SIAM J. COMPUT.
Vol. 42, No. 1, pp. 230–264

CIRCUMVENTING THE PRICE OF ANARCHY: LEADING DYNAMICS TO GOOD BEHAVIOR*

MARIA-FLORINA BALCAN†, AVRIM BLUM‡, AND YISHAY MANSOUR§

Abstract. Many natural games have a dramatic difference between the quality of their best and worst Nash equilibria, even in pure strategies. Yet, nearly all results to date on dynamics in games show only convergence to *some* equilibrium, especially within a polynomial number of steps. In this work we initiate a theory of how well-motivated multiagent dynamics can make use of global information about the game—which might be common knowledge or injected into the system by a helpful central agency—and show that in a wide range of interesting games this can allow the dynamics to quickly reach (within a polynomial number of steps) states of cost comparable to the *best* Nash equilibrium. We present several natural models for dynamics that can use such additional information and analyze their ability to reach low-cost states for two important and widely studied classes of potential games: network design with fair cost-sharing and party affiliation games (which include consensus and cut games). From the perspective of a central agency, our work can be viewed as analyzing how a public service advertising campaign can help "nudge" behavior into a good state, when players cannot be expected to all blindly follow along but instead view the information as an additional input into their dynamics. We show that in many cases, this additional information is sufficient for natural dynamics to quickly reach states of cost comparable to the best Nash equilibrium of the game.

Key words. game theory, dynamics in games, potential games, quality of equilibria

AMS subject classifications. 68Q25, 68T05, 91A06, 91A40

DOI. 10.1137/110821317

1. Introduction.

Understanding the *quality* of Nash equilibria in a game has been a major focus of algorithmic game theory. The main motivation is to understand the additional cost that is incurred when we assume that agents are behaving in a strategic manner, rather than performing a global optimization. Koutsoupias and Papadimitriou [39] proposed the notion of *price of anarchy (PoA)* as the ratio of the cost of the worst Nash equilibrium to the social optimum [45]. The PoA has been studied for a large variety of games, including routing, network design with cost-sharing, job scheduling, and network creation (see [49, 24, 28, 3, 25, 33]). While the PoA takes a worst-case view, the *price of stability (PoS)* [5, 23] instead considers the ratio of the *best* Nash equilibrium to the social optimum. In fact, for many natural games, the PoA may be very large while the PoS is quite low. For example, for

*Received by the editors January 18, 2011; accepted for publication (in revised form) November 16, 2012; published electronically January 29, 2013. This paper combines and extends results appearing in "Improved Equilibria via Public Service Advertising," SODA 2009, and "Circumventing the Price of Anarchy: Leading Dynamics to Good behavior," ICS 2010.

http://www.siam.org/journals/sicomp/42-1/82131.html

†School of Computer Science, Georgia Institute of Technology, Atlanta, GA 30332 (ninamf@cc.gatech.edu). This author's work was supported in part by ONR grant N00014-09-1-0751, AFOSR grant FA9550-09-1-0538, and NSF Career grant CCF-0953192.

‡School of Computer Science, Carnegie Mellon University, Pittsburgh, PA 15213 (avrim@cs.cmu.edu). This author's work was supported in part by NSF grants CCF-0514922, CCF-0830540, and CCF-1101215.

§Blavatnik School of Computer Science, Tel Aviv University, Tel Aviv, Israel (mansour@tau.ac.il). This author's work was supported in part by the IST Programme of the European Community, under the PASCAL Network of Excellence, by a grant from the Israel Science Foundation, by a grant from United States-Israel Binational Science Foundation (BSF), by a grant from the Israeli Ministry of Science (MoS), and by the Israeli Centers of Research Excellence (I-CORE) program (Center 4/11). This publication reflects the authors' views only.

consensus games with n players, the PoA is $\Theta(n^2)$, whereas there is always a Nash equilibrium that is also socially optimal (and hence the PoS is 1). Another example is fair cost-sharing games, where the PoA is $\Theta(n)$ and the PoS is $\Theta(\log n)$.

However, equilibria are only part of the story. Players in multiagent systems are typically dynamic, with behavior that depends on their knowledge and understanding of the game and their beliefs about other players in the system. As a consequence, there has been substantial interest in analyzing dynamics as well. For example, regret-minimizing behavior [34, 41, 16, 6, 54] is known to approach the set of coarse-correlated equilibria; internal-regret minimizing behavior [30, 31, 32, 35, 36, 51] approaches the set of correlated equilibria; and in potential games, best-response behavior will reach a pure Nash equilibrium. Yet these results in general show only convergence to *some* equilibrium, which could be the worst equilibrium.[1] This is especially worrisome if the worst equilibrium can indeed be much worse than the best one, such as in games where the PoA is high and yet the PoS is low. A key problem is that these dynamics are highly myopic, viewing the environment either as static or as an adversarial sequence, but not as composed of entities with their own interests in mind and not using any information about the overall structure of the game. In fact, one can show that for fair cost-sharing games, any stochastic dynamics that depends only on the observed cost of actions (such as best-response, noisy best-response, multiplicative weighting, etc.) will not be able to reach states of low cost within a polynomial number of steps; see Appendix C for a formal model and lower bound. In order to overcome this barrier, in this work we initiate a theory of how well-motivated multiagent dynamics can make use of global information about the game—which might be common knowledge or injected into the system by a helpful planner—and show that in a wide range of interesting games this can allow the dynamics to reach states of cost comparable to the *best* Nash equilibrium.

In particular, in this work we address the following question: for games with equilibria of wildly different social costs, how can players who are selfish but who also possess information about overall socially beneficial behaviors reach states of cost comparable to the best equilibrium or the PoS of the game? Having this kind of information is in many cases very natural. To take a specific example, a canonical class of games with a wide gap between the quality of their equilibria are *fair cost-sharing* games. Here, each player i has a source s_i and sink t_i and must choose an s_i-t_i path in an underlying weighted directed graph; players then split the cost of edges used with others using those edges. Even in the simplest nontrivial case of a two-node graph with two parallel links of costs 1 and k, $k \leq n$, the worst equilibrium (all players on the more expensive edge) is a factor k more expensive than the best equilibrium (all players on the cheaper edge).[2] Yet given the graph, in many cases such as when all players have the same source and sink, it is easy to compute an overall social optimum. In others (in general this is the NP-hard directed Steiner forest problem) a central planner after substantial computation may be able to announce one. Of course, the players might not care—after all, they are selfish, and these

[1] The set of coarse-correlated equilibria contains the set of correlated equilibria which contains the set of Nash equilibria which contains the set of pure Nash equilibria. In some cases, specific regret-minimizing algorithms can be shown to perform better than the worst coarse-correlated equilibrium but still at the level of the worst pure Nash equilibrium [37].

[2] One can view this case as n people who want to split the cost of a van service from the airport to their hotel. Two equally good services are available, one a factor of k more costly than the other. For $1 < k < n$, sharing the expensive service would be an equilibrium, but not one that reasonable people would reach if the overall costs of each van were known.

behaviors might only be good for an individual *if others follow it as well*. But, if this information is placed as an extra "untrusted" input into a natural dynamics, can that help behavior to approach a near-optimal equilibrium? To take another example, consider *cut games*, where players are nodes in a network who can be in one of two states and incur cost proportional to the number of their neighbors who are in the same state (modeling wireless radio transmission frequencies, for instance). Suppose a central planner proposes some assignment of states and players individually decide whether to follow the advice using some natural learning rule—when is this enough for the system to reach a near-optimal configuration?

In this work, we propose and analyze several natural models for this type of scenario and dynamics: a simple model we call the "basic advertising" model, a more individually motivated setting we call the "learn-then-decide" model, and our most general model we call the "smoothly adaptive" model. We show that for many interesting games, under these models the system will indeed reach states of cost comparable to that given by the PoS of the game. This is noteworthy because as mentioned above, for fair cost-sharing, one of our main games of interest, purely myopic dynamics, cannot in general reach states of cost $o(\mathbf{OPT} \cdot n/\log n)$ in a polynomial number of steps, thus performing much worse than the PoS (see Appendix C). In contrast, we will be able to achieve states of cost $O(\mathbf{OPT} \cdot \log n)$.

Our work can be thought of both from the perspective of the individual players and from the perspective of a helpful but untrusted central agency or planner. From the latter perspective, the question here is, can such an agency help to "nudge" players from a bad equilibrium state into a good one by just injecting a small amount of information into the system? In this context, the central agency can be viewed as proposing a behavior of good social cost (a "public service advertising campaign") that the players incorporate into their dynamics. In our most basic model, the players only once select randomly between following the proposed strategy (with some probability α) or not (with probability $1-\alpha$). Players who chose to follow the proposed strategy temporarily stick with their decision, while the others settle on some conditional equilibrium for themselves. Then, in a final phase all players perform a best-response dynamics to converge to a final equilibrium. (The convergence is guaranteed since the discussion is limited to potential games.) We show that natural games in fact exhibit an interesting structure: for some such as fair cost-sharing, even a very small value of α is sufficient to bring the system into a low-cost equilibrium, whereas for others such as consensus and cut games there is a threshold phenomenon—a tipping point in the fraction α needed to reach low-cost states. In our more learning-based models, the players instead repeatedly reselect between following the proposed strategy or performing a best response and potentially adapt their probabilities using a learning rule. Conceptually, the benefit of these models is that the players are "symmetric" and can continually switch between the two alternatives, which better models selfish behavior. This repeated process is what enables the players to both explore and exploit the two alternative actions.[3]

For some of the games we consider, *finding* an optimal solution is NP-hard. However, all our results apply given a good approximation to the optimal as well. One can also imagine that agents or central planners who care sufficiently about the instance at hand have put in the effort to find a good solution. In either case, one way to view our results is to separate the *computational* problem of finding a near-socially-optimal

[3] This frequent randomization makes the analysis of these dynamics technically more challenging, however.

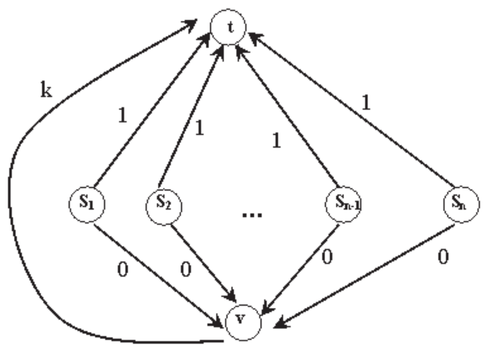

FIG. 1. *A directed cost-sharing game with multiple sources and a common sink that models a setting where each player can choose either to drive its own car to work at a cost of* 1 *or share public transportation with others, splitting an overall cost of* k*. Note that for any* $1 < k < n$*, if players arrive one at a time and each greedily chooses a path minimizing its cost, then the equilibrium obtained has cost* n*, whereas* **OPT** *has cost only* k*.*

behavior from the *game-theoretic* problem of analyzing natural dynamics under the assumption that such behavior is known. Overall, we hope to lay foundations for exploring how various forms of knowledge of a game's structure can be used by natural learning agents to reach overall socially beneficial states.

1.1. Our results. We consider two main classes of potential games: fair cost-sharing in networks and party affiliation games. In *fair cost-sharing*, n players choose routes in a network and split the cost of edges they take with others using the same edge; these games can model scenarios such as whether to drive one's own car to work or to share public transportation with others (see Figure 1 for a simple example). These games have a PoA of n and PoS of $\Theta(\log n)$. In *party affiliation games*, players are nodes in a network where each edge is labeled $+$ or $-$; players choose colors and incur a cost for each $+$-neighbor of different color and each $-$-neighbor of the same color. Two special cases are *consensus* and *cut* games, corresponding to only $+$-edges or only $-$-edges, respectively. These games have PoA of $\Theta(n^2)$ and PoS of 1. See sections 2.1 and 3.2 for formal definitions.

We start by analyzing a clean model in which a helpful planner can guide dynamics by broadcasting a socially beneficial behavior—a "public service advertising campaign"—and having some fraction of the population follow this advice temporarily. In particular, after the advertising campaign is announced, we assume a random fraction $0 < \alpha < 1$ of players will be receptive to it (the *receptive* players) and temporarily follow along, playing their part of the advertised behavior. The others will instead behave selfishly and reach an equilibrium for themselves conditioned on the fixed behavior of the receptive players. After this equilibrium is reached, the campaign wears off and the entire system evolves using best-response dynamics. For this model, we prove that for fair cost-sharing games, for any $\alpha > 0$, the expected cost of the new equilibrium is at most an $O((\log n)/\alpha)$ factor larger than the optimal solution, much closer to the $\Theta(\log n)$ PoS than the factor n PoA. These results extend to the case where we add to the cost model a linear latency function which depends on the load observed on the edges.[4] On the other hand, for party affiliation games where

[4]The proof for this case uses a shadow-game argument and is technically more challenging. We remark that after modifying a constant fraction of the players' actions the social cost can still be quite high, so the result cannot be derived by the standard potential-function argument.

players have degree $\omega(\log n)$ we show a threshold property: any value of $\alpha < 1/2$ is not sufficient to improve the equilibrium beyond the $\Omega(n^2)$ price of anarchy, while any value $\alpha > 1/2$ is sufficient to produce behavior within an $O(1)$ factor of optimal. We also point out there do exist classes of "stubborn" games such as job scheduling on unrelated machines where no constant $\alpha < 1$ is sufficient to guarantee good behavior.

We next consider a more adaptive learning-style model where rather than having separate receptive and nonreceptive players, each player i individually decides on each round between following the advertised behavior or choosing a best-response action with some probability p_i. In the simplest *learn-then-decide* version of this model, these p_i are fixed up to some common time T^* (the end of the "exploration phase"), at which point players commit to one option or the other based on their experience. In the more flexible (but more difficult to analyze) *smoothly adaptive* version, players instead slowly adjust their probabilities p_i over time using some arbitrary learning rule of their choosing. The only requirement is a bound on the rate at which the p_i may change (i.e., on the learning rate). The goal again is in both models to analyze when this process results in a low-cost state. For fair cost-sharing in graphs with m edges, we show that if there are sufficiently many players of each type (i.e., having the same source and sink), then indeed in both models the system will, with high probability, approach states of cost only a factor $O(\log mn)$ larger than optimal—even $O(1)$ if the number of each type is high enough. We also show a logarithmic bound for the learn-then-decide model that applies in general, no matter how many players of each type there are. For party affiliation games, we show that all the positive results in the basic advertising model carry over to the learn-then-decide model as well. Moreover, for consensus games—as well as more general party affiliation games in the case that every node experiences low cost in the advertised solution—we show that in both models the bound on the minimum degree in the graph can be removed and the system will still, with high probability, quickly reach a good state. Thus, these models perform even better than the basic advertising model. These results are described in section 4.

Conceptually, we show how information about socially beneficial behaviors—which in many games is readily available and in others can be injected into the system by a central planner—can be used by self-interested players in order to reach states of cost comparable to the best equilibrium. We point out that the dynamics we study are not complicated and make sense from the point of the individual. Even in the basic advertising model, the receptive players only follow the advice temporarily, and in the models of section 4 all players are on an equal footing, testing out the advice along with best-response behavior as the game dynamics progresses.

1.2. Related work. *Dynamics and convergence to equilibria.* All the games we consider in this paper are potential games [43]. It is well known that in potential games, best-response dynamics, in which players take turns each making a best-response move to the current state of all the others, is guaranteed to converge to a pure-strategy Nash equilibrium [47, 43]. Significant effort has been spent recently on the convergence time of these dynamics [45, 1], with both examples of games in which such dynamics can take exponential time to converge, and results on fast convergence to states with cost comparable to the PoA of the game [29, 7]. However, there is in general no guarantee on the quality of the equilibrium reached, and even for natural initial conditions, the ratio of the cost to optimal could be as high as the PoA for the game.

Another well-known general result that applies to any finite game is that if players each follow a "no-internal-regret" strategy, then the empirical distribution of play is guaranteed to approach the set of correlated equilibria of the game [30, 32, 35]. In particular, for good no-regret algorithms, the empirical distribution will be an ϵ-correlated equilibrium after $O(1/\epsilon^2)$ rounds of play [13]. Kleinberg, Piliouras, and Tardos [37] analyze a version of the weighted-majority algorithm in congestion games, showing that behavior converges to the set of weakly stable equilibria. This implies performance better than the worst correlated equilibrium, but even in this case the guarantee is no better than the worst pure-strategy Nash equilibrium.

Noisy best-response dynamics has been shown in the limit to reach states of minimum global potential [14, 15, 42] and thus provide strong positive results in games with a small gap between potential and cost. However, this convergence may take exponential time, even for fair cost-sharing games. We provide a formal lower bound in Appendix C.

An alternative line of work assumes that the system starts empty and players join one at a time. Charikar et al. [17] analyze fair cost-sharing in this setting on an *undirected* graph where all players have a common sink. Their model has two phases: in the first, players enter one at a time and use a greedy algorithm to connect to the current tree, and in the second phase the players undergo best-response dynamics. They show that in this case a good equilibrium (one that is within only a polylog(n) factor of optimal) is reached. We discuss this result further in Appendix B. We remark here that for directed graphs, it is easy to construct simple examples where this process reaches an equilibrium that is $\Omega(n)$ from optimal (see Figure 1).

Stackelberg strategies and the value of altruism. A Stackelberg leader strategy is a strategy by which one player (called the leader) aims to achieve the highest utility for itself under the assumption that the other player(s) will reach a Nash equilibrium in the game conditioned on its behavior [52, 46, 38]. Building on this concept, Sharma and Williamson [50] define the notion of the *value of altruism* in which a central authority controls some fraction of the players (or flow) in a symmetric network game, with the aim of maximizing social welfare under the assumption that others will reach a Nash equilibrium conditioned on this behavior. Our basic advertisement model is related to this work in that the central planner would like the equilibria of nonreceptive players given the behavior of receptive players to have as low a cost as possible. However, the central planner cannot control *which* players will follow the given advice (a distinction when the game is not symmetric), and moreover we assume that after this equilibrium is reached, the advice wears off and the overall play then follows best-response dynamics to an overall pure Nash equilibrium. In particular, note that the notion of dynamics is crucial to our model, because otherwise play could just revert to the initial (equilibrium) state.

Taxation. There has also been work on using taxes to improve the quality of behavior [22, 21]. Here the aim is to adjust the *utilities* of each player via taxes so that the only Nash equilibria in the new game correspond to optimal or near-optimal behavior in the original game. In contrast, our focus is on *information*: how can natural dynamics with the aid of global information about the game avoid being stuck in poor equilibria and reach states of near-optimal cost?

Strong price of anarchy. It is also worthwhile to compare and contrast our basic advertising model with that of the *strong price of anarchy* [4]. The k-strong price of anarchy [4] focuses on those equilibria such that no subset of at most k players can deviate and *all* strictly benefit; thus it is like a model of stability when players can intelligently form coalitions. In contrast, we consider players that are more myopic in the usual Nash sense, except some are willing to give the advertised behavior a try. The two solution concepts are incomparable in their final guarantees: for job scheduling on two unrelated machines, the 2-strong price of anarchy is $O(n)$ [4], while in our advertising model the PoA is still unbounded (see Theorem 31). On the other hand, one can also show a reverse example, of a high-cost strong equilibrium [4], where modifying the action of any single job would lead to an optimal outcome.

Subsequent work. Subsequent to the initial conference publication of this work [8, 10], the models described here have been used to analyze other natural multiagent settings. Demaine and Zadimoghaddam [26] consider a network formation game in which each player may construct up to k edges and wants to minimize its total distance to all other nodes. They show that in our basic advertising model, any fraction $\alpha > 0$ of receptive players is sufficient to achieve cost within $O(1/\alpha)$ of optimal. Balcan et al. [11] consider packing and covering games, modeling a number of on/off distributed decision-making problems such as awake/sleep decisions in low-power sensor networks or decisions in the placement of costly facilities. They show that in both the basic advertising model and the learn-then-decide model, the system can be ensured to reach states within an $O(\log n)$ factor of optimal for a broad class of such games and even an $O(1)$ factor of optimal for vertex-cover and independent-set games.

1.3. Organization. We begin with general notation and definitions in section 2. We then formally describe the advertisement model in section 3 and analyze two classes of games in this model: cost-sharing games in section 3.1 and party affiliation games in section 3.2. We also discuss extensions to the model and further issues in section 3.3. We then describe our learning based models in section 4. In section 4.3 we analyze fair cost-sharing in these models, and in section 4.4 we analyze party affiliation games. Interestingly, in some cases the results are even more positive than for the basic advertisement model; e.g., see sections 4.4.2 and 4.4.3. We finish with a discussion in section 5.

2. Preliminaries. We denote a game by a tuple $\mathcal{G} =< N, (\mathcal{S}_i), (\texttt{cost}_i) >$, where N is a set of n players, $\mathcal{S}_i$ is the finite action space of player $i \in N$, and $\texttt{cost}_i$ is the cost function of player i. The joint action space of the players is $\mathcal{S} = \mathcal{S}_1 \times \cdots \times \mathcal{S}_n$. For a joint action $s \in \mathcal{S}$ we denote by s_{-i} the actions of players $j \neq i$, i.e., $s_{-i} = (s_1, \ldots, s_{i-1}, s_{i+1}, \ldots, s_n)$. The cost function of player i maps a joint action $s \in \mathcal{S}$ to a real nonnegative number, i.e., $\texttt{cost}_i : \mathcal{S} \to \mathbb{R}^+$. Every game has an associated social cost function $\texttt{cost} : \mathcal{S} \to \mathbb{R}$ that maps a joint action to a real value. In this paper we concentrate on the case where the social cost function is the sum of the player's costs, i.e., $\texttt{cost}(s) = \sum_{i=1}^{n} \texttt{cost}_i(s)$. The optimal social cost is $\mathbf{OPT}(\mathcal{G}) = \min_{s\in\mathcal{S}} \texttt{cost}(s)$. We sometimes overload notation and use $\mathbf{OPT}$ for a joint action s that achieves cost $\mathbf{OPT}(\mathcal{G})$.

Given a joint action s, the *best response (BR)* of player i is the set of actions $BR_i(s)$ that minimizes its cost, given the other players actions s_{-i}, i.e., $BR_i(s_{-i}) = \arg\min_{a\in\mathcal{S}_i} \texttt{cost}_i(a, s_{-i})$.

A joint action $s \in \mathcal{S}$ is a *pure Nash equilibrium (NE)* if no player $i \in N$ can benefit from unilaterally deviating to another action, namely, every player is playing a best-response action in s, i.e., $s_i \in BR_i(s_{-i})$ for every $i \in N$. A best-response dynamics is a process in which at each time step, some player which is not playing a best response switches its action to a best-response action, given the current joint action. In this paper we will concentrate on games in which any best-response dynamics converges to a pure Nash equilibrium (which are equivalent to the class of ordinal potential games [43]).

Let $\mathcal{N}(\mathcal{G})$ be the set of Nash equilibria of the game $\mathcal{G}$. The PoA is defined as the ratio between the maximum cost of any Nash equilibrium and the social optimum, i.e., $(\max_{s\in\mathcal{N}(\mathcal{G})} \texttt{cost}(s))/\mathbf{OPT}(\mathcal{G})$. The PoS is the ratio between the minimum-cost Nash equilibrium and the social optimum, i.e., $(\min_{s\in\mathcal{N}(\mathcal{G})} \texttt{cost}(s))/\mathbf{OPT}(\mathcal{G})$.

2.1. Classes of games studied. One of the main classes of games we study in this paper, because of their rich structure and wide gap between PoA and PoS, is the class of *fair cost-sharing games*. These games are defined as follows. We are given a graph $G = (V, E)$ on m edges, which can be directed or undirected, where each edge $e \in E$ has a nonnegative cost $c_e \geq 0$. There is a set $N = \{1, \dots, n\}$ of n players, where player i is associated with a source s_i and a sink t_i. The strategy set of player i is the set $\mathcal{S}_i$ of $s_i - t_i$ paths. In an outcome of the game, each player i chooses a single path $P_i \in \mathcal{S}_i$. Given a vector of players' strategies $s = (P_1, \dots, P_n)$, let x_e be the number of agents whose strategy contains edge e. In fair cost-sharing, the cost of each edge e is shared equally among the x_e players using that edge, so the cost to agent i is $\texttt{cost}_i(s) = \sum_{e\in P_i} \frac{c_e}{x_e}$ and the goal of each agent is to connect its terminals with minimum total cost. The social cost of an outcome $s = (P_1, \dots, P_n)$ is defined to be $\texttt{cost}(P_1, \dots, P_n) = \sum_{e\in\cup_i P_i} c_e = \sum_i \texttt{cost}_i(s)$. It is well known that fair cost-sharing games are potential games [43, 5] and the PoA in these games is $\Theta(n)$, while the PoS is $H(n)$ [5], where $H(n) = \sum_{i=1}^{n} 1/i = \Theta(\log n)$.

The other main class of games we consider is that of party affiliation games [19]. Here, players are vertices in an n-vertex graph, where each edge has label $+$ or $-$. Players must choose a color, red or blue, and incur a cost equal to the number of positive neighbors (neighbors along edges labeled $+$) of different color, plus the number of negative neighbors of the same color. When all edges are positive these are called *consensus* games, and when all edges are negative these are called *cut* games. We define social cost to be the sum of player costs, plus 1 in order to keep all ratios finite. These games have a PoA of $\Theta(n^2)$ and a PoS of 1. For further details, see section 3.2.

3. The advertisement model. The first model we introduce and study in this paper is one that we call the *advertising model*. In this model, a central planner first suggests to each player an alternative action, and each player accepts the proposed action with some (constant) probability α. The players that accept the new action are called receptive players, and they stay with the new action long enough so that the nonreceptive players converge to some Nash equilibrium, given the fixed actions of the receptive players.[5] Then all players follow an arbitrary best-response dynamics and converge to some pure Nash equilibrium. (We will only consider games where best-response dynamics is guaranteed to converge to a pure Nash equilibrium.) We define this model formally in Figure 2.

When we refer later to an *advertising strategy* we mean a joint action $s^{ad} = \texttt{advertise}(\mathcal{G})$. Given a set of receptive players R, there is a set $U(s^{ad}, R)$ which includes all the equilibria s^f that the dynamics could reach having an advertisement s^{ad} and a set of receptive players R. The *expected cost of the final equilibrium* given s^{ad} is $E_R[\max_{s^f\in U(s^{ad},R)} \texttt{cost}(s^f)]$. When we say that *for game $\mathcal{G}$ there exists a strategy for the advertising model which has an expected cost of the final equilibrium at most X*, we mean that there exists a joint action s^{ad} for $\mathcal{G}$ such that $E_R[\max_{s^f\in U(s^{ad},R)} \texttt{cost}(s^f)] \leq X$.

3.1. Cost-sharing games. We show below that in fair cost-sharing games there exists a strategy for the advertising model producing an equilibrium whose expected

[5] In section 3.3.1 we consider a variation in which this is relaxed to a small number of best-response moves by the nonreceptive players.

Input: Game $\mathcal{G}$, parameter $\alpha > 0$.
Initially players are playing some joint action $s^{ini} \in \mathcal{S}$.

1. An advertising campaign proposes an action to each player. Formally, let $s^{ad} = \texttt{advertise}(\mathcal{G})$ be the advertised joint action.
2. Each player i independently decides to follow or not to follow the proposed action s_i^{ad} by flipping a coin of bias α. Let R be set of players who decide to follow the proposal - we will call them the *receptive players*. Each player $i \in R$ now switches to playing s_i^{ad}.
3. The non-receptive players (player in $N \setminus R$) settle on a Nash equilibrium s^{nr} for themselves, given that the receptive players play s^{ad}. Namely, for each player $j \in N \setminus R$ we have that $s_j^{nr} \in BR_j(s_{-j}^{nr}; s_R^{ad})$. The equilibrium s^{nr} for players in $N \setminus R$ might be *adversarially selected*. Let $s^{med} = (s^{nr}; s_R^{ad})$ be the behavior at this point.
4. The players in R stop following the advertising campaign, and the entire set of players N undergoes a best response dynamics until a Nash equilibrium s^f for the whole game $\mathcal{G}$ is reached.

FIG. 2. *Advertising model.*

cost is at most an $O((1/\alpha)\log n)$ factor larger than the cost of the optimal solution. Thus we get significant benefit from advertising in these games.

Before presenting the proof of our main results we first give two useful lemmas. The first is a well-known characterization of the potential function $\Phi(s) = \sum_{e \in E} \sum_{x=1}^{x_e} f_e(x)$ of these games [43, 5], where $f_e(x) = c_e/x$ is the cost of edge e as a function of the number of players x using it.

LEMMA 1. *In fair cost-sharing games, for any joint action $s \in \mathcal{S}$ we have* $\texttt{cost}(s) \leq \Phi(s) \leq H(n) \cdot \texttt{cost}(s)$.

The second lemma, whose proof is in Appendix A, states the following useful property of a binomial random variable.

LEMMA 2. *Let X be a binomial random variable distributed $Bi(n,p)$; then* $\mathbf{E}_X[\frac{c}{X+1}] = O(\frac{c}{p \cdot n})$.

We start with our main result concerning fair cost-sharing games.

THEOREM 3. *For fair cost-sharing games, there exists a strategy for the advertising model which has an expected cost of the final equilibrium at most $O((1/\alpha)(\log n) \cdot$* $\texttt{cost}(\mathbf{OPT}))$.

Proof. Fix some optimal solution $\mathbf{OPT}$. Let $s^{ad} = \mathbf{OPT}$, so the advertising strategy will be to tell each player i to use his path P_i^{OPT} in $\mathbf{OPT}$. Let R be the set of receptive players and $\bar{R} = N \setminus R$. For every edge e, let n_e^{opt} denote the number of people who use edge e in $\mathbf{OPT}$, and let us decompose this quantity into the number $n_{e,R}^{\text{opt}}$ of those in set R and the number $n_{e,\bar{R}}^{\text{opt}}$ of those not in R (both of which are random variables); so

$$n_e^{\text{opt}} = n_{e,R}^{\text{opt}} + n_{e,\bar{R}}^{\text{opt}}.$$

We start by bounding the expected worst-case cost of the behavior s^{med} produced at the end of step 3, that is, $\mathbf{E}_R[\max_{s^{med}=(s^{nr},s_R^{ad})} \texttt{cost}(s^{med})]$, where the max is taken over all behaviors s^{nr} that are Nash equilibria for players in $N \setminus R$ given the behavior of players in R. We will call this $\mathbf{E}[\texttt{cost}(s^{med})]$ for short. For every edge e let B_e denote the cost of edge e for a nonreceptive player $i \in \bar{R}$ under the joint action s^{med}, and let A_e denote the cost of edge e for a receptive player $i \in R$ under the joint action s^{med}. Let

$$X_e = c_e/(1 + n^{\text{opt}}_{e,R}) \text{ and } X'_e = c_e/n^{\text{opt}}_{e,R}.$$

We clearly have $B_e \leq X_e$ and $A_e \leq X'_e$. So, for any player $i \notin R$ we have $\texttt{cost}_i(P_i^{OPT}, s^{med}_{-i}) \leq \sum_{e \in P_i^{OPT}} X_e$ and for any player $i \in R$ we have $\texttt{cost}_i(P_i^{OPT}, s^{med}_{-i}) \leq \sum_{e \in P_i^{OPT}} X'_e$. Since s^{med} is an equilibrium for the nonreceptive players we have $s^{med}_i \in BR_i(s^{med}_{-i})$ for $i \in \bar{R}$. This implies the following upper bound on the expected total cost at the end of step 3:

$$\mathbf{E}[\texttt{cost}(s^{med})] \leq \mathbf{E}\left[\sum_{i \notin R} \sum_{e \in P_i^{OPT}} X_e\right] + \mathbf{E}\left[\sum_{i \in R} \sum_{e \in P_i^{OPT}} X'_e\right].$$

Rearranging the sum over players to be a sum over edges we get

$$\mathbf{E}[\texttt{cost}(s^{med})] \leq \mathbf{E}\left[\sum_e X_e \cdot n^{\text{opt}}_{e,\bar{R}}\right] + \mathbf{E}\left[\sum_e X'_e \cdot n^{\text{opt}}_{e,R}\right].$$

Note that $X'_e \leq 2X_e$ when $n^{\text{opt}}_{e,R} > 0$. This implies that

$$\mathbf{E}[\texttt{cost}(s^{med})] \leq 3\mathbf{E}\left[\sum_e X_e \cdot n^{\text{opt}}_e\right] = 3\sum_e \mathbf{E}[X_e] \cdot n^{\text{opt}}_e.$$

So, to complete the proof we need only to analyze $\mathbf{E}[X_e]$. Lemma 2 implies that $\mathbf{E}[X_e]$ is $O\left(c_e/(\alpha \cdot n^{\text{opt}}_e)\right)$ for $n^{\text{opt}}_e \geq 1$. This implies that the expected cost at the end of step 3 satisfies

$$\mathbf{E}[\texttt{cost}(s^{med})] = O\left((1/\alpha) \sum_{e \in \mathbf{OPT}} c_e\right) = O((1/\alpha)\ \texttt{cost}(\mathbf{OPT})). \tag{3.1}$$

Equation (3.1) together with Lemma 1 implies that the expected value of the potential function Φ at the end of step 3 is only an $O((1/\alpha)\log n)$ factor larger than the cost of **OPT**, i.e.,

$$\mathbf{E}[\Phi(s^{med})] = O((1/\alpha)(\log n) \cdot \texttt{cost}(\mathbf{OPT})).$$

This implies that the expected cost of the final equilibrium at the end of step 4, i.e., $\mathbf{E}[\texttt{cost}(s^f)]$, is at most that large, as desired. ∎

Note that in the proof of Theorem 3, no special properties of **OPT** are used, only its cost. Replacing **OPT** with any other valid solution F simply replaces $\texttt{cost}(\mathbf{OPT})$ with $\texttt{cost}(\text{F})$ in the bound. Therefore, we immediately have the following.

THEOREM 4. *Consider fair cost-sharing games and a joint action* F. *Using* $s^{ad} = F$ *for the advertising model yields a final equilibrium with expected cost at most* $O((1/\alpha)(\log n) \cdot \texttt{cost}(\text{F}))$.

3.1.1. Extensions. A well-studied extension of the fair cost-sharing game is one where instead of a constant cost c_e, each edge has a cost $c_e(x)$ that is a nondecreasing but concave function of the number of players x using that edge [5]. For example, this can model a buy-at-bulk economy of scale for buying edges that can be used by more players. Notice that the cost of an edge $c_e(x)$ might increase with the number of players using it, but the cost per player $f_e(x) = c_e(x)/x$ decreases if $c_e(x)$ is concave. We can extend our result to this case as well.

THEOREM 5. *For fair cost-sharing with nondecreasing concave cost functions $c_e(x)$, there exists a strategy for the advertising model yielding a final equilibrium of expected cost at most $O((1/\alpha)(\log n) \cdot \texttt{cost}(\mathbf{OPT}))$.*

Proof. The argument is similar to that in the proof of Theorem 3. First note that these games remain potential games [43, 5] with potential $\Phi(s) = \sum_{e \in E} \sum_{x=1}^{x_e} c_e(x)/x$.

As in the proof of Theorem 3, to analyze $\mathbf{E}[\texttt{cost}(s^{med})]$, for every edge e let B_e denote the cost of edge e for a nonreceptive player $i \in \bar{R}$ under the joint action s^{med}, and let A_e denote the cost of edge e for a receptive player $i \in R$ under the joint action s^{med}. Let $X_e = c_e(1 + n_e^{\text{opt}})/(1 + n_{e,R}^{\text{opt}})$ and let $X'_e = c_e(n_e^{\text{opt}})/n_{e,R}^{\text{opt}}$. Since by assumption the cost per player $f_e(x) = c_e(x)/x$ decreases with x, while $c_e(x)$ increases with x, we have $B_e \le c_e(1 + n_{e,R}^{\text{opt}})/(1 + n_{e,R}^{\text{opt}}) \le X_e$ and $A_e \le c_e(n_{e,R}^{\text{opt}})/n_{e,R}^{\text{opt}} \le X'_e$. In addition, $X'_e \le 2X_e$, for $n_{e,R}^{\text{opt}} > 0$. So as in the proof of Theorem 3 an upper bound on the expected cost at the end of step 3 is

$$\mathbf{E}[\texttt{cost}(s^{med})] \le 3 \sum_e \mathbf{E}[X_e] \cdot n_e^{\text{opt}} = O((1/\alpha)\texttt{cost}(\mathbf{OPT})).$$

Furthermore, as shown in [5], Lemma 1 holds for this game as well. These then imply that the expected value of the potential function Φ at the end of step 3 is only an $O((1/\alpha) \log n)$ factor larger than the cost of **OPT**, which implies the cost of the equilibrium at the end of step 4 is at most that large, as desired. ∎

A *cost-sharing game with delays* [5] is an extension of the fair cost-sharing game where each edge has both a cost function $c_e(x)$ and a latency function $d_e(x)$, where $c_e(x)$ is the cost of building the edge e for x users which the users will share, while $d_e(x)$ is the delay suffered by each user on edge e if x users are sharing the edge. The goal of each user will be to minimize the sum of his cost and his latency. If we assume that both the cost and the latency for each edge depend only on the number of players using that edge, then the total cost felt by each user on the edge is $f_e(x) = c_e(x)/x + d_e(x)$. These games remain potential games [43, 5] (they are particular cases of congestion games). One can prove a lemma similar to Lemma 1 relating the cost and the value of the potential function for any given joint action. In particular, for linear delays we have the following.

LEMMA 6. *Consider the cost-sharing game with delays where the cost function on edge e is $f_e(x) = c_e/x + l_e \cdot x$. For any joint action $s \in \mathcal{S}$ we have*

$$\frac{1}{2}\texttt{cost}(s) \le \Phi(s) \le H(n) \cdot \texttt{cost}(s).$$

Proof. In this game we have $\texttt{cost}(s) = \sum_{e:x_e \ge 1}(c_e + l_e x_e^2)$ and $\Phi(s) = \sum_{e:x_e \ge 1} \sum_{x=1}^{x_e}(\frac{c_e}{x} + l_e \cdot x) = \sum_{e:x_e \ge 1}[c_e(1 + 1/2 + \cdots + 1/x_e) + l_e x_e(x_e + 1)/2]$. One can see that the desired inequalities hold separately for each edge, and therefore they hold also for the sum. ∎

We now show how to extend our results to deal with linear delays. The extension is not immediate though since the part of our argument in Theorem 3 that states that after step 2, every nonreceptive player i has a reasonably cheap option to try (namely, its path in **OPT**), is not clear anymore: since the original behavior s^{ini} was *arbitrary* there could exist edges with a much higher number of players on them under the joint action $(s_{\bar{R}}^{ini}, s_R^{ad})$ than in **OPT**. In order to prove the desired result we instead argue the existence of a related "shadow" game, whose price of anarchy is not too large, and then relate performance of the behaviors as well as the optimum values between the two games.

THEOREM 7. *For the cost-sharing game with delays where the cost function on edge e is $f_e(x) = c_e/x + l_e \cdot x$ there exists a strategy for the advertising model which has an expected cost of the final equilibrium at most $O((1/\alpha)(\log n) \cdot \texttt{cost}(\mathbf{OPT}))$.*

Proof. Fix some optimal solution **OPT** and let $s^{ad} = \texttt{advertise}(\mathcal{G}) = \mathbf{OPT}$. Namely, the advertising strategy will be to tell each player i to use his path P_i^{OPT} in **OPT**. Let R be the set of receptive players. Let $n_{e,R}^{\text{opt}}$ denote the number of people in R who use edge e and let n_e^{opt} denote the number of people in **OPT** who use e.

By assumption, in step 3 all the users not in R settle on some equilibrium (given s_R^{ad}). Let n_e denote the number of users who are now on edge e. So, $n_e \geq n_{e,R}^{\text{opt}}$. We now define a new game $\mathcal{G}'$ with respect to the users in $\bar{R} = N \setminus R$, which is a congestion game with a linear latency function $h_e(n) = a_e + l_e \cdot n$, where $a_e = c_e/(1 + n_e)$. Let $\mathbf{OPT}'$ denote the optimal cost for this game $\mathcal{G}'$.

We first claim that the behavior at the end of step 3 is also an equilibrium for users in $\bar{R}$ if we use the cost h_e instead of f_e on all edges e. In particular, suppose this was not the case. So, some user i currently using a set S of edges would prefer switching to some set $S - A + B$, i.e.,

$$\sum_{e \in A} \left(\frac{c_e}{n_e + 1} + l_e n_e \right) > \sum_{e \in B} \left(\frac{c_e}{n_e + 1} + l_e(n_e + 1) \right). \tag{3.2}$$

However, if we replace $\frac{c_e}{n_e+1}$ with $\frac{c_e}{n_e}$ on the left-hand side of (3.2), then the gap only gets larger. This means that i is not at equilibrium under f since it can benefit from switching.

Now we use the fact that the new game $\mathcal{G}'$ has a price of anarchy of 5/2 [18]. So, the total cost in $\mathcal{G}'$ of the behavior of the nonreceptive players at the end of step 3 is $O(\mathbf{OPT}')$. Note now that the following hold:

(a) The cost of the nonreceptive players at the end of step 3 using cost functions f_e is at most twice their cost using functions h_e (since adding 1 to the denominator in moving from f to h at most reduces the cost by a factor of 2).

(b) $\mathbf{E}(\mathbf{OPT}') = O((1/\alpha)\,\texttt{cost}(\mathbf{OPT}))$. This is because one option for $\mathbf{OPT}'$ is to use the same paths as in **OPT**, in which case,
   (i) the $l_e \cdot n$ terms are the same as in **OPT**, and
   (ii) since $n_e \geq n_{e,R}^{\text{opt}}$ and as we argued in Theorem 3 we have $\mathbf{E}[c_e/(1 + n_{e,R}^{\text{opt}})] = O(c_e/(\alpha \cdot n_e^{\text{opt}}))$, this means the expected sum of the a_e terms in $\mathbf{OPT}'$ is $O(\sum_{e:n_e^{\text{opt}}>0} n_{e,\bar{R}}^{\text{opt}} \cdot c_e/(\alpha \cdot n_e^{\text{opt}}))$.

These imply that the expected cost under f of the nonreceptive players $\bar{R}$ at the end of step 3 satisfies

$$\mathbf{E}[\texttt{cost}(s_{\bar{R}}^{med})] = O((1/\alpha)\,\texttt{cost}(\mathbf{OPT})).$$

We now argue that the expected cost in the original game for the receptive players at the end of step 3 is also $O((1/\alpha)\,\texttt{cost}(\mathbf{OPT}))$. In particular, the key issue is the latency term, since there could potentially be more nonreceptive players on any given edge than in **OPT**. However, if on a given edge there are more receptive players than nonreceptive players in s^{med}, then we lose at most a factor of two compared to the latency cost in **OPT**; on the other hand, if there are more nonreceptive players than receptive ones, then we can charge the cost of the receptive players to the cost of the nonreceptive players, which we already bounded. This implies that the expected cost at the end of step 3 satisfies

$$\mathbf{E}[\texttt{cost}(s^{med})] = O((1/\alpha)\ \texttt{cost}(\mathbf{OPT})). \tag{3.3}$$

This together with Lemma 6 implies that the expected value of the potential function Φ at the end of step 3 is at most an $O((1/\alpha)\log n)$ factor larger than the cost of **OPT**, which implies the cost of any final equilibrium at the end of step 4 is at most that large, as desired. □

Remarks. Note that in all the variants of the cost-sharing game studied in this section, the cost of the final equilibrium reached is $O((1/\alpha)\log n)$ from the optimal cost, while the price of stability is $\Theta(\log n)$ as shown in [5]. This implies that the difference in guarantee is only a factor $O(1/\alpha)$. Second, as with Theorem 4, in all our proofs we can replace **OPT** with any other solution F as s^{ad} and get a pure Nash equilibrium whose cost is at most $O((1/\alpha)\log n)$ from $\texttt{cost}(\text{F})$. For example, if we let F be the best Nash equilibrium, then since the price of stability is $O(\log n)$, the expected cost of the final equilibrium is within $O((1/\alpha)\log^2 n)$ of the optimal cost.

3.2. Consensus games, cut games, and party affiliation games. In this section we consider three related classes of games, played by users who are viewed as vertices in an undirected simple graph $G = (N, E)$ with n vertices, where $N = \{1, \ldots, n\}$. We will first describe the most general case of *party affiliation* games and then discuss the special cases of *consensus games* and *cut games*.

In *party affiliation games* [19] the set of edges E is partitioned into positive and negative edges, denoted by PE and NE, respectively. Each player i has two actions r (red) or b (blue), i.e., $\mathcal{S}_i = \{r, b\}$. A player has cost 1 for each incident positive edge on which he disagrees with his neighbor and cost 1 for each negative edge on which he agrees with his neighbor, i.e., $\texttt{cost}_i(s) = \sum_{(i,j)\in PE} \mathbf{I}_{(\mathbf{s_i}\neq\mathbf{s_j})} + \sum_{(i,j)\in NE} \mathbf{I}_{(\mathbf{s_i}=\mathbf{s_j})}$. The overall social cost is the sum of the costs of all the users plus 1, i.e., $\texttt{cost}(s) = 1 + \sum_{i\in N} \texttt{cost}_i(s)$.[6] It is straightforward to show that $\Phi(s) = (\texttt{cost}(s) - 1)/2$ is an exact potential function [43] for the party affiliation game. Also, in any party affiliation game the social optimum is a Nash equilibrium, and thus the PoS is 1.

Consensus games are a special case of party affiliation games where all the edges in G are positive edges, i.e., $NE = \emptyset$. The two social optimal solutions in a consensus game are "all blue" and "all red," both of which are also Nash equilibria. On the other hand, consensus games may also have high-cost equilibria. For n even, let G be the clique K_n with a perfect matching $\{(2i-1, 2i)\}_{i=1}^{n/2}$ removed. Consider the joint action s in which even players play r and odd players play b, i.e., $s_{2i} = r$ and $s_{2i-1} = b$. This is a Nash equilibrium, since each player has exactly half its neighbors the same color and exactly half the opposite color. This results in a social cost of $\Omega(n^2)$ and thus the PoA for consensus games is $\Omega(n^2)$.[7]

Cut games (see [19]) are a special case of party affiliation games where all the edges in G are negative edges, i.e., $PE = \emptyset$. Thus cut games have the opposite objective from consensus games. As mentioned above, in a cut game the optimal solution is a Nash equilibrium and so the PoS is 1. However, the problem of *finding* an (approximately) optimal solution is the Min-UnCut problem [2] for which the best

[6] The "+1" is just to ensure the cost is nonzero so that all ratios are well-defined.

[7] Intuitively one can think of this example as two countries, one using English units and one using the metric system, neither wanting to switch.

efficient approximation algorithm known has approximation ratio $O(\sqrt{\log n})$. As with consensus games, the PoA for cut games is $\Omega(n^2)$. For instance, if G is the complete bipartite graph $K_{n/2,n/2}$, then coloring half the nodes on the left and half the nodes on the right blue, and coloring half the nodes on the left and half the nodes on the right red, is a Nash equilibrium with cost $\Omega(n^2)$; yet the optimal solution has cost 1 since the graph is bipartite.

We first show that if all nodes have degree $\omega(\log n)$, then in the advertising model all these games have a sharp threshold at $\alpha = 1/2$: any constant $\alpha > 1/2$ is sufficient to produce an optimal or a near-optimal solution, and yet there exist families of graphs for which any constant $\alpha < 1/2$ yields a solution of cost nearly as bad as possible (a factor $\Omega(n^2)$ worse than optimal).

THEOREM 8. *For consensus games over graphs in which each node has degree $\omega(\log n)$, there is a sharp threshold at $\alpha = 1/2$ in the advertising model. Specifically, for any $\alpha > 1/2$, for any graph G of minimum degree $\geq \ln(n)/(\alpha - 1/2)^2$, there exists an advertising strategy such that with high probability the final equilibrium is the optimal solution (ratio of 1); on the other hand, for any constant $\alpha < 1/2$ there exist graphs such that for any advertising strategy, with high probability the final equilibrium will be a factor $\Omega(n^2)$ worse than optimal.*

Proof. For the upper bound, the advertising strategy is simply to tell all nodes to become red, i.e., $s^{ad} = (r, \dots, r)$. Hoeffding bounds (e.g., see Appendix A) imply that a node with degree $d \geq \ln(n)/(\alpha - 1/2)^2$ has more than half its neighbors in the receptive set R with probability at least $1 - e^{-2(\alpha-1/2)^2 d} \geq 1 - 1/n^2$. Since we have assumed *all* nodes have degree at least this large, by the union bound, all nodes have more than half their neighbors in R with probability at least $1 - 1/n$. So, with high probability, at the end of step 3 all nodes are red, i.e., $s^{med} = (r, \dots, r)$, which is optimal.

For the lower bound, let $\gamma = 1/2 - \alpha$ and consider a graph consisting of two cliques of size $n/2$, where each vertex has $\gamma n/8$ neighbors in the other clique. Suppose initially we have one clique red and the other clique blue. Since γ is a positive constant, for sufficiently large n we have that with high probability each node has at most a $1/2 - \gamma/2$ fraction of its neighbors in R. However, since each node initially has only a $\gamma/4$ fraction of its neighbors of the other color, this will not be sufficient to cause any of the nodes not in R to change color in step 3. Therefore, in step 4, all nodes in R will simply revert to their original color and we again have $(\gamma n/8)(n/2) = \gamma n^2/16 = \Omega(n^2)$ badly colored edges. ∎

The key to the upper bound above is that with high probability the set R satisfies the property that every vertex not in R has more than half its neighbors in R. For cut games and more generally party affiliation games, we will need a bit more (in particular because **OPT** no longer necessarily has zero cost for every player). Specifically let us say that a set S is a β-dominating set if every vertex not in S has more than a $1/2 + \beta$ fraction of its neighbors in S. We can then say the following.

LEMMA 9. *For any party affiliation game in which each node has degree $\omega(\log n)$, for any constant $\alpha > 1/2 + 2\beta$, with probability $1 - o(1)$ the set of receptive players is a β-dominating set in the advertising model. Specifically, for any given α this holds for any graph with minimum degree $\geq \ln(n)/(\alpha - 1/2 - \beta)^2$.*

Proof. By Hoeffding bounds, the probability a given node of degree d does not have more than a $1/2 + \beta$ fraction of its neighbors in R is at most $e^{-2(\alpha-1/2-\beta)^2 d} \leq 1/n^2$ for $d \geq \ln(n)/(\alpha - 1/2 - \beta)^2$. Thus, by the union bound, with high probability R is a β-dominating set. ∎

We now show the following property of β-dominating sets in party affiliation games.

LEMMA 10. *For party affiliation games, if the set R following the advertising strategy is a β-dominating set, then using $s^{ad} = \mathbf{OPT}$ will produce a solution within an $O(1/\beta)$ factor of optimal.*

Proof. We argue by considering two kinds of nodes: those with less than a β fraction of their incident edges incurring a cost (of one) in **OPT** (call those "low-cost" nodes), and those with more than a β fraction of their incident edges incurring a cost in **OPT** (call those "high-cost" nodes). The advertising strategy is to tell nodes to behave according to **OPT**, i.e., $s^{ad} = \mathbf{OPT}$. Since R is a β-dominating set, each low-cost node will change in step 3 to its color in **OPT** (because that color minimizes its cost with a majority of its neighbors). For the high-cost nodes, we may not produce the desired behavior; however, no matter how the high-cost nodes behave, they cannot incur a cost that is more than a $1/\beta$ factor worse than their cost in **OPT** (by definition of "high cost"). Therefore, the total cost by the end of step 3 is at most $(1 + 1/\beta)\mathbf{OPT}$. Finally, the cost can only improve via the best-response process in step 4. ∎

For constant $\alpha > 1/2$, by setting $\beta = (\alpha - 1/2)/2$, Lemmas 9 and 10 imply that the cost of the final equilibrium is within an $O(1)$ factor of optimal. For $\alpha < 1/2$, the $\Omega(n^2)$ bound from Theorem 8 still applies, thus exhibiting a sharp threshold at $\alpha = 1/2$. We summarize these results in Theorem 11 below.

THEOREM 11. *For party affiliation games in which each node has degree $\omega(\log n)$, for any constant $\alpha > 1/2$ there exists a strategy for the advertising model such that with high probability the final equilibrium has cost $O(\mathbf{OPT})$. Moreover, for any constant $\alpha < 1/2$ there exist party affiliation games such that for any advertising strategy with high probability the final equilibrium will be a factor $\Omega(n^2)$ worse than optimal.*

The above analysis shows that for $\alpha > 1/2$ and minimum degree $\omega(\log n)$, advertising the optimal strategy will produce a final state of cost $O(\mathbf{OPT})$ with high probability. However, in some cases the optimal strategy may be computationally hard to find, so one may instead want to use an approximation algorithm. For example, for cut games, there is an $O(\sqrt{\log n})$-approximation for the associated Min-UnCut problem [2]. In this case, the only change to the analysis needed is in Lemma 10. In particular, given some solution **ALG**, one simply defines high-cost and low-cost nodes as those with more or less than a β fraction of the incident edges incurring cost in **ALG**. The remainder of the argument proceeds as before. We thus have the following.

THEOREM 12. *For party affiliation games in which each node has degree $\omega(\log n)$, for any constant $\alpha > 1/2$, advertising a strategy* **ALG** *will with high probability produce a final equilibrium of cost $O(\texttt{cost}(\mathbf{ALG}))$.*

3.3. Extensions and further considerations.

3.3.1. Extensions to the model. In this section we show that many of our results are robust to a natural variation on the basic advertising model in which we replace step 3 (nonreceptive players settling to a Nash equilibrium) with a small number of rounds of best-response dynamics. This then leads us to consider in section 4 a more adaptive model in which we completely remove the distinction between receptive and nonreceptive players and instead view each player as a self-interested learning algorithm, deciding in each step between the two options of following the advertised strategy or following best-response dynamics based on some sort of learning rule. We then give positive results for both a "batch" and an adaptive version of such learning behavior.

Formally, we consider here the following *random best-response* variation on the basic advertising model of Figure 2. In this version, rather than nonreceptive players moving to a Nash equilibrium in step 3, they instead undergo random best-response dynamics for a polynomial number of time steps. That is, at each time step a random player in $N - R$ is selected to move, and this player then moves to its best response given the current state of all the players. Finally, in step 4, all players undergo best-response dynamics as before. We say that an advertising strategy is successful if a polynomial number of rounds of this modified step 3 is sufficient to produce a solution of low expected cost. We call this the *random best-response advertising model*.

The simplest game to consider in this model is the case of consensus games. Here, the proof of Theorem 8 shows that for $\alpha > 1/2$, with high probability all nonreceptive nodes in fact have as their only best response to play the advertised strategy, regardless of the behavior of the other nonreceptive players. Thus, with high probability, after each nonreceptive player has moved at least once, all will be following the (optimal) advertised strategy, and with high probability this requires only $O(n \log n)$ random best-response moves. More generally, for party affiliation games, Lemma 10 (if the receptive players form a β-dominating set, then the solution will have low social cost) applies with the modified step 3 as well. Therefore, the bounds proved in section 3.2 apply.

For fair cost-sharing, the situation is more complicated and it may not be sufficient for each nonreceptive player to simply move once in order to reach a low-cost state. Nonetheless, a polynomial number of rounds is sufficient, as shown below.

THEOREM 13. *For the fair cost-sharing game in the random best-response advertising model, a polynomial number of best-response moves is sufficient for the expected cost of the final equilibrium to be $O((1/\alpha)(\log n) \cdot \texttt{cost}(\mathbf{OPT}))$.*

Proof. Following the notation of Theorem 3, let $X_e = c_e/(1 + n_{e,R}^{opt})$ denote the guaranteed maximum cost of edge e to a nonreceptive player. Also, for player $i \notin R$, let $Y_i = \sum_{e \in P_i^{OPT}} X_e$ be the random variable denoting the cost of player i following the advertised strategy, ignoring any savings possible due to the behavior of other nonreceptive players. Let $Y = \sum_{i \notin R} Y_i$. If the current cost of the nonreceptive players is C, then since each such player i has the option of playing the advertised strategy at cost at most Y_i, the expected drop in potential Φ due to a random best-response move is at least $(C - Y)/n$. Note that if $C > 2Y$, then this expected drop is at least $C/(2n)$. Since the total contribution of nonreceptive players to the potential is at most $O(C \log n)$, this means that by Chernoff bounds, with high probability the cost C will drop to at most $2Y$ within $O(n \log n)$ steps. Also, as shown in the proof of Theorem 3, $\mathbf{E}[Y] = O((1/\alpha)\mathbf{OPT})$ as is the expected cost of the receptive players. Finally, since the potential is at most a logarithmic factor larger than the total cost and the potential never increases (neither in step 3 nor in step 4), the expected cost of the final equilibrium is $O((1/\alpha)(\log n) \cdot \texttt{cost}(\mathbf{OPT}))$ as desired. ∎

3.3.2. Stubborn games. Given our results, one might wonder if all potential games have some constant $\alpha < 1$ under which the system will in expectation reach a near-optimal state. In Appendix D we show that unfortunately there exist certain "stubborn" games where this is not the case, even in the basic advertising model.

4. A learning-based model. We now consider a more adaptive learning-style model where rather than having receptive and nonreceptive players, each player instead views best-response or following the advertised strategy as two possible options of behavior (or "experts") and aims to learn which of these is most suitable for himself. We specifically consider two instantiations of this kind of learning: a *learn-then-decide* model, where players initially follow an "exploration" phase where they put roughly equal probability on each expert, followed by a "commitment" phase where based on their experience they choose one to use from then on, and a *smoothly adaptive* model, where they slowly change their probabilities over time. In both models, the next player to move is always chosen uniformly at random: the difference is how the players choose to act when their turn arrives. We show that for both fair cost-sharing and consensus games, these processes lead to good quality behavior.

Conceptually, the aim is to understand how information about socially beneficial behaviors—which may be readily available or otherwise injected into the system by a central planner—can be used by self-interested players in order to reach states of cost comparable to the best equilibrium. We point out that the dynamics we study are not complicated and make sense from the point of view of the individual.

4.1. The model. We now describe the formal model that we consider. As before, players begin in some arbitrary state, which could be a high-cost equilibrium or even a state that is not an equilibrium at all. Next, the advertiser proposes global behavior s^{ad}. Now, players move one at a time in a random order. Each player, when it is his turn to move, chooses among two options. The first is to make a best-response move to the current configuration. The second option is to follow his part of s^{ad}. These two high-level strategies (best-response or follow s^{ad}) are viewed as two "experts" and the player then runs some learning algorithm aiming to learn which of these is most suitable for himself. Note that best response is an *abstract* option—the specific action it corresponds to may change over time.

Because moves occur in an asynchronous manner, there are multiple reasonable ways to model the feedback each player gives to its learning algorithm: for instance, does it consider the average cost since the player's previous move or just the cost when it is the player's turn to go, or something in between? In addition, does it get to observe the cost of the action it did not take (the full information model) or only the action it chose (the bandit model)? To abstract away these issues, and even allow different players to address them differently, we consider here two models that make only very mild assumptions on the kind of learning and adaptation made by players.

Learn-then-decide model. In this model, players follow an "exploration" phase where each time it is their turn to move, they flip a coin to decide whether to follow the proposed behavior s^{ad} or to do a best-response move to the current configuration. We assume that the coin gives probability at least α to s^{ad} for some constant $\alpha > 0$. Finally, after some common time T^*, all players switch to an "exploitation" phase where they each commit in an arbitrary way based on their past experience to follow s^{ad} or perform best response from then on. (The time T^* is selected in advance.)

The above model assumes some degree of coordination: a fixed time T^* after which all players make their decisions. One could imagine instead each player i having its own time T_i^* at which it commits to one of the experts, perhaps with the time itself depending on the player's experience. In the smoothly adaptive model below we even more generally allow players to smoothly adjust their probabilities over time as they like, subject only to a constraint on the amount by which probabilities may change between time steps.

Smoothly adaptive model. In this model, there are no separate exploration and exploitation phases. Instead, each player i maintains and adjusts a value p_i over time. When player i is chosen to move, it flips a coin of bias p_i to select between the proposed behavior or a best-response move, choosing s^{ad} with probability p_i. We allow the players to use arbitrary adaptive learning algorithms to adjust these probabilities with the sole requirement that learning proceed slowly. Specifically, using p_i^t to denote the value of p_i at time t, we require that

$$|p_i^t - p_i^{t+1}| \leq \Delta$$

for a sufficiently (polynomially) small quantity Δ and furthermore that for all i, the initial probability $p_i^0 \geq p^0$ for some overall constant $0 < p^0 < 1$. Note that the algorithm may update p_i even in time steps at which it does not move. The learning algorithm may use any kind of feedback or weight-updating strategy it wants to (e.g., gradient descent [53, 54], multiplicative updating [40, 41, 16]) subject to this bounded step-size requirement.

We say that the probabilities are (T, α)-good if for any time $t \leq T$ we have for all i, $p_i^t > \alpha$. (Note that if $\Delta < (p^0 - \alpha)/T$, then clearly the probabilities are (T, α)-good.)

We point out that while one might at first think that any natural adaptive algorithm would learn to favor best response (always decreasing p_i), this depends on the kind of feedback it uses. For instance, if the algorithm considers only its cost immediately after it moves, then indeed by definition best response will appear better. However, if it considers its cost immediately *before* it moves (comparing that to what its cost would have been had it chosen the other alternative) or even the sum total cost since its previous move, then s^{ad} might appear better. Our model allows users to update in any way they wish, so long as the updates are sufficiently gradual.

Finally, in both our models it is an easy observation that for any game, if the proposed solution is a good equilibrium, then in the limit (as $T^* \to \infty$ in the learn-then-decide model or as $\Delta \to 0$ in the smoothly adaptive model) the system will *eventually* reach the equilibrium and stay there indefinitely. Our interest, however, is in polynomial-time behavior.

4.2. Main results. For these models, we have the following results. For fair cost-sharing, we show that in the learn-then-decide model, so long as the exploration phase has sufficient (polynomial) length and the proposed strategy is near-optimal, the expected cost of the system reached is $O(\log(n)\log(nm)\mathbf{OPT})$. Thus, this is only slightly larger than that given by the price of stability of the game. If there are many players of each type (i.e., associated to each (s_i, t_i) pair), then we show that for both learn-then-decide and smoothly adaptive models, we do even better: with high probability the system achieves cost $O(\log(nm)\mathbf{OPT})$, or even $O(\mathbf{OPT})$ if the number of players of each type is high enough. Note that with many players of each type the price of anarchy remains $\Omega(n)$, though as implied by Theorem 17, the price of stability becomes $O(1)$. For party affiliation games, we show that all the positive results of section 3.2 carry over to the learn-then-decide model as well. In addition, we show that for consensus games—as well as party affiliation games with no high-cost nodes—we can do even *better* than in the advertisement model. In particular, for both learn-then-decide and smoothly adaptive models, the restriction on minimum node degree can be removed: for *any* graph G, so long as $\alpha > 1/2$, with high probability the system will reach the optimum solution in a polynomial number of steps. This holds even in graphs where the advertisement model can be shown to fail.

4.3. Fair cost-sharing. For ease of notation, we assume in this section that the proposed strategy s^{ad} is the socially optimal behavior $\mathbf{OPT}$, so we can identify $P_i^{OPT} = P_i^{s^{ad}}$ as the behavior proposed by s^{ad} to player i. If s^{ad} is different from $\mathbf{OPT}$, then we simply lose the corresponding approximation factor.

The high level idea of the analysis is that we first prove that so long as each player randomizes with probability near to 50/50, with high probability the overall cost of the system will drop to within a logarithmic factor of **OPT** in a polynomial number of steps; moreover, at that point both the best response and the proposed actions are pretty good strategies from the individual players' point of view. To finish the analysis in the learn-then-decide model, we show that in the remaining steps of the exploration phase the *expected* cost does not increase by much; using properties of the potential function, we then show that in the final "decision" round T^*, the overall potential cannot increase substantially either, which in turn implies a bound on the increase in overall cost.

For the adaptive model, one key difficulty in the analysis is to show that if the system reaches a state where the social cost is low and both abstract actions are pretty good for most players, the cost never goes high again. We are able to show that this indeed is the case as long as there are many players of each type, no matter how the players adjust their probabilities p_i^t or make their choice between best response and following the proposed behavior.

We begin with the following key lemma, which is useful in the analysis of both learn-then-decide and smoothly adaptive models.

LEMMA 14. *Consider a fair cost-sharing game. If the probabilities are* $(2n^3, \alpha)$*-good for constant* $\alpha > 0$*, then with high probability the cost at some time* $T \leq 2n^3$ *will be at most* $O(\mathbf{OPT}\log(mn))$*. Moreover, if we have at least* $c\log(nm)$ *players of each* (s_i, t_i) *pair for sufficiently large constant* c*, then with high probability the cost at some time* $T \leq 2n^3$ *will be* $O(\mathbf{OPT})$*.*

Proof. We begin with the general case. Let n_e^{opt} denote the number of players who use edge e in **OPT**. We partition edges into two classes. We say an edge is a "high-traffic" edge if $n_e^{opt} > c\log(nm)$, where c is a sufficiently large constant ($c = 32/\alpha$ suffices for the argument below). We say it is a "low-traffic" edge otherwise.

Define $T_0 = 2n\log n$. With high probability, by time T_0 each player has had a chance to move at least once. We assume in the following that this indeed is the case. Note that as a crude bound, at this point the cost of the system is at most $n^2 \cdot \mathbf{OPT}$. (Each player will move to a path of cost-share at most **OPT** and therefore of actual cost at most $n \cdot \mathbf{OPT}$.) Next, by Chernoff bounds (see Appendix A), for any given high-traffic edge e and any given time step T such that probabilities are (T, α)-good, the probability that e has at least $\alpha n_e^{opt}/2$ players on it is at least $1 - e^{-\alpha n_e^{opt}/8} \geq 1 - 1/(mn)^4$. Therefore, by the union bound, *every* high-traffic edge has at least $\alpha n_e^{opt}/2$ players on it at all time steps $T \in [T_0, T_0 + n^3]$ with probability at least $1 - mn^3/(mn)^4 \geq 1 - 1/n$. In the remaining analysis, we assume this indeed is the case as well.

Let $\mathbf{OPT}_i$ denote the cost of player i in **OPT**, so that $\mathbf{OPT} = \sum_i \mathbf{OPT}_i$. Our assumption above implies that for any time step T under consideration, if player i follows the proposed strategy P_i^{OPT}, its cost will be at most $c\log(nm)\mathbf{OPT}_i$. In particular, its cost on the low-traffic edges in P_i^{OPT} can be at most a factor $c\log(nm)$ larger than its cost on those edges under **OPT**, and its cost on high-traffic edges is at most a $2/\alpha$ factor larger than its cost on those edges under **OPT**.

We now argue as follows. Let $\texttt{cost}_T$ denote the cost of the system at time T. If

$$\texttt{cost}_T \geq 2c\log(nm)\mathbf{OPT},$$

then the expected cost of a random player is at least $2c\log(nm)\mathbf{OPT}/n$. On the other hand, if player i is chosen to move at time T, from the above analysis its cost after the move (whether it chooses s^{ad} or best response) will be at most $c\log(nm)\mathbf{OPT}_i$.

The expected value of this quantity over players i chosen at random is at most $c\log(nm)\mathbf{OPT}/n$. Therefore, if $\texttt{cost}_T \geq 2c\log(nm)\mathbf{OPT}$, the expected drop in potential at time T is at least $c\log(nm)\mathbf{OPT}/n$.

Note additionally (this fact will be needed later) that this guarantee on the expected drop holds even if we cap all decreases at $2c\log(nm)\mathbf{OPT}$, that is, if we consider $\mathbf{E}[\min(\Phi_{T-1} - \Phi_T, 2c\log(nm)\mathbf{OPT})]$. This is because the expected drop is now at least (using $\texttt{cost}_{iT}$ to denote the cost of player i at time T):

$$\begin{aligned}
&\frac{1}{n}\sum_i[\min(\texttt{cost}_{iT}, 2c\log(nm)\mathbf{OPT}) - c\log(nm)\mathbf{OPT}_i] \\
&\geq \frac{1}{n}(2c\log(nm)\mathbf{OPT} - c\log(nm)\mathbf{OPT}) \\
&= c\log(nm)\mathbf{OPT}/n.
\end{aligned}$$

Finally, since the cost at time T_0 was at most $n^2 \cdot \mathbf{OPT}$, which implies by Lemma 1 the value of the potential was at most $n^2(1 + \log(n))\mathbf{OPT}$, with high probability this cannot continue for more than $O(n^3)$ steps. Formally, we can apply Hoeffding–Azuma bounds for supermartingales (see Appendix A) as follows: let us define $Q = c\log(nm)\mathbf{OPT}$ and

$$\Delta_T = \max(\Phi_T - \Phi_{T-1} + Q/n, -2Q)$$

and consider running this process stopping when $\texttt{cost}_T < 2Q$. Let

$$X_T = \Phi_0 + \Delta_1 + \cdots + \Delta_T.$$

Then throughout the process we have

$$\mathbf{E}[X_T|X_1, \ldots, X_{T-1}] \leq X_{T-1} \quad \text{and} \quad |X_T - X_{T-1}| \leq 2Q,$$

where the first inequality holds because (as shown above) our analysis showing an expected decrease in potential of at least Q/n is true even if we cap all decreases to a maximum of $2Q$ as in the definition of Δ_T. So, by Hoeffding–Azuma, after n^3 steps in the nonstopped process with high probability we would have

$$X_T - X_0 \leq \frac{1}{2}n^2 Q,$$

which is not possible since by definition of X_T we have

$$\Phi_T \leq \Phi_0 + (X_T - X_0) - TQ/n,$$

which would be negative. Therefore, with high probability stopping must occur before this time as desired.

Finally, if we have at least $c\log(mn)$ players of each type, then there are no low-traffic edges and so we do not need to lose the $c\log(mn)$ factor in the argument. ∎

The above lemma shows that with high probability, the social cost will reach a low value within a polynomial number of steps. However, it does not show that the cost *stays* low: in principle, the cost could "bounce" and be low only once every n^3 steps. We now present a second lemma showing that in the learn-then-decide model, if the cost is low at some time T_1, then the expected cost remains low up to time T^* so long as $T^* = poly(n)$.

LEMMA 15. *Consider fair cost-sharing in the learn-then-decide model. If the cost of the system at time T_1 is $O(\mathbf{OPT}\log(mn))$, and $T = T_1 + poly(n) < T^*$, then the expected value of the potential at time T is $O(\mathbf{OPT}\log(mn)\log(n))$.*

Proof. First, as argued in the proof of Lemma 14, with high probability for any player i and any time $t \in [T_1, T]$, the cost for player i to follow the proposed strategy at time t is at most $c\log(nm)\mathbf{OPT}_i$ for some constant c. Let us assume below that this is indeed the case.

Next, the above bound implies that if the cost at time $t \in [T_1, T]$ is $\texttt{cost}_t$, then the expected decrease in potential caused by a random player moving (whether following the proposed strategy or performing best response) is at least

$$(\texttt{cost}_t - c\log(nm)\mathbf{OPT})/n;$$

in particular, $\texttt{cost}_t/n$ is the expected cost of a random player before its move, and $c\log(nm)\mathbf{OPT})/n$ is an upper bound on the expected cost of a random player after its move. Note that this is an expectation over randomness in the choice of player at time t, conditioned on the value of $\texttt{cost}_t$. In particular, since this holds true for any value of $\texttt{cost}_t$, we can take expectation over the entire process from time T_1 up to t, and we have that if

$$\mathbf{E}[\texttt{cost}_t] \geq c\log(nm)\mathbf{OPT},$$

then

$$\mathbf{E}[\Phi_{t+1}] \leq \mathbf{E}[\Phi_t],$$

where Φ_t is the value of the potential at time step t.

Finally, since $\Phi_t \leq \log(n)\texttt{cost}_t$, this implies that if

$$\mathbf{E}[\Phi_t] \geq c\log(nm)\log(n)\mathbf{OPT},$$

then

$$\mathbf{E}[\Phi_{t+1}] \leq \mathbf{E}[\Phi_t].$$

Since for any value of $\texttt{cost}_t$ we always have

$$\mathbf{E}[\Phi_{t+1}] \leq \mathbf{E}[\Phi_t] + c\log(nm)\mathbf{OPT}/n,$$

this in turn implies by induction on t that

$$\mathbf{E}[\Phi_t] \leq c\log(nm)\log(n)\mathbf{OPT} + c\log(nm)\mathbf{OPT}/n$$

for all $t \in [T_1, T]$, as desired. ∎

Our main result in the learn-then-decide model is the following.

THEOREM 16. *For fair cost-sharing games in the learn-then-decide model, a polynomial number of exploration steps T^* is sufficient so that the expected cost at any time $T' \geq T^*$ is $O(\log(n)\log(nm)\mathbf{OPT})$.*

Proof. From Lemma 14, there exists $T = poly(n)$ such that with high probability the cost of the system will be at most $O(\mathbf{OPT}\log(mn))$ at some time $T_1 \in [T^* - T, T^*]$. From Lemma 15, this implies the expected value of the potential at time T^* right before the final exploitation phase is $O(\mathbf{OPT}\log(mn)\log(n))$. Finally, we consider the decisions made at time T^*. When a player chooses to make a best-response move, this can only decrease the potential. When a player chooses s^{ad}, this could *increase* the potential. However, for any edge e in the proposed solution s^{ad}, the total increase in the potential caused by edge e over *all* players who have e in their proposed solution is at most

$$c_e \cdot H(n_e^{opt}) = O(c_e \log n).$$

This is because whenever a new player makes a decision to commit to following the proposed strategy and using edge e, all previous players who made that commitment after time T and whose proposed strategy uses edge e are still there. Thus, the total increase in potential after time T^* is at most $O(\mathbf{OPT} \log n)$.

Since after all players have committed, potential can only decrease, this implies that the expected value of the potential at any time $T' \geq T^*$ is $O(\mathbf{OPT} \log(mn) \log(n))$. Therefore, the expected cost at time T' is at most this much as well, as desired. ∎

We now use Lemma 14 to analyze fair cost-sharing games in the smoothly adaptive model when the number of players n_i of each type (i.e., associated to each (s_i, t_i) pair) is large.

THEOREM 17. *Consider a fair cost-sharing game in the smoothly adaptive model satisfying $n_i = \Omega(m)$ for all i. There exists a $T_1 = poly(n)$ such that if the probabilities are (T_1, α)-good for constant $\alpha > 0$, then with high probability for* all $T \geq T_1$ *the cost at time T is $O(\log(nm)\mathbf{OPT})$. Moreover, there exists constant c such that if $n_i \geq \max\{m, c \log(mn)\}$, then with high probability, for all $T \geq T_1$ the cost at time T is $O(\mathbf{OPT})$.*

Proof. First, by Lemma 14 we have that with high probability at some time $T_0 \leq T_1$, the cost of the system reaches $O(\mathbf{OPT} \log(mn))$. The key to the argument now is to prove that once the cost becomes low, it will *never* become high again. To do this we use the fact that n_i is large for all i. Our argument, which follows the proof of Theorem 4.3 in [9], is as follows.

Let U be the set of all edges in use at time T_0 along with all edges used in s^{ad}. In general, we will insert an edge into U if it is ever used from then on in the process, and we never remove any edge from U, even if it later becomes unused. Let c^* be the total cost of all edges in U. So, c^* is an upper bound on the cost of the current state and the only way c^* can increase is by some player choosing a best-response path that includes some edge not in U. Now, notice that any time a best-response path for some (s_i, t_i) player uses such an edge, the total cost of all edges inserted into U due to that path is at most c^*/n_i, because the current player can always choose to take the path used by the least-cost player of his type and those n_i players are currently sharing a total cost of at most c^*. Thus, any time new edges are added into U, c^* increases by at most a multiplicative $(1 + 1/n_i)$ factor. We can insert edges into U at most m times, so the final cost is at most

$$(\mathtt{cost}(T_0) + \mathbf{OPT})(1 + 1/n_{i^*})^m,$$

where $n_{i^*} = \min_i n_i$. This implies that as long as $n_{i^*} = \Omega(m)$ we have $\mathtt{cost}(T) = O(\mathtt{cost}(T_0))$ for all $T \geq T_0$. Thus, overall the total cost remains $O(\mathbf{OPT} \log(mn))$.

If $n_i \geq \max\{m, c \log(mn)\}$, we simply use the improved guarantee provided by Lemma 14 to say that with high probability the cost of the system reaches $O(\mathbf{OPT})$ within $T_0 \leq T_1$ time steps and then apply the charging argument as above in order to get the desired result. ∎

Variations. One could imagine a variation on our framework where rather than proposing a fixed behavior s^{ad}, one can propose a more complex strategy such as "follow s^{ad} until time step T^* and then follow new strategy $s^{ad'}$" or "follow s^{ad} until time step T^* and then perform best response." In the latter case, the smoothly adaptive model with (T^*, α)-good probabilities becomes essentially a special case of the learn-then-decide model and all results above for the learn-then-decide model go through immediately. On the other hand, this type of strategy requires a form of global coordination that one would prefer to avoid.

4.4. Consensus games, cut games, and party affiliation games. We now consider consensus games, cut games, and, more generally, party affiliation games. We first show that the positive results of section 3.2 for the advertisement model carry over directly to the learn-then-decide model as well. Specifically, for any constant $\alpha > 1/2$, for $T^* = \mathbf{poly}(n)$ (in fact, $T^* = O(n \log n)$ is sufficient), if nodes have degree $\omega(\log n)$, then with high probability the cost of the state at all time steps $T \geq T^*$ will be $O(\mathbf{OPT})$. We then show that for consensus games—as well as party affiliation games with no high-cost nodes—the restriction on minimum degree can be removed. In particular, for both learn-then-decide and smoothly adaptive models, we show that with high probability the system will reach the optimum solution in a polynomial number of steps.

The latter result is particularly interesting because in the advertisement model, a lower bound on minimum degree is *necessary* for a positive result, even for the simplest case of consensus games. For instance, consider a graph G consisting of $n/4$ 4-cycles each in the initial equilibrium configuration $red, red, blue, blue$. In that case, in the advertisement model, with high probability a constant fraction of the 4-cycles contain no receptive players and will remain in a high-cost state forever. On the other hand, it is easy to see that this specific example will perform well in both our learn-then-decide and smoothly adaptive models. Our main result in this context (Theorem 20 below) is that in fact *all* graphs G perform well, though this requires care in the argument due to correlations that may arise.

We begin with analysis of the high-degree case, then give our analysis for general graphs. Throughout this section, unless otherwise specified, we assume that $s^{ad} = \mathbf{OPT}$.

4.4.1. The high-degree case.

THEOREM 18. *For party affiliation games in the learn-then-decide model with constant $\alpha > 1/2$, if the proposed behavior $s^{ad} = \mathbf{OPT}$ and each node has degree $\omega(\log n)$, then with high probability the state will have cost $O(\mathbf{OPT})$ for all $T \geq T^*$ for $T^* > cn \log n$ for sufficiently large constant c.*

Proof. Let T_1 be a random variable that denotes the first time by which all players have moved at least once. Let $T_2 > T_1$ be a random variable that denotes the first time by which all players have moved at least once since T_1. Let T^* be a bound such that with high probability $T^* > T_2$. Clearly, it is enough for $T^* > cn \log n$ for sufficiently large constant c, and we assume below that indeed this high-probability event occurs.

As in the proof of Lemma 10, call a node "low-cost" if at most a β fraction of its incident edges incur cost in $\mathbf{OPT}$, where $\beta = (\alpha - 1/2)/2$.

We begin by arguing that with high probability, all low-cost nodes will have the same color as in $\mathbf{OPT}$ at time T^*. In particular, let v be an arbitrary low-cost node and consider the *last* time T_v that v moves before T^*. By assumption, $T_v > T_1$, so all neighbors of v have moved before time T_v. Each neighbor of v, when it last moved, chose s^{ad} with probability α and chose to do best response with probability $1-\alpha$, independently of the choices of all other neighbors of v. (The actual realizations of best-response behaviors as colors may exhibit complex dependencies, but not the choice of s^{ad} versus best response.) Therefore, since v is a low-cost node of degree

$\omega(\log n)$, by Hoeffding bounds we have that with probability $1 - 1/n^{\omega(1)}$, more than an $\alpha - 2\beta = 1/2$ fraction of v's neighbors are both following **OPT** *and* are such that their incident edges incur zero cost in **OPT**. Therefore, with probability $1 - 1/n^{\omega(1)}$, node v will have **OPT** as its best response and so will switch to its color in **OPT** regardless of its coin flip. Thus, with probability $1 - 1/n^{\omega(1)}$, all low-cost nodes will have the same color as in **OPT** at time T^*. Note that this implies that with high probability, the overall *cost* at the start of the decide phase is $O(\mathbf{OPT})$.

We now consider the decide phase. Here, notice that even if an adversary controls the order in which players move as well as their decisions (**OPT** or best response), if node v was the same color as **OPT** at time T^*, then the cost of the system *cannot* increase in v's move. The reason is that if v chooses **OPT**, then the state stays the same, and if v chooses best response, then the potential can only decrease, which implies the cost can only decrease as well because cost and potential for party affiliation games are linearly related. Therefore, with high probability, cost can only increase with moves of high-cost nodes. This means the total increase in cost during the decide phase will be at most the sum of degrees of all high-cost nodes, which is $O(\mathbf{OPT})$. Adding this to the $O(\mathbf{OPT})$ cost at the start of the decide phase yields an $O(\mathbf{OPT})$ at the end of the phase. Finally, after the decide phase is over, the potential (and therefore the cost) can only decrease. ∎

As in the analysis of section 3.2, if $s^{ad} \neq \mathbf{OPT}$, then we simply define high-cost and low-cost nodes with respect to s^{ad} and then the same analysis as above implies we reach a state of cost $O(\texttt{cost}(s^{ad}))$.

4.4.2. General graphs: Consensus games. We now consider the case of general graphs with no lower bound on node degrees. We show that for consensus games, for any graph G, for constant $\alpha > 1/2$, the system will in polynomial time reach optimal behavior (if $s^{ad} = \mathbf{OPT}$) and then remain there forever.

Specifically, we assume that the proposed behavior s^{ad} is "all blue" and prove that with high probability the configuration will reach this state after $O(n \log^2 n)$ steps. We begin with a simple preliminary lemma.

LEMMA 19. *Let $X_1, \ldots, X_d$ be $\{0,1\}$-valued random variables with $\mathbf{Pr}(X_i = 1) \geq p$, and let p_{MAJ} denote the probability that a strict majority of the X_i are equal to* 1. *Then no matter how the X_i are correlated, $p_{MAJ} \geq 2p - 1$.*

Proof. The expected number of 1's is at least pd and at most $d \cdot p_{MAJ} + (d/2) \cdot (1 - p_{MAJ})$. Solving, we get $p_{MAJ} \geq 2p - 1$. This can also be seen as an immediate consequence of Markov's inequality applied to the random variable $\sum_{i=1}^{d} (1 - X_i)$. ∎

We now present our main result of this section.

THEOREM 20. *For consensus games in both learn-then-decide and smoothly adaptive models, for any constant $\alpha > 1/2$, if the proposed behavior s^{ad} is optimal, then with high probability play will become optimal in $O(n \log^2 n)$ steps. For the smoothly adaptive model we assume behavior is $(cn \log^2 n, \alpha)$-good for sufficiently large constant c.*

Proof. In both models the dynamics contain two random processes: the random order in which players move and the random coins flipped by the players to determine how they want to move. In order to prove the theorem it will be helpful to separate these two processes and consider them each in turn.

First, consider and fix a random sequence S of players to move. Let T_1 denote the time by which all players have moved at least once according to S, and more generally let T_{t+1} denote the time by which all players have moved at least once since time T_t.

Since players move in a random order, with high probability $T_{t+1} \leq T_t + 3n\log(n)$ for all $t = 1, \ldots, n$. We assume in the following that this indeed is the case; in fact, we can allow S to be an arbitrary, adversarially chosen sequence subject to this constraint.

Fixing S, we now consider the coin flips of the individual players. We will prove by induction that each player has probability at least q_t of being blue at time T_t (not necessarily independently) for

$$q_t = (1-\gamma)q_{t-1} + \gamma \quad \text{for} \quad \gamma = 2\alpha - 1.$$

Equivalently, we can write this as

$$1 - q_t = (1 - q_{t-1})(1-\gamma).$$

Since $\gamma > 0$ (because $\alpha > 1/2$), this in turn implies that $t = O(\log n)$ is sufficient to reach $1 - q_t \leq 1/n^2$, meaning that with high probability all nodes are blue as desired.

We prove this bound on q_t as follows. Consider the nodes that move at times $T_{t-1}+1, T_{t-1}+2, \ldots, T_t$ in order. When some node v moves, by induction each neighbor w of v has probability at least q_t of being blue (though these may not be independent). By assumption, v chooses s^{ad} with some probability $\alpha' \geq \alpha > 1/2$ and chooses best response with probability $1 - \alpha'$. Therefore, the probability v becomes blue is at least

$$\begin{aligned}
&\alpha' + (1-\alpha')\mathbf{Pr}(\text{majority of neighbors of } v \text{ are blue}) \\
&\quad \geq \alpha' + (1-\alpha')(2q_{t-1} - 1) \qquad \text{(by Lemma 19)} \\
&\quad = (1 - (2\alpha' - 1))q_{t-1} + (2\alpha' - 1) \\
&\quad \geq (1-\gamma)q_{t-1} + \gamma \qquad \text{(for } \gamma = 2\alpha - 1)
\end{aligned}$$

as desired. Thus, with high probability all nodes are blue by time T_t for $t = O(\log n)$, and by our assumption on S this occurs within $O(n \log^2 n)$ steps. □

The general result above requires $\alpha > 1/2$, and as noted earlier there exist graphs G and initial configurations such that the process will fail for any constant $\alpha < 1/2$. On the other hand, for several "nice" graphs such as the line or grid, *any* constant $\alpha > 0$ is sufficient. For this we assume that best response will only ask to switch color if the new color is *strictly* better than the current color.

THEOREM 21. *For the line and d-dimensional grid graphs (constant d), for any constant $\alpha > 0$, if the proposed action s^{ad} is optimal then in both learn-then-decide and smoothly adaptive models, with high probability, play will reach optimal in poly(n) steps. For the case of the smoothly adaptive model we assume behavior is (T, α)-good for T a sufficiently large polynomial in n.*

Proof. Assume s^{ad} is "all blue." On the line, if any two neighbors become blue, they will remain blue indefinitely. Similarly in the grid, if any d-dimensional cube becomes blue, the nodes in the cube will also remain blue indefinitely. On the line, any neighbors, in any two consecutive steps, have probability at least α^2/n^2 of becoming blue, and on the d-dimensional grid, any cube, in any 2^d consecutive steps, has probability at least $(\alpha/n)^{2^d}$ of becoming all blue. Therefore with high probability all nodes become blue in a polynomial number of steps. □

4.4.3. General graphs: Party affiliation with no high-cost nodes. The above results for the case of constant $\alpha > 1/2$ can be directly extended to the case of general party affiliation games in the case that all nodes have less than a $\beta = \alpha - 1/2$

fraction of their incident edges incurring cost in **OPT**. Obviously in consensus games, all nodes have this property. To achieve this extension, it is convenient to state a generalization of Lemma 19 as follows (which reduces to Lemma 19 when $\theta = 1/2$).

LEMMA 22. *Let $X_1, \dots, X_d$ be $\{0,1\}$-valued random variables such that $\mathbf{Pr}(X_i = 1) \geq p$. Let p_θ denote the probability that more than a θ fraction of the X_i are equal to 1. Then $p_\theta \geq (p-\theta)/(1-\theta)$.*

Proof. The expected number of 1's is at least dp and is at most $dp_\theta + \theta d(1-p_\theta)$. Solving, we get $dp_\theta(1-\theta) + \theta d \geq pd$ which implies $p_\theta \geq (p-\theta)/(1-\theta)$ as desired. □

We now show the following.

THEOREM 23. *For party affiliation games in both the learn-then-decide and smoothly adaptive models, if $s^{ad} = \mathbf{OPT}$, α is a constant greater than 1/2, and all nodes have less than a $\beta = \alpha - 1/2$ fraction of their incident edges incurring cost in **OPT**, then with high probability play will become optimal in $O(n \log^2 n)$ steps. For the smoothly adaptive model we assume behavior is $(cn \log^2 n, \alpha)$-good for sufficiently large constant c.*

Proof. Define sequence S and times T_t exactly as in the proof of Theorem 20. Following that proof, we define quantities $q_t \to 1$ and show by induction that each player has probability at least q_t of playing their strategy in **OPT** at time T_t (not necessarily independently). Specifically, we show this holds for q_t given by

$$1 - q_t \; = \; (1-q_{t-1})(3/2-\alpha), \tag{4.1}$$

where $q_1 = \alpha$. Since α is a constant greater than 1/2, this implies that $t = O(\log n)$ is sufficient to reach $1 - q_t \leq 1/n^2$, meaning that with high probability all nodes are following **OPT** as desired.

As in the proof of Theorem 20, we prove (4.1) by considering the nodes that move at times $T_{t-1}+1, T_{t-1}+2, \dots, T_t$ in order. When node v moves, it chooses s^{ad} with some probability $\alpha' \geq \alpha > 1/2$ and chooses best response with probability $1-\alpha'$. If it chooses best response, this is guaranteed to be equal to its action in **OPT** if more than a θ fraction *of its neighbors whose incident edges do not incur a cost in* **OPT** are also following **OPT** themselves, for $\theta(1-\beta) = 1/2$ or equivalently $\theta = \frac{1}{3-2\alpha}$. So, we have

$$q_t \geq \alpha + (1-\alpha)(q_{t-1} - \theta)/(1-\theta),$$

which implies

$$\begin{aligned} 1 - q_t &\leq 1 - \alpha - (1-\alpha)(q_{t-1}-\theta)/(1-\theta) \\ &= (1-\alpha)(1 - (q_{t-1}-\theta)/(1-\theta)) \\ &= (1-\alpha)(1-q_{t-1})/(1-\theta)) \\ &= (1-q_{t-1})(3/2-\alpha), \end{aligned}$$

yielding (4.1) as desired. □

In fact, the above analysis immediately applies to any advertised strategy s^{ad} even if not optimal *if* s^{ad} has no high-cost nodes. In particular, if all nodes have at most a β fraction of their edges incurring cost in s^{ad}, then the exact argument given above shows that with high probability, the behavior will approach s^{ad}.

Allowing high-cost nodes. If the graph has high-cost nodes (nodes with more than a $\beta = 1/2 - \alpha$ fraction of their incident edges incurring cost in **OPT**), then the above analysis breaks down. However, if the graph has the property that for each vertex v, there are at most $O(r^d)$ nodes and edges within radius r of v for some constant d (such as in a d-dimensional mesh—this d is sometimes called the pointwise dimension of the graph [20]), then one can show behavior will reach states of expected cost $O(\mathbf{OPT})$ in the learn-then-decide model.

THEOREM 24. *For party affiliation games in the learn-then-decide model, if $s^{ad} = \mathbf{OPT}$, α is a constant greater than $1/2$, and all balls of radius r contain at most $O(r^d)$ nodes and edges for some constant d, then $T^* = O(n \log^2 n)$ is sufficient so that the expected cost at any time $T' \geq T^*$ is $O(\mathbf{OPT})$.*

Proof. Let $r_H(v)$ denote the distance from v to its nearest high-cost node (so if v is high-cost itself, then $r_H(v) = 0$). Then, the analysis in the proof of Theorem 23 can be used to show that each node v will have probability at least $q_{r_H(v)+1}$ of playing its strategy in $\mathbf{OPT}$ at any time $T \geq T_{r_H(v)+1}$. Specifically, assume by induction (on t) that all nodes w have probability at least $q_{\min(r_H(w)+1,t-1)}$ of playing their strategy in $\mathbf{OPT}$ at time T_{t-1} (base case $t-1 = 1$ is immediate). Now, consider the nodes that move at time $T_{t-1}+1, T_{t-1}+2, \ldots, T_t$ in turn. When some node v moves, since all neighbors w of v satisfy $r_H(w) \geq r_H(v)-1$, by induction we have that w has probability at least $q_{\min(r_H(v),t-1)}$ of playing its strategy in $\mathbf{OPT}$. So, by the analysis in the proof of Theorem 23, v will have probability at least $q_{\min(r_H(v),t-1)+1} = q_{\min(r_H(v)+1,t)}$ of playing its strategy in $\mathbf{OPT}$ at time T_t, as desired. Now, by (4.1), this in turn implies the probability of a node v *not* playing its strategy in $\mathbf{OPT}$ at any given time $T \geq T_{r_H(v)+1}$ is $O((3/2-\alpha)^{r_H(v)})$. Let us now "charge" this probability to the nearest high-cost node to v (breaking ties arbitrarily). Since the graph has constant dimension d, we have that the total charge to each high-cost node (the total influence of each high-cost node on the overall expected cost) is at most $O(\sum_{r=1}^{\infty}(3/2-\alpha)^r r^d) = O(1)$. Thus, the expected cost at the start of the decide phase is $O(\mathbf{OPT})$, as is the expected number of nodes not following $\mathbf{OPT}$. This means that by the same reasoning as in the proof of Theorem 18, the expected cost after the decide phase will be $O(\mathbf{OPT})$ as well. ∎

5. Discussion and conclusions. In this paper we consider natural game dynamics that can incorporate global information about socially beneficial behaviors (which players might have already or which might be given via a public service advertising campaign). From the perspective of a central agency injecting such information into the system, the question is to what extent can this advice help to bring these dynamics to states comparable to the *best* equilibrium of the game? Unlike the notion of price of stability we do not assume everyone immediately takes the advice, and unlike the notion of value of altruism we assume everyone ultimately is acting in their own interest.

We begin with our simplest "basic advertising" model, showing that for some natural games (cost-sharing), it is enough for a small random fraction α of the players to follow the advice in order to reach a good equilibrium (and performance degrades gracefully with α), whereas others (party affiliation) have a threshold property. We next consider two more adaptive models, where rather than a fixed random subset of players following the given advice, the set of players following it changes over time as players individually test out the advice and use adaptive learning behavior. We show a number of positive results in this setting. Among them, we show for fair cost-sharing that so long as $n_i = \Omega(m)$ for all i, there exists $T_0 = \mathbf{poly}(n, m)$ such that with high probability, for all $T \geq T_0$, the cost at time T is $O(\log(nm)\mathtt{cost}(\mathbf{OPT}))$, even in our most adaptive model. We leave as an open problem extending this result for adaptive learning to the case of only a few players of each type. We also show for consensus and certain other party affiliation games, these adaptive models can produce good behavior even in cases where the basic advertising model is not sufficient.

Our analysis suggests a new angle on what it means to learn in a multiagent environment. Rather than learning over the immediate concrete action set $\mathcal{S}_i$ (as in traditional analysis of, e.g., regret-minimizing algorithms or noisy best-response dynamics [12, 37, 48, 14, 15, 42]), our framework instead involves agents learning in a *meta-game* over *abstract* actions. In our case these correspond to (1) playing best response or (2) following the known socially beneficial strategy. We then show that learning over these abstract actions in the meta-game results in good behavior in the original game. The consideration of abstract actions is especially natural in multiagent settings because the quality of the immediate concrete actions depends greatly on what other players are currently doing; so, learning over the concrete actions (as done, for instance, in standard regret-minimization analysis) can be ineffective both from the players' perspective and from the social welfare perspective when the game has equilibria of greatly different qualities.[8] By contrast, the use of abstract actions allows players instead to reason at a higher level and better models the kind of strategic thinking that players aware of the structure of the game could be expected to perform. As we show, in interesting games this leads to much more desirable social behavior. The specific meta-actions we consider can be conceptually viewed as "act selfishly" versus "act cooperatively." One can also envision multiple high-level strategies—e.g., "act defensively" as if some fraction of agents will be adversarial, "act optimistically" as if other agents will behave in a socially optimal way, or even multiple competing advertising campaigns or multiple alternative suggested behaviors. In this case, it would be very interesting to analyze learning dynamics that can run on top of this wider range of high-level strategies.

Appendix A. Useful lemmas.

LEMMA 2. *If X is a binomial random variable distributed $Bi(n,p)$, then* $\mathbf{E}_X[\frac{c}{X+1}] = O(\frac{c}{p\cdot n})$.

Proof. We have

$$\mathbf{E}_X\left[\frac{n+1}{X+1}\right] = \sum_{i=0}^{n} \frac{n+1}{i+1}\binom{n}{i}p^i(1-p)^{n-i} = \frac{1}{p}\sum_{i=0}^{n}\binom{n+1}{i+1}p^{i+1}(1-p)^{n-i}$$
$$= \frac{1}{p} - \frac{(1-p)^{n+1}}{p}.$$

This implies that

$$\mathbf{E}_X\left[\frac{c}{X+1}\right] = \frac{c}{p\cdot(n+1)} - \frac{c(1-p)^{n+1}}{p\cdot(n+1)} = \frac{c}{n+1}\left[\frac{1}{p} - \frac{(1-p)^{n+1}}{p}\right] = O\left(\frac{c}{p\cdot n}\right),$$

as desired. ∎

LEMMA 25 (Azuma–Hoeffding). *Let $Z_1, Z_2, \ldots, Z_n$ be a supermartingale sequence such that $|Z_k - Z_{k-1}| \le c_k$. Then*

$$\mathbf{Pr}[Z_n - Z_0 \ge t] \le e^{-t^2/(2\sum_k c_k^2)}.$$

[8]Or even when the game has poor-quality correlated equilibria, in the case of general regret minimization.

Corollary 26 (Hoeffding bounds). *Let $X_1, X_2, \ldots, X_n$ be independent $\{0,1\}$ random variables and let $Z_n = X_1 + \cdots + X_n$. Then*

$$\mathbf{Pr}[Z_n - \mathbf{E}[Z_n] \geq \gamma n] \leq e^{-2n\gamma^2}.$$

Lemma 27 (Chernoff bounds). *Let $X_1, X_2, \ldots, X_n$ be independent $\{0,1\}$ random variables and let $Z_n = X_1 + \cdots + X_n$. Then for $\gamma \in [0,1]$,*

$$\mathbf{Pr}[Z_n > (1+\gamma)\mathbf{E}[Z_n]] \leq e^{-\mathbf{E}[Z_n]\gamma^2/3}$$

and

$$\mathbf{Pr}[Z_n > (1-\gamma)\mathbf{E}[Z_n]] \leq e^{-\mathbf{E}[Z_n]\gamma^2/2}.$$

Appendix B. Further discussion of other dynamics: Players entering one at a time. As mentioned in section 1.2, one form of natural dynamics that have been studied in potential games is where the system starts empty and players join one at a time. Charikar et al. [17] analyze this setting for fair cost-sharing on an *undirected* graph where all players have a common sink. They consider a two-phase process. In phase 1, the players arrive one by one and each connects to the root by greedily choosing a path minimizing its cost, i.e., each selects a greedy (best response) path relative to the selection of paths by the previous players. In phase 2, players are allowed to change their paths in order to decrease their costs, namely, in the second step players play best response dynamics. Charikar et al. [17] show that interestingly the sum of the players' costs at the end of the first step will be within an $O(\log^2 n)$ factor of the cost of a socially optimal solution (which in this case is defined to be a minimum Steiner tree connecting the players to the root). This then implies that the cost of the Nash equilibrium achieved in the second step, as well as all states reached along the way, is $O(\log^3 n)$ close to **OPT**.

Note that in the directed case the result above does not hold; in fact such dynamics can lead to very poor equilibria. Figure 1 shows an example where if the players arrive one by one and each connects to the root by greedily choosing a path minimizing its cost, then the cost of the equilibrium obtained can be much worse than the cost of **OPT**. The optimal solution which is also a Nash equilibrium in this example is $(P_1, \ldots, P_n)$, where $P_i = s_i \to v \to t$ for each i; however, the solution obtained if the players arrive one at a time and each connects to the root by greedily choosing a path minimizing its cost is $(P'_1, \ldots, P'_n)$, where $P'_i = s_i \to t$ for each player i. Clearly, $\mathtt{cost}(P'_1, \ldots, P'_n) = n$, which is much worse than $\mathtt{cost}(\mathbf{OPT}) = k$. Moreover, if one modifies the example by making many copies of the edge of cost k, then even if behavior begins in a *random* initial configuration, with high probability each edge of cost k will have few players on it and so best-response behavior will lead to the equilibrium of cost n.

Appendix C. A lower bound for natural uncoupled dynamics. In noisy best-response dynamics (also called log-linear learning [42]), when it is player i's turn to move, it probabilistically chooses an action with a probability that decreases exponentially with the gap between the cost of that action and the cost of the best-reponse action. The rate of decrease is controlled by a temperature term τ, much like in simulated annealing. In fact, the dynamics can be viewed as a form of simulated annealing with the global potential as the objective function. At zero temperature, the dynamics is equivalent to standard (nonnoisy) best response, and at infinite temperature the dynamics is completely random. While it is known that with appropriate temperature control this process in the limit will stabilize at states of optimum global potential, we show here there exist cost-sharing instances such that no dynamics of this form can achieve expected cost $o(n \cdot \mathbf{OPT}/\log n)$ within a polynomial number of time steps.

We begin with a definition capturing a broad range of dynamics of this form.

DEFINITION 1. *Let* generalized noisy best response dynamics *be any dynamics where starting from a random initial state, players move one at a time in a random order, and when it is a given player i's turn to move, it probabilistically selects among its available actions. The sole requirement is that if action a is a worse response for player i than action b in not only the current state S but also in all past states S', then the probability of choosing a should be at most the probability of choosing b.*

The above definition captures almost any natural individual learning-based dynamics for players who move asynchronously. We now show that no dynamics of this form can achieve expected cost $o(n \cdot \mathbf{OPT}/\log n)$ within a polynomial number of time steps for fair cost-sharing. (We discuss the easier case of simultaneous movement dynamics afterward).

THEOREM 28. *For the fair cost-sharing game, no generalized noisy best response dynamics can achieve expected cost $o(n \cdot \mathbf{OPT}/\log n)$ within a polynomial number of time steps.*

Proof. We consider a version of the "cars or public transit" example of Figure 1, but where each player has $n-1$ cars (options of cost 1 that cannot be shared by others) in addition to the common edge of cost k, where we will use $k = \Theta(\log n)$. For this problem, we can describe the evolution of the system as a random walk on a line, where the current position t indicates the number of players currently using public transit (the shared edge of cost k). The exact probabilities in this walk depend on specifics of the dynamics and may even change over time, but one immediate fact is that so long as the walk has never reached $t \geq k$, the shared edge of cost k is a worse response to any player than its edges of cost 1. Furthermore, with high probability the initial random state has fewer than k players on the shared edge so we may assume $t < k$ at the start. Therefore, by definition of generalized noisy best response, until we reach $t = k$, each player has at most a $1/n$ chance of choosing the shared edge, and at least a $1 - 1/n$ chance of choosing a private edge, when it is that player's turn to move. Since position t corresponds to a t/n fraction of players on the shared edge and a $1 - t/n$ fraction on private edges, this in turn implies that given that the random walk is in position $1 \leq t \leq k-1$,

1. the probability $p_{t,t+1}$ of moving to position $t+1$ is at most $\frac{1}{n}(1 - \frac{t}{n})$ and
2. the probability $p_{t,t-1}$ of moving to position $t-1$ is at least $(1 - \frac{1}{n})\frac{t}{n}$.

(With the remaining probability $1 - p_{t,t+1} - p_{t,t-1}$ the walk remains in position t.) In particular,

$$\frac{p_{t,t+1}}{p_{t,t-1}} \leq \frac{n-t}{t(n-1)} \leq \frac{1}{t}.$$

We now argue that the expected time for this walk to reach position k is superpolynomial in n for $k = \log n$. In particular, consider a simplified version of the above Markov chain where $p_{t,t+1}/p_{t,t-1} = 1/t$ (rather than $\leq 1/t$) and we delete all self-loops except at the origin (so $p_{t,t+1} = 1/(t+1)$ and $p_{t,t-1} = t/(t+1)$ for $1 \leq t \leq k-1$). Deleting self-loops can only *decrease* the expected time to reach position k since it corresponds to simply ignoring time spent in self-loops, and the same for setting $p_{t,t+1}/p_{t,t-1} = 1/t$. So, it suffices to show the expected time for this simplified walk to reach position k is superpolynomial in n.

For convenience, set $p_{k,k-1} = 1$. We can now solve for the stationary distribution π of this chain. In particular, the simplified Markov chain is now equivalent to a random walk on an undirected multigraph with vertices $v_0, v_1, \ldots, v_k$ having one edge between v_k and v_{k-1}, $(k-1)$ edges between v_{k-1} and v_{k-2}, $(k-1)(k-2)$ edges between v_{k-2} and v_{k-3}, and in general $(k-1)(k-2)\cdots t$ edges between v_t and v_{t-1} for $1 \le t \le k-1$. In addition, node v_0 has $(n-1)\cdot(k-1)!$ edges in self-loops since the probability $p_{0,0}$ is at least $(n-1)/n$. Therefore, since the stationary distribution of an undirected random walk is proportional to the degree of each node [44], we have that

$$\pi_k \le \frac{1}{(n\cdot(k-1)!)} < \frac{1}{k!}.$$

Last, since the expected time h_{kk} between consecutive visits to node v_k satisfies $h_{kk} = 1/\pi_k$ by the fundamental theorem of Markov chains, the expected time to reach v_k from a start state v_t, $t < k$, is at least $1/\pi_k$ as well. So, the expected time to reach v_k is at least $k!$, which is superpolynomial in n for $k = \log(n)$ (or even $k = \omega(\log n/\log\log n)$).

Finally, the fact that the expected time to reach position k is superpolynomial in n implies that the probability of reaching position k within a polynomial number of time steps is less than $1/poly(n)$: specifically, if the walk has probability p of reaching position k in T time steps starting from position 0, then by the Markov property the expected time to reach position k is at most T/p. Moreover, so long as the walk has not yet reached position k, the cost of the system is $\Theta(n) = \Theta(n \cdot \mathbf{OPT}/k)$. Thus, the expected cost of the system within a polynomial number of time steps is $\Omega(n\cdot\mathbf{OPT}/\log n)$ as desired. ∎

We remark that the above analysis can be extended to dynamics in which all players move simultaneously (in fact, this case is even easier). In particular, so long as the system has not yet reached a state with k or more players on the shared edge, the probability of transitioning to such a state is at most the probability of observing k or more heads when flipping n coins (one per player) of bias $1/n$ (the maximum probability each player can place on the shared edge). This in turn is less than $1/poly(n)$ for $k = \omega(\log n/\log\log n)$. Thus, for $k = \log n$, with high probability play remains at states of cost $\Omega(n\cdot\mathbf{OPT}/\log n)$ for a superpolynomial number of time steps.

Appendix D. Stubborn games. Given our results, one might wonder if for all potential games, there is some constant $\alpha < 1$ under which the system will in expectation reach a near-optimal state. Here, we show that unfortunately certain games can be very stubborn, even in the basic advertising model. In particular, we consider here the game of load balancing on unrelated machines [45]. In this game, there are n jobs and m unrelated machines. The jobs are the players, and for each job j, its set of feasible actions is to place itself on some machine i, $i \in \{1,\ldots,m\}$. Each job j has associated a cost $c_{i,j}$ for running on machine i. Given an assignment of jobs to machines, the load of machine i is the sum of the costs of the jobs that are assigned to that machine, i.e., $L_i(s) = \sum_{j\in B_i(s)} c_{i,j}$, where $B_i(s)$ is the set of jobs assigned to machine i, i.e., $B_i(s) = \{j : s_j = i\}$. The cost of a player j is the load on the machine that player j selected, i.e., $\mathtt{cost}_j(s) = L_{s_j}(s)$. The traditional social cost considered for this problem, which we consider as well, is the *makespan*, which is the load on the most loaded machine, i.e., $\mathtt{cost}(s) = \max_i L_i(s)$; however, our lower bound applies equally well to the sum-of-costs social cost function (see Remark 30). The price of stability in this game is 1, since there is always a pure Nash equilibrium which is also socially optimal [27].

We show that for these games, so long as there are at least two players who do not follows the advertised strategy, the ratio of the cost of the final state to the social optimum can be unbounded.

THEOREM 29. *For any* $\epsilon > 0$ *there is a load balancing game such that for any advertised joint action* s^{ad} *and any set* R *of at most* $n-2$ *receptive players, there is an equilibrium* $s^f \in U(s^{ad}, R)$ *such that* $\texttt{cost}(s^f) \geq 1$ *while* $\mathbf{OPT} = \epsilon$*. Moreover, there exist initial joint actions* s^{ini} *such that this can occur through equilibria* s^{nr} *that are reachable via better-response dynamics from the initial state.*

Proof. Consider the following load balancing game $\mathcal{G}$ with n jobs and $m = n$ machines. Job j has cost ϵ on machine j and cost 1 on any other machine. The social optimum assigns job j to machine j and has cost ϵ. Note that a joint action in this game is a Nash equilibrium if and only if it allocates a single job to each machine. Let an *empty* machine be a machine with no job assigned to it.

Consider an initial joint action s^{ini} which assigns an even job $2k$ to machine $2k-1$ and an odd job $2k-1$ to machine $2k$. Note that this initial assignment s^{ini} is a Nash equilibrium and has cost 1. Let R be an arbitrary set of $n-2$ jobs and let s^{ad} be the recommendation for players in R. Let j_1 and j_2 be the two players not in R and $i_1 = s^{ini}_{j_1}$ and $i_2 = s^{ini}_{j_2}$ the machines on which they run in s^{ini}, respectively. Given our initial joint action s^{ini} we know that $j_1 \neq i_1$ and $j_2 \neq i_2$. Also, if $j_1 = i_2$, then $j_2 = i_1$ and vice versa. Hence, either $j_1 = i_2$ and $j_2 = i_1$ or $j_1 \neq i_2$ and $j_2 \neq i_1$.

If there is a player k in R which s^{ad} assigns to a machine $s^{ad}_k \neq k$, then consider the following dynamics. Let all the jobs except k reach any equilibrium s' for them (say, using a best-response dynamics). In s' there is no empty machine, since otherwise some job $j \neq k$ can improve its cost by moving to the empty machine. This implies that s' is a pure Nash equilibrium of the game $\mathcal{G}$ and therefore $s^f = s'$. Since the cost of job k in s^f is 1, we have that $\texttt{cost}(s^f) = 1$, and we are done. Therefore, we may assume that $s^{ad}_k = k$ for all k.

We have two remaining cases to analyze. The first case is when $j_1 \neq i_2$ and $j_2 \neq i_1$. In this case s^{ad} assigns job i_1 to machine i_1 and job i_2 to machine i_2, and each machine has a load of $1+\epsilon$, while machines j_1 and j_2 are empty. Jobs j_1 and j_2 can then undergo a better-response process and select the following equilibrium: job j_1 selects machine j_2 and job j_2 selects machine j_1, having a cost of 1 for each. Since this is an equilibrium we also reached s^f which has a cost of 1. In the second case $j_1 = i_2$ and $j_2 = i_1$. In this case after s^{ad} each machine has a single job, and hence we are at an equilibrium which has a cost of 1. □

Remark 30. Note that because $\max_i L_i(s) \leq \sum_j \texttt{cost}_j(s) \leq n \cdot \max_i L_i(s)$, by choosing $\epsilon' = \epsilon/n$, the conclusion of Theorem 29 applies to the sum social cost function as well. Specifically, we have $\sum_j \texttt{cost}_j(s^f) \geq \max_i L_i(s^f) \geq 1$, while $\mathbf{OPT} \leq n\epsilon' = \epsilon$. More generally, any upper or lower bound for makespan applies up to a factor of n for the sum social cost function.

For the case of two machines and an arbitrary number of jobs we can derive the following result.

THEOREM 31. *There is a load balancing game with two unrelated machines and* n *jobs such that for any set* R *of at most* $n/2-1$ *players, the cost of the final Nash equilibrium might be unbounded with respect to the social optimum.*

Proof. Assume that n is even, i.e., $n = 2k$. We have k jobs of type I, defined such as their cost on machine 1 is ϵ and on machine 2 is 1; we also have k jobs of type II, defined such as their cost on machine 1 is 1 and on machine 2 is ϵ. In s^{ini} all the jobs of type I are on machine 2 and all the jobs of type II are on machine 1, which is a Nash equilibrium that has a cost of k (compared to $\mathbf{OPT}$ which has a cost of $k\epsilon$).

Suppose that R includes k_1 jobs of type I and k_2 jobs of type 2 and advertises s^{ad} for them. Consider the following Nash equilibrium s^{med} for the players not in R. We select k_1 jobs from type II and k_2 jobs of type I and pair them with the jobs in R where in each pair one job is of type I and the other of type II. (Since R is strictly less than half the jobs, i.e., $|R| \leq k-1$, we can do it, and there will be at least one type I job and one type II job remaining.) For each pair of matched jobs $j_1 \in R$ and $j_2 \notin R$ we set the action of job j_2 to be the opposite machine of j_1, i.e., $s^{med}_{j_2} = 3 - s^{ad}_{j_1}$. This implies that the pair's contribution on each machine is identical. Therefore at s^{med} we have that the load on both machines is identical, and hence it is a Nash equilibrium. Since there is a pair of jobs that did not move from their action in s^{ini} the cost is at least 1, while the optimal cost is $k\epsilon$. ∎

Acknowledgments. We would like to thank the anonymous referees for their helpful comments.

REFERENCES

[1] H. Ackermann, H. Röglin, and B. Vöcking, *On the impact of combinatorial structure on congestion games*, in Proceedings of the 47th FOCS, 2006.

[2] A. Agarwal, M. Charikar, K. Makarychev, and Y. Makarychev, *$O(\sqrt{\log n})$ approximation algorithms for Min UnCut, Min 2CNF deletion, and directed cut problems*, in Proceedings of the 37th STOC, 2005.

[3] S. Albers, S. Eilts, E. Even-Dar, Y. Mansour, and L. Roditty, *On Nash equilibria for a network creation game*, in Proceedings of the 17th ACM-SIAM SODA, 2006.

[4] N. Andelman, M. Feldman, and Y. Mansour, *Strong price of anarchy*, in Proceedings of the 18th ACM-SIAM SODA, 2007.

[5] E. Anshelevich, A. Dasgupta, J. Kleinberg, E. Tardos, T. Wexler, and T. Roughgarden, *The price of stability for network design with fair cost allocation*, in Proceedings of FOCS, 2004.

[6] P. Auer, N. Cesa-Bianchi, Y. Freund, and R. E. Schapire, *The nonstochastic multiarmed bandit problem*, SIAM J. Comput., 32 (2002), pp. 48–77.

[7] B. Awerbuch, Y. Azar, A. Epstein, and A. Skopalik, *Fast convergence to nearly optimal solutions in potential games*, in Proceedings of ACM Conference on Electronic Commerce, 2008.

[8] M.-F. Balcan, A. Blum, and Y. Mansour, *Improved equilibria via public service advertising*, in Proceedings of ACM-SIAM Symposium on Discrete Algorithms, 2009.

[9] M.-F. Balcan, A. Blum, and Y. Mansour, *The price of uncertainty*, in Proceedings of the 10th ACM Conference on Electronic Commerce, 2009.

[10] M.-F. Balcan, A. Blum, and Y. Mansour, *Circumventing the price of anarchy: Leading dynamics to good behavior*, in Proceedings of ICS, 2010, pp. 200–213.

[11] M.-F. Balcan, S. Krehbiel, G. Piliouras, and J. Shin, *Near Optimality in Covering and Packing Games by Exposing Global Information*, CoRR abs/1109.3606, 2011.

[12] A. Blum, Hajiaghayi, K. Ligett, and A. Roth, *Regret minimization and the price of total anarchy*, in Proceedings of STOC, 2008.

[13] A. Blum and Y. Mansour, *Learning, regret-minimization, and equilibria*, in Algorithmic Game Theory, N. Nisan, T. Roughgarden, E. Tardos, and V. V. Vazirani, eds., Cambridge University Press, Cambridge, UK, 2007.

[14] L. Blume, *The statistical mechanics of strategic interaction*, Games Econom. Behav., 5 (1993), pp. 387–424.

[15] L. Blume, *How noise matters*, Games and Econom. Behav., 44 (2003), pp. 251–271.

[16] N. Cesa-Bianchi, Y. Freund, D. P. Helmbold, D. Haussler, R. E. Schapire, and M. K. Warmuth, *How to use expert advice*, J. ACM, 44 (1997), pp. 427–485.

[17] M. Charikar, H. Karloff, C. Mathieu, J. Naor, and M. Saks, *Online multicast with egalitarian cost sharing*, in Proceedings of the 20th SPAA, 2008.

[18] G. Christodoulou and E. Koutsoupias, *The price of anarchy of finite congestion games*, in Proceedings of the 37th STOC, 2005.

[19] G. Christodoulou, V. S. Mirrokni, and A. Sidiropoulos, *Convergence and approximation in potential games*, in Proceedings of the 23rd STACS, 2006.

[20] K. L. Clarkson, *Nearest-neighbor searching and metric space dimensions*, in Nearest-Neighbor Methods in Learning and Vision: Theory and Practice, G. Shakhnarovich, T. Darrell, and P. Indyk, eds., MIT Press, Cambridge, MA, 2006, pp. 15–59.

[21] R. Cole, Y. Dodis, and T. Roughgarden, *How much can taxes help selfish routing?*, in Proceedings of ACM-EC, 2003.

[22] R. Cole, Y. Dodis, and T. Roughgarden, *Pricing network edges for heterogeneous selfish users*, in Proceedings of STOC, 2003.

[23] J. R. Correa, A. S. Schulz, and N. E. Stier-Moses, *Selfish routing in capacitated networks*, Math. Oper. Res., 29 (2004), pp. 236–259.

[24] A. Czumaj and B. Voecking, *Tight bounds for worst-case equilibria*, in Proceedings of the 13th ACM-SIAM SODA, 2002.

[25] E. D. Demaine, M.T. Hajiaghayi, H. Mahini, and M. Zadimoghaddam, *The price of anarchy in network creation games*, in Proceedings of the 17th SODA, 2006.

[26] E. D. Demaine and M. Zadimoghaddam, *Constant price of anarchy in network-creation games via public-service advertising*, Internet Math., 8 (2012), pp. 29–45.

[27] E. Even-Dar, A. Kesselman, and Y. Mansour, *Convergence time to Nash equilibrium in load balancing*, ACM Trans. Algorithms, 3 (2007).

[28] A. Fabrikant, A. Luthra, E. Maneva, C. H. Papadimitriou, and S. Shenker, *On a network creation game*, in Proceedings of the 22nd PODC, 2003.

[29] A. Fanelli, M. Flamminia, and L. Moscardelli, *Speed of convergence in congestion games under best responce dynamics*, in Proceedings of the 35th ICALP, 2008.

[30] D. Foster and R. Vohra, *Calibrated learning and correlated equilibrium*, Games Econom. Behav., 21 (1997), pp. 40–55.

[31] D. Foster and R. Vohra, *Asymptotic calibration*, Biometrika, 85 (1998), pp. 379–390.

[32] D. Foster and R. Vohra, *Regret in the on-line decision problem*, Games Econom. Behav., 29 (1999), pp. 7–36.

[33] D. Fotakis, S. Kontogiannis, and P. Spirakis, *Selfish unsplittable flows*, Theoret. Comput. Sci., 348 (2005), pp. 226–239.

[34] J. Hannan, *Approximation to bayes risk in repeated plays*, in Contributions to the Theory of Games, M. Dresher, A. Tucker, and P. Wolfe, eds., Princeton University Press, Princeton, NJ, 1957, pp, 97–139.

[35] S. Hart and A. Mas-Colell, *A simple adaptive procedure leading to correlated equilibrium*, Econometrica, 68 (2000), pp. 1127–1150.

[36] S. Hart and A. Mas-Colell, *A reinforcement procedure leading to correlated equilibrium*, in Economic Essays, W. Hildenbrand, G. Debreu, W. Neuefeind, and W. Trockel, eds., Springer, Berlin, 2001, pp. 181–200.

[37] R. Kleinberg, G. Piliouras, and E. Tardos, *Multiplicative updates outperform generic no-regret learning in congestion games*, in Proceedings of the 41th STOC, 2009.

[38] D. Korzhyk, V. Conitzer, and R. Parr, *Complexity of computing optimal stackelberg strategies in security resource allocation games*, in Proceedings of the 24th AAAI, 2010.

[39] E. Koutsoupias and C. H. Papadimitriou, *Worst-case equilibria*, in Proceedings of the 16th STACS, 1999, pp. 404–413.

[40] N. Littlestone, P. M. Long, and M. K. Warmuth, *On-line learning of linear functions*, in Proceedings of the 23rd Symposium on Theory of Computing, ACM Press, New York, 1991, pp. 465–475.

[41] N. Littlestone and M. K. Warmuth, *The weighted majority algorithm*, Inform. Comput., 108 (1994), pp. 212–261.

[42] J. R. Marden and J. S. Shamma, *Revisiting log-linear learning: Asynchrony, completeness, and payoff-based implementation*, Games Econom. Behav., 72 (2012), pp. 788–808.

[43] D. Monderer and L. S. Shapley, *Potential games*, Games Econom. Behav., 14 (1996), pp. 124–143.

[44] R. Motwani and P. Raghavan, *Randomized Algorithms*, Cambridge University Press, Cambridge, UK, 1997.

[45] N. Nisan, T. Roughgarden, E. Tardos, and V. Vazirani, eds., *Algorithmic Game Theory*, Cambridge University Press, Cambridge, UK, 2007.

[46] J. Pita, M. Jain, F. Ordóñez, C. Portway, M. Tambe, C. Western, P. Paruchuri, and S. Kraus, *Using game theory for Los Angeles airport security*, AI Magazine, 30 (2009), pp. 43–57.

[47] R. W. Rosenthal, *A class of games possessing pure-strategy Nash equilibria*, Internat. J. Game Theory, 2 (1973).

[48] T. Roughgarden, *Intrinsic robustness of the price of anarchy*, in Proceedings of STOC, 2009.

[49] T. Roughgarden and E. Tardos, *How bad is selfish routing?*, J. ACM, 49 (2002), pp. 236–259.

[50] Y. Sharma and D. P. Williamson, *Stackelberg thresholds in network routing games or the value of altruism*, in Proceedings of ACM Conference on Electronic Commerce, 2007.
[51] G. Stoltz and G. Lugosi, *Internal regret in on-line portfolio selection*, Machine Learning J., 59 (2005), pp. 125–159.
[52] H. von Stackelberg, *The Theory of the Market Economy*, Oxford University Press, Oxford, UK, 1952.
[53] B. Widrow and M. E. Hoff, *Adaptive switching circuits*, 1960 IRE WESCON Convention Record, 1960, pp. 96–104.
[54] M. Zinkevich, *Online convex programming and generalized infinitesimal gradient ascent*, in Proceedings of the 20th International Conference on Machine Learning, 2003.