the cost a mobile pays for deviant reporting if the error exceeds c_m. Through pricing, we obtain a separating PBE for the modified game.

Theorem 2: With Δ satisfying (6), the M_1+1-noncooperative game with the modified utilities has the following separating PBE:

$$\mu_m^*(s_m|h_m) = \delta(s_m - h_m) \text{ for all } m \leq M_1 \text{ and } h_m \in \mathcal{H}_m,$$

$$\pi_m^*(h_m|s_m) = \delta(h_m - s_m) \text{ for all } m \leq M_1 \text{ and } s_m \in \mathcal{H}_m,$$

and with $A^*(\mathbf{s}) = \arg\max_j f(s_j)$, $\beta^*(\cdot|\mathbf{s})$ is any probability measure with support set $A^*(\mathbf{s})$.

Proof: See Appendix B. ■

We thus achieve a TRE using this method. However it is important to note here that one may not be able to estimate the instantaneous channel error of (5) even after the transmission is complete. We propose in the following subsection another (impractical) method based on signal statistics with the intention of introducing our ideas on robustification. More practical policies based on average throughput error will be dealt in the next section.

B. 'Predicting' the Signal Statistics

Data transmissions are not just one-shot, but occur over several slots. This enables the BS to learn the statistics of the signals sent by mobiles. To explore this idea we begin with the simplifying restriction that mobile strategies are stationary. This enables us to study, yet again, a one-shot game where mobile signaling statistics are known to the BS. This leads to a strategic form game with mobile actions as before whereas the BS's action depends not only on the signals $\mathbf{s}$, but also on the (learned and therefore assumed perfectly known) statistics of signals, $p_{\mathbf{S}} := (p_{S_m}, m \in [M])$, i.e., $(\mathbf{s}, p_{\mathbf{S}}) \mapsto \beta(\cdot \mid \mathbf{s}, p_{\mathbf{S}}) \in \mathcal{P}([M])$. (Recall that $p_{S_m} = \sum_h \mu_m(\cdot|h) p_{H_m}(h)$). The payoff for the mobiles and BS are as before.

Consider the following BS policy denoted β_p^*. Find the set of mobiles m whose signaled statistics p_{S_m} equals p_{H_m}. The strategy β_p^* makes an equiprobable choice among those mobiles in this subset that have the largest signal amplitudes. If the set is empty, the BS does not allocate the channel to any of the mobiles.

Some remarks are in order. First, "p_{S_m} equals p_{H_m}" assumes knowledge of statistics of the signals. This is not available in practice, must be estimated, and will therefore have estimation errors. The term "equals" should therefore be interpreted in practice as "approximately equals" to within a desired level accuracy. Second, β_p^* is a punitive strategy in that only those mobiles whose signaling statistics match the channel's true statistics obtain a strictly positive utility. Third, mobiles may deceive and yet obtain a strictly positive utility so long as signaling statistics match. But inflationary signaling for lower levels have to be compensated by deflationary signaling at higher levels.

Clearly $\mu_m^*(s_m|h_m) = \delta(s_m - h_m),$ for all $s_m \in \mathcal{H}_m, m \in [M_1]$ and β_p^* constitute a NE, i.e. a TRE, with BS utility

$$U_{\max}^* := \mathbb{E}_{\mathbf{h}} \left[\max_{m \leq M} f(h_m) \right], \tag{7}$$

the maximum possible. Multiuser diversity gains are thus obtained but under simplifying assumptions.

V. Stochastic approximation

The BS policies of the previous section, though yielding a TRE, are based on an artificial assumption that the BS has perfect knowledge of either the signal statistics or the (delayed) deviation of the signaled rate from the true rate. The aim there was to motivate a method to get a TRE. We now develop that idea and describe a realistic policy based on the technique of stochastic approximation (SA). Briefly, the policy works as follows. It continuously (i) estimates the average throughput that each mobile gets; (ii) estimates the excess utility that each mobile accumulates beyond its share when in a cooperative setting; (iii) applies a "correction" based on the excess utility. The resulting estimates are then used to make scheduling decisions.

The policy of a BS is now a time-varying function prescribing it's actions at every time point. The action at time k depends on mobile signals up to and including time k. Throughout this section, $\mathcal{H}_m$ is a compact subset of $\mathbb{R}$ for each $m \in [M]$. We restrict attention to stationary and memoryless policies for mobiles, i.e., $s_m : \mathcal{H}_m \to \mathcal{H}_m$ maps the current channel state $h_{m,k}$ in a deterministic and stationary fashion to $s_m(h_{m,k})$ for any slot index k. For convenience, we define $s_m(h_{m,k}) = h_{m,k}$ for the cooperative mobiles $m > M_1$. We make the following additional assumptions for mathematical tractability.

A.1 The processes $\{h_{m,k}\}_{k\geq 1}$ is an independent and identically distributed (IID) sequence for every m and is further independent across mobiles. For each m, the distribution of the random variable $h_{m,1}$ has a bounded density. The range of the function f, $f(\mathcal{H}_m)$, is bounded.

A.2 The function $f : \mathcal{H}_m \to \mathbb{R}_+$ is continuously differentiable and invertible. So is f^{-1}.

A.3 The induced random variables $s_m(h_{m,1})$ have bounded densities for each m.

Instead of **A.**2 and **A**.3, the following assumption may be used.

A.4 The induced random variables $f(s_m(h_{m,1}))$, representing the reported rates, have bounded densities for each m.

We now define the utilities of all the players. Let $\phi_{m,k}$ be the slot-level utility derived by mobile m in slot k. Obviously $0 \leq \phi_{m,k} \leq f(h_{m,k})$ and $\phi_{m,k} = 0$ if the channel is not assigned to mobile m in slot k. Then if the following limits exist, set

$$U_m = \lim_{k\to\infty} \frac{1}{k} \sum_{l\leq k} \phi_{m,l} \text{ and } U_{BS} = \sum_{m=1}^{M} U_m. \tag{8}$$

In the following we describe a BS policy which along with truthful signals from the mobiles constitutes a NE, i.e., a TRE, for the game described by the above utilities.

Under true signaling, the BS achieves its maximum utility $U^*_{\max}$ given in (7) while the mth mobile gets

$$U^*_m = \theta^0_m := \mathbb{E}_{\mathbf{h}} \left[f(h_m) 1_{\{f(h_m)\geq f(h_j) \text{ for all } j\}} \right].$$

With the information available, the BS can calculate $\Theta^0 := (\theta^0_1, \cdots, \theta^0_M)$, which we will refer henceforth by *cooperative shares*.

The (deterministic) BS policy β^* is defined using the following set of recursive updates. Let $\Delta \in \mathbb{R}_+$ be a parameter and let $\epsilon_k = 1/(k+1)$. Initialize $\theta_{m,0} = \theta^0_m$ for all m and let

$$\textbf{Policy } \beta^* : \qquad \theta_{m,k+1} = \theta_{m,k} + \epsilon_k \left(\tilde{f}_{m,k+1} I_{m,k+1} - \theta_{m,k} \right), \text{ where} \tag{9}$$

$$\tilde{f}_{m,k+1} = f(s_m(h_{m,k+1})) - \left(\theta_{m,k} - \theta^0_m\right) \Delta, \tag{10}$$

$$I_{m,k+1} = 1_{\{f(s_m(h_{m,k+1}))\geq f(s_j(h_{j,k+1})) \text{ for all } j\}} 1_{\{\tilde{f}_{m,k+1}\geq 0\}}. \tag{11}$$

In words, the BS (i) tracks average reported utility via $\theta_{m,k}$ (see (9)), (ii) computes excess utility $\theta_{m,k} - \theta^0_m$ relative to the mobiles' cooperative shares and subtracts the excess from the instantaneous signaled utility after magnification by Δ (see (10)), and (iii) uses updated values to make a current scheduling decision : the allocation distribution at time $k+1$ is $(I_{m,k+1}, m \in [M])$, a delta function (with probability 1) that places all its mass on the unique mobile with the largest reported value $f(s_m(h_{m,k+1}))$; however BS allocates only the corrected value $\tilde{f}_{m,k+1}$ to the scheduled mobile. The choice of Δ determines the proximity of the converged solution to Θ^0 (as will be seen later).

If BS schedules mobile m in slot k, the latter gets a utility

$$\bar{f}_{m,k} \;\; := \;\; \max\left\{0, \min\left\{\tilde{f}_{m,k}, f(h_{m,k})\right\}\right\} \tag{12}$$

because, if $\tilde{f}_{m,k} < 0$ for the selected mobile no transmission is made, and if $\tilde{f}_{m,k} < f(h_{m,k})$ transmission is made at a lesser rate to get a slot-level utility $\tilde{f}_{m,k}$, and if $\tilde{f}_{m,k} \geq f(h_{m,k})$ then the obtained slot-level utility is only $f(h_{m,k})$ (see section II-1 for justification of this utility). Consequently, $\phi_{m,k} = \bar{f}_{m,k} I_{m,k}$, and the limiting utility for mobile m may be written as $U_m = \lim_{k\to\infty} \bar{\theta}_{m,k}$ with $\bar{\theta}_{m,k} := \frac{1}{k}\sum_{l\leq k}\phi_{m,l}$. Incidentally, $\bar{\theta}_{m,k}$ satisfies the following recursive update equation:

$$\bar{\theta}_{m,k+1} \;\; = \;\; \bar{\theta}_{m,k} + \epsilon_k\left(\bar{f}_{m,k+1} I_{m,k+1} - \bar{\theta}_{m,k}\right) \text{ with initialization } \bar{\theta}_{m,0} = \theta^0_m. \tag{13}$$

We employ the commonly used ordinary differential equations (ODE) approximation technique (see [17] or [8]) to analyze utilities (8) and obtain optimality properties of policy β^* (9)-(11).

Let $S := (s_1, \cdots s_M)$, $\Theta_k := (\theta_{1,k}, \cdots, \theta_{M,k})$ and $\bar{\Theta}_k := (\bar{\theta}_{1,k}, \cdots, \bar{\theta}_{M,k})$. Also let $\Theta := (\theta_1, \cdots, \theta_M)$ and $\bar{\Theta} := (\bar{\theta}_1, \cdots, \bar{\theta}_M)$. Further define the following for all $m \leq M$:

$$\tilde{f}^S_m(h_m, \theta_m) \;\; := \;\; f(s_m(h_m)) - (\theta_m - \theta^0_m)\Delta \tag{14}$$

$$I^S_m(\mathbf{h}, \theta_m) \;\; := \;\; 1_{\left\{f(s_m(h_m)) \geq f(s_j(h_j)) \text{ for all } j \neq m\right\}} 1_{\left\{\tilde{f}^S_m(h_m, \theta_m) \geq 0\right\}}, \text{ and} \tag{15}$$

$$\bar{f}^S_m(h_m, \theta_m) \;\; := \;\; \max\left\{0, \min\left\{\tilde{f}^S_m(h_m, \theta_m), f(h_m)\right\}\right\}. \tag{16}$$

We will show that the trajectories of $(\Theta_k, \bar{\Theta}_k)$ given by (9) and (13), for any finite time, converge to the solution of the system of ODEs

$$\begin{aligned} \dot{\Theta}(t) &= H^S(\Theta(t)) - \Theta(t); \quad H^S_m(\Theta) = \mathbb{E}_{\mathbf{h}}\left[\tilde{f}^S_m(h_m,\theta_m) I^S_m(\mathbf{h},\theta_m)\right] \text{ with } \Theta(t_0) = \Theta^0; \quad (17)\\ \dot{\bar{\Theta}}(t) &= \bar{H}^S(\Theta(t)) - \bar{\Theta}(t); \quad \bar{H}^S_m(\Theta) = \mathbb{E}_{\mathbf{h}}\left[\bar{f}^S_m(h_m,\theta_m) I^S_m(\mathbf{h},\theta_m)\right] \text{ with } \bar{\Theta}(t_0) = \Theta^0. \quad (18) \end{aligned}$$

By Lemma 1 given below, the ODE system (17) - (18) has a unique bounded solution $(\Theta(t), \bar{\Theta}(t))$ for any finite time interval $t \in [0,T]$ where $0 < T < \infty$. Further the system has a unique global exponentially stable attractor.

Lemma 1: The function $H^S(\cdot)$ is continuously differentiable, while the function $\bar{H}^S(\cdot)$ is globally Lipschitz. The ODE system (17) - (18) thus has a unique solution. Moreover, for any strategy profile S, the ODE system has a unique global exponentially stable attractor $(\Theta^{*,S}, \bar{H}^S(\Theta^{*,S}))$ satisfying

$$\begin{aligned} |\Theta(t) - \Theta^{*,S}| &\leq |\Theta^0 - \Theta^{*,S}| e^{-t}, \quad (19)\\ |\bar{\Theta}(t) - \bar{H}^S(\Theta^{*,S})| &\leq e^{-t}\left(|\Theta^0 - \bar{H}^S(\Theta^{*,S})| + \Delta|\Theta^0 - \Theta^{*,S}|t\right) \text{ for all } t, \end{aligned}$$

where $\Theta^{*,S}$ is the unique fixed point of the function $H^S(\cdot)$.

Proof: See Appendix C-A ■

Let us define

$$\begin{aligned} t(r) &:= \sum_{k=0}^{r} \epsilon_k,\\ \Psi_k &:= \left(\Theta_k, \bar{\Theta}_k\right),\\ m(n,T) &:= \arg\max_{r \geq n}\{t(r) - t(n) \leq T\}. \end{aligned}$$

Let $\Psi(t, t_0, (\Theta_0, \bar{\Theta}_0))$ represent the solution of the ODE system (17) - (18) with $\Psi(t_0) = (\Theta_0, \bar{\Theta}_0)$. We then have the following ODE approximation theorem.

Theorem 3: Assume **A**.1-**A**.3 and $\{\epsilon_k = (k+1)^{-1}\}$. Fix any $T > 0$, any $\delta > 0$, any $\mathbf{h}$, and let $(\mathbf{h}_n, \Psi_n)$ be initialized to $(\mathbf{h}, \Phi) = (\mathbf{h}, (\Theta, \bar{\Theta}))$. Let $P_{n:\mathbf{h},\Phi}$ denote the distribution of

$\{(\mathbf{h}_{n+k}, \Psi_{n+k})\}_{k\geq 0}$ with initializations $\mathbf{h}_n = h$, $\Psi_n = \Phi$. Then, as $n \to \infty$, we have

$$P_{n:\mathbf{h},\Phi}\left\{\sup_{\{n\leq r\leq m(n,T)\}} |\Psi_r - \Psi\left(t(r),\ t(n),\ \Phi\right)| \geq \delta\right\} \to 0 \text{ uniformly for all } \Phi \in Q_1, \quad (20)$$

where Q_1 is any compact set. □

Proof: See Appendix C-C. ■

Theorem 3 approximates the trajectories on (any) bounded time interval; see (20). However for analyzing the (time) limits of the trajectories using the attractor of the ODE system, Theorem 3 is not sufficient. In [13], we study the extension of the above algorithm for generalized α-fairness, and prove the ODE approximation theorem for infinite time using weak convergence methods of Kushner et al. ([17]). The error between the tail of actual trajectory and that of the ODE trajectory is the object of study there. The algorithms (13) and (9) are special cases. In fact in [13], the algorithm is studied under much general conditions, which do not require the IID conditions. Hence the results described below are applicable under stationary channel conditions given in Section VII of [13].

We thus analyze the utilities (8) corresponding to (13) by replacing the limits of the trajectories with the unique attractor of the ODE (18). Using this, with $S = I$ representing the truth revealing strategy profile ($s_m(h_m) = h_m$ for all h_m, m), we next show below that the policy β^* along with truth revealing signals I forms an ϵ-NE (see Section II-4) and hence a TRE.

Step 1: When $S \neq I$, we claim that $U_m \leq \theta_m^0 + O(1/\Delta)$, in probability. Indeed, the unique attractor $\Theta^{*,S}$ of the ODE (17) is the unique fixed point of function $H^S(\cdot)$ (Lemma 1) and hence satisfies

$$\begin{aligned}
\theta_m^{*,S} - \theta_m^0 &= \frac{\mathbb{E}_{\mathbf{h}}\left[f(s_m(h_m))I_m^S(\mathbf{h}, \theta_m^{*,S})\right] - \theta_m^0}{1 + \Delta\mathbb{E}\left[I_m^S(\mathbf{h}, \theta_m^{*,S})\right]} \\
&\leq \frac{C_f\mathbb{E}\left[I_m^S(\mathbf{h}, \theta_m^{*,S})\right]}{1 + \Delta\mathbb{E}\left[I_m^S(\mathbf{h}, \theta_m^{*,S})\right]},
\end{aligned}$$

where C_f is the upper bound on f. It follows that $\theta_m^{*,S} \leq \theta_m^0 + O(1/\Delta)$. Further, the unique attractor of ODE (18) satisfies $\bar{\theta}_m^* = \bar{H}_m^S(\Theta^{*,S}) \leq \theta_m^{*,S}$ for all m. The time limit of the utility

U_m converges weakly to the constant $\bar{\theta}^*_m$ (the weak convergence is shown in [13]), and therefore in probability. Hence

$$U_m \to \bar{\theta}^*_m \leq \theta^{*,S}_m \leq \theta^0_m + O(1/\Delta), \tag{21}$$

where the first convergence is "in probability".

Step 2: When $S = I$, it is easy to see that the time limit U_m equals the cooperative shares in probability, i.e., $U_m = \theta^0_m$, the mth component of Θ_0 that constitutes the unique attractor for the ODE system (Θ_0, Θ_0), in probability.

Step 3: Under $S = I$, the optimal allocation strategy β^* for the BS attains the maximum possible BS utility, in probability.

The above three steps show that $S = I$ together with β^* constitutes a TRE.

We conclude this section with an interesting observation. For large values of Δ, we have

$$\begin{aligned}(\theta^{*,S}_m - \theta^0_m)\Delta &= \frac{\Delta(\mathbb{E}_{\mathbf{h}}\left[f(s_m(h_m))I^S_m(\mathbf{h}, \theta^{*,S}_m)\right] - \theta^0_m)}{1 + \Delta\mathbb{E}\left[I^S_m(\mathbf{h}, \theta^{*,S}_m)\right]} \\ &\approx \frac{\mathbb{E}_{\mathbf{h}}\left[f(s_m(h_m))I^S_m(\mathbf{h}, \theta^{*,S}_m)\right] - \theta^0_m}{\mathbb{E}\left[I^S_m(\mathbf{h}, \theta^{*,S}_m)\right]}\end{aligned}$$

which can be significant but is bounded (independently of Δ) because of the boundedness of f. If any mobile reports much larger than its true value, i.e., if $f(h_m) \ll f(s_m(h_m))$, and if in fact it is large enough that $f(h_m) \ll f(s_m(h_m)) - (\theta^{*,S}_m - \theta^0_m)\Delta$, then

$$U_m \ll \mathbb{E}_{\mathbf{h}}[(f(s_m(h_m)) - (\theta^*_m - \theta^0_m)\Delta)I^S_m(\mathbf{h}, \theta^*_m)] = \theta^*_m.$$

Hence $U_m \ll \theta^0_m$, i.e., that particular mobile's utility is much lesser than θ^0_m, its own cooperative share. *Hence a mobile that deviates more loses more* (see Figure 3).

A. Further Robustification of the SA Policy

The policy β^* (9) induces a truth revealing equilibrium. The robustness in (9) is achieved by reducing the allocation to the selected mobile, based on its deviation from the cooperative share. This however does not rule out the possibility of non-robust scheduling decisions which can result in a significant loss for other truthful mobiles, as can be seen in Figure 2(d), and in a significant loss for the BS.

In the following we propose a better variant of the policy (9) where robustness is also built into decision making, i.e, we use decisions $\{\hat{I}_{m,k}\}$ given below in place of $\{I_{m,k}\}$ of (11) :

$$\textbf{Policy } \hat{\beta}^* : \qquad \theta_{m,k+1} = \theta_{m,k} + \epsilon_k \left(\tilde{f}_{m,k+1} \hat{I}_{m,k+1} - \theta_{m,k} \right), \text{ with} \tag{22}$$

$$\hat{I}_{m,k+1} = 1_{\left\{ \tilde{f}_{m,k+1} \geq \tilde{f}_{j,k+1} \text{ for all } j \right\}} 1_{\left\{ \tilde{f}_{m,k+1} \geq 0 \right\}} \tag{23}$$

and the corresponding true utility adaptation (the actual utility obtained by the user) is given by

$$\bar{\theta}_{m,k+1} = \bar{\theta}_{m,k} + \epsilon_k \left(\bar{f}_{m,k+1} \hat{I}_{m,k+1} - \bar{\theta}_{m,k} \right) \tag{24}$$

with the initialization $\bar{\theta}_{m,0} = \theta_{m,0} = \theta_m^0$.

Using the assumptions **A.**1-**A.**3 or **A.**1 and **A.**4, and using Lemma 3 of Appendix C-D, the policy $\hat{\beta}^*$ can be analyzed in exactly the same way as we analyzed β^*. As a first step it is approximated using the ODE system

$$\dot{\Theta}(t) = \hat{H}^S(\Theta(t)) - \Theta(t); \quad \hat{H}_m^S(\Theta) = \mathbb{E}_{\mathbf{h}} \left[\tilde{f}_m^S(h_m, \theta_m) \hat{I}_m^S(\mathbf{h}, \Theta) \right]; \tag{25}$$

$$\dot{\bar{\Theta}}(t) = \hat{\bar{H}}^S(\Theta(t)) - \bar{\Theta}(t); \quad \hat{\bar{H}}_m^S(\Theta) = \mathbb{E}_{\mathbf{h}} \left[\bar{f}_m^S(h_m, \theta_m) \hat{I}_m^S(\mathbf{h}, \Theta) \right]; \tag{26}$$

$$\hat{I}_m^S(\mathbf{h}, \Theta) := 1_{\left\{ \tilde{f}_m^S(h_m,\theta_m) \geq \tilde{f}_j^S(h_j,\theta_j) \text{ for all } j \neq m \right\}} 1_{\left\{ \tilde{f}_m^S(h_m,\theta_m) \geq 0 \right\}},$$

with the initializations $(\Theta(t_0), \bar{\Theta}(t_0)) = (\Theta^0, \Theta^0)$.

However in contrast to the previous section, the question of uniqueness of the attractor for the new ODE system remains open, let alone global asymptotic stability. Step 1 of the previous section, but without the time limiting argument, holds for any (single point) attractor of the new ODE system. Similarly Step 2 and further remarks on the properties of the attractors also hold for the new ODE system. The analysis would be complete if it can be shown that under truthful strategies, i.e., $S = I$, the new ODE system has a global asymptotically stable, and hence unique, attractor. This would then show that the time limits of the utilities are indeed given by the components of the attractor. While numerical results (given in the next subsection) support this on the examples studied, a proof remains elusive.

B. Examples

We start this section with an example that reinforces our observation that ODE attractors are good approximations for time limits of almost all trajectories, under the true utility adaptations (13) or (24). In Figure 1, we consider an example with two statistically identical and cooperative mobiles. Let $f_Z(z;\sigma^2) = ze^{-z^2/2\sigma^2}, z \geq 0$ represent the density of the Rayleigh distributed random variable $Z(\sigma^2)$. The channel gains h_m of the two mobiles are conditional Rayleigh distributed, i.e., for both $m = 1, 2$, we have

$$h_m \sim f_Z(z;3)1_{\{z\leq 2\}}dz/Prob(Z(3) \leq 2).$$

The utilities as before are the achievable rates $f(h) = \log(1 + h^2)$. We plot two independent trajectories (sample paths) of the two mobiles, $\{\bar{\theta}_{m,k}\}_{m=1,2 \quad k\geq 1}$ initialized away from their cooperative shares; the initial values are set to $\theta_1^0 = \theta_2^0 = 0.456$. We set $\Delta = 100$. Figure 1(a) is for the policy $\hat{\beta}^*$ (22) while Figure 1(b) is for the policy β^* (9). All the trajectories converge close to the attractors of the ODE thus corroborating theory.

We present another example in Figure 2 to illustrate the robustness and comparison of both the BS policies. Here too we consider two statistically identical mobiles, but now with

$$h_m \sim f_Z(z;1)1_{\{z\leq 2\}}dz/Prob(Z(1) \leq 2).$$

The first mobile can be noncooperative with $s_1(h) = h + (2 - h)\delta$. We consider the policy $\hat{\beta}^*$ in Figures 2(a), (c) and the policy β^* in Figures 2(b), (d). In Figures (a)-(b) we plot trajectories corresponding to cooperative behavior ($\delta = 0$) while the curves in Figures (c)-(d) are for the case when the first mobile is noncooperative with $\delta = 0.95$ and with $\Delta = 100$.

All the cooperative curves (Figures 2 (a)-(b)) converge towards the cooperative share (which is the same for both the mobiles because they are statistically identical). The true utility of the

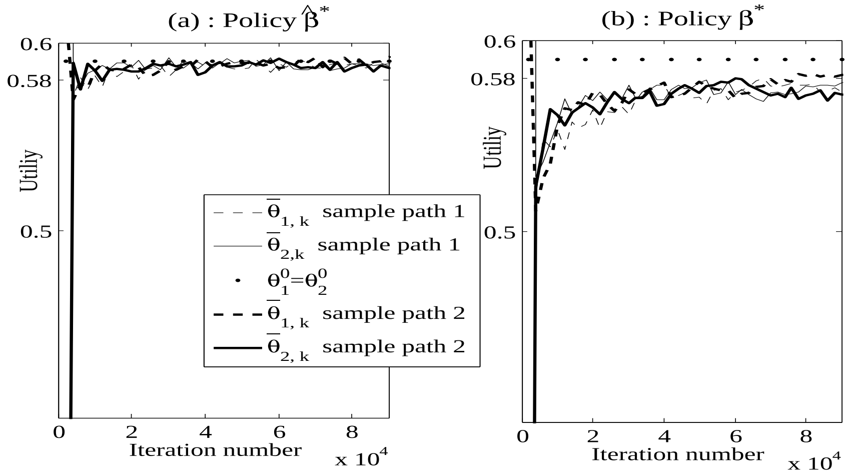

Fig. 1. Time limits. For both policies $\hat{\beta}^*$ and β^*, the ODE attractors (cooperative shares in this case) approximate the time limits of the adaptations (13) and (24). This is demonstrated via two independent set of trajectories (thick lines form one set while the thin ones form the other set). Each set has two trajectories corresponding to the true utility adaptations $\{\bar{\theta}_{m,k}\}_k, m = 1, 2$ of the two mobiles. The cooperative shares for both mobiles are equal and the dotted straight lines in both the figures represent the common cooperative share.

noncooperative mobile (mobile 1) under both the BS policies converges to a value less than the cooperative share; this confirming the theory of the previous sections. However, there are two important differences in the behavior of the two policies under noncooperation. 1) The utility of the noncooperative mobile (thin dash lines) under policy $\hat{\beta}^*$ (which is close to 0.41 in Figure 2(c)) is lesser than that under policy β^* (which is close to 0.45 in Figure 2(d)). Of course, both are less than the cooperative shares. 2) The utility of the cooperative mobile (that of mobile 2, given by thick lines) under policy $\hat{\beta}^*$ is closer to its cooperative share, but is close to zero under the policy β^*. These two observations go well with the extra robustification built into the strategy $\hat{\beta}^*$ via the decisions in (23).

We conclude this section with another example in Figure 3 to illustrate the further properties of both policies. The settings of this figure remain same as that in Figure 1, except that we now use the Rayleigh random variable $Z(0.01)$ for channel amplitude gain. We see that the more a mobile deviates from cooperative behavior, the more it loses. This is clearly visible under both policies, as the limit of the true utility of the noncooperative mobile deviates most from its cooperative share when $\delta = 0.95$. Further, the policy $\hat{\beta}^*$ penalizes the deviant mobile more than the policy β^*, and hence is more robust.

Some final remarks on the simulation results are the following. Simulation results showed that

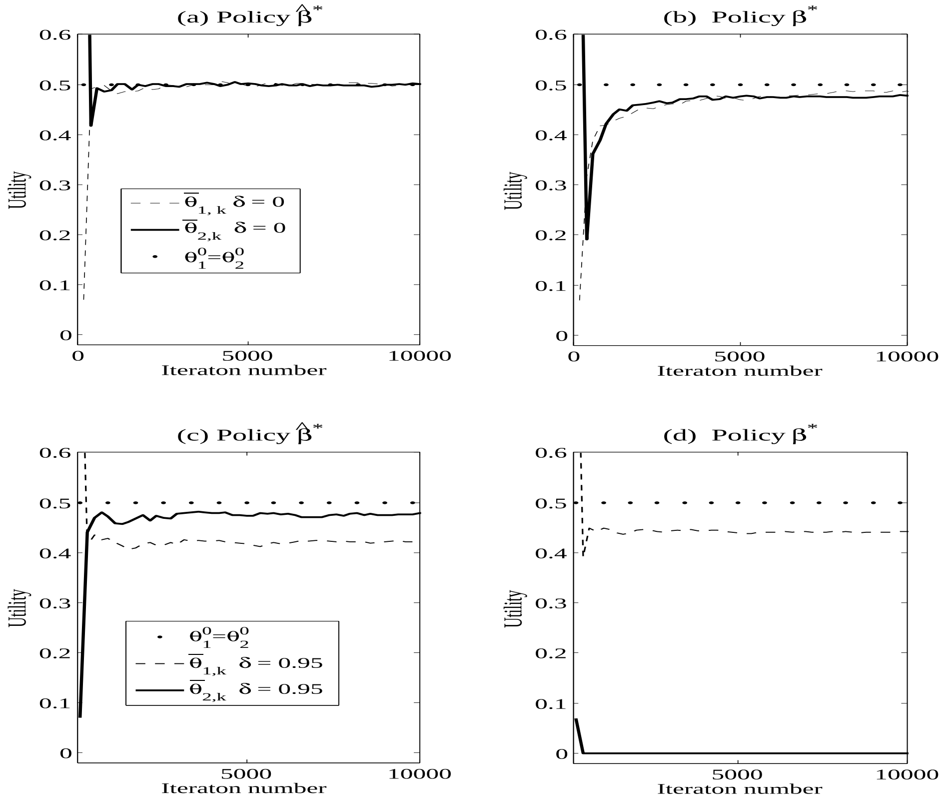

Fig. 2. Robustness and comparison of SA based BS Policies. True utility adaptations are plotted for two cases. Cooperative case (corresponding to $\delta = 0$) is given in figures (a)-(b), while the case with the first mobile being noncooperative (with $\delta = 0.95$) is given in figures (c)-(d). The cooperative shares for both the mobiles are equal and the dotted straight lines in all the figures represent the common cooperative share.

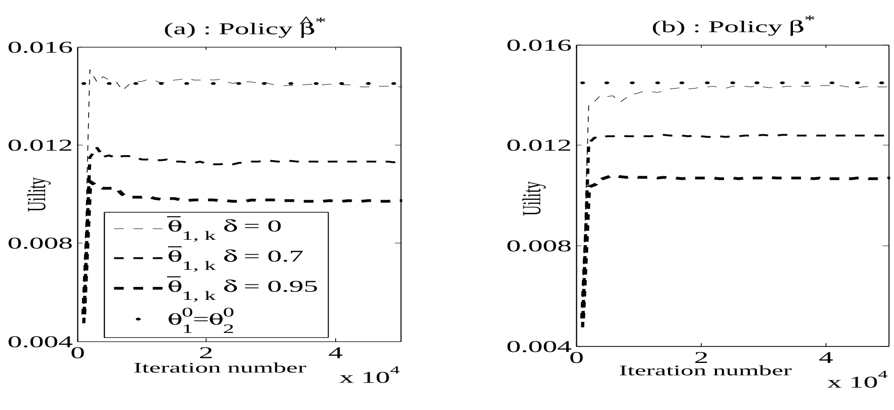

Fig. 3. More the noncooperation, the more the loss. For both policies $\hat{\beta}^*$ and β^*, the asymptotic true utility is maximum and equal to cooperative share when the mobile is cooperative ($\delta = 0$). The utility reduces as the level of noncooperation increases (i.e., as δ increases).

the reported rate trajectories $\{\theta_{m,k}\}$ (see 9) and (22) tend to the cooperative shares θ_m^0 much faster than the true rate trajectories. They however are not relevant and are not presented. In Figure 1, the step sizes are larger than those in the other two figures, and hence the curves in the latter two figures are smoother. Convergence, on the other hand, is faster in Figure 1 as one would expect.

VI. Existence of other Nash Equilibria

Thus far we obtained two types of NE. Under the first type (babbling equilibrium of Theorem 1) BS schedules using only the signals from cooperative mobiles and channel statistics of noncooperative mobiles. The BS utility is minimum among all the possible equilibrium utilities and equals U^*_{cop} given in (4). The second type (equilibria of Sections IV and V) constitutes truth revealing equilibria (TRE). BS achieved these equilibria by using ITR (incentives for truth revealing) policies. When in a TRE, the BS achieves the maximum possible equilibrium utility $U^*_{\max}$ given in (7).

Clearly $U^*_{cop} \leq U^*_{\max}$. This raises a natural question on the existence of other NE with BS's equilibrium utility taking values in the interval $[U^*_{cop}, U^*_{\max}]$. In this section we further study "predictive" policies, $\beta(.|\mathrm{s}, p_{\mathbf{S}})$, of Section IV-B and prove the existence of other NE (in Theorem 4).

A. Motivating Example continued

We first return to the motivating example of Section II to describe the main ideas.

Optimal policy for mobile 2 : The mobile uses the policy $\bar{\mu}_2$ described as follows. It declares to be in state h_2^1 yielding utility 10 when in the same state, i.e., $\bar{\mu}_2(h_2^1 \mid h_2^1) = 1$ and in addition declares with probability ρ to be in h_2^1 whenever in state h_2^2, i.e., $\bar{\mu}_2(h_2^1|h_2^2) = \rho = 1 - \bar{\mu}_2(h_2^2|h_2^2)$. Finally $\bar{\mu}_2(h_2^3|h_2^3) = 1$. Choose ρ such that the best response of BS to this policy is to allocate to mobile 2 whenever state h_2^1 is declared. For this to hold, ρ should be such that the utility of the BS is at least $U^*_{cop} = 8$ obtained by always allocating to the cooperative mobile 1. For such

ρ, the utility of BS and mobile 2 are given by

$$\begin{aligned} U_{BS} &= \frac{1}{4} \cdot 10 + \rho \cdot \frac{1}{2} \cdot 2 + (1-\rho) \cdot \frac{1}{2} \cdot 8 + \frac{1}{4} \cdot 8 \\ &= 8.5 - 3\rho \end{aligned}$$

and

$$U_2 = 2 \cdot \frac{1}{2} \cdot \rho + 10 \cdot \frac{1}{4}.$$

The ρ that maximizes the utility of mobile 2 and yet keeps the BS utility above U^*_{cop} (i.e., which satisfies constraint $8.5 - 3\rho \geq 8$) is $\rho = \frac{1}{6}$. The probability that mobile 2 declares that its channel is in state h_2^1 is $p_{S_2}(h_2^1) = \frac{1}{4} + \frac{1}{2} \cdot \frac{1}{6} = \frac{7}{24}$. Thus with the above policy $\bar{\mu}_2$ for mobile 2, BS's best response among the simple policies is to select mobile 2 whenever it declares a h_2^1. Denote it by $\bar{\beta}$. The couple $(\bar{\mu}_2, \bar{\beta})$ is not an equilibrium because mobile 2's best response against $\bar{\beta}$ is simply to declare h_2^1 always. Thus the BS should allocate channel to mobile 2 upon hearing the signal h_2^1, only if it is guaranteed a utility of 8 or more. This can be done in a way similar to that in Section IV-B by allocating the channel to mobile 2 after further verifying that the mobile 2 declares to be in h_2^1 for not more than $\frac{7}{24}$ of time. More precisely, the BS chooses the following "signal predictive" policy[2] $\hat{\beta}$: whenever mobile 2 declares h_2^1 allocate channel to mobile 2 with probability $q(p_{S_2})$ where

$$q(p_{S_2}) := \min \left\{ 1, \frac{p_{S_2}(h_2^1)}{7/24} \right\}.$$

One can verify that $(\bar{\mu}_2, \widehat{\beta})$ is an equilibrium that guarantees a rate of $8/3$ (resp. $16/3$) to mobile 2 (resp. 1) and a rate of 8 to the BS.

Infinitely many equilibria in feedback policies : In the sequel, we show that there exists a continuum of NE where the BS gets a utility greater than 8. We use the same type of policy $\bar{\mu}_2$ for mobile 2, but we choose $\rho < 1/6$. Then the probability that mobile 2 declares that it is in state 10 is $\bar{p}_{S_2}(h_2^1) = 1/4 + \rho/2$. Consider the BS policy $\tilde{\beta}$: BS selects mobile 2 with probability

[2]This policy knows a priori the signal probabilities of mobile 2 and uses it for decision making.

q whenever the mobile declares that it in state h_2^1 where

$$q = \min\Big(1, \frac{p_{S_2}(h_2^1)}{1/4 + \rho/2}\Big).$$

Thus the utilities of BS and mobile 2 are $8 + (1/2)q(1 - 6\rho)$ and $(10/4)q + q\rho$, respectively. It is easy to show that the couple $(\bar{\mu}_2, \tilde{\beta})$ is a NE for each $\rho \in [0, 1/6]$. In Figure 4, we plot the utility of BS, mobile 1 and mobile 2 at equilibrium as function of ρ.

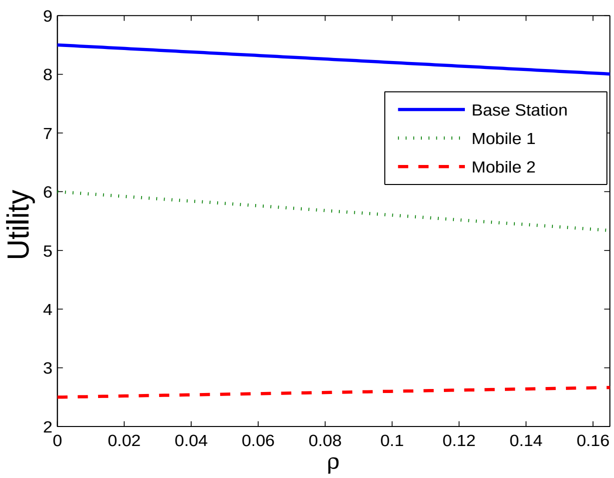

Fig. 4. Infinitely many NE. The utility of the BS, mobile 1, and mobile 2 at equilibrium as function of ρ.

B. Main result : A generalization

In this subsection we generalize the example of the previous subsection to an arbitrary number of players and states. *We assume that signal statistics of all the mobiles* $p_{\mathbf{S}}$ *is known to the BS.* Hence the BS's policy is given by $\beta(.|\mathbf{s}, p_{\mathbf{S}})$ as in Section IV-B.

Let $\mathbb{E}^{\mu_m}[f(h_m)|s_m]$ represent the conditional expectation of the mobile's utility conditioned on the signal s_m when mobile m uses strategy μ^m, i.e., for every $s_m \in \mathcal{H}_m$

$$\mathbb{E}^{\mu_m}[f(h_m) \mid s_m] \quad := \quad \sum_{h_m \in \mathcal{H}_m} \frac{p_{H_m}(h_m)\mu_m(s_m \mid h_m)}{\sum_{\tilde{h}_m \in \mathcal{H}_m} p_{H_m}(\tilde{h}_m)\mu_m(s_m|\tilde{h}_m)} f(h_m).$$

With $\mathbb{E}_{\mathbf{s}}$ representing the expectation w.r.t. $p_{\mathbf{S}}$, the payoff for mobile m is

$$U_m(\mu_m, \beta) \quad = \quad \mathbb{E}_{\mathbf{s},\mathbf{h}}\left[\beta(m|\mathbf{s}, p_{\mathbf{S}})f(h_m)\right] = \mathbb{E}_{\mathbf{s}}\left[\beta(m|\mathbf{s}, p_{\mathbf{S}})\mathbb{E}^{\mu_m}\left[f(h_m)|s_m\right]\right]. \tag{27}$$

Given a signal probability M-tuple $p_{\mathbf{S}}$, let μ^* (or more appropriately $\mu^*(p_{\mathbf{S}})$) represent what we shall call *best* mobile strategy that satisfies the condition

$$f(h^i) \geq f(h^j) \implies \mathbb{E}^{\mu_m^*}[f(h_m)|s_m = h^i] \geq \mathbb{E}^{\mu_m^*}[f(h_m)|s_m = h^j].$$

for every $h^i, h^j \in \mathcal{H}_m$ and for all $m \leq M$

Construction of μ^* **:** Consider mobile 1 without loss of generality. Let $\mathcal{H}_1 = \{h^1, h^2, \cdots, h^{N_1}\}$ and assume $f(h^1) > f(h^2) > \cdots > f(h^{N_1})$. In the following few lines we omit subscript 1 to improve readability, i.e., h_1, s_1 etc. are represented by h, s, etc. Strategy μ_1^* is defined in a iterative way as follows.

We first define $\{\mu_1^*(s = h^1|h); h \in \mathcal{H}_1\}$, i.e., the conditional probability that mobile 1 declares that it is in its best state h^1 when it is actually in state $h \in \mathcal{H}_1$. Find the minimum index j_1^* such that the probability that the channel is in one of the top j_1^* states is greater than or equal to $p_{S_1}(h^1)$, i.e., let,

$$j_1^* := \arg\min_j \left\{ \sum_{i=1}^{j} p_{H_1}(h^i) \geq p_{S_1}(h^1) \right\}.$$

Declare state h^1 whenever the true channel is one among the top $j_1^* - 1$ states, i.e., for all h with $f(h) > f(h^{j_1^*})$, set $\mu_1^*(s = h^1|h) := 1$. When $h = h^{j_1^*}$, signal the best state h^1 for a fraction of time, where the fraction is chosen so that the overall probability of the signal $s = h^1$ equals $p_{S_1}(h^1)$, i.e.,

$$\mu_1^*(s = h^1|h = h^{j_1^*}) = \frac{p_{S_1}(h^1) - \sum_{l<j_1^*} p_{H_1}(h^l)}{p_{H_1}(h^{j_1^*})}.$$

Set $\mu_1^*(s = h^1|h^i) = 0$ whenever $i > j_1^*$. Now we define $\{\mu_1^*(s = h^2|h); h \in \mathcal{H}_1\}$. Let

$$j_2^* := \arg\min_j \left\{ \sum_{l=1}^{j} p_{H_1}(h^l) - p_{S_1}(h^1) \geq p_{S_1}(h^2) \right\}.$$

The definitions below are for $j_2^* > j_1^*$; if not, one can appropriately modify the definitions. Define

$$\begin{aligned}
\mu_1^*(s = h^2 \mid h^{j_1^*}) &= 1 - \mu_1^*(s = h^1 | h^{j_1^*}), \\
\mu_1^*(s = h^2 \mid h^{i}) &= 1 \text{ whenever } j_1^* < i < j_2^*, \\
\mu_1^*(s = h^2 \mid h^{j_2^*}) &= \frac{p_{S_1}(h^1) + p_{S_1}(h^2) - \sum_{l<j_2^*} p_{H_1}(h^l)}{p_{H_1}(h^{j_2^*})} \text{ and} \\
\mu_1^*(s = h^2 \mid h) &= 0 \text{ for the remaining } h.
\end{aligned}$$

With the above,

$$\mathbb{E}^{\mu_1^*}[f(h_1)|s = h^1] \geq f(h^{j_1^*}) \geq \mathbb{E}^{\mu_1^*}[f(h_1)|s = h^2].$$

Continue in the same way to obtain

$$\mathbb{E}^{\mu_1^*}[f(h_1)|s = h^1] \geq \mathbb{E}^{\mu_1^*}[f(h_1)|s = h^2] \geq \cdots \geq \mathbb{E}^{\mu_1^*}[f(h_1)|s = h^{N_1}]. \tag{28}$$

For $m > M_1$, we set $\mu^*(h_m \mid s_m) = 1_{\{h_m = s_m\}}$. These strategies are called 'best' because mobile m gets the best payoff for masquerading a signal probability p_{S_m}. More precisely, if mobile 1 uses any other strategy μ_1 that results in the same signal probability p_{S_1} of $p_{\mathbf{S}}$ while all other mobiles use their 'best' strategy, and the BS uses the policy

$$\beta(m|\mathbf{s}, p_{\mathbf{S}}) := 1_{\{m = \arg\max_m \mathbb{E}[f(h_m)|s_m]\}},$$

then by Lemma 2 given below, we have $U_1(\mu_1^*, \beta) \geq U_1(\mu_1, \beta)$.

Lemma 2: Fix signal probabilities at $\bar{p}_{\mathbf{S}}$, consider the *best* mobile policies $\bar{\mu}^* = \{\bar{\mu}_m^*\}$ associated with $\bar{p}_{\mathbf{S}}$, and the BS policy $\bar{\beta}^*$ given by (30) and (31) below. Amongst all strategies that preserve the signaling probabilities, mobile 1's best response to $\bar{\mu}_{-1}^*$ and $\bar{\beta}^*$ is $\bar{\mu}_1^*$.

Proof: See Appendix D. ■

With the help of the best strategies we obtain the existence of other NE.

Theorem 4: For every signal probability vector $\bar{p}_{\mathbf{S}}$ with the associated best strategies $\bar{\mu}^*$, we have

$$U_{cop}^* \leq \mathbb{E}_{\mathbf{h},\mathbf{s}}\left[f\left(h_{m^*}\right)\right] \tag{29}$$

with

$$m^* := \arg \max_{1 \le m \le M} \mathbb{E}^{\bar{\mu}_m^*}[f(h_m)|s_m].$$

The ordered pair $\left(\bar{\mu}^*, \bar{\beta}^*\right)$ is a NE at which the BS obtains the right-hand side of (29) as its utility, where the feedback policy $\bar{\beta}^*$ of the BS is given by the following.

Let $\mu = (\mu_1, \cdots, \mu_{M_1})$ be any signaling policy of the mobiles and let $p_{\mathbf{S}} = \{p_{S_m}; m \le M_1\}$ be the signaling probabilities resulting from μ. Define

$$q_m(p_{\mathbf{S}}, s_m) := \min \left\{ 1, \frac{\bar{p}_{S_m}(s_m)}{p_{S_m}(s_m)} \right\}$$

for all $s_m \in \mathcal{H}_m$, and for all $m \le M_1$. For $m > M_1$ define $q_m(\cdot, \cdot) = 1$. For any given signal vector $\mathbf{s}$, define

$$m_1^*(\mathbf{s}) := \arg \max_{m \le M} \quad \mathbb{E}^{\bar{\mu}_m^*}[f(h_m)|s_m],$$

the best among all the mobiles, and

$$m_2^*(\mathbf{s}) := \arg \max_{m > M_1} \quad f(s_m),$$

the best among the cooperative mobiles. Then define $\bar{\beta}^*(m|\mathbf{s}, p_{\mathbf{S}}) = 0$ for all $m \ne m_1^*, m_2^*$, and finally

$$\bar{\beta}^*(m_1^*|\mathbf{s}, p_{\mathbf{S}}) \quad = \quad q_{m_1^*}\left(p_{\mathbf{S}}, s_{m_1^*}\right), \tag{30}$$

$$\bar{\beta}^*(m_2^*|\mathbf{s}, p_{\mathbf{S}}) \quad = \quad \left(1 - q_{m_1^*}\left(p_{\mathbf{S}}, s_{m_1^*}\right)\right). \tag{31}$$

Proof: If all the noncooperative mobiles are fixed with signaling policy $\bar{\mu}^*$ then the signaling probabilities will be given by $\bar{p}_{\mathbf{S}}$ and we have, $q_m(\bar{p}_{\mathbf{S}}, s_m) = 1$ for all $s_m \in \mathcal{H}_m$ and for all $m \le M$. Hence $\bar{\beta}^*(m|\mathbf{s}, \bar{p}_{\mathbf{S}}) = 1_{\{m = m_1^*(\mathbf{s})\}}$. From (27), the total payoff of the BS with signal probabilities fixed at $\bar{p}_{\mathbf{S}}$, when it uses some arbitrary channel allocation say $\beta(.|\mathbf{s})$, is given by

$$U_{BS} = \mathbb{E}_{\mathbf{s}} \left[\sum_{m=1}^{M} \beta(m|\mathbf{s}) \mathbb{E}^{\bar{\mu}_m^*}[f(h_m)|s_m] \right].$$

Clearly, the BS achieves the maximum with $\bar{\beta}^*$.

Assume now that the base station uses the policy $\bar{\beta}^*$. Without loss of generality assume mobile 1 unilaterally deviates from strategy $\bar{\mu}_1^*$ and signals instead using μ_1 such that the signal probabilities remain the same. Then by the Lemma 2 mobile 1 gets lesser utility than before. If now μ_1 is such that even the signal probabilities are different from $\bar{p}_{S_1}$ then the payoff of the mobile 1 is further reduced as is seen from (30) and (31), as now it is possible that $q_1(\mu_1, s_1) < 1$ for some values (note that $(1 - q_1(\mu_1, s_1))$ fraction of the time channel is allocated to a cooperative mobile) and the remaining steps are as in the proof of Lemma 2 given in the Appendix D. ■

Remarks : The above theorem establishes the existence of NE, other than TRE, at which a noncooperative mobile's utility can be greater than that at a TRE while the utility of the BS though less than that at a TRE, is greater than U^*_{cop} under noncooperation.

CONCLUDING REMARKS

We studied centralized downlink transmissions in a cellular network in the presence of noncooperative mobiles. We modeled this as a signaling game with several players serving as leaders that send signals and with the BS serving as a follower. In the absence of extra intelligence, only the babbling equilibrium is obtained where both the BS and the noncooperative players make no use of the signaling opportunities. We then proposed three approaches to obtain an efficient equilibrium (TRE), all of which required extra intelligence from the BS but resulted in the mobiles signaling truthfully. We further showed the existence of other inefficient equilibria at which a noncooperative mobile achieves a better utility than at a TRE; the BS achieves better utility than that at a babbling equilibrium but a lower value than that at a TRE.

We see several avenues open for further research on scheduling under noncooperation. We recall that we assumed that a player is either cooperative or not. What if the player can choose? Preliminary research show that there is no clear answer: it depends on the channel statistics of the player as well as that of others. Another related question is, what if the BS does not know whether a mobile cooperates or not?

The objective of throughput maximization used by base station favor a few strong users with "relatively best" channel, thereby resulting in unfair resource allocation. In [17], Kushner and Whiting studied a stochastic approximation based algorithm that achieves generalized alpha fairness. In [13], we show that this algorithm is not robust to noncooperation. We further proposed a modification of the corrective SA algorithm to make it robust once again to noncooperation and yet be fair to all users.

Appendix A
Proof of Theorem 1

By definition, at any PBE, for any $m \leq M_1$, and for any $h_m \in \mathcal{H}_m$,

$$\mu_m^*(\cdot \mid h_m) \in \arg \max_{\alpha \in \mathcal{P}(\mathcal{H}_m)} \sum_{s_m \in \mathcal{H}_m} \alpha(s_m) \left[\sum_{\mathbf{s}_{-m}} \beta^*(m \mid \mathbf{s}) \sum_{\mathbf{h}_{-m}} p_{H_{-m}}(\mathbf{h}_{-m}) \mu_{-m}^*(\mathbf{s}_{-m} | \mathbf{h}_{-m}) f(h_m) \right].$$

Since $f(h_m)$ is independent of $s_m, \mathbf{h}_{-m}, \mathbf{s}_{-m}$ (true for every m),

$$\mu_m^*(\cdot \mid h_m) \in \arg \max_{\alpha \in \mathcal{P}(\mathcal{H}_m)} \sum_{s_m \in \mathcal{H}_m} \alpha(s_m) \left[\sum_{\mathbf{s}_{-m}} \beta^*(m \mid \mathbf{s}) \sum_{\mathbf{h}_{-m}} p_{H_{-m}}(\mathbf{h}_{-m}) \mu_{-m}^*(\mathbf{s}_{-m} | \mathbf{h}_{-m}) \right], \quad (32)$$

which is independent of h_m. Thus $\mu_m^*(\cdot \mid h_m) = \mu_m^*(\cdot)$ for some probability distribution μ_m^* on $\mathcal{H}_m$, for all $m \leq M_1$, h_m, i.e., the optimal signaling policy does not depend upon h_m. However $\mu_m^*(\cdot)$ can depend on m as p_{H_m} need not be identical across mobiles.

With the above, for any $1 \leq m \leq M_1$ and for any $s_m \in \mathcal{H}_m$ with $\mu_m^*(s_m) \neq 0$, Definition 2.2 yields

$$\pi_m^*(h_m \mid s_m) = \frac{p_{H_m}(h_m)\mu_m^*(s_m)}{\sum_{h'} p_{H_m}(h')\mu_m^*(s_m)} = p_{H_m}(h_m).$$

When $\mu_m^*(s_m) = 0$, the denominator is zero, but we may set $\pi_m^*(h_m \mid s_m) = p_{H_m}(h_m)$ for such s_m. This implies that in equilibrium the posterior beliefs cannot be improved.

For any $\mathbf{s} = (s_1, \cdots, s_M)$, the first optimization in the definition of PBE becomes

$$\beta^*(\cdot \mid \mathbf{s}) \in \arg \max_{\gamma \in \mathcal{P}([M])} \left[\sum_{j=1}^{M_1} \sum_{h_j} p_{H_j}(h_j) f(h_j) \gamma(j) + \sum_{j > M_1} f(s_j)\gamma(j) \right].$$

The above optimization is independent of $\{s_m; \quad m \leq M_1\}$ and hence the optimization reduces to maximizing

$$\sum_{j=1}^{M_1} \gamma(j)\mathbb{E}\left[f(h_j)\right] + \sum_{j>M_1} \gamma(j) f(s_j),$$

justifying the definition of $\beta^*(\cdot \mid \cdot)$ in the statement of the theorem.

Since $\beta^*(\cdot \mid \mathbf{s}) = \beta^*(\cdot|s_{M_1+1}, \cdots, s_M)$ for all $\mathbf{s}$, the optimization in (32) can be rewritten as

$$\begin{aligned} \mu_m^*(\cdot \mid h_m) &\in \arg\max_{\alpha \in \mathcal{P}(\mathcal{H}_m)} \sum_{s_m \in \mathcal{H}_m} \alpha(s_m) \left[\sum_{s_{M_1+1}, \cdots, s_M} \beta^*(m \mid s_{M_1+1}, \cdots, s_M) \prod_{l>M_1} p_{H_l}(s_l) \right] \\ &= \arg\max_{\alpha \in \mathcal{P}(\mathcal{H}_m)} \left[\sum_{s_{M_1+1}, \cdots, s_M} \beta^*(m \mid s_{M_1+1}, \cdots, s_M) \prod_{l>M_1} p_{H_l}(s_l) \right], \end{aligned}$$

where the last inequality follows because the term within square brackets does not depend on s_m and $\sum_{s_m} \alpha(s_m) = 1$. The objective function is thus a constant over the variable of optimization α, and therefore $\mu_m^*(\cdot \mid h_m)$ can be any fixed $\mu_m^* \in \mathcal{P}(\mathcal{H}_m)$. This concludes the proof. ■

APPENDIX B
PROOF OF THEOREM 2

Let μ_m^* and π_m^* be as specified in the theorem statement for every m. Then,

$$\beta^*(\cdot|\mathbf{s}) \in \arg\max_{\gamma} \left[\sum_{k=1}^{M} f(s_k)\gamma(k) \right] \text{ for all } \mathbf{s}.$$

Thus $\beta^*(m|\mathbf{s})$ is any probability measure with support set $A^*(\mathbf{s})$ as specified in the theorem statement.

Fix m, h_m. With β^*, μ_{-m}^* as given in the theorem statement, we have

$$\mu_m^*(\cdot \mid h_m) \in \arg\max_{\alpha \in \mathcal{P}(\mathcal{H}_m)} \sum_{s_m \in \mathcal{H}_m} \alpha(s_m) \left[\sum_{\mathbf{s}_{-m}} p_{H_{-m}}(\mathbf{s}_{-m}) \beta^*(m|\mathbf{s}) U_m(h_m, s_m, m) \right]. \quad (33)$$

By the choice of $\{c_m\}$ and Δ as in (6), for any α, we have

$$\begin{aligned}
&\sum_{s_m} \alpha(s_m) \left[\sum_{\mathbf{s}_{-m}} p_{H_{-m}}(\mathbf{s}_{-m}) \beta^*(m|\mathbf{s}) U_m(h_m, s_m, m) \right] \\
&= \sum_{\{s_m:(s_m-h_m)^2 > c_m\}} \alpha(s_m) \left[\sum_{\mathbf{s}_{-m}} p_{H_{-m}}(\mathbf{s}_{-m}) \beta^*(m|\mathbf{s}) U_m(h_m, s_m, m) \right] \\
&\quad \sum_{\{s_m:(s_m-h_m)^2 \le c_m\}} \alpha(s_m) \left[\sum_{\mathbf{s}_{-m}} p_{H_{-m}}(\mathbf{s}_{-m}) \beta^*(m|\mathbf{s}) U_m(h_m, s_m, m) \right] \\
&\le \alpha(h_m) \left[\sum_{\mathbf{s}_{-m}} p_{H_{-m}}(\mathbf{s}_{-m}) \beta^*(m|(h_m, \mathbf{s}_{-m})) U_m(h_m, h_m, m) \right] \\
&\le \left[\sum_{\mathbf{s}_{-m}} p_{H_{-m}}(\mathbf{s}_{-m}) \beta^*(m|(h_m, \mathbf{s}_{-m})) U_m(h_m, h_m, m) \right].
\end{aligned}$$

Hence the maximum in (33) is achieved by the choice of μ_m^* given in theorem statement. This completes the proof. ∎

APPENDIX C

PROOFS OF RESULTS ON STOCHASTIC APPROXIMATION BASED POLICIES

A. Proof of Lemma 1

Define $A_j(h_m) := \{h_j : f(s_m(h_m)) \ge f(s_j(h_j))\}$. After substitution of (14) and (15) in (17), and after using the independence of h_m across mobiles, we get

$$\begin{aligned}
H_m^S(\Theta) &= \mathbb{E}_{\mathbf{h}} \left[\tilde{f}_m^S(h_m, \theta_m) I_m^S(\mathbf{h}, \theta_m) \right] \\
&= \mathbb{E}_{\mathbf{h}} \left[\tilde{f}_m^S(h_m, \theta_m) 1_{\{\tilde{f}_m^S(h_m,\theta_m) \ge 0\}} \prod_{j \ne m} \Pr\{A_j(h_m)\} \right]. \qquad (34)
\end{aligned}$$

Similarly, starting from (18), we get

$$\bar{H}_m^S(\Theta) = \mathbb{E}_{\mathbf{h}} \left[\bar{f}_m^S(h_m, \theta_m) 1_{\{\bar{f}_m^S(h_m,\theta_m) \ge 0\}} \prod_{j \ne m} \Pr\{A_j(h_m)\} \right].$$

The definitions of $\tilde{f}_m^S$ and $\bar{f}_m^S$ in (10) and (16), respectively, indicate that these are continuous and piecewise linear functions of θ_m. The maximum magnitude of the slope is Δ. It follows that the terms that depend on θ_m inside the expectations above have the following property:

$$\begin{aligned} |\tilde{f}_m^S(h_m,\theta_m)1_{\{\tilde{f}_m^S(h_m,\theta_m)\geq 0\}} - \tilde{f}_m^S(h_m,\theta'_m)1_{\{\tilde{f}_m^S(h_m,\theta'_m)\geq 0\}}| &\leq \Delta|\theta_m - \theta'_m| \\ |\bar{f}_m^S(h_m,\theta_m)1_{\{\tilde{f}_m^S(h_m,\theta_m)\geq 0\}} - \bar{f}_m^S(h_m,\theta'_m)1_{\{\tilde{f}_m^S(h_m,\theta'_m)\geq 0\}}| &\leq \Delta|\theta_m - \theta'_m|. \end{aligned}$$

Since this is true for each m, the global Lipschitz property of H^S and $\bar{H}^S$ is an immediate consequence.

Observe next that $\tilde{f}_m^S(h_m,\theta_m)1_{\{\tilde{f}_m^S(h_m,\theta_m)\geq 0\}}$ are almost surely differentiable, thanks to assumption **A.**2. Moreover, assumption **A.**3. is that the distribution of the induced random variable $f(s(h_m))$ has bounded density. These two observations and the dominated convergence theorem imply that the expectation in the right-hand side of (34) is differentiable. Thus H^S is differentiable.

We next address the global exponential stability property of the solution to the ODE system. Fix S. It is easy to see that for any m, the function $H_m^S(\cdot)$ depends on Θ only through θ_m. Furthermore, it is non-increasing in θ_m because each of the functions $\tilde{f}_m(h_m,\theta_m)$ and $1_{\{\tilde{f}_m(h_m,\theta_m)\geq 0\}}$ are non-increasing in θ_m for each h_m. Consequently,

$$\left(H_m^S(\theta_m) - H_m^S(\theta'_m)\right)(\theta_m - \theta'_m) \leq 0$$

for any θ_m and θ'_m, and therefore

$$\left\langle H^S(\Theta) - H^S(\Theta'),\ \Theta - \Theta' \right\rangle \leq 0 \tag{35}$$

for any Θ and Θ'.

We claim that the function H^S is bounded. To see this, consider (17). From the definition of $I_m^S(\mathbf{h},\theta_m)$ in (15), we have that $\tilde{f}_m^S(h_m,\theta_m)I_m^S(\mathbf{h},\theta_m)$ is bounded below by 0. We now show an upper bound. The cooperative share θ_m^0 is easily seen to be bounded between 0 and the bound on f. Furthermore, the iteration in (9) is such that the iterates are always nonnegative, and therefore θ_m may be restricted to nonnegative values. It follows that $\tilde{f}_m^S(h_m,\theta_m)$ in (14) is upper bounded by twice the bound on f. The same bounds hold for its expectation, yielding that H^S is bounded.

Since H^S is continuous and bounded, the Brouwer fixed point theorem[3] yields that H^S has a fixed point Θ^* in the positive Orthant. This may of course depend on the strategy profile S.

Define the error $\mathbf{e}(t) = \Theta(t) - \Theta^*$ (with $\mathbf{e} = [e_1, \cdots, e_M]^T$). Then, we can write $\dot{\mathbf{e}} = g(\mathbf{e})$, where

$$g(\mathbf{e}) := H^S(\mathbf{e} + \Theta^*) - (\mathbf{e} + \Theta^*) = H^S(\mathbf{e} + \Theta^*) - H^S(\Theta^*) - \mathbf{e}.$$

From (35), we get

$$\left\langle \dot{\mathbf{e}}, \mathbf{e} \right\rangle = \left\langle H^S(\mathbf{e} + \Theta^*) - H^S(\Theta^*), \quad \mathbf{e} \right\rangle - |\mathbf{e}|^2 \leq -|\mathbf{e}|^2.$$

A standard argument (see [23, pp. 169-170]) then shows that $|\mathbf{e}(t)| \leq |\mathbf{e}(0)| e^{-t}$ for all $t \geq 0$, i.e.,

$$|\Theta(t) - \Theta^*| \leq |\Theta(0) - \Theta^*| e^{-t} \text{ for all } t \geq 0. \tag{36}$$

We shall have occasion to use this argument a few times in the sequel. It follows from (36) that Θ^* is the unique fixed point for H^S, and is a global exponentially stable attractor for the ODE (17).

Let us turn to the actual rate trajectory $\bar{\Theta}(t)$. Define the error for the actual rate trajectory as $\bar{e}_m(t) := \bar{\theta}_m(t) - \bar{H}^S_m(\Theta^*)$ for all m, and the actual rate error vector as $\bar{\mathbf{e}} = [\bar{e}_1, \cdots, \bar{e}_M]^T$. Then

$$\dot{\bar{\mathbf{e}}}\,(t) = \bar{H}^S(\Theta(t)) - \bar{H}^S(\Theta^*) - \bar{\mathbf{e}}(t).$$

Using the Cauchy-Schwarz inequality, the global Lipschitz property of $\bar{H}^S$, with say the Lipschitz constant L, and the upper bound (36), we obtain

$$\begin{aligned}
\left\langle \dot{\bar{\mathbf{e}}}\,(t), \bar{\mathbf{e}}(t) \right\rangle &= \left\langle \bar{H}^S(\Theta(t)) - \bar{H}^S(\Theta^*) - \bar{\mathbf{e}}(t), \bar{\mathbf{e}}(t) \right\rangle \\
&\leq L\,|\Theta(t) - \Theta^*|\,|\bar{\mathbf{e}}(t)| - |\bar{\mathbf{e}}(t)|^2 \\
&\leq L\,|\Theta(0) - \Theta^*|\,e^{-t}|\bar{\mathbf{e}}(t)| - |\bar{\mathbf{e}}(t)|^2.
\end{aligned}$$

[3] *Brouwer fixed point theorem:* Every continuous function f from a closed ball of a Euclidean space to itself has a fixed point, i.e., an x^* that satisfies $x^* = f(x^*)$.

By the standard argument in [23, pp. 169-170], we have $|\bar{\mathbf{e}}(t)| \leq \bar{k}(t)$, where $\bar{k}(t)$ is the solution of the ODE

$$\dot{\bar{k}}\,(t) = L\,|\Theta(0) - \Theta^*|\,e^{-t} - \bar{k}(t)$$

with the initial condition $\bar{k}(0) = |\bar{\mathbf{e}}(0)|$. It is easy to verify that the solution to the above ODE is

$$\bar{k}(t) = e^{-t}\left(\bar{k}(0) + L\,|\Theta(0) - \Theta^*|\,t\right).$$

Thus

$$|\bar{\Theta}(t) - \bar{H}^S(\Theta^*)| \leq e^{-t}\left(|\bar{\Theta}(0) - \bar{H}^S(\Theta^*)| + L|\Theta(0) - \Theta^*|t\right),$$

and hence $\bar{H}^S(\Theta^*)$ is the unique global asymptotically stable attractor of the ODE (18). This concludes the proof. ■.

B. The ODE Approximation Theorem

Benveniste et al. obtain the ODE approximation [8, Th.9, p.232] for the system

$$\Psi_{k+1} \quad = \quad \Psi_k + \mu_k \quad Z(\Psi_k, G_{k+1}). \tag{37}$$

We reproduce the result here in a form suitable for use in this paper.

Let Ψ take values in an open subset D of $\mathcal{R}^m$. We make the following assumptions:

B.0 $\{\mu_k\}$ is a decreasing sequence with $\sum_k \mu_k = \infty$ and $\sum_k \mu_k^{1+\delta} < \infty$ for some $\delta > 0$.

B.1 There exists a family $\{P_\psi\}$ of transition probabilities $P_\psi(g, \mathbf{A})$ such that, for any Borel set $\mathbf{A}$,

$$P[G_{n+1} \in \mathbf{A}|\mathcal{F}_n] = P_{\Psi_n}(G_n, \mathbf{A})$$

where $\mathcal{F}_k := \sigma(\Psi_0, G_0, G_1, \cdots, G_k)$. Thus $\{G_k, \Psi_k\}$ forms a Markov chain.

B.2 For any compact $Q \subset D$, there exist C_1, q_1 such that $|Z(\psi, g)| \leq C_1(1 + |g|^{q_1})$ uniformly for all $\psi \in Q$.

B.3 There exists a function z on D, and for each $\psi \in D$ a function $\nu_\psi(\cdot)$ such that the following hold:

a) z is locally Lipschitz on D.

b) $(I - P_\psi)\nu_\psi(g) = Z(\psi, g) - z(\psi)$ where $P_\psi \nu_\psi(g) = \mathbb{E}[\nu_\psi(G_1)|G_0 = g, \Psi_0 = \psi]$.

c) For all compact subsets Q of D, there exist constants C_3, C_4, q_3, q_4 and $\lambda \in [0.5, 1]$, such that for all $\psi, \psi^{'} \in Q$

(i) $|\nu_\psi(g)| \leq C_3(1 + |g|^{q_3})$,

(ii) $|P_\psi \nu_\psi(g) - P_{\psi^{'}} \nu_{\psi^{'}}(g)| \leq C_4 \left(1 + |g|^{q_4}\right) \left|(\psi) - (\psi^{'})\right|^{\lambda}$.

B.4 For any compact set Q in D and for any $q > 0$, there exists a $\mu_q(Q) < \infty$, such that for all n, g, and $\psi \in \mathcal{R}^d$ (with $\mathbb{E}_{g,\psi}$ representing the expectation taken with $(G_0, \Psi_0) = (g, \psi)$),

$$\mathbb{E}_{g,\psi} \left\{ I_{\{\Psi_k \in Q; \forall k \leq n\}} \left(1 + |G_{n+1}|^q\right) \right\} \quad \leq \quad \mu_q(Q) \left(1 + |g|^q\right).$$

Define $t(r) := \sum_{k=0}^{r} \mu_k$ and $m(n,T) := \arg\max_{r \geq n} \{t(r) - t(n) \leq T\}$. Let $\Psi(t, t_0, \psi)$ represent a solution of

$$\dot{\Psi}\ (t) = z(\Psi(t)),$$

with initial condition $\Psi(t_0) = \psi$. Let Q_1 and Q_2 be any two compact subsets, such that $Q_1 \subset Q_2$ and such that we can choose a $T > 0$ and a $\delta_0 > 0$ satisfying

$$d\left(\Psi(t, 0, \psi), Q_2^c\right) \geq \delta_0, \tag{38}$$

for all $\psi \in Q_1$ and all $t \in [0, T]$. Let $P_{n:g,\psi}$ denote the distribution of $\{(G_{n+k}, \Psi_{n+k})\}_{k \geq 0}$ with $G_n = g, \Psi_n = \psi$. We then have the following theorem:

Theorem 5: Assume **B**.0–**B**.4. Pick Q_1, Q_2, T and δ_0 satisfying (38). Then for all $\delta \leq \delta_0$, for any (ψ, g) and when (G_n, Ψ_n) is initialized with (g, ψ),

$$P_{n:g,\psi} \left\{ \sup_{\{n \leq r \leq m(n,T)\}} \left|\Psi_r - \Psi\left(t(r),\ t(n),\ \psi\right)\right| \geq \delta \right\} \to 0 \text{ as } n \to \infty,$$

uniformly for all $\psi \in Q_1$. $\square$

C. *Proof of Theorem 3*

Results of Benveniste et al. ([8]) (see Appendix C-B) are used in this proof.

Let $G_k := \mathbf{h}_k = (h_{1,k}, h_{2,k}, \cdots, h_{M,k})$. The equations (9) and (13) can be rewritten in the format of Benveniste et al. ([8]; see Appendix C-B) as follows. Let us define

$$\begin{aligned} W_m(G_{k+1}, \Theta_k) &= \left(\tilde{f}_{m,k+1} I_{m,k+1} - \theta_{m,k}\right), \\ \bar{W}_m(G_{k+1}, \Theta_k, \bar{\Theta}_k) &= \left(\bar{f}_{m,k+1} I_{m,k+1} - \bar{\theta}_{m,k}\right). \end{aligned}$$

Then

$$\Theta_{k+1} = \Theta_k + \epsilon_k W(G_{k+1}, \Theta_k) \tag{39}$$

$$\bar{\Theta}_{k+1} = \bar{\Theta}_k + \bar{\epsilon}_k \bar{W}(G_{k+1}, \Theta_k, \bar{\Theta}_k). \tag{40}$$

We obtain the proof using Theorem 5 of Appendix C-B and towards this, we first verify the assumptions **B.**0 - **B.**4 of Appendix C-B. Let, $\Psi := (\Theta, \bar{\Theta})$ and $Z = (W, \bar{W})$. By assumption **A.**1, G_k is a IID sequence and hence $P_\Psi(g, .) = P(.)$, with P the probability distribution of G_1 for all Ψ and all initial conditions g. Thus assumption **B.**1 is satisfied. By boundedness of function f, variables $\tilde{f}$, $\bar{f}$ are also bounded and hence the assumption **B**.2 is satisfied. For assumption **B.**3, we define the following four quantities, which readily satisfy assumption **B.**3(b):

$$\begin{aligned} w(\Theta) &:= \mathbb{E}_{\mathbf{h}}[W(G, \Theta)], \\ \bar{w}(\Theta, \bar{\Theta}) &:= \mathbb{E}_{\mathbf{h}}[\bar{W}(G, \Theta, \bar{\Theta})] = \bar{H}^S(\Theta) - \bar{\Theta}, \\ \nu_\Psi(G) = \nu_\Theta(G) &:= W(G, \Theta) - w(\Theta), \\ \bar{\nu}_\Psi(G) = \bar{\nu}_{\Theta,\bar{\Theta}}(G) &= \bar{W}(G, \Theta, \bar{\Theta}) - \bar{w}(\Theta, \bar{\Theta}). \end{aligned}$$

By Lemma 1, the functions $z := (w, \bar{w})$ satisfy assumption **B**.3(a). Assumption **B**.3(c).(i) is satisfied by boundedness while the assumption **B**.3(c).(ii) is satisfied because $P_\psi \nu_\psi = 0$ and $P_\psi \bar{\nu}_\psi = 0$. Assumption **B**.4 is satisfied again by boundedness.

By Lemma 1, the two ODE system have bounded solution for any finite time. Hence, for any compact set Q_1 and any finite T, we can find a compact Q_2 that satisfies assumption (38) of Appendix C-B for all $t \leq T$ and all $\Psi \in Q_1$. Thus the theorem follows from Theorem 5 of Appendix C-B. ■

D. The Modified SA Policy Analysis

The following Lemma establishes the properties needed to show that the ODE system (25) - (26) has a unique solution.

Lemma 3: The function $\hat{H}^S(\cdot)$ in (25) is continuously differentiable, and the function $\hat{\hat{H}}^S(\cdot)$ in (26) is locally Lipschitz. Consequently, the ODE system (25) - (26) has a unique solution for any finite time. Furthermore, the solution is bounded uniformly for any finite time.

Proof: The initial part of the proof proceeds as in proof of Lemma 1 in Appendix C-A, with some modifications. Define the set

$$A_j(h_m, \Theta) := \left\{h_j : f(s_m(h_m)) - f(s_j(h_j)) \geq \Delta(\theta_m - \theta_j - \theta_m^0 + \theta_j^0)\right\}.$$

By independence of h_m across the mobiles,

$$\hat{H}_m^S(\Theta) = \mathbb{E}_{\mathbf{h}}\left[\tilde{f}_m^S(h_m, \theta_m)\hat{I}_m^S(\mathbf{h}, \Theta)\right] = \mathbb{E}_{h_m}\left[\tilde{f}_m^S(h_m, \theta_m)1_{\left\{\tilde{f}_m^S(h_m,\theta_m)\geq 0\right\}}\Pi_{j\neq m}\Pr(A_j(h_m, \Theta))\right].$$

The first part of the lemma follows from the bounded convergence theorem if we show that the functions $\{\Pr(A_j(h_m, \Theta))\}_{j\neq m}$ and

$$\tilde{f}_m^S(h_m, \theta_m)1_{\left\{\tilde{f}_m^S(h_m,\theta_m)\geq 0\right\}}$$

are, almost surely, continuously differentiable (with respect to Θ) with uniformly bounded derivatives, for almost all h_m. This is immediately evident for $\tilde{f}_m^S 1_{\left\{\tilde{f}_m^S\geq 0\right\}}$ by assumptions **A.**2 - **A.**3. The same holds for $\{\Pr(A_j(h_m, \Theta))\}_{j\neq m}$ by assumptions **A.**2 - **A.**3 because

$$\begin{aligned}\frac{\partial \Pr(A_j(h_m, \Theta))}{\partial \theta_l} &= (-1)^{\delta(l-m)} g_{s_j}\left(f^{-1}\left(f(s_m(h_m)) - \Delta\left(\theta_m - \theta_j - \theta_m^0 + \theta_j^0\right)\right)\right) \\ &\quad \cdot f^{-1\prime}\left(f(s_m(h_m)) - \Delta\left(\theta_m - \theta_j - \theta_m^0 + \theta_j^0\right)\right)\Delta \end{aligned} \quad (41)$$

for $l = m, j$, where g_{s_j} is the (bounded) density of signal $s_j(h_j)$. In the above the continuous derivative $f^{-1\prime}$ will also be uniformly bounded for all Θ coming from a compact set, because of the boundedness of f. It is easy to see that one can achieve the result instead using assumption **A.**4 instead of **A.**2 - **A.**3. Since

$$\bar{f}_m^S(h_m, \theta_m) - \bar{f}_m^S(h_m, \theta'_m) \leq \Delta \left|\theta_m - \theta'_m\right|,$$

with C_f representing the upper bound on f, we have

$$\begin{aligned}
&\hat{\bar{H}}_m^S(\Theta) - \hat{\bar{H}}_m^S(\Theta') \\
&\quad = \mathbb{E}_{\mathbf{h}}\left[\left(\bar{f}_m^S(h_m, \theta_m) - \bar{f}_m^S(h_m, \theta'_m)\right) \hat{I}_m^S(\mathbf{h}, \Theta)\right] + \mathbb{E}_{\mathbf{h}}\left[\bar{f}_m^S(h_m, \theta'_m)\left(\hat{I}_m^S(\mathbf{h}, \Theta) - \hat{I}_m^S(\mathbf{h}, \Theta')\right)\right] \\
&\quad \leq \Delta|\theta_m - \theta'_m|\mathbb{E}_{\mathbf{h}}\left[\hat{I}_m^S(\mathbf{h}, \Theta)\right] + C_f \mathbb{E}_{\mathbf{h}}\left|\hat{I}_m^S(\mathbf{h}, \Theta) - \hat{I}_m^S(\mathbf{h}, \Theta')\right|.
\end{aligned}$$

Let $\mathcal{D}$ be any compact set. Let C_{gS} represent the common upper bound on density g_{s_m} for all m, and let $C_{f^{-1}}(\mathcal{D})$ represent the upper bound (which is independent of $\mathbf{h}$) on $f^{-1\prime}$ for all $\Theta \in \mathcal{D}$. From (41), it follows that there is a constant $C(M)$ such that

$$\mathbb{E}_{\mathbf{h}}\left|\hat{I}_m^S(\mathbf{h}, \Theta) - \hat{I}_m^S(\mathbf{h}, \Theta')\right| \leq C(M) C_{gS} C_{f^{-1}}(\mathcal{D}) \Delta \left|\Theta - \Theta'\right|$$

and hence $\hat{\bar{H}}^S(\Theta)$ is locally Lipschitz. From standard results in ODE systems [23, pp. 169-170], it follows that the system of ODEs has a unique solution for any finite time. Furthermore, the solution is bounded for any finite time. Indeed, one has

$$\begin{aligned}
\left\langle \dot{\Theta}(t), \Theta \right\rangle &\leq C_f|\Theta| - |\Theta|^2, \\
\left\langle \dot{\bar{\Theta}}(t), \bar{\Theta} \right\rangle &\leq C_f|\bar{\Theta}| - |\bar{\Theta}|^2,
\end{aligned}$$

with C_f being the upper bound on function f, and hence,

$$\begin{aligned}
|\Theta(t)| &\leq |\Theta_0| e^{-t} + C_f, \\
|\bar{\Theta}(t)| &\leq |\bar{\Theta}_0| e^{-t} + C_f.
\end{aligned}$$

This concludes the proof of the lemma. ■

Appendix D
Proof of Lemma 2

Under the given hypothesis, the BS's policy from (30) and (31) would be $1_{\{m=m_1^*(\mathbf{s})\}}$. From (27), the payoff for mobile 1 under policy μ_1 is

$$U_1(\mu_1, \bar{\beta}^*) = \mathbb{E}_{\mathbf{s}}\left[1_{\{m=m_1^*(\mathbf{s})\}}\mathbb{E}^{\mu_1}\left[f(h_1) \mid s_1\right]\right].$$

Define the probability measure

$$\tilde{p}(u) := \Pr\left(\mathbf{s}_{-1} : u = \max_{1<m\leq M} \mathbb{E}^{\bar{\mu}_m^*}[f(h_m) \mid s_m]\right)$$

for all possible u resulting from $\bar{\mathcal{P}}_s$ and $\{\bar{\mu}_m^*; m > 1\}$. Also define

$$\bar{E}(s) := \mathbb{E}^{\bar{\mu}_1^*}[f(h_1)|s_1 = s] = \sum_{\tilde{h}\in\mathcal{H}_1} \frac{\bar{\mu}_1^*(s|\tilde{h})p_{H_1}(\tilde{h})}{\bar{p}_{S_1}(s)} f(\tilde{h}).$$

By independence, we are interested in the constrained optimization problem:

$$\begin{aligned}
\max_{\{\mu_1\}} \quad & \sum_{\tilde{u}} \sum_{s\in\mathcal{H}_1} \tilde{p}(\tilde{u}) 1_{\{\bar{E}(s)>\tilde{u}\}} \sum_{h\in\mathcal{H}_1} p_{H_1}(h)\mu_1(s|h)f(h) \\
\text{subject to} \quad & \sum_{s\in\mathcal{H}_1} \mu_1(s|h) = 1 \text{ for all } h \in \mathcal{H}_1, \\
& \sum_{h\in\mathcal{H}_1} p_{H_1}(h)\mu_1(s|h) = \bar{p}_{S_1}(s) \text{ for all } s \in \mathcal{H}_1.
\end{aligned}$$

If $\mathcal{H}_1 = \{h^1, h^2, \cdots, h^{N_1}\}$ with $f(h^1) \geq f(h^2) \cdots \geq f(h^{N_1})$, then clearly by definition of the best strategies $\bar{\mu}^*$, we have

$$\bar{E}(h^1) \geq \bar{E}(h^2) \geq \cdots \geq \bar{E}(h^{N_1}). \tag{42}$$

Define $r(h,s) := p_{H_1}(h)\mu_1(s|h)$. One can rewrite the objective function in the above optimization problem as

$$\sum_s \sum_h r(h,s) f(h) \left(\sum_{\tilde{u}} 1_{\{\bar{E}(s)>\tilde{u}\}} \tilde{p}(\tilde{u}) \right). \tag{43}$$

From (42), we have

$$\left(\sum_{\tilde{u}} 1_{\{\bar{E}(h^1)>\tilde{u}\}} \tilde{p}(\tilde{u}) \right) \geq \left(\sum_{\tilde{u}} 1_{\{\bar{E}(h^2)>\tilde{u}\}} \tilde{p}(\tilde{u}) \right) \geq \cdots \geq \left(\sum_{\tilde{u}} 1_{\{\bar{E}(h^{N_1})>\tilde{u}\}} \tilde{p}(\tilde{u}) \right).$$

Hence, the maximum of the objective function in (43) is achieved under the required constraints by first maximizing the term (the constraints on this alone will be less strict) $\sum_h r(h,h^1) f(h)$ subject to

$$\sum_h r(h,h^1) = \bar{p}_{S_1}(h^1), r(h,h^1) \leq p_{H_1}(h) \text{ for all } h,$$

to obtain $\{r^*(h,h^1), \mu_1^*\}$, and then maximizing the term (while ensuring that these variables and the optimal variables from the previous step jointly satisfy the required constraints)

$$\sum_h r(h,h^2) f(h) \text{ subject to } \sum_h r(h,h^2) \;=\; \bar{p}_{S_1}(h^2) \text{ and } r(h,h^2) \leq (1-\mu_1^*(h,h^1)) p_{H_1}(h),$$

for all h, and so on. Assuming condition $\bar{p}_{S_1}(h^1) \geq p_{H_1}(h^1)$, it is easily seen that,

$$\begin{aligned}
r^*(h^1,h^1) &= p_{H_1}(h^1), \\
r^*(h^2,h^1) &= \min\{(p_S(h^1) - p_{H_1}(h^1)), \quad p_{H_1}(h^2)\}, \\
r^*(h^k,h^1) &= \max\left\{0, \min\left\{ \left(p_S(h^1) - \sum_{l<k} p_{H_1}(h^l) \right), \quad p_{H_1}(h^k) \right\}\right\} \text{ for all } k>2.
\end{aligned}$$

Further,

$$\begin{aligned}
r^*(h^1,h^2) &= 0, \\
r^*(h^2,h^2) &= \min\{p_S(h^2), (1-p^*_{h,s}(h^2,h^1)) p_{H_1}(h^2)\}, \\
r^*(h^k,h^2) &= \min\left\{ p_S(h^2) - \sum_{l<k} r^*(h^l,h^2), \;\; (1-p^*_{h,s}(h^k,h^1)) p_{H_1}(h^k) \right\} \text{ for all } k>2.
\end{aligned}$$

The above defines the joint distribution $r^*(h_1, s_1)$ with prescribed marginals. It is now straightforward to see that the conditional distribution $r^*(h_1, s_1)/p_{H_1}(h_1)$ is indeed $\mu_1^*(s_1|h_1)$. This completes the proof. ■

REFERENCES

[1] Qualcomm, Inc., "1xEV:1x Evolution IS-856 TIA/EIA Standard Airlink Overview", Nov. 2001.

[2] 3GPP TS 25.308, Technical Specification 3rd Generation Partnership Project; Technical Specification Group Radio Access Network; High Speed Downlink Packet Access (HSDPA); *Overall description*; Stage 2, Release 8, V8.3.0, Sept. 2008.

[3] S. Mare, D. Kotz, and A. Kumar, "Experimental validation of analytical performance models for IEEE 802.11 networks," *Proceedings of the Workshop on WIreless Systems: Advanced Research and Development (WISARD 2010)*, pp. 1-8, January, 2010. IEEE Computer Society Press

[4] G. Bianchi, A. Di Stefano, C. Giaconia, L. Scalia, G. Terrazzino, and I. Tinnirello, "Experimental Assessment of the Backoff Behavior of Commercial IEEE 802.11b Network Cards," in Proc. of the IEEE INFOCOM 2007, May 2007, pp. 1181-1189.

[5] R. Agrawal, A. Bedekar, R. J. La, and V. Subramanian, "Class and channel condition based weighted proportional fair scheduler," in *Teletraffic Engineering in the Internet Era, Proc. ITC-17*, S. da Bahia, J. M. de Souza, N. L. S. da Fonseca, and E. A. de Souza e Silva, Eds. Amsterdam, The Netherlands: North-Holland, 2001, pp. 553-565.

[6] D. M. Andrews, K. Kumaran, K. Ramanan, A. L. Stolyar, R. Vijayakumar, and P. A. Whiting, "Scheduling in a queueing system with asynchronously varying service rates," *Prob. Eng. Inf. Sci.*, vol. 18, pp. 191-217, 2004.

[7] P. Bender, P. Black, M. Grob, R. Padovani, N. Sindhushayana, and A. Viterbi, "CDMA/HDR: A bandwidth-efficient high-speed wireless data service for nomadic users," *IEEE Commun. Mag.*, vol. 38, no. 7, pp. 70-77, July 2000.

[8] A. Benveniste, M. Métivier and P. Priouret, *Adaptive algorithms and stochastic approximation*. Springer-Verlag, April 1990.

[9] S. C. Borst and P. A. Whiting, "Dynamic rate control algorithms for HDR throughput optimization," in *Proc. of the IEEE INFOCOM*, 2001, pp. 976-985.

[10] E. F. Chaponniere, P. J. Black, J. M. Holtzman, and D. N. C. Tse, "Transmitter Directed code division multiple access system using path diversity to equitably maximize throughput," U.S. Patent 6,449,490, Sep. 10, 2002.

[11] J. Farrell, "Meaning and credibility in cheap-talk games," *Games and Economic Behavior*, vol. 5, pp. 514-531, 1993.

[12] D. Garg and Y. Narahari, "Foundations of Mechanism Design," *Technical Report*, Department of Computer Science and Automation, Indian Institute of Science, Bangalore, India, Nov. 2006. Weblink: `http://lcm.csa.iisc.ernet.in/gametheory/survey/md_nov06.pdf`

[13] V. Kavitha, E. Altman, R. El-Azouzi and R. Sundaresan, "Fair scheduling in cellular systems in the presence of noncooperative mobiles," *Proc. of the IEEE INFOCOM 2010*, Mar. 15-19, 2010, San Diego, USA.

[14] F. P. Kelly, "Charging and rate control for elastic traffic," *Euro. Trans. on Telecommun.*, vol. 8, pp. 33-37, 1997.

[15] Z. Kong, Y-K. Kwok, and J. Wang, "On game theoretic rate-maximizing packet scheduling in non-cooperative wireless networks," in *Proc. IEEE Int. Symp. on a World of Wireless, Mobile and Multimedia Networks 2007*, WoWMoM 2007, pp. 1-4, June 2007.

[16] D. M. Kreps and J. Sobel, "Signaling", Chapter 25 in *Handbook of Game Theory*, vol. 2, R. J. Aumann and S. Hart, Elsevier Science B.V., 1994, pp. 849-867.

[17] H. J. Kushner, P. A. Whiting, "Convergence of proportional-fair sharing algorithms under general conditions," *IEEE Trans. Wireless Commun.* vol. 3, no. 4, pp. 1250-1259, Jul. 2004.

[18] H. Kushner and G. Yin, *Stochastic Approximation Algorithms and Applications*. Springer-Verlag New York, 1997.

[19] X. Liu, E. K. P. Chong, and N. B. Shroff, "A framework for opportunistic scheduling in wireless networks," *Comput. Netw.*, vol. 41, pp. 451-474, 2003.

[20] Novatel, Novatel Merlin u870 PC Card, http://www. novatelwireless.com/products/merlin/merlin-u870.html, 2008.

[21] P. Nuggehalli, M. Sarkar, K. Kulkarni, R. R. Rao, "A game-theoretic analysis of QoS in wireless MAC," *Proc. of the IEEE INFOCOM 2008*, pp. 1903-1911, Apr. 2008.

[22] A. Ozdaglar and R. Srikant, "Incentives and pricing in communications networks," Chapter 22 in *Algorithmic Game Theory*, N. Nisan, T. Roughgarden, E. Tardos, and V. V. Vazirani, pp. 571-591, 2007.

[23] L. C. Piccinini, G. Stampacchia and G. Vidossich, *Ordinary Differential Equations in* $\mathbb{R}^n$. Springer-Verlag, New York, Volume 39, 1978.

[24] J. Price and T. Javidi, "Leveraging downlink for efficient uplink allocation in a single-hop wireless network," *IEEE Trans. Inform. Th.*, vol. 53, pp. 4330-4339, Nov. 2007.

[25] S. Shakkottai and A. L. Stolyar, "Scheduling algorithms for a mixture of real-time and non-real-time data in HDR," in *Teletraffic Engineering in the Internet Era, Proc. ITC-17*, S. da Bahia, J. M. de Souza, N. L. S. da Fonseca, and E. A. de Souza e Silva, Eds. Amsterdam, The Netherlands: North-Holland, 2001, pp. 793-804.

[26] J. Sobel, "Signaling games,", to appear in *Encyclopedia of Complexity and System Science*. M. Sotomayor (ed.), Springer, forthcoming, 2009.

[27] K. Tuna, Hacking EVDO, Proc. Defcon 15, 2007.