Repeated Games with Intervention: Theory and Applications in Communications

Yuanzhang Xiao, Jaeok Park, and Mihaela van der Schaar, *Fellow, IEEE*

***Abstract*—In communication systems where users share common resources, selfish behavior usually results in suboptimal resource utilization. There have been extensive works that model communication systems with selfish users as one-shot games and propose incentive schemes to achieve Pareto-optimal outcomes. However, in many communication systems, due to strong negative externalities among users, the sets of feasible payoffs in one-shot games are nonconvex. Thus, it is possible to expand the set of feasible payoffs by having users choose different action profiles in an alternating manner. In this paper, we formulate a model of repeated games with intervention. First, by using repeated games we can convexify the set of feasible payoffs in one-shot games. Second, by using intervention in repeated games we can achieve a larger set of equilibrium payoffs and loosen requirements for users' patience to achieve a target payoff. We study the problem of maximizing a welfare function defined on users' payoffs. We characterize the limit set of equilibrium payoffs. Given the optimal equilibrium payoff, we derive the sufficient condition on the discount factor and the intervention capability to achieve it, and design corresponding equilibrium strategies. We illustrate our analytical results with power control and flow control.**

***Index Terms*—Repeated games, intervention, power control, flow control.**

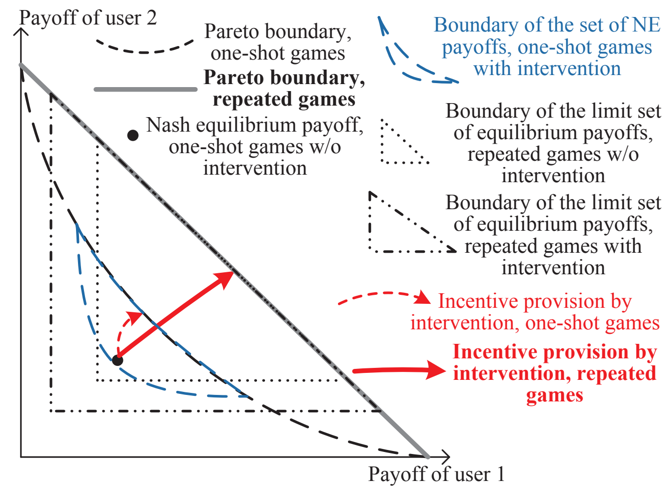

Fig. 1. Sets of equilibrium payoffs in a two-user game under different incentive schemes. We use the same intervention device (thus the same intervention capability) in the games with intervention. The (limit) sets of equilibrium payoffs in repeated games are obtained with the discount factor arbitrarily close to 1. We can see that the set of equilibrium payoffs of the repeated game with intervention includes the set of equilibrium payoffs in the one-shot game with intervention and that in the repeated game without intervention.

I. INTRODUCTION

GAME theory is a formal framework to model and analyze the interaction of selfish users. It has been used in the literature to study communication networks with selfish users [1][2]. Most works modeled communication systems as one-shot games, studied the inefficiency of noncooperative outcomes, and proposed incentive schemes, such as pricing [3]–[12], to improve the inefficient outcomes towards the Pareto boundary. Recently, a new incentive scheme, called "intervention", has been proposed in [13] in the one-shot game framework, with applications to medium access control (MAC) games [14][15] and power control games [16].[1] In an intervention scheme, the designer places an *intervention device*, which can monitor the user behavior and intervene in their interaction, in the system. The intervention device observes a signal about the users' actions and then chooses an intervention action depending on the observed signal. In this way, it can deter misbehavior of a user by exerting punishment following a signal that suggests a deviation.

In communication systems where users create severe congestion or interference, increasing a user's payoff requires a significant sacrifice of others' payoffs. This feature is reflected by a nonconvex set of feasible payoffs in some systems studied in the aforementioned works [3]–[17] using one-shot game models. For example, in one-shot power control games, the set of feasible payoffs is nonconvex when the cross channel gains are large [18][19]. In one-shot MAC games based on the collision model, the set of feasible payoffs is also nonconvex, because transmissions from multiple users cause packet loss [10][20]. Moreover, the sets of feasible payoffs of some one-shot flow control games are also nonconvex. To sum up, the sets of feasible payoffs are nonconvex in many communication scenarios, and when the set of feasible payoffs is nonconvex, its Pareto boundary can be dominated by a convex combination of different payoffs (see Fig. 1 for illustration). In one-shot games, such convex combinations cannot be achieved unless a public correlation device is used.[2]

When users interact in a system for a long period of

Paper approved by T. T. Lee, the Editor for Switching Architecture Performance of the IEEE Communications Society. Manuscript received September 17, 2011; revised February 27, 2012.

Y. Xiao and M. van der Schaar are with the Electrical Engineering Department, University of California, Los Angeles, CA 90095 USA (e-mail: {yxiao, mihaela}@ee.ucla.edu).

J. Park is with the School of Economics, Yonsei University, Seoul 120-749, Korea (e-mail: jaeok.park@yonsei.ac.kr). He was previously with the Electrical Engineering Department, University of California, Los Angeles, CA 90095 USA.

Digital Object Identifier 10.1109/TCOMM.2012.071612.110626

[1] With the same philosophy as intervention, a packet-dropping incentive scheme was proposed for flow control games in [17].

[2] Public correlation devices are used in game theory literature as a simple method to achieve payoffs that are convex combinations of pure-action payoffs. Such devices may not be available in communication systems. Even if such devices are available, there are additional costs on broadcasting random signals generated by a public correlation device.

time, we can model their interaction as a repeated game, in which users play a stage game repeatedly over time obtaining information about the actions of other users. In a repeated game, users can choose different actions in an alternating manner, obtaining a convex combination of different stage-game payoffs as a repeated-game payoff. A repeated-game strategy specifies actions to be taken given a history of past observations and thus can be interpreted as a protocol. If a repeated game strategy constitutes a subgame perfect equilibrium (SPE), then no user can gain from deviation at any occasion. Hence, a SPE can be considered as a deviation-proof protocol.

In this paper, we consider the problem of designing a protocol to maximize some welfare function defined on users' payoffs among (subgame perfect) equilibrium protocols. First, we characterize the set of equilibrium payoffs when users are sufficiently patient. This set determines the feasible set of the welfare maximization problem. Consequently, a larger set of equilibrium payoffs can result in higher social welfare. In this paper, we will characterize the set of equilibrium payoffs in repeated games with intervention, and show that using intervention can yield a larger set of equilibrium payoffs than the corresponding set without intervention (Fig. 1 illustrates the promising gain by using repeated games with intervention.).

Given the feasible set determined by the set of equilibrium payoffs, we can obtain the target payoff that is in the feasible set and maximizes the welfare function. Then we will analytically determine the range of the intervention capability and the discount factors under which the target payoff can be achieved at an SPE. The intervention capability reflects the extent to which the intervention device can intervene in the users' interaction. The discount factor is affected by the user patience and the network dynamics. It represents the rate at which users discount future payoffs; a less patient user has a smaller discount factor. The discount factor can also model the probability of users remaining in the network in each period; a more dynamic network results in a smaller discount factor. A lower discount factor is desirable in the sense that with a lower requirement, a protocol is effective in a wider variety of users and more dynamic networks. We will show that using intervention can lower the requirement on the discount factor compared to the case without intervention. Moreover, we obtain a trade-off between the discount factor and the intervention capability. Hence, given a discount factor, we can calculate the minimum intervention capability required to support the target payoff, and thus determine which intervention device to use. Conversely, given the intervention capability available, we can calculate the minimum discount factor required, and thus determine the types of users and networks that can be supported. Finally, given a target payoff and the discount factor, we show how to construct an equilibrium strategy, namely the deviation-proof protocol. Simulate results validate our theoretical analysis and demonstrate the performance gain over existing incentive schemes.

The rest of this paper is organized as follows. Section II provides a summary of related works. Section III describes a repeated game model generalized by intervention and formulates the protocol design problem. In Section IV, we characterize the set of equilibrium payoffs when the discount factor is arbitrarily close to one, and specify the structure of equilibrium strategies. Then we solve the design problem in Section V. Simulation results are presented in Section VI. Finally, Section VII concludes the paper.

II. Related Works

To improve the inefficient noncooperative outcomes, most existing works proposed incentive schemes based on pricing [3]–[12] under one-shot game models. As discussed before, these schemes are inefficient when the multi-user interference is so strong that the set of feasible payoffs in one-shot games is nonconvex. In addition, one problem in pricing is to determine how a user's payoff changes with the amount of payment. Although it is usually assumed that the payoff is linearly decreasing in the amount of payment, this assumption may not be true. For example, a user's payoff may be a concave decreasing function of the amount of payment, reflecting the fact that a user's payoff decreases faster when the payment is larger. One advantage of intervention is that the designer knows how the users' payoffs are affected by the actions of the intervention device according to some physical laws. For example, in power control, the designer knows that by transmitting at certain power level, the intervention device can cause some interference to the users, and the amount of interference will be added to the denominator of each user's signal-to-interference-and-noise ratio (SINR). Another advantage of intervention is that the intervention device directly interacts with the users in the system, instead of using outside instruments such as monetary payments in pricing. As a result, intervention can provide more robust incentives in the sense that users cannot avoid punishment.

Intervention was originally proposed in [13] under one-shot game models, and was applied to one-shot power control games [16] and one-shot MAC games [14][15]. Although it has several advantages over pricing, intervention in one-shot games is inefficient in the scenarios with strong multi-user interference. This motivates us to study intervention under repeated game models.

In economics, the study of repeated games focuses on proving folk theorems [21]. Specifically, they characterize the set of equilibrium payoffs when users are sufficiently patient and the network is static (i.e., the discount factor can be arbitrarily close to 1). For most games, they prove that any equilibrium payoff can be sustained under some sufficiently large discount factor, but they do not know how large the discount factor has to be to sustain a particular equilibrium payoff. In contrast, we are more interested in the case where the users are impatient or the network is dynamic with users leaving with certain probability. Hence, given an equilibrium payoff, we want to analytically determine how large a discount factor we need in order to achieve the given equilibrium payoff. A computational method was proposed in [22]. Analytical characterization is available for a few special cases, such as games with transferable payoffs in [23] and the repeated prisoners' dilemma game in [24]. For the class of games we consider, our result (Theorem 1) is the first that analytically determines how large the discount factor needs to be to achieve a given Pareto-optimal equilibrium payoff.

Following the game theory literature, most works applying repeated games in communication systems assume that the discount factor can be arbitrarily close to 1 [25]–[32]. Moreover, almost all the works use strategies in which the users play the same action profile at the equilibrium [25]–[35]. Since the users play the same action profile, they can only achieve the feasible payoffs in one-shot games. Hence, applying such strategies to the scenarios with strong multi-user interference will still result in inefficient outcomes. In contrast, we allow the users to choose different action profiles in an alternating manner to achieve Pareto optimality, although such repeated-game strategies are much harder to analyze.

III. Model of Repeated Games With Intervention

A. The Stage Game With Intervention

We consider a system with N users. The set of the users is denoted by $\mathcal{N} \triangleq \{1, 2, \ldots, N\}$. Each user i chooses an action a_i from the set $A_i \subset \mathbb{R}^{k_i}$ for some integer $k_i > 0$. An action profile of the users is denoted by $\mathbf{a} = (a_1, \ldots, a_N)$, and the set of action profiles of the users is denoted by $\mathcal{A} = \prod_{i=1}^N A_i$. We use $\mathbf{a}_{-i}$ to denote the action profile of the users except for user i. In addition to the N users, there exists an intervention device in the system, indexed by 0. The intervention device chooses an action a_0 from the set $A_0 \subset \mathbb{R}^{k_0}$ for some integer $k_0 > 0$. We refer to $(a_0, \mathbf{a}) \in A_0 \times \mathcal{A}$ as an action profile (of the intervention device and the users). The payoffs of the users are determined by the action profile, and user i's payoff function is denoted by $u_i : A_0 \times \mathcal{A} \to \mathbb{R}$. Given the payoff functions of the users, the set A_0 determines the intervention capability of the intervention device, since it reflects the extent to which the intervention device can affect the users' payoffs. We assume that there exists a null intervention action, denoted by $\underline{a}_0 \in A_0$, which corresponds to the case where there is no intervention device. We further assume that an intervention action can only decrease the payoffs of the users, i.e., $u_i(a_0, \mathbf{a}) \leq u_i(\underline{a}_0, \mathbf{a})$ for all $a_0 \in A_0$, all $\mathbf{a} \in \mathcal{A}$, and all $i \in \mathcal{N}$. In other words, intervention can provide only punishment to users, not reward. A (strategic-form) game with intervention is summarized by the tuple $\langle \mathcal{N}, (A_0, (A_i)_{i\in\mathcal{N}}), (u_i)_{i\in\mathcal{N}} \rangle$. For technical reasons, we assume that each of A_0 and A_i is a compact and convex set and that u_i is continuous for all i.

An important feature of the intervention device is that it does not have its own objective and can be programmed in the way that the protocol designer desires. Hence, when discussing equilibrium, we need to take only the users' incentives into account. An action profile $(a_0^*, \mathbf{a}^*)$ is a Nash equilibrium (NE) of the game $\langle \mathcal{N}, (A_0, (A_i)_{i\in\mathcal{N}}), (u_i)_{i\in\mathcal{N}} \rangle$ if

$$u_i(a_0^*, \mathbf{a}^*) \geq u_i(a_0^*, a_i, \mathbf{a}_{-i}^*), \ \forall a_i \in A_i, \ \forall i \in \mathcal{N}. \quad (1)$$

We call an action profile $(\underline{a}_0, \mathbf{a}^*)$ a Nash equilibrium without intervention if (1) is satisfied with $a_0^* = \underline{a}_0$.

B. The Repeated Game With Intervention

In a repeated game with intervention, the users play the same game $\langle \mathcal{N}, (A_0, (A_i)_{i\in\mathcal{N}}), (u_i)_{i\in\mathcal{N}} \rangle$ in every period $t = 0, 1, 2, \ldots$. The game played in each period is called the stage game of the repeated game. At the end of period t, all the users and the intervention device observe the action profile chosen in period t, which is denoted by $(a_0^t, \mathbf{a}^t)$. That is, we assume perfect monitoring. As a result, the users and the intervention device share the same history at each period, and the history at the beginning of period t is the collection of all the action profiles before period t. We denote the history in period $t \geq 1$ by $h^t = (a_0^0, \mathbf{a}^0; a_0^1, \mathbf{a}^1; \ldots; a_0^{t-1}, \mathbf{a}^{t-1})$, while we set the initial history, i.e., the history in period 0, as $h^0 = \varnothing$. The set of possible histories in period t is denoted by $\mathscr{H}^t$, and the set of all possible histories by $\mathscr{H} = \bigcup_{t=0}^{\infty} \mathscr{H}^t$.

The (pure) strategy of user i in the repeated game is a mapping from the set of all possible histories to its action set, written as $\sigma_i : \mathscr{H} \to A_i$. When user i uses strategy σ_i, its action at history h^t is determined by $a_i^t = \sigma_i(h^t)$. The strategy profile of the users is denoted by $\sigma = (\sigma_1, \ldots, \sigma_N)$, and the strategy profile of the users except for user i by σ_{-i}. The set of all strategies of user i is denoted by Σ_i, while the set of all strategy profiles is denoted by $\Sigma = \prod_{i=1}^N \Sigma_i$. The intervention device chooses its actions following an intervention rule, which is represented by a mapping $\sigma_0 : \mathscr{H} \to A_0$.[3] When the intervention device uses the intervention rule σ_0, its action at history h^t is determined by $a_0^t = \sigma_0(h^t)$. The set of all intervention rules is denoted by Σ_0. When the intervention rule is constant at $\underline{a}_0$, namely $\sigma_0(h) = \underline{a}_0$ for all $h \in \mathscr{H}$, the repeated game with intervention reduces to the conventional repeated game without intervention.

To compute the payoff of a user in the repeated game, we use the discounted average payoff of the user. We assume that all the users have the same discount factor $\delta \in [0, 1)$, as commonly assumed in the literature [25]–[35]. Then the payoff function of user i in the repeated game is

$$U_i(\sigma_0, \sigma) = (1-\delta) \sum_{t=0}^{\infty} \delta^t u_i(a_0^t(\sigma_0, \sigma), \mathbf{a}^t(\sigma_0, \sigma)), \quad (2)$$

where $(a_0^t(\sigma_0, \sigma), \mathbf{a}^t(\sigma_0, \sigma))$ is the action profile in period t induced by (σ_0, σ). The path of plays under (σ_0, σ), $\{(a_0^t(\sigma_0, \sigma), \mathbf{a}^t(\sigma_0, \sigma))\}_{t=0}^{\infty}$, can be calculated recursively by $(a_0^0(\sigma_0, \sigma), \mathbf{a}^0(\sigma_0, \sigma)) = (\sigma_0(\varnothing), \sigma(\varnothing))$ and for all $t \geq 1$,

$$\begin{aligned} &(a_0^t(\sigma_0, \sigma), \mathbf{a}^t(\sigma_0, \sigma)) \quad (3) \\ &= \big(\sigma_0(a_0^0(\sigma_0, \sigma), \mathbf{a}^0(\sigma_0, \sigma); \ldots; a_0^{t-1}(\sigma_0, \sigma), \mathbf{a}^{t-1}(\sigma_0, \sigma)), \\ &\quad \sigma(a_0^0(\sigma_0, \sigma), \mathbf{a}^0(\sigma_0, \sigma); \ldots; a_0^{t-1}(\sigma_0, \sigma), \mathbf{a}^{t-1}(\sigma_0, \sigma))\big). \end{aligned}$$

User i's continuation strategy induced by any history $h^t \in \mathscr{H}$, denoted $\sigma_i|_{h^t}$, is defined by $\sigma_i|_{h^t}(h^\tau) = \sigma_i(h^t h^\tau), \forall h^\tau \in \mathscr{H}$, where $h^t h^\tau$ is the concatenation of the history h^t followed by the history h^τ. Similarly, we can define the continuation intervention rule, $\sigma_0|_{h^t}$. By convention, we use $\sigma|_{h^t}$ and $\sigma_{-i}|_{h^t}$ to denote the continuation strategy profile of the users and the continuation strategy profile of the users except for user i, respectively. Then a pair of the intervention rule and the strategy profile (σ_0, σ) is a subgame perfect equilibrium (SPE) of the repeated game with intervention if for all $h^t \in \mathscr{H}$

$$U_i(\sigma_0|_{h^t}, \sigma|_{h^t}) \geq U_i(\sigma_0|_{h^t}, \sigma_i', \sigma_{-i}|_{h^t}), \ \forall \sigma_i' \in \Sigma_i, \ \forall i \in \mathcal{N}.$$

If the users use a strategy profile that constitutes a SPE together with an intervention rule, no user can gain from choosing a different strategy starting at any history.

[3] We call σ_0 an "intervention rule" instead of an "intervention strategy" because the intervention device is not a strategic player but can be programmed to follow a rule prescribed by the designer.

C. Problem Formulation

We use the term "protocol" to refer to (σ_0, σ), which prescribes the behavior of the intervention device and the users in the repeated game. There is a protocol designer who chooses an intervention rule σ_0 and recommends a strategy profile σ to the users, having complete information about the repeated game described above. In order to make the protocol "deviation-proof" following any history, the designer should choose a protocol that constitutes a SPE. The designer has an objective function, which is a welfare function defined over the repeated-game payoffs of the users, denoted by $W : \mathbb{R}^N \to \mathbb{R}$. Then the protocol design problem can be formally written as

$$\max_{\sigma_0,\sigma} \quad W(U_1(\sigma_0,\sigma),\ldots,U_N(\sigma_0,\sigma)) \tag{4}$$
$$\text{subject to} \quad (\sigma_0,\sigma) \text{ is subgame perfect equilibrium.}$$

Two widely-used welfare functions are sum payoff $\sum_{i=1}^N U_i$ (the utilitarian social welfare function) and max-min fairness $\min_{i\in\mathcal{N}} U_i$ (the egalitarian social welfare function) [36]. Sometimes, the designer may want to guarantee a minimum payoff γ_i to each user i. This minimum payoff guarantee can be incorporated into the welfare function.

IV. Characterizing the Limit Set of Equilibrium Payoffs

A. Folk Theorem for Repeated Games With Intervention

In order for the designer to solve the problem (4), it is important to identify the set of payoffs achievable at a SPE. In general, finding the set of equilibrium payoffs for a given discount factor is a challenging task. However, folk theorems characterize the limit set of equilibrium payoffs as the discount factor goes to 1. We will adapt conventional folk theorems for repeated games with intervention. Our proof is constructive and thus provide an equilibrium protocol that achieves a desired equilibrium payoff. Our results show that intervention can enlarge the limit set of equilibrium payoffs by strengthening punishment for deviation.

We first provide some preliminaries for our folk theorem.

Definition 1 (Minmax Payoff With Intervention): User i's (pure-action) minmax payoff with intervention is defined as

$$\underline{v}_i^w \triangleq \min_{(a_0,\mathbf{a}_{-i})\in A_0\times\mathcal{A}_{-i}} \max_{a_i\in A_i} u_i(a_0,a_i,\mathbf{a}_{-i}). \tag{5}$$

User i's minmax payoff is the payoff it can guarantee by playing a best response to others' actions at every history. We say that a payoff vector $\mathbf{v} = (v_1,\ldots,v_N)$ is strictly individually rational if $v_i > \underline{v}_i^w$ for all $i \in \mathcal{N}$. Let $\mathbf{u}(a_0,\mathbf{a}) = (u_1(a_0,\mathbf{a}),\ldots,u_N(a_0,\mathbf{a}))$. Then the set of feasible and strictly individually rational payoffs is

$$\mathscr{V}_w^\dagger = \text{co}\{\mathbf{v}\in\mathbb{R}^N : \exists(a_0,\mathbf{a})\in A_0\times\mathcal{A} \text{ s.t. } \mathbf{v}=\mathbf{u}(a_0,\mathbf{a})\} \cap \{\mathbf{v}\in\mathbb{R}^N : v_i > \underline{v}_i^w,\ \forall i\in\mathcal{N}\},$$

where $\text{co}X$ denotes the convex hull of a set X. When the action profile to achieve a user's minmax payoff does not depend on the minmaxed user, we obtain a mutual minmax profile, as formally defined below.

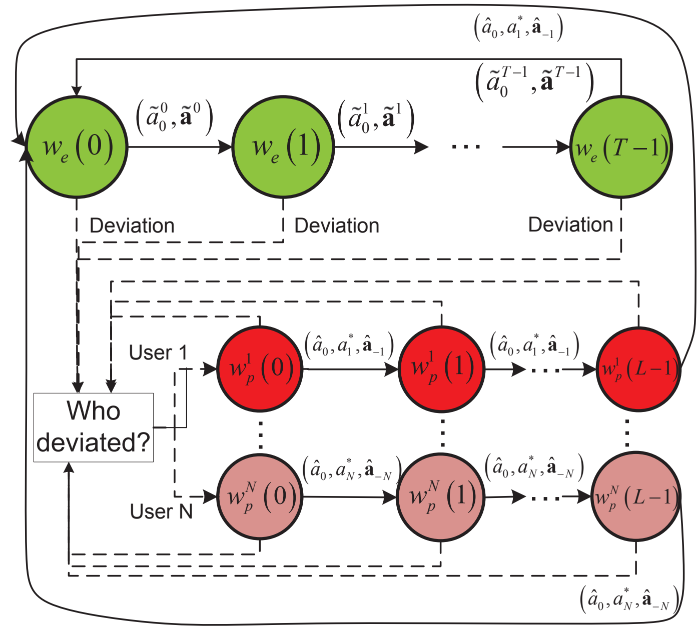

Fig. 2. The automaton of the equilibrium strategy of the game with a mutual minmax profile $(\hat{a}_0, \hat{\mathbf{a}})$. Circles are states, where $\{w_e(\tau)\}_{\tau=0}^{T-1}$ is the set of states in the equilibrium path, and $\{w_p^i(\ell)\}_{\ell=0}^{L-1}$ is the set of states in the punishment phase for user i. The initial state is $w_e(0)$. Solid arrows are the prescribed state transitions labeled by the action profiles leading to the transitions. Dashed arrows are the state transitions when deviation happens. a_i^* is user i's best response to $\hat{a}_0$ and $\hat{\mathbf{a}}_{-\mathbf{i}}$.

Definition 2 (Mutual Minmax Profile): An action profile $(\hat{a}_0, \hat{\mathbf{a}})$ is a mutual minmax profile if for all $i \in \mathcal{N}$,

$$\underline{v}_i^w = \max_{a_i} u_i(\hat{a}_0, a_i, \hat{\mathbf{a}}_{-i}).$$

Now we state the folk theorem for repeated games with intervention.

Proposition 1 (Minmax Folk Theorem): Suppose that there exists a mutual minmax profile $(\hat{a}_0, \hat{\mathbf{a}})$. Then for every feasible and strictly individually rational payoff $\mathbf{v} \in \mathscr{V}_w^\dagger$, there exists $\underline{\delta} < 1$ such that for all $\delta \in (\underline{\delta}, 1)$, there exists a subgame perfect equilibrium with payoffs $\mathbf{v}$, of the repeated game with intervention.

Proof: See [37, Appendix A]. ■

We briefly describe the structure of equilibrium protocols that are used to prove the folk theorem. Suppose that we want to achieve a payoff vector $\mathbf{v}$. By [21, Lemma 3.7.2], when the discount factor is sufficiently close to 1, we can find a sequence of action profiles $\{(\tilde{a}_0^t, \tilde{\mathbf{a}}^t)\}_{t=0}^\infty$ whose discounted average payoffs are $\mathbf{v}$ and whose continuation payoffs at any time t are close to $\mathbf{v}$. In the equilibrium protocol constructed in the proof of Proposition 1, the intervention device and the users start from following the sequence $\{(\tilde{a}_0^t, \tilde{\mathbf{a}}^t)\}_{t=0}^\infty$ and keep following it as long as there is no unilateral deviation. If user i deviates unilaterally, the intervention device chooses $\hat{a}_0$ and the other users choose $\hat{\mathbf{a}}_{-i}$ for L periods as a punishment on user i, where L is chosen sufficiently large to make any deviation from the sequence unprofitable. After the L periods of punishment, the intervention device and the users return to the sequence starting from the beginning. Any unilateral deviation during the L periods of punishment will trigger another punishment on the deviating user for L periods. Fig. 2 shows an automaton representation of the equilibrium protocol described above assuming that the sequence is a repetition of a cycle of finite length T.

B. Roles of Intervention in the Folk Theorem

We discuss two roles of intervention in obtaining the folk theorem and provide two examples to illustrate these roles. First, intervention lowers the minmax payoff of each user and thus enlarges the limit set of equilibrium payoffs. In other words, by utilizing intervention, we can provide a stronger threat to deter a deviation, and thus we can achieve more outcomes as a SPE. Second, intervention may turn the mutual minmax profile into a NE of the stage game when the mutual minmax profile without intervention is not a NE. Having a mutual minmax profile that is a NE simplifies the structure of equilibrium protocols. When we have a mutual minmax profile that is a NE, permanent reversion to the mutual minmax profile provides the most severe punishment on every user. Thus, we can focus on grim trigger protocols, where any deviation from a prescribed sequence of action profiles results in permanent reversion to the mutual minmax profile. On the contrary, when the mutual minmax profile is not a NE, there is a user that can increase its stage-game payoff by deviation, and we use the promise of returning to the equilibrium path after a finite number of periods to deter users from deviating in a punishment phase. As a result, having a mutual minmax profile that is a NE makes it simpler for the designer to construct an equilibrium protocol and for the users to execute the protocol.

Example 1 (Power Control): Consider a network where N users and the intervention device transmit power in a wireless channel. Each user i's action is its transmit power $a_i \in A_i = [0, \bar{a}_i]$, and the action of the intervention device is its transmits power $a_0 \in A_0 = [0, \bar{a}_0]$. We assume that $\bar{a}_i > 0$ for all i and $\bar{a}_0 > 0$. User i's SINR is calculated by $h_{ii}a_i/(h_{i0}a_0 + \sum_{j\neq i} h_{ij}a_j + n_i)$, where h_{ij} is the channel gain from user j's transmitter to user i's receiver, and n_i is the noise power at user i's receiver. Each user i's stage-game payoff is its throughput [4][19][39], namely

$$u_i(a_0, a_i, \mathbf{a}_{-i}) = \log_2\left(1 + \frac{h_{ii}a_i}{h_{i0}a_0 + \sum_{j\neq i} h_{ij}a_j + n_i}\right). \quad (6)$$

Using any other increasing function of the SINR as the payoff function does not change the analysis here.

In this power control game, the null intervention action is $\underline{a}_0 = 0$. Without intervention (i.e., a_0 is fixed at $\underline{a}_0$), the only NE of the stage game is $(\underline{a}_0, \bar{\mathbf{a}}) = (0, \bar{a}_1, \ldots, \bar{a}_N)$, where every user transmits at its maximum power, and the NE is the mutual minmax profile with payoff $\underline{\mathbf{v}}^o = (\underline{v}_1^o, \ldots, \underline{v}_N^o) = \mathbf{u}(0, \bar{\mathbf{a}})$. With intervention, the mutual minmax profile is $(\bar{a}_0, \bar{\mathbf{a}})$, which is also a NE, with payoff $\underline{\mathbf{v}}^w = (\underline{v}_1^w, \ldots, \underline{v}_N^w) = \mathbf{u}(\bar{a}_0, \bar{\mathbf{a}}) < \mathbf{u}(0, \bar{\mathbf{a}})$. Note that $\underline{v}_i^w < \underline{v}_i^o \ \forall i$, and that each $\underline{v}_i^w$ reduces as $\bar{a}_0$ increases. Hence, in this example, intervention plays only the first role, expanding the limit set of equilibrium payoffs as the maximum intervention power $\bar{a}_0$ increases.

Example 2: Consider a network where N users and the intervention device transmit packets through a single server, which can be modeled as an M/M/1 queue [41]–[43]. User i's action is its transmission rate $a_i \in A_i = [0, \bar{a}_i]$. The intervention device also transmits packets at the rate of $a_0 \in A_0 = [0, \bar{a}_0]$. We assume that $\bar{a}_i > 0$ for all i and $\bar{a}_0 > 0$. User i's payoff is a function of its transmission rate a_i and its delay $1/(\mu - a_0 - \sum_{j=1}^N a_j)$ [40]–[43], defined by

$$u_i(a_0, a_i, \mathbf{a}_{-i}) = a_i^{\beta_i} \max\left\{0,\ \mu - a_0 - \sum_{j=1}^N a_j\right\}, \quad (7)$$

where $\mu > 0$ is the server's service rate, and $\beta_i > 0$ is the parameter reflecting the trade-off between the transmission rate and the delay. Here the "max" function indicates the fact that the payoff is zero when the total arrival rate is larger than the service rate. We assume that the service rate is no smaller than the maximum total arrival rate without intervention, i.e., $\mu \geq \sum_{j=1}^N \bar{a}_j$.

Without intervention, the mutual minmax profile is $(0, \bar{a}_1, \ldots, \bar{a}_N)$, where every user transmits at its maximum rate. When there is no intervention, user i's best response to $\mathbf{a}_{-i}$ is given by $a_i^* = \min\left\{\frac{\beta_i}{1+\beta_i}\left(\mu - \sum_{j\neq i} a_j\right),\ \bar{a}_i\right\}$. Thus, the mutual minmax profile $(0, \bar{a}_1, \ldots, \bar{a}_N)$ is a NE if and only if

$$\bar{a}_i \leq \frac{\beta_i}{1+\beta_i}\left(\mu - \sum_{j\neq i} \bar{a}_j\right), \forall i \in \mathcal{N}. \quad (8)$$

Hence, without intervention, the mutual minmax profile is not a NE if $\bar{a}_i/\beta_i > \mu - \sum_{j=1}^N \bar{a}_j$ for some i. However, with intervention, the mutual minmax profile $(\bar{a}_0, \bar{a}_1, \ldots, \bar{a}_N)$ is a NE, as long as the maximum rate $\bar{a}_0$ of the intervention device is high enough to yield

$$\mu - \sum_{j\neq i} \bar{a}_j - \bar{a}_0 \leq 0,\ \forall i \in \mathcal{N}. \quad (9)$$

Now we analyze, in the context of the above flow control game, the conditions for a protocol to support a constant action profile $(0, \tilde{\mathbf{a}})$ as a SPE outcome using punishments of playing the mutual minmax profile for L periods. As derived in the proof of Proposition 1, the protocol using L as the punishment length is a SPE if the discount factor δ satisfies the following two sets of inequalities: first, for all $i \in \mathcal{N}$,

$$\delta + \cdots + \delta^L \geq \frac{(\tilde{a}_i^*)^{\beta_i}\left(\mu - \tilde{a}_i^* - \sum_{j\neq i} \tilde{a}_j\right) - \tilde{a}_i^{\beta_i}\left(\mu - \sum_{j=1}^N \tilde{a}_j\right)}{\tilde{a}_i^{\beta_i}\left(\mu - \sum_{j=1}^N \tilde{a}_j\right) - \bar{a}_i^{\beta_i} \cdot \left(\mu - \bar{a}_0 - \sum_{j=1}^N \bar{a}_j\right)^+}, \quad (10)$$

where $\tilde{a}_i^* = \min\left\{\bar{a}_i, \frac{\beta_i}{1+\beta_i}\left(\mu - \sum_{j\neq i} \tilde{a}_j\right)\right\}$ and $(x)^+ \triangleq \max\{0, x\}$; second, for all $i \in \mathcal{N}$,

$$\delta^L \geq \frac{(\bar{a}_i^*)^{\beta_i} \cdot \left(\mu - \bar{a}_0 - \bar{a}_i^* - \sum_{j\neq i} \bar{a}_j\right)^+}{\tilde{a}_i^{\beta_i}\left(\mu - \sum_{j=1}^N \tilde{a}_j\right) - \bar{a}_i^{\beta_i} \cdot \left(\mu - \bar{a}_0 - \sum_{j=1}^N \bar{a}_j\right)^+} - \frac{\bar{a}_i^{\beta_i} \cdot \left(\mu - \bar{a}_0 - \sum_{j=1}^N \bar{a}_j\right)^+}{\tilde{a}_i^{\beta_i}\left(\mu - \sum_{j=1}^N \tilde{a}_j\right) - \bar{a}_i^{\beta_i} \cdot \left(\mu - \bar{a}_0 - \sum_{j=1}^N \bar{a}_j\right)^+}, \quad (11)$$

where $\bar{a}_i^* = \min\left\{\bar{a}_i, \frac{\beta_i}{1+\beta_i}\left(\mu - \bar{a}_0 - \sum_{j\neq i} \bar{a}_j\right)\right\}$. Equation (10) guarantees that no user can gain from deviating from the equilibrium path, while equation (11) ensures that no user can gain from deviating in a punishment phase. If the mutual

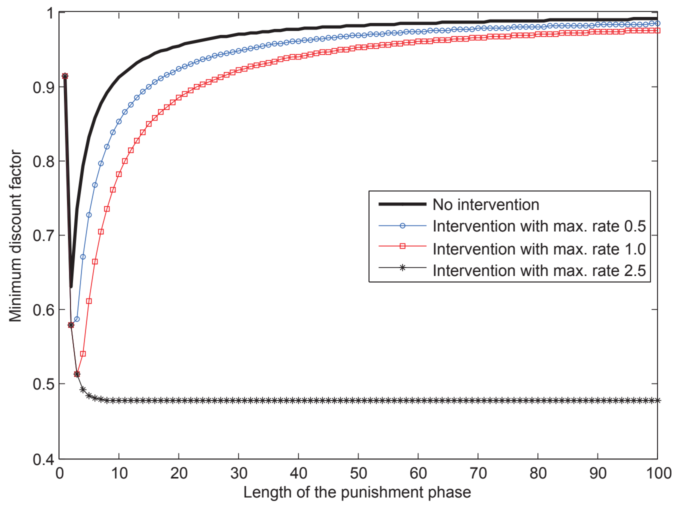

Fig. 3. Minimum discount factors to support the pure-action profile that achieves maximum sum payoff, under different punishment lengths and maximum intervention flow rates. $N = 4$. The service rate is $\mu = 10$ bits/s. The maximum flow rates for all the users are 2.5 bits/s. The trade-off factors are $\beta_1 = \beta_2 = 2$ and $\beta_3 = \beta_4 = 3$.

minmax profile is a NE, we can use perpetual reversion to the mutual minmax profile as the most severe punishment. In this case, executing punishment is self-enforcing, and we require only (10) with the left-hand side replaced by $\delta/(1-\delta)$.

Fig. 3 plots the minimum discount factor required for the protocol to be a SPE under different punishment lengths L and maximum intervention flow rates $\bar{a}_0$ in a scenario where $N = 4$, $\mu = 10$ bits/s, $\bar{a}_i = 2.5$ bits/s for all i, $\beta_1 = \beta_2 = 2$ and $\beta_3 = \beta_4 = 3$. The target payoff is chosen to maximize the proportional fairness (defined as the sum of the logarithms of payoffs), which is achieved at the action profile $(0, \tilde{\mathbf{a}})$ where $\tilde{a}_1 = \tilde{a}_2 = 1.43$ bits/s and $\tilde{a}_3 = \tilde{a}_4 = 2.14$ bits/s. When the maximum intervention flow rate is 2.5 bits/s, the mutual minmax profile is a NE, and by using permanent reversion to it we obtain the minimum discount factor 0.48. When there is no intervention or when the maximum intervention rate is 0.5 bits/s or 1.0 bits/s, the mutual minmax profile is not a NE. In this case, given a punishment length, the minimum discount factor decreases as the the maximum intervention flow rate increases. In other words, intervention make the protocol a SPE for a wider range of the discount factors by making punishment stronger. Thus, this example exhibits both roles of intervention. We can see that the minimum discount factor increases with the punishment length beyond a threshold. This is because the users need to be more patient to carry out a longer punishment, reflected in (11).

V. Design of Equilibrium Protocols

In Section IV, we have characterized the limit set of equilibrium payoffs $\mathscr{V}_w^\dagger$. With this set, we can obtain the optimal equilibrium payoff $\mathbf{v}^\star$ that maximizes the welfare function by solving $\mathbf{v}^\star = \arg\max_{\mathbf{v} \in \mathscr{V}_w^\dagger} W(\mathbf{v})$. Now we will derive the minimum discount factor under which $\mathbf{v}^\star$ can be achieved by an equilibrium protocol of the structure described in Fig. 2, and construct the corresponding equilibrium protocol.

For analytical tractability, we impose the following three assumptions on the stage game.

Assumption 1: With intervention, the stage game has a mutual minmax profile $(\hat{a}_0, \hat{\mathbf{a}})$, which is also a NE of the stage game.

Assumption 2: For each user i, there exists an action profile $\check{\mathbf{a}}^i$ of the users such that for any $j \neq i$,

$$\bar{v}_i \triangleq \max_{\mathbf{a}} u_i(\underline{a}_0, \mathbf{a}) = u_i(\underline{a}_0, \check{\mathbf{a}}^i) \text{ and } u_j(\underline{a}_0, \check{\mathbf{a}}^i) = 0. \quad (12)$$

Assumption 3: The set of feasible payoffs is $\mathscr{V}_w = \text{co}\left\{(0,\ldots,0), \mathbf{u}(\underline{a}_0, \check{\mathbf{a}}^1), \ldots, \mathbf{u}(\underline{a}_0, \check{\mathbf{a}}^N)\right\}$.

Assumption 1 holds in many resource sharing scenarios. Suppose that each user's action is a scalar representing its usage level and that increasing a user's usage level increases its own payoff while decreasing the payoffs of all the other users. Then each user's choosing its maximum usage level is a NE, while it is a mutual minmax profile when the intervention device also chooses its maximum intervention level. The power control game discussed in Example 1 satisfies Assumption 1. The flow control game in Example 2 also satisfies Assumption 1 when the intervention flow rate is large enough to make the mutual minmax profile a NE. The action profile $\check{\mathbf{a}}^i$ in Assumption 2 is the one at which user i takes the most advantage of resources. Suppose that there is interference among users and that each user has an action not to utilize the resources at all. Then full utilization by a user can be achieved only when all the other users do not use the resources at all, leading to $u_j(\underline{a}_0, \check{\mathbf{a}}^i) = 0$ for all $j \neq i$, for all i. When the minimum usage level of some user is positive, we may have $u_j(\underline{a}_0, \check{\mathbf{a}}) > 0$ for some i and j. Our analysis can be carried over to this scenario, but for notational simplicity we choose to assume $u_j(\underline{a}_0, \check{\mathbf{a}}^i) = 0$ for all $j \neq i$, for all i. Assumption 2 holds for the power control and flow control games in Examples 1 and 2. Lastly, Assumption 3 is likely to hold when there is strong interference among the users.

When the welfare function W is increasing in the payoffs of the users, a solution to the protocol design problem (4) occurs on the Pareto boundary of the limit set of equilibrium payoffs $\mathscr{V}_w^\dagger$. Due to Assumption 3, we can write the Pareto boundary of $\mathscr{V}_w^\dagger$ by $\mathscr{P} = \left\{\mathbf{v} \in \mathscr{V}_w^\dagger : \sum_{i=1}^N (v_i/\bar{v}_i) = 1\right\}$. For a target payoff $\mathbf{v}^\star \in \mathscr{P}$, we derive the sufficient condition on the discount factor under which $\mathbf{v}^\star$ can be achieved by an equilibrium protocol.

Theorem 1: Given a target payoff $\mathbf{v}^\star \in \mathscr{P}$, there exists a discount factor $\bar{\delta}(\mathbf{v}^\star)$, such that for any discount factor $\delta \geq \bar{\delta}(\mathbf{v}^\star)$, there exists an equilibrium protocol that achieves the target payoff $\mathbf{v}^\star$. In particular, $\bar{\delta}(\mathbf{v}^\star)$ can be calculated analytically by

$$\bar{\delta}(\mathbf{v}^\star) = \max\left\{\max_{j\neq i} \frac{w_j - v_j^\star}{w_j - \underline{v}_j^w}, \right. \quad (13)$$

$$\left. \frac{2(N-1)}{(N-R) + \sqrt{(N-R)^2 + 4(R-S)(N-1)}}\right\},$$

where $w_j = \max_{i\neq j} \max_{a_j} u_j(\underline{a}_0, a_j, \check{\mathbf{a}}_{-j}^i)$ is the maximum stage-game payoff that user j can obtain by deviating from any of the profiles $\{(\underline{a}_0, \check{\mathbf{a}}^i)\}_{i\neq j}$, $R = \sum_{i=1}^N (w_i/\bar{v}_i)$, and $S = \sum_{i=1}^N (\underline{v}_i^w/\bar{v}_i)$.

Proof: See [37, Appendix C]. ■

TABLE I
DISTRIBUTED ALGORITHM RUN BY USER i

Require: The target payoff $\mathbf{v}^\star$ and the discount factor δ
Initialization: $\tau = 0$, $\mathbf{v}(0) = \mathbf{v}^\star$, $\nu_i = w_i - (w_i - \underline{v}_i)\cdot\delta, \forall i.$
Repeat
calculates $v_i(\tau)/\nu_i, \forall i$
calculates $i^* \triangleq \arg\max_{i\in\mathcal{N}} v_i(\tau)/\nu_i$
$\tilde{\mathbf{a}}^\tau = \tilde{\mathbf{a}}^{i^*}$, $\mathbf{v}(\tau+1) = \frac{1}{\delta}\mathbf{v}(\tau) - \frac{1-\delta}{\delta}\cdot\mathbf{u}(\underline{a}_0, \tilde{\mathbf{a}}^{i^*})$
$\tau \leftarrow \tau + 1$
Until $\mathbf{v}(\tau) = \mathbf{v}^\star$

Note that $\bar{\delta}(\mathbf{v}^\star)$ is increasing with respect to the minmax payoffs $\underline{v}_i^w$ for all $i \in \mathcal{N}$. Hence, a larger intervention capability decreases the minmax payoffs, which results in a smaller $\bar{\delta}(\mathbf{v}^\star)$, and thus allows a wider range of discount factors. By Theorem 1, we can determine the minimum intervention capability required to achieve a target payoff $\mathbf{v}^\star$ under a given discount factor δ, by finding the minimum intervention capability such that $\delta = \bar{\delta}(\mathbf{v}^\star)$. When A_0 is one-dimensional ($k_0 = 1$), the intervention capability is determined by the maximum intervention action $\bar{a}_0 = \arg\max_{a_0\in A_0} a_0$. Since $\bar{\delta}(\mathbf{v}^\star)$ is decreasing in $\bar{a}_0$, we can determine the minimum $\bar{a}_0$ such that $\delta = \bar{\delta}(\mathbf{v}^\star)$ efficiently by bisection methods.

Next, for a given target payoff $\mathbf{v}^\star \in \mathscr{P}$ and a discount factor $\delta \geq \bar{\delta}(\mathbf{v}^\star)$, we construct an equilibrium protocol that achieves $\mathbf{v}^\star$. Due to Assumption 1, we can use a trigger protocol in which the punishment phase lasts forever ($L = \infty$). Due to Assumptions 2 and 3, every action profile $(\tilde{a}_0^t, \tilde{\mathbf{a}}^t)$ in the equilibrium path $\{(\tilde{a}_0^t, \tilde{\mathbf{a}}^t)\}_{t=0}^\infty$ should be one of the action profiles $\{(\underline{a}_0, \tilde{\mathbf{a}}^i)\}_{i=1}^N$.

Theorem 2: Given a target payoff $\mathbf{v}^\star \in \mathscr{P}$, let $\{\tilde{\mathbf{a}}^t\}_{t=0}^\infty$ be the sequence of action profiles generated by the algorithm in Table I, and let (σ_0, σ) be the trigger protocol in which the intervention device and the users follow $\{(\underline{a}_0, \tilde{\mathbf{a}}^t)\}_{t=0}^\infty$, and any deviation leads to a perpetual play of $(\hat{a}_0, \hat{\mathbf{a}})$. Then for any discount factor $\delta \geq \bar{\delta}(\mathbf{v}^\star)$, the protocol (σ_0, σ) is a SPE of the repeated game with intervention and achieves $\mathbf{v}^\star$.

Proof: See [37, Appendix D]. ■

The algorithm in Table I can be run by each user in a distributed fashion. In each period, each user does N divisions and finds the largest one among the results of the N divisions. The overall complexity is $O(N)$, which is fairly small.

Remark 1: Note that any action profile $\tilde{\mathbf{a}}(t)$ in the sequence is from the set of $\{\tilde{\mathbf{a}}^i\}_{i=1}^N$. In other words, only one user takes nonzero action in each period. This greatly simplifies the monitoring burden of the intervention device and the users, when the users are not strategic in reporting deviation. Actually, at each period, the inactive users, who take zero actions, do not need to monitor, because the active user, who takes nonzero action at that period, can sense the interference and report to the intervention device the detection of deviation. Then the intervention device can broadcast the detection of deviation to trigger the punishment. In addition, no inactive user can gain from sending false report to the intervention device, because the intervention device only trusts the report from the active user. Note that if the users are strategic in reporting deviation, an active user may not report deviation voluntarily, if the loss from low future payoffs in the punishment phase is larger than the loss from low current payoffs due to interference from the deviating user. In this case, having reports from multiple users may be helpful. However, the strategic behavior in reporting deviation is out of the scope of this paper.

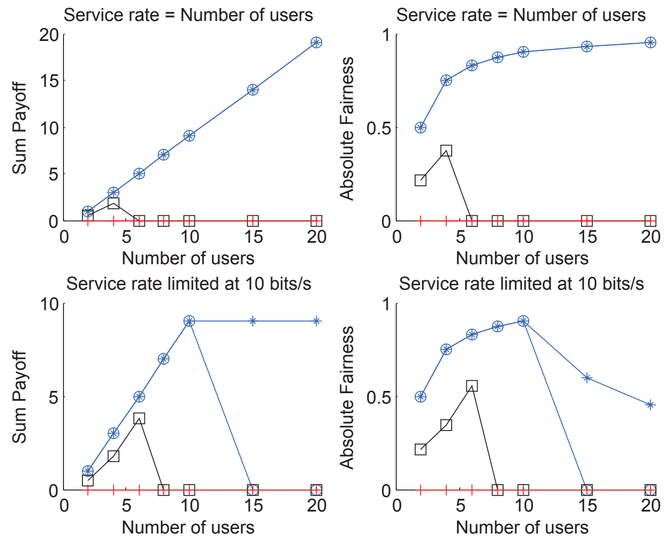

Fig. 4. Performance comparison among the four schemes with the number of users increasing. Lines with asterisks: repeated games with intervention, lines with circles: repeated games without intervention, lines with squares: one-shot games with incentive schemes, lines with crosses: Nash equilibrium of the one-shot game.

VI. SIMULATION RESULTS

In this section, we consider the flow control game in Example 2 to illustrate the performance gain of using intervention in repeated games and the analytical result on the trade-off between the discount factor and the intervention capability.

A. Performance Gain

First, we compare the performance when the protocol designer solves the protocol design problem (4) using four different schemes, namely greedy algorithms [41]–[43], one-shot games with incentive schemes [13][17][40], repeated games without intervention, and repeated games with intervention. The two welfare functions we use are the sum payoff $\sum_{i\in\mathcal{N}} U_i$ and the max-min fairness $\min_{i\in\mathcal{N}} U_i$. When we use the sum payoff as the welfare function, we impose a minimum payoff guarantee γ_i for each user i to prevent the users from receiving extremely low payoffs at the solution to the sum payoff maximization problem. Greedy algorithms achieve the NE, which may not satisfy the minimum payoff guarantees. For one-shot games with incentive schemes, we assume that the entire Pareto boundary of the set of pure-action payoffs can be achieved as NE by using appropriate incentive schemes, in order to get the best performance achievable by this scheme.

1) Impact of the number of users: We compare how the performance scales with the number of users under the four schemes. We study the symmetric case, where all the users have the same throughput-delay trade-off $\beta = 3$ and the same maximum flow rate normalized to 1 bits/s, such that the performance is affected only by the number of users, not by the heterogeneity of the users. We consider two scenarios:

first, the server's capacity (service rate) increases linearly with the number of users, i.e. $\mu = N$ bits/s, and second, the server's capacity is limited at 10 bits/s, i.e. $\mu = \min\{N, 10\}$ bits/s. The maximum intervention flow rate is $\bar{a}_0 = \max\{\mu-(N-1), 0\}$ bits/s to ensure the mutual minmax profile is an NE. For each user i, we set the minimum payoff guarantee as 10% of the maximum stage-game payoff $\bar{v}_i$, namely $\gamma_i = 10\% \cdot \bar{v}_i$ bits/s.

Fig. 4 shows the sum payoff and the fairness achieved by different schemes. When the capacity increases linearly with the number of users N, both the sum payoff and the fairness increase with N by using repeated games. In contrast, when using one-shot games with incentive schemes, the sum payoff and fairness increase initially when the number of users is small, and decrease when $N > 4$. A sum payoff or fairness of the value 0 means that the minimum payoff cannot be guaranteed. This happens in one shot games with incentive schemes when $N > 5$. The NE payoff does not satisfy the minimum payoff guarantee with any number of users.

When the capacity is limited at 10 bits/s, for repeated games with intervention, the sum payoff reaches the bottleneck of 10, and due to congestion, the fairness decreases when $N > 10$. For repeated games without intervention, the sum payoff and the fairness decrease more rapidly, and the minimum payoff guarantee cannot be met when $N \geq 15$. For one-shot games, the trend of the sum payoff and fairness is similar to the case with increasing capacity.

In conclusion, using repeated games with intervention has a large performance gain over the other three schemes, in terms of both the sum payoff and the fairness.

2) Impact of minimum payoff guarantees: We compare the performance of the four schemes under different minimum payoff guarantees. The system parameters are the same as those in Fig. 3. The maximum intervention flow rate is $\bar{a}_0 = 2.5$ bits/s to ensure the mutual minmax profile is an NE. We set the same minimum payoff guarantee for all the users, namely $\gamma_i = \gamma_j, \forall i, j \in \mathcal{N}$.

Table II shows the sum payoff and the fairness achieved by the four schemes. We write "N/A" when the minimum payoff guarantee cannot be satisfied. For repeated games, we show the minimum discount factors allowed to achieve the optimal performance in the parenthesis next to the performance metric. An immediate observation is the inefficiency of the NE, as expected.

When using one-shot game model with incentive schemes, the performance loss compared to using repeated games is small when the minimum payoff guarantee is small ($\gamma_i = 1$). However, when the minimum payoff guarantee increases, using one-shot games is far from optimality in terms of both the sum payoff and the fairness. Note that using one-shot games fails to satisfy the large minimum payoff guarantee ($\gamma_i = 14$). In summary, using repeated games has large performance gain over using one-shot games in most cases, and is necessary when the minimum payoff guarantee is large.

Under the specific parameters in this simulation, if we allow the discount factor to be sufficiently close to 1, using intervention in repeated games has little performance gain over repeated games without intervention, especially when the minimum payoff guarantee is large. This is because the minmax payoff without intervention is already small, such

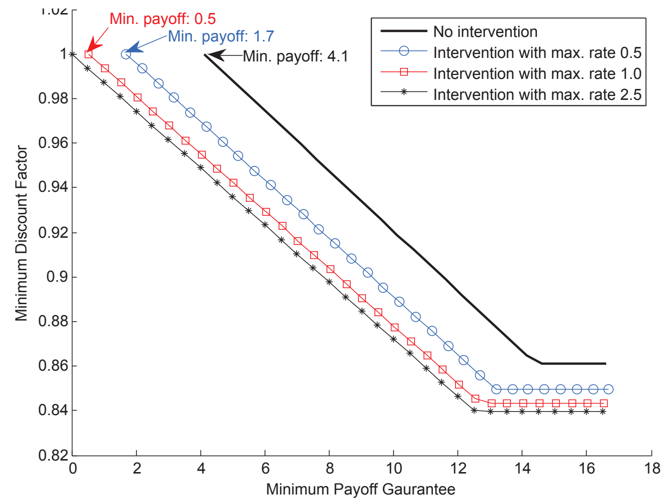

Fig. 5. The trade-offs between the minimum discount factor and the minimum payoff guarantee under different maximum intervention flow rates. At the beginning of each curve, we mark the smallest minimum payoff guarantees we can impose, which indicates the largest feasible set of the protocol design problem with each maximum intervention flow rate.

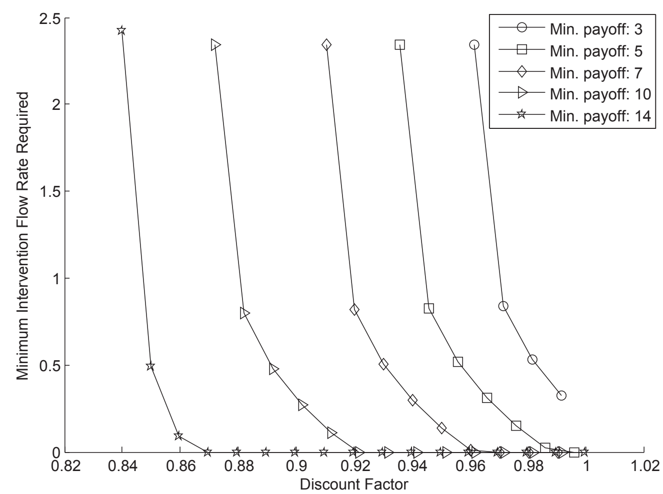

Fig. 6. The trade-offs between the required maximum intervention flow rate and the discount factor under different minimum payoff guarantees.

that its is Pareto dominated by some large minimum payoff guarantees. In this case, the advantage of using intervention in repeated games is that it allows smaller discount factors to achieve the same or better performance, compared to using repeated games without intervention. This is also confirmed by Fig. 5, which shows the minimum discount factor allowed to achieve the target payoff that maximizes the sum payoff, under different minimum payoffs γ and different maximum intervention flow rates.

B. Trade-off among δ, γ, and $\bar{a}_0$

Consider the same system as that in Fig. 3 and Fig. 5. Suppose that the target payoff is the one that maximizes the sum payoff. In Fig. 6, we plot the trade-off between the required maximum intervention flow rate and the discount factor under different minimum payoff guarantees. In Fig. 7, we plot the trade-off between the required maximum intervention rate and the minimum payoff guarantees under different discount

TABLE II
PERFORMANCE COMPARISON AMONG DIFFERENT SCHEMES UNDER DIFFERENT MINIMUM PAYOFFS GUARANTEES
(DISCOUNT FACTORS FOR REPEATED GAMES SHOWN IN THE PARENTHESIS)

Min. payoff	Metrics	NE [41]–[43]	One-shot [13][17][40]	Repeated w/o intervention	Repeated with intervention
$\gamma_i = 1$	Sum Payoff	39.3	110.4	110.2 (1.000)	114.2 (0.987)
	Absolute Fairness	4.0	9.6	16.7 (0.861)	16.7 (0.840)
$\gamma_i = 3$	Sum Payoff	39.3	85.8	108.2 (1.000)	108.2 (0.962)
	Absolute Fairness	4.0	10.6	16.7 (0.861)	16.7 (0.840)
$\gamma_i = 7$	Sum Payoff	N/A	64.4	96.2 (0.960)	96.2 (0.910)
	Absolute Fairness	N/A	10.3	16.7 (0.861)	16.7 (0.840)
$\gamma_i = 14$	Sum Payoff	N/A	N/A	75.2 (0.861)	75.2 (0.840)
	Absolute Fairness	N/A	N/A	16.7 (0.861)	16.7 (0.840)

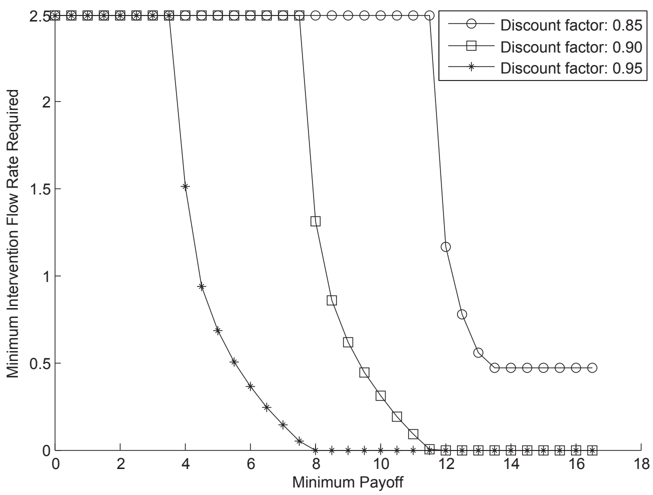

Fig. 7. The trade-offs between the required maximum intervention flow rate and the minimum payoff guarantee under different discount factors.

factors. The protocol designer can use these trade-off curves as guidelines for determining the maximum intervention flow rate required under different discount factors and minimum payoff guarantees.

VII. CONCLUSION

In this paper, we proposed a repeated game model generalized by intervention, which can be applied to a large variety of communication systems. We use repeated games to achieve the equilibrium payoffs that Pareto dominate some payoffs on the Pareto boundary of the nonconvex set of feasible payoffs in one-shot games. In addition, we apply intervention in repeated games, and show that intervention enlarges the limit set of equilibrium payoffs and simplifies the equilibrium protocol. Then we consider the protocol design problem of maximizing the welfare function. For any target payoff, we derive sufficient conditions on the discount factor and the intervention capability under which the target payoff can be achieved at a SPE. Under the discount factor and the intervention capability satisfying the sufficient condition, we construct an equilibrium protocol to achieve the target payoff. Simulation results show the great performance gain, in terms of sum payoff and max-min fairness, by using intervention in repeated games.

REFERENCES

[1] A. MacKenzie and S. Wicker, "Game theory and the design of self-configuring, adaptive wireless networks," *IEEE Commun. Mag.*, vol. 39, no. 11, pp. 126–131, Nov. 2001.

[2] E. Altman, T. Boulogne, R. El-Azouzi, T. Jimenez, and L. Wynter, "A survey on networking games in telecommunications," *Computers and Operations Research*, vol. 33, no. 2, pp. 286–311, Feb. 2006.

[3] J. Huang, R. A. Berry, and M. L. Honig, "Distributed interference compensation for wireless networks," *IEEE J. Sel. Areas Commun.*, vol. 24, no. 5, pp. 1074–1084, May 2006.

[4] T. Alpcan, T. Basar, R. Srikant, and E. Altman, "CDMA uplink power control as a noncooperative game," *Wireless Networks*, vol. 8, pp. 659–670, 2002.

[5] C. U. Saraydar, N. B. Mandayam, and D. J. Goodman, "Efficient power control via pricing in wireless data networks," *IEEE Trans. Commun.*, vol. 50, no. 2, pp. 291–303, Feb. 2002.

[6] J. Huang, R. A. Berry, and M. L. Honig, "Auction-based spectrum sharing," *Mobile Networks and Applications*, vol. 11, pp. 405–418, 2006.

[7] C. Long, Q. Zhang, B. Li, H. Yang, X. Guan, "Non-cooperative power control for wireless ad hoc networks with repeated games," *IEEE J. Sel. Areas Commun.*, vol. 25, no. 6, pp. 1101–1112, Aug. 2007.

[8] A. Schmeink, "On fair rate adaption in interference-limited systems," *European Trans. Telecommun.*, vol. 22, no. 5, pp. 200–210, Aug. 2011.

[9] D. Wang, C. Comaniciu, and U. Tureli, "Cooperation and fairness for slotted Aloha," *Wireless Personal Commun.*, vol. 43, no. 1, pp. 13–27, 2007.

[10] L. Yang, H. Kim, J. Zhang, M. Chiang, and C. W. Tan, "Pricing-based spectrum access control in cognitive radio networks with random access," in *Proc. 2011 IEEE INFOCOM*, pp. 2228–2236.

[11] T. Basar and R. Srikant, "Revenue-maximizing pricing and capacity expansion in a many-users regime," in *Proc. 2002 IEEE INFOCOM*, pp. 1556–1563.

[12] H. Shen and T. Basar, "Optimal nonlinear pricing for a monopolistic network service provider with complete and incomplete information," *IEEE J. Sel. Areas Commun.*, vol. 25, pp. 1216–1223, Aug. 2007.

[13] J. Park and M. van der Schaar, "The theory of intervention games for resource sharing in wireless communications," *IEEE J. Sel. Areas Commun.*, vol. 30, no. 1, pp. 165–175, Jan. 2012.

[14] J. Park and M. van der Schaar, "Stackelberg contention games in multiuser networks," *EURASIP J. Advances in Signal Process.*, vol. 2009, pp. 1–15, 2009.

[15] J. Park and M. van der Schaar, "Designing incentive schemes based on intervention: the case of imperfect monitoring," in *Proc. 2011 GameNets*.

[16] Y. Xiao, J. Park, and M. van der Schaar, "Intervention in power control games with selfish users," *IEEE J. Sel. Topics Signal Process.*, vol. 6, no. 2, pp. 165–179, Apr. 2012.

[17] Y. Gai, H. Liu, and B. Krishnamachari, "A packet dropping-based incentive mechanism for M/M/1 queues with selfish users," in *Proc. 2011 IEEE INFOCOM*, pp. 2687–2695.

[18] S. Stańczak and H. Boche, "On the convexity of feasible QoS regions," *IEEE Trans. Inf. Theory*, vol. 53, no. 2, pp. 779–783, Feb. 2007.

[19] E. G. Larsson, E. A. Jorswieck, J. Lindblom, and R. Mochaourab, "Game theory and the flat-fading Gaussian interference channel," *IEEE Signal Process. Mag.*, vol. 26, no. 5, pp. 18–27, Sep. 2009.

[20] J. Park and M. van der Schaar, "Medium access control protocols with memory," *IEEE/ACM Trans. Netw.*, vol. 18, no. 6, pp. 1921–1934, Dec. 2010.

[21] G. J. Mailath and L. Samuelson, *Repeated Games and Reputations: Long-run Relationships*. Oxford University Press, 2006.

[22] K. L. Judd, S. Yeltekin, and J. Conklin, "Computing supergame equilibria," *Econometrica*, vol. 71, no. 4, pp. 1239–1254, July 2003.
[23] S. Goldlücke and S. Kranz, "Infinitely repeated games with public monitoring and monetary transfers," *J. Economic Theory*, forthcoming.
[24] G. J. Mailath, I. Obara, and T. Sekiguchi, "The maximum efficient equilibrium payoff in the repeated prisoners' dilemma," *Games and Economic Behavior*, vol. 40, pp. 99–122, 2002.
[25] R. J. La and V. Anantharam, "Optimal routing control: repeated game approach," *IEEE Trans. Autom. Control*, vol. 47, no. 3, pp. 437–450, Mar. 2002.
[26] R. Etkin, A. Parekh, and D. Tse, "Spectrum sharing for unlicensed bands," *IEEE J. Sel. Areas Commun.*, vol. 25, no. 3, pp. 517–528, Apr. 2007.
[27] G. Theodorakopoulos and J. S. Baras, "Game theoretic modeling of malicious users in collaborative networks," *IEEE J. Sel. Areas Commun.*, vol. 26, no. 7, pp. 1317–1327, Sep. 2008.
[28] C. Pandana, Z. Han, and K. J. R. Liu, "Cooperation enforcement and learning for optimizing packet forwarding in autonomous wireless networks," vol. 7, no. 8, pp. 3150–3163, Aug. 2008.
[29] Y. Wu, B. Wang, K. J. R. Liu, and T. C. Clancy, "Repeated open spectrum sharing game with cheat-proof strategies," *IEEE Trans. Wireless Commun.*, vol. 8, no. 4, pp. 1922–1933, Apr. 2009.
[30] Z. Ji, W. Yu, and K. J. R. Liu, "A belief evaluation framework in autonomous MANETs under noisy and imperfect observation: vulnerability analysis and cooperation enforcement," *IEEE Trans. Mobile Comput.*, vol. 9, no. 9, pp. 1242–1254, Sep. 2010.
[31] L. Jiang, V. Anantharam, and J. Walrand, "How bad are selfish investments in network security?" *IEEE/ACM Trans. Netw.*, vol. 19, no. 2, pp. 549–560, Apr. 2011.
[32] B. Niu, H. V. Zhao, and H. Jiang, "A cooperation stimulation strategy in wireless multicast networks," *IEEE Trans. Signal Process.*, vol. 59, no. 5, pp. 2355–2369, May 2011.
[33] M. Felegyhazi, J. P. Hubaux, and L. Buttyan, "Nash equilibria of packet forwarding strategies in wireless ad hoc networks," *IEEE Trans. Mobile Comput.*, vol. 5, no. 5, pp. 463–476, May 2006.
[34] Z. Han, Z. Ji, and K. J. R. Liu, "A cartel maintenance framework to enforce cooperation in wireless networks with selfish users," *IEEE Trans. Wireless Commun.*, vol. 7, no. 5, pp. 1889–1899, May 2008.
[35] M. LeTreust and S. Lasaulce, "A repeated game formulation of energy-efficient decentralized power control," *IEEE Trans. Wireless Commun.*, vol. 9, no. 9, pp. 2860–2869, Sep. 2010.
[36] R. B. Myerson, "Utilitarianism, egalitarianism, and the timing effect in social choice problems," *Econometrica*, vol. 49, no. 4, pp. 883-897, June 1981.
[37] Y. Xiao, J. Park, and M. van der Schaar, "Repeated games with intervention: theory and applications in communications," Tech. Report. Available: http://arxiv.org/abs/1111.2456.
[38] J. E. Hopcroft, R. Motwani, and J. D. Ullman, *Introduction to Automata Theory, Languages, and Computation*. Addison Wesley, 2006.
[39] M. Chiang, P. Hande, T. Lan, and C. W. Tan, "Power control in wireless cellular networks," *Foundations and Trends in Networking*, vol. 2, no. 4, pp. 381–533, Apr. 2008.
[40] Y. Su and M. van der Schaar, "Linearly coupled communication games," *IEEE Trans. Commun.*, vol. 59, no. 9, pp. 2543–2553, Sep. 2011.
[41] K. Bharath-Kumar and J. M. Jaffe, "A new approach to performance oriented flow control," *IEEE Trans. Commun.*, vol. 29, pp. 427–435, 1981.
[42] C. Douligeris and R. Mazumdar, "A game theoretic perspective to flow control in telecommunication networks," *J. Franklin Institute*, vol. 329, no. 2, pp. 383–402, 1992.
[43] Z. Zhang and C. Douligeris, "Convergence of synchronous and asynchronous greedy algorithm in a multiclass telecommunications environment," *IEEE Trans. Commun.*, vol. 40, no. 8, pp. 1277–1281, Aug. 1992.

Yuanzhang Xiao received the B.E. and M.E. degree in electrical engineering from Tsinghua University, Beijing, China, in 2006 and 2009, respectively. He is currently pursuing the Ph.D. degree in the Electrical Engineering Department at University of California, Los Angeles. His research interests include game theory, optimization, communication networks, and network economics.

Jaeok Park received the B.A. degree in economics from Yonsei University, Seoul, Korea, in 2003, and the M.A. and Ph.D. degrees in economics from the University of California, Los Angeles, in 2005 and 2009, respectively. He is currently an Assistant Professor in the School of Economics at Yonsei University, Seoul, Korea. From 2009 to 2011, he was a Postdoctoral Scholar in the Electrical Engineering Department at the University of California, Los Angeles. From 2006 to 2008, he served in the Republic of Korea Army. His research interests include game theory, mechanism design, network economics, and wireless communication.

Mihaela van der Schaar is Chancellor's Professor of Electrical Engineering at University of California, Los Angeles. She is an IEEE Fellow, a Distinguished Lecturer of the Communications Society for 2011-2012 and the Editor in Chief of IEEE Transactions on Multimedia. She received an NSF CAREER Award (2004), the Best Paper Award from IEEE Transactions on Circuits and Systems for Video Technology (2005), the Okawa Foundation Award (2006), the IBM Faculty Award (2005, 2007, 2008), the Most Cited Paper Award from EURASIP: Image Communications Journal (2006), the Gamenets Conference Best Paper Award (2011) and the 2011 IEEE Circuits and Systems Society Darlington Award Best Paper Award. She holds 33 granted US patents. For more information about her research visit: http://medianetlab.ee.ucla.edu/