Carnegie Mellon University
Research Showcase @ CMU

Computer Science Department | School of Computer Science

3-7-2008

Regret Minimization and the Price of Total Anarchy

Avrim Blum
Carnegie Mellon University, avrim@cs.cmu.edu

MohammedTaghi Hajiaghayi
AT&T Labs

Katrina Ligett
Carnegie Mellon University

Aaron Roth
Carnegie Mellon University

Follow this and additional works at: http://repository.cmu.edu/compsci

Published In
Proceedings of the 40th Annual ACM Symposium on theory of Computing . STOC '08., 373- 382.

This Conference Proceeding is brought to you for free and open access by the School of Computer Science at Research Showcase @ CMU. It has been accepted for inclusion in Computer Science Department by an authorized administrator of Research Showcase @ CMU. For more information, please contact research-showcase@andrew.cmu.edu.

Regret Minimization and the Price of Total Anarchy *

Avrim Blum
Carnegie Mellon
Pittsburgh, PA
avrim@cs.cmu.edu

MohammadTaghi Hajiaghayi †
AT&T Labs–Research
Florham Park, NJ
hajiagha@research.att.com

Katrina Ligett‡
Carnegie Mellon
Pittsburgh, PA
katrina@cs.cmu.edu

Aaron Roth
Carnegie Mellon
Pittsburgh, PA
alroth@cs.cmu.edu

March 7, 2008

Abstract

We propose weakening the assumption made when studying the price of anarchy: Rather than assume that self-interested players will play according to a Nash equilibrium (which may even be computationally hard to find), we assume only that selfish players play so as to minimize their own regret. Regret minimization can be done via simple, efficient algorithms even in many settings where the number of action choices for each player is exponential in the natural parameters of the problem. We prove that despite our weakened assumptions, in several broad classes of games, this "price of total anarchy" matches the Nash price of anarchy, even though play may never converge to Nash equilibrium. In contrast to the price of anarchy and the recently introduced price of sinking [15], which require all players to behave in a prescribed manner, we show that the price of total anarchy is in many cases resilient to the presence of Byzantine players, about whom we make no assumptions. Finally, because the price of total anarchy is an upper bound on the price of anarchy even in mixed strategies, for some games our results yield as corollaries previously unknown bounds on the price of anarchy in mixed strategies.

1 Introduction

Computer systems increasingly involve the interaction of multiple self-interested agents. The designers of these systems have objectives they wish to optimize, but by allowing selfish agents to interact in the system, they lose the ability to directly control behavior. How much is lost by this lack of centralized control? Much as the study of approximation algorithms aims to understand what is lost when computation is limited, and the field of online algorithms aims to understand what is lost when information is limited, the study of the price of anarchy has aimed to understand what is lost when central organization is limited.

In order to study the cost incurred when coordination is lost, we must make some assumption about how selfish agents behave. Traditionally, the assumption has been that selfish agents will play Nash equilibrium strategies, and the *price of anarchy* of a game is defined to be the ratio of the value of the objective function in the worst Nash equilibrium to the social optimum value.

It does not seem realistic, however, to assume that all agents in a system will necessarily play strategies that form a Nash equilibrium. Even with centralized control, Nash equilibria can be computationally hard to find. Moreover, even when Nash equilibria are easy to find computationally, there is no reason in general to believe that distributed

*Supported in part by the National Science Foundation under grant CCF-0514922.
†Work completed while the author was a postdoctoral fellow at Carnegie Mellon.
‡Supported in part by an AT&T Labs Graduate Research Fellowship and an NSF Graduate Research Fellowship.

self-interested agents, often with limited information about the overall state of the system, will necessarily converge to them. In addition, for games with only mixed-strategy equilibria, we would have to assume that rational agents not only play so as to maximize their own utility, but also so as to preserve the stability of the system. Since a game may have many Nash equilibria, and agents may individually prefer different equilibria, it is not clear why agents would want to preserve the stability of a Nash equilibrium, even if they managed to reach one.

In this paper, we study the value obtained in games with selfish agents when we make a much weaker and more realistic assumption about their behavior. We consider repeated play of the game and allow agents to play any sequence of actions with only the assumption that this action sequence has low regret with respect to the best fixed action in hindsight. This "price of total anarchy" is strictly a generalization of price of anarchy, since in a Nash equilibrium, all players have zero regret. Regret minimization is a realistic assumption because there exist a number of efficient algorithms for playing games that guarantee regret that tends to zero, because it requires only localized information, and because in a game with many players in which the actions of any single player do not greatly affect the decisions of other players (as is often studied in the network setting), players can only improve their situation by switching from a strategy with high regret to a strategy with low regret.

We consider four classes of games: Hotelling games, in which players compete with each other for market share, valid games [30] (a broad class of games that includes among others facility location, market sharing [14], traffic routing, and multiple-item auctions), linear congestion games with atomic players and unsplittable flow [1] [7], and parallel link congestion games [24]. We prove that in the first three cases, the price of total anarchy matches the price of anarchy exactly even if the play itself is not approaching equilibrium; for parallel link congestion we get an exact match for $n = 2$ links but an exponentially greater price for general n when the social cost function is the makespan. When we consider average load instead, we prove that if the machine speeds are relatively bounded, that the price of total anarchy is $1 + o(1)$, matching the price of anarchy. For linear congestion games and average cost load balancing, the price of anarchy bounds were previously only known for pure strategy Nash equilibria, and as a corollary of our price of total anarchy bounds, we prove the corresponding price of anarchy bound for mixed Nash equilibria as well.

Most of our results further extend to the case in which only some of the agents are acting to minimize regret and others are acting in an arbitrary (possibly adversarial) manner. When studying *anarchy*, it is vital to consider players who behave unpredictably, and yet this has been largely ignored up until now. Since Nash equilibria are stable only if all players are participating, and sink equilibria [15] are defined over state graphs that assume that all players play rationally, such guarantees are not possible under the standard price of anarchy model or the price of sinking model [15].[1]

1.1 Regret minimization and the price of total anarchy

The *regret* of a sequence of actions in a repeated game is defined as the difference between the average cost incurred by those actions and the average cost the best fixed solution would have incurred, where the best is chosen with the benefit of hindsight. An algorithm is called regret-minimizing, or no-regret, if the expected regret it incurs goes to zero as a function of time.

Regret-minimizing algorithms have been known since the 1950's, when Hannan [16] gave such an algorithm for repeated two-player games. Recent work on regret minimization has focused on algorithmic efficiency and convergence rates as a function of the number of actions available, and has broadened the set of situations in which no-regret algorithms are known. Kalai and Vempala [20] show that Hannan's algorithm can be used to solve online linear optimization problems with regret approaching 0 at a rate $O(1/\sqrt{T})$, given access to an exact best-response oracle. Zinkevich [31] developed a regret-minimizing algorithm for online *convex* optimization problems. So-called *bandit* algorithms have also been developed [2, 26, 10, 22], which achieve low regret even in the situation where the algorithm receives very limited information after each round of play. Kakade et al. [19] show how to use an α-approximate best-response oracle to achieve online performance in linear optimization problems that is close to α times that of the best static solution. Those results provide efficient algorithms for many situations in which the number of strategies for each player is exponential in the size of the natural representation of the game. In cases where each player has only a polynomial number of strategies, Littlestone and Warmuth's weighted majority algorithm [25] can be used to

[1] Babaioff et al. [3] propose a model of network congestion with "malicious" players. Their model defines malicious behavior as optimizing a specific function, however, and is not equivalent to arbitrary play.

minimize regret.

In this paper, we propose regret-minimization as a reasonable definition of self-interested behavior and study the outcome of such behavior in a variety of classes of repeated games. We introduce the term "price of total anarchy" to describe the ratio between the optimum social cost and the social welfare achieved in a game where the players minimize regret. We note that the guarantees we prove using the no-regret property are strictly stronger than minimax guarantees.

1.2 Related work: Price of anarchy

Economists have long studied games with self-interested players. A *Nash equilibrium* in such a game is a profile of strategies for each player such that, given the strategies of the other players, no player prefers to deviate from her strategy in the profile. A Nash equilibrium can be *pure* or *mixed*, depending on whether the players all play pure, deterministic strategies, or they randomize over pure strategies to give a mixed strategy.

The study of the effect of selfishness in games has been quite popular in computer science for much of the past decade. In 1999, Koutsoupias and Papadimitriou [24] introduced the notion of the *price of anarchy* as a measure of this effect: they studied the ratio between the social welfare of the optimum solution and that of the worst Nash equilibrium. Since then, researchers have studied the price of anarchy in a wide variety of games (for example [28, 30, 11]).

Unfortunately, Nash equilibria are not necessarily the best definition of selfish behavior. In 2-player, n-action games, Nash equilibria are PPAD-hard to compute [4], but in *any* game with a polynomial number of actions, one can run regret-minimizing algorithms. (One can also do so efficiently in many settings with even an exponential number of actions.) Many games only admit mixed Nash equilibria, and there is no immediate incentive for players to play their given mixed strategy as opposed to any one of the pure strategies in the support of the mixed strategy. In addition, there is no reason to assume in general games that agents demonstrating selfish behavior should *converge* to a Nash equilibrium.

Our work is most similar in spirit to that of Mirrokni and Vetta [27] and Goemans et al. [15], who also question the plausibility of selfish agents converging to Nash equilibria. They introduce the notion of sink equilibria, which generalize Nash equilibria in a different way than we do. In doing so, they abandon simultaneous play, and instead consider sequential myopic best response plays. They analyze sink equilibria in the class of valid games and show that valid games have a price of sinking of between n and $n + 1$. In contrast, we prove that valid games have a price of total anarchy of 2, matching the (Nash) price of anarchy. One reason for this gap is that myopic best responses provide no guarantee about the payoff of any individual player. Indeed, the example in [15] of a valid game with price of sinking n demonstrates that myopic best response is not always rational: In their example, myopic best response players each expect average payoff tending to zero as the number of players increases, whereas they could each easily guarantee themselves payoffs of one on every turn (and would do so if they minimized regret). Additionally, because sink equilibria rely on play entering and never leaving sinks of a best response graph, the price of sinking is brittle to *Byzantine* players who may not be playing best responses. In contrast, we show that valid games have a price of total anarchy of 2 even in the presence of arbitrarily many Byzantine players, about whom we make no assumptions.

1.3 Related work: Correlated equilibria

Foster and Vohra [12] show that any algorithm that minimizes a stronger notion of regret known as "internal regret" will result in an empirical distribution of play that converges to a weaker notion of equilibrium, a correlated equilibrium. In addition, several polynomial-time internal-regret-minimizing algorithms are known for settings in which action choices are explicitly given [17]. Because we place a weaker assumption on the agents' algorithms, there are more algorithms, simpler algorithms, and more efficient algorithms for regret minimization than for internal regret minimization. In addition, we are able to prove guarantees even in Byzantine settings, where not all players behave rationally; such settings need not correspond to correlated equilibria.

1.4 Our results

In this paper, we study the price of total anarchy in four classes of games. We emphasize that our analysis does not presume that players play according to any particular class of algorithms; our results hold whenever players happen to

experience low regret, which is a strictly weaker assumption than that players play according to a Nash equilibrium. than In Section 3 we examine a class of generalized Hotelling games, where sellers select locations on a graph and achieve revenues that depend on their own locations as well as the locations chosen by the other sellers. We prove that for such games (and an even broader class, see Section 3.3), any regret minimizing player gets at least half of her fair share of the sales, regardless of how the other (Byzantine) players behave.[2] This result exactly matches the price of anarchy in these games.

Valid games, introduced by Vetta [30], model games where the social utility is submodular, the private utility of each player is at least her Vickrey utility (the amount her presence contributes to the overall welfare), and where the sum of the players' private utilities is at most the total social utility. In Section 4 we prove that the price of total anarchy in valid games with nondecreasing social utility functions exactly matches the (Nash) price of anarchy, even if Byzantine players are added to the system.

Finally, in Section 5, we analyze atomic congestion games with two types of social welfare functions. First, we consider unweighted atomic congestion games with player-summed social welfare functions, and in both the linear cost and the polynomial cost case, we show price of total anarchy results that match the price of anarchy [7, 1]. Next, we consider a parallel link congestion game with social welfare equal to makespan, the game that initiated the study of the price of anarchy [24], and show that the price of total anarchy of the parallel link congestion game with two links is $3/2$, exactly matching the price of anarchy. We also show that the price of total anarchy in the parallel link game with n links is $\Omega(\sqrt{n})$, which is strictly worse than the price of anarchy. Finally, we show a price of total anarchy matching the known price of anarchy in the load balancing game with the sum social utility function. In the case of load balancing with sum social utility, our price of total anarchy results also yield previously unknown price of anarchy results for mixed strategies.

In Section 6, we discuss techniques for minimizing regret in each of these settings.

2 Preliminaries

In this paper, we consider k-player games. For each player i, we denote by $\mathcal{A}_i$ the set of pure strategies available to that player. A mixed strategy is a probability distribution over actions in $\mathcal{A}_i$; we denote by $\mathcal{S}_i$ the set of mixed strategies available to player i. Let $\mathcal{A} = \mathcal{A}_1 \times \mathcal{A}_2 \times \ldots \times \mathcal{A}_k$ and $\mathcal{S} = \mathcal{S}_1 \times \mathcal{S}_2 \times \ldots \times \mathcal{S}_k$. Every game has an associated social utility function $\gamma : \mathcal{A} \to \mathbb{R}$ that takes a set containing an action for each player to some real value. Each player i has an individual utility function $\alpha_i : \mathcal{A} \to \mathbb{R}$.

We often want to talk about the social or individual utility of a strategy profile $S = \{s_1, \ldots, s_k\} \in \mathcal{S}$. To this end, we denote by $\bar{\gamma} : \mathcal{S} \to \mathbb{R}$ the expected social utility over randomness of the players and by $\bar{\alpha}_i : \mathcal{S} \to \mathbb{R}$ the expected value of the utility of a strategy profile to player i. We denote the social value of the socially optimum strategy profile by $\mathbf{OPT} = \max_{S \in \mathcal{S}} \bar{\gamma}(S)$ in maximization problems. Correspondingly, $\mathbf{OPT} = \min_{S \in \mathcal{S}} \bar{\gamma}(S)$ in minimization problems.

We also sometimes wish to talk about a modification of a particular strategy profile; let $S \oplus s_i'$ be the strategy set obtained if player i changes her strategy from s_i to s_i'. Let $\emptyset_i$ be the null strategy for player i (player i takes no action). We use superscripts to denote time, so S^t is the strategy profile at time t; s_i^t is player i's strategy at time t.

We consider both maximization and minimization games in this paper. In *maximization* games the goal is to *maximize* the social utility function and the players wish to *maximize* their individual utility functions; in *minimization* games, both quantities minimized. We define the price of anarchy and the price of total anarchy so that their values are always greater than or equal to one, regardless of whether we are discussing a maximization or a minimization game:

Definition 2.1. *The price of anarchy for an instance of a maximization game is defined to be $\frac{\mathbf{OPT}}{\bar{\gamma}(S)}$, where S is the worst Nash equilibrium for the game (the equilibrium that maximizes the price of anarchy). The price of anarchy for an instance of a minimization game is defined to be $\frac{\bar{\gamma}(S)}{\mathbf{OPT}}$, where S is the worst Nash equilibrium for the game (the equilibrium that maximizes the price of anarchy).*

[2] We note that robustness to Byzantine players is not inherent in our model. Indeed, there exist games for which the addition of Byzantine players can make the social welfare, as well as the utilities of individual regret-minimizing players, arbitrarily bad.

Definition 2.2. *The regret of player i in a maximization game given action sets $A^1, A^2, \ldots, A^T$ is*

$$\max_{a_i \in \mathcal{A}_i} \frac{1}{T} \sum_{t=1}^{T} \alpha_i(A^t \oplus a_i) - \frac{1}{T} \sum_{t=1}^{T} \alpha_i(A^t).$$

The regret of player i in a minimization game given action sets $A^1, A^2, \ldots, A^T$ is

$$\frac{1}{T} \sum_{t=1}^{T} \alpha_i(A^t) - \min_{a_i \in \mathcal{A}_i} \frac{1}{T} \sum_{t=1}^{T} \alpha_i(A^t \oplus a_i).$$

A regret-minimizing algorithm is one with low expected regret.

Property 2.3. *When a player i uses a regret-minimizing algorithm or achieves low regret, for any sequence $A^1, \ldots, A^T$, she achieves the property*

$$\max_{a_i \in \mathcal{A}_i} \frac{1}{T} \sum_{t=1}^{T} \alpha_i(A^t \oplus a_i) \leq R(T) + E\left[\frac{1}{T} \sum_{t=1}^{T} \alpha_i(A^t)\right]$$

for maximization games and

$$E\left[\frac{1}{T} \sum_{t=1}^{T} \alpha_i(A^t)\right] \leq R(T) + \min_{a_i \in \mathcal{A}_i} \frac{1}{T} \sum_{t=1}^{T} \alpha_i(A^t \oplus a_i)$$

for minimization games, where expectation is over the internal randomness of the algorithm, and where $R(T) \to 0$ as $T \to \infty$. The function $R(T)$ may depend on the size of the game or a compact representation thereof. We then define T_ϵ to be the number of time steps required to get $R(T) = \epsilon$.

Note that this implies that, for any sequence $S^1, \ldots, S^T$, a player with the regret-minimizing property achieves

$$\max_{a_i \in \mathcal{A}_i} \frac{1}{T} \sum_{t=1}^{T} \bar{\alpha}_i(S^t \oplus a_i) \leq R(T) + \frac{1}{T} \sum_{t=1}^{T} \bar{\alpha}_i(S^t)$$

for *maximization* games and

$$\frac{1}{T} \sum_{t=1}^{T} \bar{\alpha}_i(S^t) \leq R(T) + \min_{a_i \in \mathcal{A}_i} \frac{1}{T} \sum_{t=1}^{T} \bar{\alpha}_i(S^t \oplus a_i)$$

for *minimization* games.

Definition 2.4. *The price of total anarchy for an instance of a maximization game is defined to be $\max \frac{\mathbf{OPT}}{\frac{1}{T}\sum_{t=1}^{T} \bar{\gamma}(S^t)}$, where the max is taken over all T and $S^1, S^2, \ldots, S^T$, where $S_1, \ldots, S_T$ are play profiles of players with the regret-minimizing property. The price of total anarchy for an instance of a minimization game is defined to be $\max \frac{\frac{1}{T}\sum_{t=1}^{T} \bar{\gamma}(S^t)}{\mathbf{OPT}}$, where the max is taken over all T and $S^1, S^2, \ldots, S^T$, where $S_1, \ldots, S_T$ are play profiles of players with the regret-minimizing property.*

Because all players have zero regret when playing a Nash equilibrium, the price of total anarchy of a game is never less than its price of anarchy. In this paper we study the price of anarchy and the price of total anarchy for general classes of games. The price of (total) anarchy for a class of games is defined to be the maximum price of (total) anarchy over any instance in that class. Bounds on the price of (total) anarchy for a class of games may not be tight for particular instances in that class.

3 Hotelling games

Hotelling games [18] are well studied in the economics literature; see, for example, [13] and [21] for surveys. Hotelling games are traditionally location games played on a line, but we generalize them to an arbitrary graph and a broad class of behaviors on the part of the customers. We prove our result first for a specific Hotelling game, and then observe that our proof still holds in a much more general setting.

3.1 Definition and price of anarchy

Imagine a set of souvenir stand owners in Paris who must decide where to set up their souvenir stands each day. Every day, n tourists buy a souvenir from whichever stand they find first. Each stand operator wishes to maximize her own sales. Every day there are n sales, and we wish to maximize fairness: The social welfare function is the minimum sales of any souvenir stand. Formally, this maximization game is defined by an n vertex graph $G = (V, E)$. Every seller i among the k sellers has strategy set $\mathcal{A}_i = V$, that is, every day she sets up her stand on some vertex of the graph. Each day, every tourist chooses a path from some private distribution over paths on the graph, and buys from the seller he encounters first (for instance, as a special case, we could have one tourist at each vertex of the graph who purchases from the nearest souvenir stand). If there is a tie between sellers, we assume the tourist splits his contribution among them equally). At any time t the social welfare is $\bar{\gamma}(S^t) = \min_i \bar{\alpha}_i(S^t)$. The social optimum is obtained by splitting all vertices equally among all k players (this can be achieved if all players play on the same vertex). Therefore $\mathbf{OPT} = n/k$.

Theorem 3.1. *The price of anarchy of the Hotelling game is* $(2k-2)/k$.

Proof. Given a strategy set S, consider the alternate set $(S \oplus \emptyset_i)$. There are $k-1$ active players in this alternate set and the total payoff is still n, so there must be some player h who achieves expected payoff $\bar{\alpha}_h(S \oplus \emptyset_i) \geq n/(k-1)$. If player i played the same strategy as player h, she would achieve expected payoff $\bar{\alpha}_i(S \oplus s_h) \geq \frac{n}{(2k-2)}$. Thus, any strategy achieving expected payoff less than $\frac{n}{(2k-2)}$ is not an equilibrium strategy, since in a Nash equilibrium, no player wishes to change her strategy.

This bound is tight: Consider a game on a graph with $k-1$ identical stars, where we identify tourists with vertices of the graph and each patronizes the nearest souvenir stand. In this example, $k-1$ of the players play deterministically at the center of their own star; player k plays uniformly at random over all $k-1$ star centers. This strategy set S is a Nash equilibrium, and the randomizing player earns $\bar{\alpha}_k(S) = n/(2k-2)$ (the other players do better), so the social welfare $\gamma(S) = n/(2k-2)$. Since $\mathbf{OPT} = n/k$, this demonstrates that the price of anarchy is $\frac{\mathbf{OPT}}{\bar{\gamma}(S)} = \frac{(2k-2)}{k}$. □

3.2 Price of total anarchy

Since at a Nash equilibrium, no player has regret, the price of total anarchy for the Hotelling game is at least $(2k-2)/k$. In this section, we show that this value is tight; that is:

Theorem 3.2. *The price of total anarchy in the Hotelling game is* $(2k-2)/k$*, matching the price of anarchy.*

The proof of this theorem relies on the symmetry of the game; this property was similarly useful to Chien and Sinclair [5] in the context of studying convergence to Nash equilibria in symmetric congestion games.

Let O_i^t be the set of plays at time t by all players *other* than player i. Let $O_i = \sum_{t=1}^{T} O_i^t$, the union with multiplicity of all plays of players other than i over all time periods.

Definition 3.3. *Let* $\Delta_i^{t \to u}$ *be the quantity such that if player* i *plays an action uniformly at random from* O_i^t *at time step* u*, she achieves expected payoff* $n/(2k-2) + \Delta_i^{t \to u}$*. Note that* $\Delta_i^{t \to t}$ *is always* 0 *because the* $k-1$ *other players have average payoff exactly* $n/(k-1)$ *when player* i *is removed.*

Lemma 3.4. *For all* i*, for all* $1 \leq t, u \leq T$*:* $\Delta_i^{u \to t} + \Delta_i^{t \to u} \geq 0$.

Proof. If $t = u$, the claim follows easily, as noted in the definition. Otherwise, imagine a $(2k-2)$-player game in which there is a time-t player and a time-u player for each original player other than i. The time-t version of a player j plays strategy s_j^t; the time u version plays s_j^u. Since the sum of all players' payoffs is n, if player i picks a random strategy from among those already being played and plays it in this imaginary game *replacing* the player she copies, i expects to have payoff $n/(2k-2)$. Half of the time, player i will select a time-t strategy and replace that time-t player. It can only improve i's payoff in this case to remove all of the other time-t players and only play against time-u players. This leaves i playing a strategy uniformly selected from O_i^t at time u. A parallel argument holds the other

half of the time, when player i selects a time-u strategy, and thus

$$\begin{aligned}\frac{n}{(2k-2)} &\leq \frac{1}{2}\left(\frac{n}{(2k-2)} + \Delta_i^{t\to u}\right) + \frac{1}{2}\left(\frac{n}{(2k-2)} + \Delta_i^{u\to t}\right) \\ &= \frac{n}{(2k-2)} + \frac{1}{2}(\Delta_i^{t\to u} + \Delta_i^{u\to t})\end{aligned}$$

as desired. □

of Theorem 3.2. Fix a sequence of plays $S^1, \ldots, S^T$. Recall that $O_i = O_i^1 + \ldots + O_i^T$. Define o_i^t to be the uniform distribution over O_i^t. Picking an action a uniformly at random from O_i is equivalent to picking a random time step u and then picking a strategy $a \in O_i^u$ uniformly at random. Player i's expected payoff had she randomly selected o_i^u and played it over all T rounds is

$$\begin{aligned}\frac{1}{T}\sum_{u=1}^{T}\sum_{t=1}^{T}\bar{\alpha}_i(S^t \oplus o_i^u) &= \frac{1}{T}\sum_{u=1}^{T}\sum_{t=1}^{T}\left(\frac{n}{(2k-2)} + \Delta_i^{u\to t}\right) \\ &= \frac{Tn}{(2k-2)} + \frac{1}{T}\sum_{u=1}^{T}\sum_{t=1}^{T}\Delta_i^{u\to t} \\ &\geq \frac{Tn}{(2k-2)},\end{aligned}$$

where the last inequality holds because of Lemma 3.4. Therefore, there must be some single fixed action $a^* \in S$ that achieves at least $\frac{Tn}{(2k-2)}$ when played over T rounds of the above game. Any regret minimizing player achieves expected total payoff at least this much (minus ϵ), and so has expected payoff at least $n/((2k-2)) - \epsilon$, proving the theorem. □

3.3 The price of total anarchy in generalized Hotelling games

We note that the proof of Theorem 3.2 made no use of the specifics of the Hotelling game described above. In particular, the same proof shows that any regret minimizing player achieves expected payoff approaching $n/(2k-2)$ *regardless* of how other players behave, and so we are able to guarantee good payoff among regret-minimizing players players even in the presence of Byzantine players making arbitrary (or adversarial) decisions.

Theorem 3.5. *Any player who minimizes regret in the Hotelling game achieves payoff approaching* $n/(2k-2)$*, regardless of how the other players play.*

The same proof also holds when the buyers use much more general rules for choosing which stand to patronize.[3] Neither do we use the fact that players' utilities are linear. In fact, our proof only makes use of three properties of the Hotelling game:

1. **Constant Sum**: The individual utilities of the players in the game always sum to the same value, regardless of play.
2. **Symmetric**: All players have the same action set, and the payoff vector is a function of the action vector that is invariant to a permutation of the names of the players.
3. **Monotone**: The game is defined for any number of players, and removing players from the game (while keeping the strategies of the remaining players fixed) does not decrease the payoff for any remaining player.

We call such games with the "fairness" social utility function $\bar{\gamma}(S) = \min_i \alpha_i(S)$ *generalized Hotelling games* and get the following theorem:

[3] One caveat is that customers may not in general base their selection rules on the actions of the players—for instance by patronizing the *second* closest souvenir stand. If we were to allow rules such as this, removing players from the game could decrease the payoff of some of the remaining players, and we rely on this not being the case.

Theorem 3.6. *In any k-player, generalized Hotelling game, the price of total anarchy among regret minimizing players is $(2k-2)/k$ even in the presence of arbitrarily many Byzantine players.*

3.4 Regret minimization need not converge

Since players may efficiently minimize regret in Hotelling games, but may not necessarily be able to compute Nash equilibria, it is notable that we are able to match standard price-of-anarchy guarantees. In fact, it is possible that regret-minimizing players in Hotelling games never converge to a Nash equilibrium:

Theorem 3.7. *Even if all players in the Hotelling game are regret minimizing, stage game play need not converge to Nash equilibrium.*

Proof. Consider k players $\{0, \ldots, k-1\}$ on a graph with $k-1$ identical n-vertex stars with centers $v_0, \ldots, v_{k-2}$ and an isolated vertex v_{k-1}. At time period t, player i plays on vertex $v_{t+i \mod k}$. Each player has expected payoff $(n(k-1)+1)/k$, but no fixed vertex has expected payoff more than $(n(k+1)/2k)$, so no player has positive regret. However, at each time period, the player at the isolated vertex v_{k-1} has incentive to deviate, so this is not a Nash equilibrium. □

A similar example shows that even if all players minimize internal regret (so that play is guaranteed to converge to the set of correlated equilibria), play can cycle forever and so need not converge to Nash equilibrium.[4]

4 Valid games

4.1 Definitions and price of anarchy

Valid games, introduced by Vetta [30], are a broad class of games that includes the market sharing game studied by Goemans et al. [14], the facility location problem, a version of the traffic routing problem of Roughgarden and Tardos [28], and multiple-item auctions [30]. When describing valid games, we slightly adapt the notation of [30]. Consider a k-player maximization game, where each player i has a groundset of actions $\mathcal{V}_i$ from which she can play some subset. Not every subset of actions is necessarily allowed. Let $\mathcal{V} = \mathcal{V}_1 \times \ldots \times \mathcal{V}_k$, and let $\mathcal{A}_i = \{a_i \subseteq \mathcal{V}_i : a_i \text{ is a feasible action}\}$. Let the game have some social utility function $\gamma : 2^{\mathcal{V}} \to \mathbb{R}$, and let each player have a private utility function $\alpha_i : 2^{\mathcal{V}} \to \mathbb{R}$. The discrete derivative of f at $X \subseteq V$ in the direction $D \subseteq V - X$ is $f'_D(X) = f(X \cup D) - f(X)$.

Definition 4.1. *A set function $f : 2^{\mathcal{V}} \to \mathbb{R}$ is submodular if for $A \subseteq B$, $f'_i(A) \geq f'_i(B)\ \forall i \in \mathcal{V} - B$.*

Note that submodular utility functions represent the economic concept of decreasing marginal utility, reflecting economies of scale.

Definition 4.2. *A game with private utility functions $\alpha_i : 2^{\mathcal{V}} \to \mathbb{R}$ and social utility function $\gamma : 2^{\mathcal{V}} \to \mathbb{R}$ is valid if γ is submodular and*

$$\bar{\alpha}_i(S) \geq \bar{\gamma}'_{s_i}(S \oplus \emptyset_i) \tag{1}$$

$$\sum_{i=1}^{k} \bar{\alpha}_i(S) \leq \bar{\gamma}(S) \tag{2}$$

Condition 1 states that each agent's payoff is at least her *Vickrey utility*—the change in social utility that would occur if agent i did not participate in the game. Condition 2 states that the social utility of the game is at least the sum of the agents' private utilities.

[4] k players play on a set of $k/2+1$ vertices. Players are divided into two equal sized groups, L and R. Every turn, there is exactly one player on $k/2$ vertices, and $k/2$ players on the remaining vertex. Players in L and R get their own vertices on alternate turns, and the crowded vertex rotates, so that each player is equally often on every vertex, and on any particular vertex she is equally often alone and crowded. Therefore no player has any incentive to swap any vertex with any other.

For example, consider the market sharing game studied by Goemans et al. [14]. The game is played on a bipartite graph $G = ((V, U), E)$. Each vertex in V is a player, and each vertex in U is a market. Each market has a value and a cost to service it, and each player has a budget. A player may enter a set of markets to which she has edges, if the sum of their costs is at most her budget. For each market that a player enters, she receives payoff equal to the value of that market divided by the number of players that chose to enter it. The social utility function is the sum of the individual player utilities, or equivalently, the sum of the values of the markets that have been entered by any player. This valid game models a situation in which cable internet providers enter different cities with values proportional to their populations and share the market equally with other local providers; the social utility is the number of people with access to high speed internet.

Vetta [30] analyzes the price of anarchy of valid games and shows that if S is a Nash equilibrium strategy and $\Omega = \{\sigma_1, \ldots, \sigma_k\}$ is a strategy profile optimizing the social utility function so that $\bar{\gamma}(\Omega) = \mathbf{OPT}$, then

$$\begin{aligned}\mathbf{OPT} \leq 2\bar{\gamma}(S) \quad &- \sum_{i:\sigma_i = s_i} \bar{\gamma}'_{s_i}(S \oplus \emptyset_i) \\ &- \sum_{i:\sigma_i \neq s_i} \bar{\gamma}'_{s_i}(\Omega \cup (S \oplus \emptyset_i \oplus \ldots \oplus \emptyset_k)).\end{aligned}$$

Thus, if γ is nondecreasing, then for any Nash equilibrium strategy S, $\gamma(S) \geq \mathbf{OPT}/2$, giving a price of anarchy of 2. In contrast, Goemans et al. [15] show that the price of sinking for valid games is larger than n.

4.2 Price of total anarchy

In this section, we show that the price of total anarchy for valid games matches the price of anarchy exactly:

Theorem 4.3. *If all players play regret-minimizing strategies for T rounds, with strategy profile S^i at time i, then*

$$\begin{aligned}\mathbf{OPT} \leq \frac{1}{T}\sum_{i=1}^{T}\Bigg(2\bar{\gamma}(S^t) - \sum_{i:\sigma_i = s_i^t} \bar{\gamma}'_{s_i^t}(S^t \oplus \emptyset_i) \\ - \sum_{i:\sigma_i \neq s_i^t} \bar{\gamma}'_{s_i^t}(\Omega \cup (S^t \oplus \emptyset_i \oplus \ldots \oplus \emptyset_k))\Bigg) + \epsilon k.\end{aligned}$$

We defer this proof to the appendix. For nondecreasing γ, we get the following corollary:

Corollary 4.4. *If γ is nondecreasing, the price of total anarchy for valid games is asymptotically 2.*

The price of anarchy and the price of sinking are both brittle to the addition of Byzantine players. In contrast, for nondecreasing social welfare functions γ, our price of total anarchy result holds even in the presence of arbitrarily many Byzantine players. In any valid game, suppose players $1, \ldots, k$ are regret minimizing. Let $\mathbf{OPT} = \gamma(\Omega)$ be the optimal value for these players playing alone. Suppose there is some additional set of Byzantine players $\mathcal{B}$ that behave arbitrarily.

Theorem 4.5. *Consider a valid game with nondecreasing social welfare function γ, where the k regret minimizing players play $S^1, \ldots, S^T$ over T time steps while the Byzantine players play $B^1, \ldots, B^T$. Then the average social welfare $1/T \sum_{t=1}^{T} \gamma(S^t \cup B^t) \geq \mathbf{OPT}/2$.*

Proof. We observe that

$$\begin{aligned}&\gamma(\Omega \cup B^t) \\ &\quad \leq \gamma(\Omega \cup S^t \cup B^t) \\ &\quad = \gamma(S^t \cup B^t) + \sum_{i:\sigma_i \neq s_i^t} \gamma'_{\sigma_i}(S^t \cup B^t \cup (\Omega \oplus \emptyset_i \oplus \ldots \oplus \emptyset_k)) \\ &\quad \leq \gamma(S^t \cup B^t) + \sum_{i:\sigma_i \neq s_i^t} \gamma'_{\sigma_i}(S^t \oplus \emptyset_i \cup B^t),\end{aligned}$$

where the first inequality follows because γ is nondecreasing, and the third follows from submodularity. We then have

$$\begin{aligned}\mathbf{OPT} &\leq \gamma(\Omega \cup B^t) \\ &\leq \gamma(S^t \cup B^t) + \sum_{i:s_i \neq \sigma_i} \gamma'_{\sigma_i}(S^t \oplus \emptyset_i \cup B^t) \\ &\leq \gamma(S^t \cup B^t) + \sum_{i:s_i \neq \sigma_i} \alpha_i(S^t \oplus \sigma_i \cup B^t)\end{aligned}$$

with the first line following because γ is nondecreasing, and the second from the Vickrey condition. Summing over T, this yields

$$T \cdot \mathbf{OPT} \leq \sum_{t=1}^{T} \gamma(S^t \cup B^t) + \sum_{t=1}^{T} \sum_{i:s_i \neq \sigma_i} \alpha_i(S^t \oplus \sigma_i \cup B^t).$$

Suppose $\sum_{t=1}^{T} \gamma(S^t \cup B^t) < T \cdot \mathbf{OPT}/2$. Since

$$\sum_{t=1}^{T} \sum_{i=1}^{k} \alpha_i(S^t \cup B^t) \leq \sum_{t=1}^{T} \sum_{i=1}^{k+|\mathcal{B}|} \alpha_i(S^t \cup B^t) \leq \sum_{t=1}^{T} \gamma(S^t \cup B^t),$$

it must be that

$$\sum_{i=1}^{k} \sum_{t=1}^{T} \alpha_i(S^t \oplus \sigma_i \cup B^t) > \sum_{i=1}^{k} \sum_{t=1}^{T} \alpha_i(S^t \cup B^t),$$

and so there is some regret minimizing player i for whom $\sum_{t=1}^{T} \alpha_i(S^t \oplus \sigma_i \cup B^t) > \sum_{t=1}^{T} \alpha_i(S^t \cup B^t)$, violating the condition that he is regret minimizing. □

Note that here we have shown that in a valid game with a nondecreasing social utility function, if k players minimize regret and an arbitrary number of Byzantine players are *added* to the system, the resulting social welfare is no worse than half the optimal social welfare for k players. This is a slightly different result than we showed for Hotelling games, where we were able to guarantee that each regret-minimizing player obtains at least half of her fair share of the entire game, regardless of what the other $k-1$ players do. On the other hand, for valid games one clearly cannot obtain half of the optimum social welfare for $k + |\mathcal{B}|$ players since the Byzantine players need not be acting in even their own interest.

5 Atomic Congestion Games

In this section, we show price of total anarchy results matching existing price of anarchy results for atomic, unweighted congestion games with social utility equal to the sum of the player utilities [7, 1]. We also consider the atomic congestion game of weighted load balancing with social utility equal to the makespan [24, 8, 23], and show matching results for two links, but demonstrate that for n links, the price of total anarchy is exponentially worse than the price of anarchy. Finally, we consider weighted load balancing with social utility equal to the sum of the player utilities [29], and show that for $k >> n$, the price of total anarchy is $1 + o(1)$. In the case of load balancing with sum social utility, our price of total anarchy results also imply previously unknown price of anarchy results for mixed strategies.

A congestion game is a minimization game consisting of a set of k players and, for each player i, a set $\mathcal{V}_i$ of facilities. Player i plays subsets of facilities from some feasible set $\mathcal{A}_i = \{a_i \subseteq \mathcal{V}_i : \ a_i \text{ is a feasible action}\}$. In *weighted* games, each player i has an associated weight w_i; in *unweighted* games, each player weight is 1. Each facility e has an associated latency function f_e. A player i playing a_i experiences cost $\alpha_i = \sum_{e \in a_i} f_e(l_e)$ where l_e is the load on facility e: $l_e = \sum j : e \in a_i w_j$.

5.1 Atomic congestion games with sum social utility

In this section, we consider unsplittable atomic selfish routing with unweighted players. The social utility function we consider in this section is the sum of the player costs, or $\gamma(A) = \sum_i \alpha_i(A)$. We write $\Omega = \{\sigma_1, \ldots, \sigma_k\}$ for a strategy profile optimizing the social utility function γ. We write l_e^t for the load on edge e at time t, and l_e^* for the load on edge e in Ω.

We first consider linear edge costs of the form $f_e(l_e) = c_e l_e + b_e$ for edge e. In this setting, Christodoulou and Koutsoupias [7] and Awerbuch et al. [1] independently showed that the price of anarchy for pure strategies is 2.5. We show a matching bound for the price of total anarchy, which also implies the matching bound shown by Christodoulou and Koutsoupias [6] for the price of anarchy for mixed strategies and for correlated equilibria. We defer the proof of the theorem to the appendix.

Theorem 5.1. *The price of total anarchy of atomic congestion games with unweighted players, sum social utility function, and linear cost functions is 2.5.*

Corollary 5.2 (Christodoulou and Koutsoupias [6]). *The price of anarchy of atomic congestion games with unweighted players, sum social utility function, and linear cost functions is 2.5, even for mixed strategies. The same bound also holds for correlated equilibria in this setting.*

We next consider polynomial latency functions and show a bound matching the price of anarchy shown by Christodoulou and Koutsoupias [7] and Awerbuch et al. [1] for mixed strategies. We defer the proof to the appendix.

Theorem 5.3. *The price of total anarchy of atomic congestion games with unweighted players, sum social utility function, and polynomial latency functions of degree d is at most $d^{d^{1-o(1)}}$.*

5.2 Parallel link congestion game with makespan social utility

The parallel link congestion game models n identical links and k weighted players (jobs) who must choose which link to use. Each player pays the sum of the weights of the jobs on the link she chose. The social cost for this game is defined as the total weight on the worst-loaded link. This game was the main focus of the Koutsoupias and Papadimitriou paper that introduced the concept of the price of anarchy [24].

More formally, this is a minimization game where for each player i, the feasible actions are $\mathcal{A}_i = \{1, \ldots n\}$. The social utility function is $\gamma(A) = \max_{j \in \{1,\ldots n\}} \sum_{i:a_i=j} w_i$.

Koutsoupias and Papadimitriou [24] proved that the price of anarchy of the parallel link congestion game with two links is $3/2$. Two groups of researchers [9, 23] later proved that the price of anarchy when there are n links is $\Theta(\log n / \log\log n)$.

In this section, we show a matching bound on the price of total anarchy for 2 links. We also show that for n links, the price of total anarchy *does not* match the price of anarchy.

Theorem 5.4. *The price of total anarchy of the parallel link congestion game with makespan social utility and two links is $3/2$, exactly matching the price of anarchy.*

The proof, which we defer to the appendix, parallels that in the original Koutsoupias and Papadimitriou paper [24]. It is subtler because regret-minimizing algorithms only give a guarantee in expectation, on average, and make no guarantees about the performance on any given day.

For the parallel link congestion game with n links, the price of total anarchy diverges from the price of anarchy. This divergence stems from the fact that in the parallel links game, the social cost function γ is defined in terms of expected maximum *link* latency, whereas individual utility is a function of average *job* latency.[5] In the single stage Nash equilibrium analyzed for price of anarchy results, the two values are related: expected job latency for player i is equal to the average link latency of every link in the support of i's mixed strategy. In a Nash equilibrium, therefore,

[5]Note that if we were to redefine the social cost function γ for the parallel links game to be the maximum expected *job* latency, it is simple to verify that the resulting price of total anarchy is 2: Rescale the weights so that $\mathbf{OPT} = 1$. Total weight is $\leq n$, and $w_i \leq 1$ for all players. Over any sequence of plays, there must be some link with average latency $l \leq 1$. Therefore, every player i is guaranteed to experience average latency in expectation at most $l + w_i + \epsilon \leq 2 + \epsilon$.

maximum expected link latency must be low, and with tail bounds, it is straightforward to argue that the expected maximum link latency cannot be too high [9]. Over an arbitrary sequence of regret-minimizing plays, however, average job latency no longer necessarily corresponds to the average latency of any link. This is demonstrated by a cycling example we use in the proof of the following theorem, which we defer to the appendix:

Theorem 5.5. *The price of total anarchy in the parallel link game with makespan social utility and n links is $\Omega(\sqrt{n})$.*

5.3 Parallel links congestion game with sum social utility

We have just shown that the price of total anarchy does not match the $O(\log n/\log\log n)$ price of anarchy for the load balancing game with the makespan social utility function. The results in Section 5.1, however, imply a price of total anarchy ≤ 2.5 for the load balancing game with the *sum* social utility function (since load balancing is a special case of routing), even for mixed strategies and different server speeds. In fact, we can show more: in this section, we show that so long as $k >> n$ and the server speeds are relatively bounded, the price of total anarchy is $1 + o(1)$. This matches a price of anarchy result shown by Suri et al. [29] for pure strategy equilibria. Our theorem below, the proof of which we defer to the appendix, implies an equivalent price of anarchy result even for mixed strategy equilibria.

Theorem 5.6. *In the load balancing game with sum social cost and linear latency functions, the price of total anarchy is $1 + o(1)$ provided that $k >> n$ and server speeds are relatively bounded.*

Corollary 5.7. *In the load balancing game with sum social cost and linear latency functions, the price of anarchy is $1 + o(1)$ provided that $k >> n$ and server speeds are relatively bounded, even for mixed strategies.*

6 Algorithmic efficiency

In the Hotelling games we analyzed in Section 3, each player has only n strategies—the n nodes in the graph. In such settings, the weighted majority algorithm [25] runs in polynomial time and minimizes regret. Similarly, in the parallel links congestion game, there are n strategies—the n links—and thus minimizing regret is relatively straightforward.

In valid games, if the set of actions available to a player is polynomial in $|\mathcal{V}_i|$, the action groundset, then once again, weighted majority can be used to minimize regret. However, in arbitrary valid games, the action space for player i could be as large as $2^{|\mathcal{V}_i|}$. In such situations, if the player's private utility is a linear function of the elements of the groundset she obtains and she can compute exact best responses in polynomial time (such as in the market sharing game of Goemans et al. [14]), then she can use results of Kalai and Vempala [20] to minimize regret in polynomial time. If her utility function is linear, but she can only compute approximate best responses, results of Kakade et al. [19] allow her to *approximately* minimize regret; that is, she obtains expected average cost close to β times the cost of the best fixed solution in hindsight, where β is the approximation ratio of her optimizer. We can modify our proof of the price of total anarchy to carry this β through and show:

Theorem 6.1. *The price of β-minimizing regret in valid games is $1 + \beta$.*

If the player's utility function is convex and well-defined over the convex hull of her pure strategies and she furthermore has the ability to project points in space onto that convex hull, then she can use an algorithm developed by Zinkevich [31] to minimize her regret. In situations where no existing techniques are a perfect fit, more specialized regret-minimizing algorithms for specific games may also be developed.

7 Conclusions

We propose regret minimization as a definition of selfish behavior in repeated games. We consider four general classes of games—generalized Hotelling games, valid games, and atomic congestion games with two different social utility functions—and show that the price of total anarchy exactly matches the price of anarchy in most cases, but there is a gap of $\Omega(\sqrt{n})$ versus $O(\frac{\log n}{\log\log n})$ in the case of n parallel links. Our results hold even in games where regret-minimizing algorithms can cycle and fail to converge to an equilibrium. We also prove results in Byzantine settings

when only some of the players achieve regret minimization and the other players are allowed to act in an arbitrary fashion. In addition, our results for weighted load balancing with player-summed social utility functions imply new price of anarchy results for mixed strategies.

Acknowledgments We thank Evangelia Pyrga for bringing [6] to our attention.

References

[1] B. Awerbuch, Y. Azar, and A. Epstein. The price of routing unsplittable flow. *STOC*, pages 57–66, 2005.

[2] B. Awerbuch and R. Kleinberg. Adaptive routing with end-to-end feedback: Distributed learning and geometric approaches. In *STOC*, 2004.

[3] M. Babaioff, R. Kleinberg, and C. H. Papadimitriou. Congestion games with malicious players. In *EC*, 2007.

[4] X. Chen and X. Deng. Settling the complexity of 2-player nash-equilibrium. In *FOCS*, 2006.

[5] S. Chien and A. Sinclair. Convergence to Approximate Nash Equilibria in Congestion Games. *SODA*, 2007.

[6] G. Christodoulou and E. Koutsoupias. On the price of anarchy and stability of correlated equilibria of linear congestion games. *ESA*, pages 59–70, 2005.

[7] G. Christodoulou and E. Koutsoupias. The price of anarchy of finite congestion games. *STOC*, pages 67–73, 2005.

[8] A. Czumaj and B. Vöcking. Tight bounds on worse case equilibria. In *SODA*, pages 413–420, 2002.

[9] A. Czumaj and B. Vöcking. Tight bounds for worst-case equilibria. *ACM Transactions on Algorithms*, 3(1), 2007.

[10] V. Dani and T. P. Hayes. Robbing the bandit: Less regret in online geometric optimization against an adaptive adversary. In *SODA*, 2006.

[11] A. Fabrikant, A. Luthra, E. Maneva, C. H. Papadimitriou, and S. Shenker. On a network creation game. In *PODC*, pages 347–351, 2003.

[12] D. P. Foster and R. V. Vohra. Calibrated learning and correlated equilibrium. *Games and Economic Behavior*, 1997.

[13] J. J. Gabszewicz and J.-F. Thisse. Location. In R. Aumann and S. Hart, editors, *Handbook of Game Theory with Economic Applications*, volume 1, chapter 9. Elsevier Science Publishers (North-Holland), 1992.

[14] M. Goemans, L. Li, V. Mirrokni, and M. Thottan. Market sharing games applied to content distribution in ad hoc networks. *Selected Areas in Communications, IEEE Journal on*, 24(5):1020–1033, 2006.

[15] M. Goemans, V. Mirrokni, and A. Vetta. Sink equilibria and convergence. In *FOCS*, pages 142–154, 2005.

[16] J. Hannan. Approximation to Bayes risk in repeated play. In M. Dresher, A. Tucker, and P. Wolfe, editors, *Contributions to the Theory of Games*, volume III, pages 97–139. 1957.

[17] S. Hart and A. Mas-Colell. A simple adaptive procedure leading to correlated equilibrium. *Econometrica*, 2000.

[18] H. Hotelling. Stability in competition. *The Economic Journal*, 39(153):41–57, 1929.

[19] S. Kakade, A. T. Kalai, and K. Ligett. Playing games with approximation algorithms. In *STOC*, 2007.

[20] A. Kalai and S. Vempala. Efficient algorithms for on-line optimization. In *COLT*, pages 26–40, 2003.

[21] M. Kilkenny and J. Thisse. Economics of Location: A Selective Survey. *Computers and Operations Research*, 26(14):1369–1394, 1999.

[22] R. Kleinberg. Anytime algorithms for multi-armed bandit problems. In *SODA*, pages 928–936, 2006.

[23] E. Koutsoupias, M. Mavronikolas, and P. Spirakis. Approximate equilibria and ball fusion. *Theory of Computing Systems*, 36(6):683–693, 2003.

[24] E. Koutsoupias and C. H. Papadimitriou. Worst-case equilibria. In *STACS*, pages 404–413, 1999.

[25] N. Littlestone and M. K. Warmuth. The weighted majority algorithm. *Information and Computation*, 108(2):212–261, 1994.

[26] H. McMahan and A. Blum. Online geometric optimization in the bandit setting against an adaptive adversary. In *COLT*, 2004.

[27] V. S. Mirrokni and A. Vetta. Convergence issues in competitive games. In *APPROX-RANDOM*, 2004.

[28] T. Roughgarden and E. Tardos. How bad is selfish routing? *Journal of the ACM*, 49(2):236–259, 2002.

[29] S. Suri, C. Toth, and Y. Zhou. Selfish Load Balancing and Atomic Congestion Games. *Algorithmica*, 47(1):79–96, 2007.

[30] A. Vetta. Nash equilibria in competitive societies, with applications to facility location, traffic routing and auctions. In *FOCS*, page 416, 2002.

[31] M. Zinkevich. Online convex programming and generalized infinitesimal gradient ascent. In *ICML*, pages 928–936, 2003.

A Proofs

of Theorem 4.3. Suppose all players use low regret strategies, so that for any player i,

$$T\epsilon + \sum_{t=1}^{T} \bar{\alpha}_i(S^t) \geq \sum_{t=1}^{T} \bar{\alpha}_i(S^t \oplus \sigma_i).$$

Expanding terms, we can rewrite this as

$$\begin{aligned} T\epsilon + \sum_{t:s_i^t=\sigma_i} \bar{\alpha}_i(S^t) + \sum_{t:s_i^t\neq\sigma_i} \bar{\alpha}_i(S^t) & \\ \geq \sum_{t:s_i^t=\sigma_i} \bar{\alpha}_i(S^t \oplus \sigma_i) + \sum_{t:s_i^t\neq\sigma_i} \bar{\alpha}_i(S^t \oplus \sigma_i). & \end{aligned}$$

We note that when $s_i^t = \sigma_i$, $\bar{\alpha}_i(S^t) = \bar{\alpha}_i(S^t \oplus \sigma_i)$, so this yields

$$\epsilon T + \sum_{t:s_i^t\neq\sigma_i} \bar{\alpha}_i(S^t) \geq \sum_{t:s_i^t\neq\sigma_i} \bar{\alpha}_i(S^t \oplus \sigma_i).$$

Summing over all players, we get

$$\begin{aligned} k\epsilon T + \sum_{i=1}^{k} \sum_{t:s_i^t\neq\sigma_i} \bar{\alpha}_i(S^t) &\geq \sum_{i=1}^{k} \sum_{t:s_i^t\neq\sigma_i} \bar{\alpha}_i(S^t \oplus \sigma_i) \\ &\geq \sum_{i=1}^{k} \sum_{t:s_i^t\neq\sigma_i} \bar{\gamma}'_{\sigma_i}(S^t \oplus \emptyset_i), \end{aligned}$$

where the second equation holds by assumption 1. Now note that

$$\begin{aligned}
\sum_{t=1}^{T}\bar{\gamma}(S^t) &\geq \sum_{t=1}^{T}\sum_{i=1}^{k}\bar{\alpha}_i(S^t) \\
&= \sum_{t=1}^{T}\sum_{i:\sigma_i=s_i^t}\bar{\alpha}_i(S^t) + \sum_{t=1}^{T}\sum_{i:\sigma_i\neq s_i^t}\bar{\alpha}_i(S^t) \\
&\geq \sum_{t=1}^{T}\sum_{i:\sigma_i\neq s_i^t}\bar{\alpha}_i(S^t) + \sum_{t=1}^{T}\sum_{i:\sigma_i=s_i^t}\bar{\gamma}'_{s_i^t}(S^t\oplus\emptyset_i) \\
&= \sum_{i=1}^{k}\sum_{t:\sigma_i\neq s_i^t}\bar{\alpha}_i(S^t) + \sum_{t=1}^{T}\sum_{i:\sigma_i=s_i^t}\bar{\gamma}'_{s_i^t}(S^t\oplus\emptyset_i),
\end{aligned}$$

where the third line holds by assumption 2 and the fourth line is a reordering of the summations. This gives us

$$\begin{aligned}
\sum_{i=1}^{k}\sum_{t:\sigma_i\neq s_i^t}\bar{\gamma}'_{\sigma_i^t}(S^t\oplus\emptyset_i) &\leq T\epsilon k + \sum_{i=1}^{k}\sum_{t:\sigma_i\neq s_i^t}\bar{\alpha}_i(S^t) \\
&\leq T\epsilon k + \sum_{t=1}^{T}\bar{\gamma}(S^t) - \sum_{t=1}^{T}\sum_{i:\sigma_i=s_i^t}\bar{\gamma}'_{s_i^t}(S^t\oplus\emptyset_i).
\end{aligned}$$

We use the following lemma proved by Vetta [30]:

Lemma A.1. *If $\Omega = \{\sigma_1,\ldots,\sigma_k\}$ is a strategy profile optimizing the social utility function γ, then for any strategy profile S*

$$\bar{\gamma}(\Omega) \leq \bar{\gamma}(S) + \sum_{i:\sigma_i\neq s_i}\bar{\gamma}'_{\sigma_i}(S\oplus\emptyset_i) - \sum_{i:\sigma_i\neq s_i}\bar{\gamma}'_{s_i}(\Omega\cup(S\oplus\emptyset_i\oplus\ldots\oplus\emptyset_k)).$$

From Lemma A.1, for any sequence of plays $S_1,\ldots,S^t$,

$$\begin{aligned}
T\bar{\gamma}(\Omega) \leq \sum_{t=1}^{T}\Bigg(&\bar{\gamma}(S^t) + \sum_{i:\sigma_i\neq s_i^t}\bar{\gamma}'_{\sigma_i}(S^t\oplus\emptyset_i) \\
&- \sum_{i:\sigma_i\neq s_i^t}\bar{\gamma}'_{s_i^t}(\Omega\cup(S^t\oplus\emptyset_i\oplus\ldots\oplus\emptyset_k))\Bigg).
\end{aligned}$$

Substituting, we get

$$\begin{aligned}
T\cdot\mathbf{OPT} \leq \sum_{i=1}^{T}\Bigg(&2\bar{\gamma}(S^t) + \epsilon k - \sum_{i:\sigma_i=s_i^t}\bar{\gamma}'_{s_i^t}(S^t\oplus\emptyset_i) \\
&- \sum_{i:\sigma_i\neq s_i^t}\bar{\gamma}'_{s_i^t}(\Omega\cup(S^t\oplus\emptyset_i\oplus\ldots\oplus\emptyset_k))\Bigg),
\end{aligned}$$

which completes the proof. □

of Theorem 5.1. Let $\Omega = \{\sigma_1,\ldots,\sigma_k\}$ be a strategy profile optimizing the social utility function so that $\bar{\gamma}(\Omega) = \mathbf{OPT}$ By the assumption of regret minimization, each player's time average cost is no more than the cost of her best

fixed action in hindsight. In particular, it is no more than if she had played her part in the optimal strategy on every timestep: For all i,

$$\begin{aligned}\sum_{t=1}^{T}\bar{\alpha}_i(S^t) &= \sum_{t=1}^{T}\sum_{e\in s_i^t} c_e l_e^t + b_e \\ &\leq \sum_{t=1}^{T}\bar{\alpha}_i(S^t \oplus \sigma_i) \\ &\leq \sum_{t=1}^{T}\sum_{e\in\sigma_i} c_e(l_e^t+1) + b_e.\end{aligned}$$

Summing over each player and rearranging the sum:

$$\begin{aligned}\sum_{t=1}^{T}\sum_{e\in E}\sum_{i \text{ s.t. } e\in s_i^t} c_e l_e^t + b_e &\leq \sum_{t=1}^{T}\sum_{e\in E}\sum_{i \text{ s.t. } e\in\sigma_i} c_e(l_e^t+1) + b_e \\ &= \sum_{t=1}^{T}\sum_{e\in E} c_e l_e^t l_e^* + c_e l_e^* + b_e l_e^*.\end{aligned}$$

We now use a lemma also used by Awerbuch et al. [1]:

Lemma A.2. *For $i, j > 0$ integers:*

1. $ij = \frac{1}{3}j^2 + \frac{3}{4}i^2 - \frac{1}{3}(j - \frac{3}{2}i)^2$
2. $\frac{9}{8}i^2 + \frac{3}{2}i - \frac{1}{2}(j - \frac{3}{2}i)^2 \leq \frac{5}{2}i^2$

We can apply part 1 of the lemma to get

$$\begin{aligned}&\sum_{t=1}^{T}\sum_{e\in E}(c_e l_e^t + b_e)l_e^t \\ &\qquad \leq \sum_{t=1}^{T}\sum_{e\in E} c_e\left(\frac{1}{3}(l_e^t)^2 + \frac{3}{4}(l_e^*)^2 - \frac{1}{3}(l_e^t - \frac{3}{2}l_e^*)^2 + l_e^*\right) + b_e l_e^*.\end{aligned}$$

This is equivalent to

$$\begin{aligned}&\sum_{t=1}^{T}\sum_{e\in E}(c_e l_e^t + \frac{3}{2}b_e)l_e^t \\ &\qquad \leq \sum_{t=1}^{T}\sum_{e\in E} c_e\left(\frac{9}{8}(l_e^*)^2 + \frac{3}{2}l_e^* - \frac{1}{2}(l_e^t - \frac{3}{2}l_e^*)^2\right) + \frac{3}{2}b_e l_e^*.\end{aligned}$$

This allows us to apply property 2 of the lemma to obtain

$$\begin{aligned}\sum_{t=1}^{T}\sum_{e\in E}(c_e l_e^t + b_e)l_e^t &\leq \sum_{t=1}^{T}\sum_{e\in E}\frac{5}{2}c_e(l_e^*)^2 + \frac{3}{2}b_e l_e^* \\ &\leq \frac{5}{2}\sum_{t=1}^{T}\sum_{e\in E}(c_e l_e^* + b_e)l_e^*,\end{aligned}$$

which proves the claim. □

of Theorem 5.3. By the no-regret property we have for each player i:

$$\sum_{t=1}^{T}\sum_{e\in a_i^t} f_e(l_e^t) \le \sum_{t=1}^{T}\sum_{e\in\sigma_i} f_e(l_e^t+1).$$

We may sum over each player:

$$\sum_{t=1}^{T}\sum_{e\in E}\sum_{i \text{ s.t. } e\in a_i^t} f_e(l_e^t) \le \sum_{t=1}^{T}\sum_{e\in E}\sum_{i \text{ s.t. } e\in\sigma_i} f_e(l_e^t+1)$$

and rearrange the sums:

$$\sum_{t=1}^{T}\sum_{e\in E} f_e(l_e^t)l_e^t \le \sum_{t=1}^{T}\sum_{e\in E} f_e(l_e^t+1)l_e^*.$$

We now apply a lemma used by Christodoulou and Koutsoupias [7]:

Lemma A.3. *For $f(x)$ a polynomial with non-negative coefficients of degree d, and for every $x, y \ge 0$:*

$$y\cdot f(x+1) \le \frac{x\cdot f(x)}{2} + \frac{C_0(d)\cdot y\cdot f(y)}{2},$$

where $C_0(d) = p^{p^{1-o(1)}}$.

Applying the lemma, we get

$$\begin{aligned}\sum_{t=1}^{T}\sum_{e\in E} f_e(l_e^t)l_e^t &\le \sum_{t=1}^{T}\sum_{e\in E} f_e(l_e^t+1)l_e^* \\ &\le \sum_{t=1}^{T}\Big(\sum_{e\in E}\frac{l_e^t f(l_e^t)}{2} + \sum_{e\in E}\frac{C_0(d)l_e^* f(l_e^*)}{2}\Big).\end{aligned}$$

Rearranging, we then get

$$\sum_{t=1}^{T}\sum_{e\in E} f_e(l_e^t)l_e^t \le C_0(d)\sum_{t=1}^{T}\sum_{e\in E} f(l_e^*)l_e^*,$$

which completes the proof. □

of Theorem 5.4. Denote by q_i the expected probability that player i is on the maximally loaded machine (breaking ties between equally loaded machines at random). Note that the expected social cost is then $\bar{\gamma}(S) = \sum_{i=1}^{k} q_i w_i$. By the regret-minimizing property, for all players, $\frac{1}{T}\sum_{t=1}^{T}\bar{\alpha}_i(S^t) \le w_i + \epsilon + \frac{1}{T}\sum_{t=1}^{T}\frac{\sum_{h\,:\,h\neq i} w_h}{2}$.

Define p_{ij} to be the expected probability that player i selects machine j; c_{ih} is the expected probability that players i and h select the same machine. Then for any fixed i,

$$\begin{aligned}\sum_{h:h\neq i}(q_i+q_h)w_h &\le \sum_{h:h\neq i}(1+c_{ih})w_h \\ &\le \sum_{h:h\neq i} w_h + \sum_{h:h\neq i} c_{ih}w_h \\ &= \sum_{h:h\neq i} w_h + \sum_{h:h\neq i}(p_{i1}p_{h1}w_h + p_{i2}p_{h2}w_h).\end{aligned}$$

Note that for any player i, regardless of her strategy, her cost is

$$\bar{\alpha}_i(S) = w_i + p_{i1}\sum_{h:h\neq i} p_{h1}w_h + p_{i2}\sum_{h:h\neq i} p_{h2}w_h$$

by definition. This relationship is essentially Lemma 1 of [24]; however they only note that it holds for Nash equilibrium strategies. This gives us $\sum_{h:h\neq i}(q_i+q_h)w_h \leq \sum_{h:h\neq i} w_h + \bar{\alpha}_i(S) - w_i$. Averaging over time, this is

$$\frac{1}{T}\sum_{t=1}^{T}\sum_{h:h\neq i}(q_i^t+q_h^t)w_h \leq \frac{1}{T}\sum_{t=1}^{T}\sum_{h:h\neq i} w_h + \frac{1}{T}\sum_{t=1}^{T}\bar{\alpha}_i(S^t) - w_i.$$

Using the fact that player i obtains low regret, we then have

$$\frac{1}{T}\sum_{t=1}^{T}\sum_{h:h\neq i}(q_i^t+q_h^t)w_h \leq \frac{1}{T}\sum_{i=1}^{T}\sum_{h:h\neq i} w_h + \epsilon + \frac{1}{T}\sum_{t=1}^{T}\frac{\sum_{h\,:\,h\neq i} w_h}{2}.$$

Rearranging, this yields for any fixed i

$$\begin{aligned}
&\frac{1}{T}\sum_{t=1}^{T}\bar{\gamma}(S^t) \\
&\quad= \frac{1}{T}\sum_{t=1}^{T}\sum_{h=1}^{k} q_h^t w_h \\
&\quad\leq \frac{1}{T}\sum_{t=1}^{T}\left(\frac{3}{2}\sum_{h:h\neq i} w_h + q_i^t w_i - \sum_{h\neq i} q_i^t w_h\right) + \epsilon \\
&\quad= \frac{1}{T}\sum_{t=1}^{T}\left(\frac{3}{2}\sum_{h=1}^{k} w_h - \frac{3w_i}{2} + q_i^t w_i - q_i^t\sum_{h=1}^{k} w_h + q_i^t w_i\right) + \epsilon \\
&\quad= \frac{1}{T}\sum_{t=1}^{T}\left(\left(\frac{3}{2} - q_i^t\right)\sum_{h=1}^{k} w_h + \left(2q_i^t - \frac{3}{2}\right)w_i\right) + \epsilon.
\end{aligned}$$

Note that $\mathbf{OPT} \geq \max\{\frac{1}{2}\sum_{h=1}^{k} w_h, w_i\}$ for any i. If for all agents i, $\frac{1}{T}\sum_{t=1}^{T} q_i^t \leq \frac{3}{4}$, then

$$\begin{aligned}
\frac{1}{T}\sum_{t=1}^{T}\bar{\gamma}(S^t) &= \frac{1}{T}\sum_{t=1}^{T}\sum_{h=1}^{k} q_h^t w_h \\
&= \sum_{h}\left(w_h\frac{1}{T}\sum_{t} q_h^t\right) \\
&\leq \frac{3}{4}\sum_{h} w_h \\
&\leq \frac{3}{2}\mathbf{OPT}\,.
\end{aligned}$$

Otherwise, there exists some agent i such that $\frac{1}{T}\sum_{t=1}^{T} q_i^t > \frac{3}{4}$ and thus

$$\begin{aligned}
&\frac{1}{T}\sum_{t=1}^{T}\bar{\gamma}(S^t) \\
&\quad\leq 2\mathbf{OPT}\,\frac{1}{T}\sum_{t}(\frac{3}{2} - q_i^t) + \mathbf{OPT}\,\frac{1}{T}\sum_{t}(2q_i^t - \frac{3}{2}) + \epsilon \\
&\quad= \frac{3}{2}\mathbf{OPT}\, + \epsilon,
\end{aligned}$$

as desired. □

of Theorem 5.5. Consider n parallel links $1, \ldots, n$, and n players all with unit weights $w_i = 1$. Clearly, $\mathbf{OPT} = 1$. Define a sequence of plays $A^1, \ldots, A^T$ as follows: Divide the players into $2\sqrt{n}$ groups $G_0, \ldots, G_{2\sqrt{n}-1}$, each of size $\sqrt{n}/2$. At time t, all players in $G_{(t \mod 2\sqrt{n})}$ play on link 1, and all other players play over links $2+(t \mod n-1), 2+(t+1 \mod n-1), \ldots, 2+(t+n-\sqrt{n}/2-1 \mod n-1)$ so that there is exactly one player on each link (ordering may be arbitrary). Then each player experiences average latency $\frac{1}{T}\sum_{t=1}^{T} \alpha_i(A^t) = \frac{1}{2\sqrt{n}} \cdot \frac{\sqrt{n}}{2} + \frac{2\sqrt{n}-1}{2\sqrt{n}} \cdot 1 = \frac{5}{4} - \frac{1}{2\sqrt{n}}$. Consider the latency experienced by player i if she were to play at any fixed node. Given the sequence of plays described above, every node $v \geq 2$ is occupied by some player $h \neq i$ on an $(n - \sqrt{n}/2 - 1)/(n-1)$ fraction of time steps. Since player i always pays for her own weight, she expects to experience latency $2 \cdot \frac{n-\frac{\sqrt{n}}{2}-1}{n-1} + 1 \cdot \frac{\sqrt{n}}{2(n-1)} = 2 - \frac{\sqrt{n}}{2(n-1)}$. Therefore, for sufficiently large n, all players experience negative regret. Nevertheless, at every time step, the maximum latency is $\Omega(\sqrt{n})$. □

of Theorem 5.6. By the no regret property, for each player i,

$$\sum_{t=1}^{T} \bar{\alpha}_i(A^t) = \sum_{t=1}^{T} \frac{l_{a_i^t}}{\pi_{a_i^t}} \leq \sum_{t=1}^{T} \bar{\alpha}_i(A^t \oplus \sigma_i) \leq \sum_{t=1}^{T} \frac{l_{\sigma_i}^t + 1}{\pi_{\sigma_i}}.$$

Summing over all players and reordering the sum, we get

$$\begin{aligned} \sum_{t=1}^{T} \sum_{e \in E} \frac{(l_e^t)^2}{\pi_e} &\leq \sum_{t=1}^{T} \sum_{e \in E} \frac{(l_e^t + 1) \cdot l_e^*}{\pi_e} \\ &\leq \sum_{t=1}^{T} \sum_{e \in E} \frac{1}{\pi_e} \left(\frac{(l_e^t)^2 + (l_e^*)^2}{2} + l_e^* \right), \end{aligned}$$

where the second inequality follows from the fact that $a \cdot b \leq \frac{a^2+b^2}{2}$. Subtracting, we get

$$\begin{aligned} \frac{1}{2} \sum_{t=1}^{T} \sum_{e \in E} \frac{(l_e^t)^2}{\pi_e} &\leq \sum_{t=1}^{T} \sum_{e \in E} \frac{1}{\pi_e} \left(\frac{(l_e^*)^2}{2} + l_e^* \right) \\ \sum_{t=1}^{T} \sum_{e \in E} \frac{(l_e^t)^2}{\pi_e} &\leq \sum_{t=1}^{T} \sum_{e \in E} \frac{(l_e^*)^2 + 2l_e^*}{\pi_e}. \end{aligned}$$

Combining these inequalities, we can bound the price of total anarchy:

$$\begin{aligned} \frac{\sum_{t=1}^{T} \sum_{e \in E} \frac{(l_e^t)^2}{\pi_e}}{\sum_{t=1}^{T} \sum_{e \in E} \frac{(l_e^*)^2}{\pi_e}} &\leq \frac{\sum_{t=1}^{T} \sum_{e \in E} \frac{(l_e^*)^2 + 2l_e^*}{\pi_e}}{\sum_{t=1}^{T} \sum_{e \in E} \frac{(l_e^*)^2}{\pi_e}} \\ &= 1 + 2 \frac{\sum_{t=1}^{T} \sum_{e \in E} \frac{l_e^*}{\pi_e}}{\sum_{t=1}^{T} \sum_{e \in E} \frac{(l_e^*)^2}{\pi_e}} \\ &\leq 1 + 2 \sum_{t=1}^{T} \frac{\sum_{e \in E} \frac{l_e^*}{\pi_e}}{\sum_{e \in E} \frac{(l_e^*)^2}{\pi_e}}. \end{aligned}$$

We then use the following technical lemma of Suri et al. [29]

Lemma A.4. *Let n, k be positive integers and $l_e \geq 0, \pi_e > 0$ be reals such that $\sum_{e \in E} l_e = k$. Then*

$$\frac{\sum_{e \in E} l_e/\pi_e}{\sum_{e \in E} l_e^2/\pi_e} \leq (1 + \sqrt{\max_{1 \leq i,j \leq n} \frac{\pi_i}{\pi_j}}) \frac{n}{2k}$$

.

This gives us

$$\frac{1}{2}\sum_{t=1}^{T}\sum_{e\in E}\frac{(l_e^t)^2}{\pi_e} \quad \leq \quad 1+2\left(1+\sqrt{\max_{1\leq i,j\leq n}\frac{\pi_i}{\sigma_j}}\right)\frac{n}{2k}.$$

This is $1+o(1)$ in k when $k >> n$, which completes the proof. $\square$