THEORY OF COMPUTING, Volume 6 (2010), pp. 179–199
www.theoryofcomputing.org

Routing Without Regret: On Convergence to Nash Equilibria of Regret-Minimizing Algorithms in Routing Games*

Avrim Blum† Eyal Even-Dar Katrina Ligett‡

Received: November 10, 2009; published: September 15, 2010.

Abstract: There has been substantial work developing simple, efficient *regret-minimizing* algorithms for a wide class of repeated decision-making problems including online routing. These are adaptive strategies an individual can use that give strong guarantees on performance even in adversarially-changing environments. There has also been substantial work on analyzing properties of Nash equilibria in routing games. In this paper, we consider the question: if each player in a routing game uses a regret-minimizing strategy, will behavior converge to a Nash equilibrium? In general games the answer to this question is known to be *no* in a strong sense, but routing games have substantially more structure.

In this paper we show that in the Wardrop setting of multicommodity flow and infinitesimal agents, behavior will approach Nash equilibrium (on most days, the cost of the flow

*A preliminary version of these results appeared in the Proceedings of the 25th Annual ACM Symposium on Principles of Distributed Computing, July 2006.

†Supported in part by the National Science Foundation under grants IIS-0121678, CCR-0122581, and CCF-0514922.

‡This material is based upon work supported in part by an AT&T Labs Graduate Fellowship, an NSF Graduate Research Fellowship, and by the National Science Foundation under Grant #0937060 to the Computing Research Association for the CIFellows program.

ACM Classification: G.2.2, J.4

AMS Classification: 68Q32, 68T05, 68W40, 91A13, 91A20, 91A43

Key words and phrases: regret minimization, Nash equilibria, routing games

DOI: 10.4086/toc.2010.v006a008

will be close to the cost of the cheapest paths possible given that flow) at a rate that depends polynomially on the players' regret bounds and the maximum slope of any latency function. We also show that Price of Anarchy results may be applied to these approximate equilibria, and also consider the finite-size (non-infinitesimal) load-balancing model of Azar (1998). Our nonatomic results also apply to the more general class of congestion games.

1 Introduction

There has been substantial work in learning theory and game theory on *regret-minimizing* algorithms for problems of repeated decision-making. Regret-minimizing algorithms have the property that in any online, repeated setting, their average loss per time step approaches that of the best fixed strategy in hindsight (or better) over time. Moreover, the convergence rates are quite good: in Hannan's original algorithm [23], the number of time steps needed to achieve a gap of ε with respect to the best fixed strategy in hindsight—the "per time step regret"—is linear in the size of the game N. This was reduced to logarithmic in N in more recent exponential-weighting algorithms for this problem [28, 8, 20] (also known as the problem of "combining expert advice"). Most recently, a number of algorithms have been developed for achieving such guarantees in a *computationally efficient* manner in many settings where the number of possible actions N is exponential in the natural description-length of the problem [25, 38, 39].

One specific setting where efficient regret-minimizing algorithms can be applied is online routing. Given a graph $G=(V,E)$ and two distinguished nodes v_{start} and v_{end}, the game for an individual player is defined as follows. At each time step t, the player's algorithm chooses a path P_t from v_{start} to v_{end}, and a set of edge costs $\{c_e^t\}_{e\in E}$ is simultaneously determined. (Many regret-minimizing algorithms can handle arbitrary adversarially chosen edge costs, but in our application, the edge costs are determined by the other players' routing decisions.) The edge costs are then revealed and the player pays the resulting cost of the path it chose. Even though the number of possible paths can be exponential in the size of a graph, regret-minimizing algorithms exist (e. g., [25, 39]) with computation costs and convergence rates to the cost of the best fixed path in hindsight that are both polynomial in the size of the graph and in the maximum edge cost. Moreover, a number of extensions [3, 29] have shown how these algorithms can be applied even to the "bandit" setting where only the cost of edges actually traversed (or even just the total cost of P_t) is revealed to the algorithm at the end of each time step t.

Along a very different line of inquiry, there has also been much recent work on the *Price of Anarchy* in games. Koutsoupias and Papadimitriou [27] defined the *Price of Anarchy*, which is the ratio of the cost of an optimal global objective function to the cost of the worst Nash equilibrium. Many subsequent results have studied the Price of Anarchy in a wide range of computational problems from job scheduling to facility location to network creation games, and especially to problems of routing in the Wardrop model such as that described above, where the cost of an edge is a function of the amount of traffic using that edge, and the individual players are infinitesimal [10, 11, 27, 34, 15]. Such work implicitly assumes that selfish individual behavior results in Nash equilibria.

Our Contribution We consider the following question: if all players in a routing game use regret-minimizing algorithms to choose their paths each day, what can we say about the overall behavior of the system? In particular, the regret-minimizing property—achieving performance close to that of the best fixed action in hindsight (also called Hannan Consistency)—can be viewed as a natural *condition* of well-reasoned self-interested behavior over time. That is, given that simple, computationally feasible regret-minimizing algorithms exist for a repeated situation faced by an agent, one might posit that the agent's self-interested behavior would be regret-minimizing. Given that, we ask, if all players are adapting their behavior in a regret-minimizing manner, can we say that the system as a whole will approach equilibrium? Our main result is that in the Wardrop network routing setting of multicommodity flow and infinitesimal agents, the flows will approach equilibrium in the sense that a $1-\varepsilon$ fraction of the daily flows will have the property that at most an ε fraction of the agents in them have more than an ε incentive to deviate from their chosen path, where ε approaches 0 at a rate that depends polynomially on the size of the graph, the regret-bounds of the algorithms, and the maximum slope of any latency function.[1] Moreover, we show that the dependence on slope (the one new parameter) is necessary.

One further consequence, as we show in Section 5, is that if all players use regret-minimizing strategies, there is no other strategy that would significantly improve any player's cost on more than a small fraction of time steps. In addition, we give stronger results for special cases such as the case of n parallel links and also consider the finite-size (atomic) load-balancing model of Azar [4]. Our results for infinitesimal players also hold for a more general class of games called congestion games, although computationally efficient regret-minimizing algorithms do not always exist for the most general of these games.

One way our result can be viewed is as follows. Regret-minimizing algorithms are very compelling from the point of view of individuals: if you use a regret-minimizing algorithm to drive to work each day, you will get a good guarantee on your performance no matter what is causing congestion (other drivers, road construction, or unpredictable events). But it would be a shame if, were everyone to use such an algorithm, this produced globally unstable behavior. Our results imply that in the Wardrop network routing model, so long as edge latencies have bounded slope, we can view equilibria as not just a stable steady-state or the result of adaptive procedures specifically designed to find them, but in fact as the natural result of individual selfishly adaptive behavior by agents that do not necessarily know (or care) what policies other agents are using. Moreover, our results do not in fact require that users follow strategies that are regret-minimizing over all possible cost scenarios, as long as their behavior satisfies the regret-minimization property over the sequence of flows actually observed.

[1] A more traditional notion of approximate Nash equilibrium requires that *no* player will have more than ε incentive to deviate from her strategy. However, one cannot hope to achieve such a guarantee using arbitrary regret-minimizing algorithms, since such algorithms allow players to occasionally try bad paths, and in fact such experimentation is even necessary in bandit settings. For the same reason, one cannot hope that *all* days will be approximate-Nash. Finally, our guarantee may make one worry that some users could always do badly, falling in the ε minority on every day, but as we discuss in Section 5, the regret-minimization property can be used in our setting to further show that no player experiences many days in which her expected cost is much worse than the best path available on that day.

1.1 Regret and Nash equilibria

At first glance, a result of this form seems that it should be obvious given that a Nash equilibrium is precisely a set of strategies (pure or mixed) that are all regret-minimizing with respect to each other. Thus if the learning algorithms settle at all, they will have to settle at a Nash equilibrium. In fact, for *zero-sum* games, regret-minimizing algorithms when played against each other will approach a minimax optimal solution [21]. However, it is known that even in small 2-player *general-sum* games, regret-minimizing algorithms need not approach a Nash equilibrium and can instead cycle, achieving performance substantially worse than any Nash equilibrium for all players. Indeed simple examples are known where standard algorithms will have this property with arbitrarily high probability [40].

1.2 Regret and correlated equilibria

It is known that certain algorithms such as that of Hart and Mas-Colell [24], as well as any algorithms satisfying the stronger property of "no internal regret" [19], have the property that the empirical distribution of play approaches the set of *correlated* equilibria. On the positive side, such results are extremely general, apply to nearly any game including routing, and do not require any bound on the slopes of edge latencies. However, such results do *not* imply that the daily flows themselves (or even the time-average flow) are at all close to equilibrium. It could well be that on each day, a substantial fraction of the players experience latency substantially greater than the best path given the flow (and we give a specific example of how this can happen when edge-latencies have unbounded slope in Section 2.4). In addition, although Neyman [32] does show that the only correlated equilibrium in potential games with strictly concave potential functions is the unique Nash equilibrium, there is no known efficient implementation for internal regret minimization for routing problems.

1.3 Related work

Sandholm [36] considers convergence in potential games (which include routing games), and shows that a very broad class of evolutionary dynamics is guaranteed to converge to Nash equilibrium.[2] Fischer and Vöcking [17] consider a specific adaptive dynamics (a particular functional form in which flow might naturally change over time) in the context of selfish routing and prove results about convergence of this dynamics to an approximately stable configuration. In more recent work, they study the convergence of a class of routing policies under a specific model of stale information [18]. Most recently, Fischer, Raecke, and Vöcking [16] gave a distributed procedure with especially good convergence properties. The key difference between that work and ours is that those results consider specific adaptive strategies designed to quickly approach equilibrium. In contrast, we are interested in showing convergence for *any* algorithm satisfying the regret-minimization property. That is, even if the players are using many different strategies, without necessarily knowing or caring about what strategies others are using, then so long as all are regret-minimizing, we show they achieve convergence. In addition, because efficient regret-minimizing algorithms exist even in the bandit setting where each agent gets feedback only about its own actions [3, 29], our results can apply to scenarios in which agents adapt their behavior based on only very limited information and there is no communication at all between different agents.

[2] Note that such dynamics do not include general regret-minimizing dynamics.

Convergence time to Nash equilibrium in load balancing has also been studied. Earlier work studied convergence time using potential functions, with the limitation that only one player is allowed to move in each time step; the convergence times derived depended on the appropriate potential functions of the exact model [30, 12]. The work of Goldberg [22] studied a randomized model in which each user can select a random delay over continuous time. This implies that only one user tries to reroute at each specific time; therefore the setting was similar to that mentioned above. Even-Dar and Mansour [13] considered a model where many users are allowed to move concurrently, and derived a logarithmic convergence rate for users following a centrally-moderated greedy algorithm. Most recently, Berenbrink et al. [6] showed weaker convergence results for a specific distributed protocol. To summarize, previous work studied the convergence time to pure Nash equilibria in situations with a centralized mechanism or specific protocol. In contrast, we present fast convergence results for approximate Nash equilibria in a non-centralized setting, and our only assumption about the player strategies is that they are all regret-minimizing.

Subsequent work Since the initial publication of these results, a number of publications have studied related questions. Blum et al. [7] and Roughgarden [33] explore the outcomes of regret-minimizing behavior in a variety of classes of games; they are able to show Price of Anarchy style bounds on the social cost, but do not prove convergence results. Kleinberg et al. [26] study agents in atomic congestion games employing a *particular* class of regret-minimization algorithms and show that in many cases, the additional assumptions on the player algorithms allow convergence to *pure* Nash equilibria. Even-Dar et al. [14] demonstrate convergence of general regret-minimizing algorithms to Nash equilibria in a general class of games they call "socially-concave" games; our games do not fit in their model, because their work considers a finite number of players.

Awerbuch et al. [2] show that a certain type of best response dynamics converges quickly to approximate Nash equilibria in congestion games. General regret-minimizing dynamics are much more complex than the dynamics they study, and, we believe, are better motivated from an individual's perspective in potentially rapidly-changing environments.

2 Preliminaries

2.1 Nonatomic congestion games

Let E be a finite ground set of elements (we refer to them as *edges*). There are k *player types* $1,2,\ldots,k$, and each player type i has an associated set of feasible paths $\mathcal{P}_i$, where $\mathcal{P}_i$ is a set of subsets of E. Elements of $\mathcal{P}_i$ are called *paths* or *strategies*. For example, player type i might correspond to players who want to travel from node u_i to node v_i in some underlying graph G, and $\mathcal{P}_i$ might be the set of all u_i-v_i paths. The continuum A_i of agents of type i is represented by the interval $[0,a_i]$, endowed with Lebesgue measure. We restrict $\sum_{i=1}^{k} a_i = 1$, so there is a total of one unit of flow. Each edge $e \in E$ has an associated traffic-dependent, non-negative, continuous, non-decreasing *latency* function ℓ_e. A *nonatomic congestion game* is defined by $(E,\ell,\mathcal{P},A)$.

A *flow* specifies a path for each player: $f_i : A_i \to \mathcal{Q}_i$ where $\mathcal{Q}_i$ is the set of $0/1$ vectors in $\mathcal{P}_i$ with exactly one 1. We write $f = (\int_{A_1} f_1, \ldots, \int_{A_k} f_k)$, where by $\int_{A_i} f_i$ we mean $(\int_{A_i}(f_i)^1, \int_{A_i}(f_i)^2, \ldots, \int_{A_i}(f_i)^{|\mathcal{P}_i|})$.

For a path P in $\mathcal{P}_i$, we write $f_P = (f_i)^P$ to represent the weight that flow f places on path P. Thus, $\sum_{P\in\mathcal{P}_i} f_P = a_i$ for all i, and f_P is the measure of the set of players selecting path P. Each flow induces a unique flow on edges such that the flow f_e on an edge e has the property $f_e = \sum_{P:e\in P} f_P$. The latency of a path P given a flow f is $\ell_P(f) = \sum_{e\in P} \ell_e(f_e)$, i. e., the sum of the latencies of the edges in the path, given that flow. The cost incurred by a player is simply the latency of the path she plays.

We define $|E| = m$. We will assume all edge latency functions have range $[0,1]$, and the latency of a path is always between 0 and n. We use L to denote n times the maximum slope of any latency function. Let $f^1, f^2, \ldots, f^T$ denote a series of flows from time 1 up to time T. We use $\hat{f}$ to denote the time-average flow, i. e., $\hat{f}_e = (1/T)\sum_{t=1}^{T} f_e^t$.

Remark 2.1. Network routing (Wardrop) games are a special case of nonatomic congestion games, where there is an underlying graph G and players of type i have a start node u_i and a destination node v_i, and $\mathcal{P}_i$ is the set of all u_i-v_i paths.

2.2 Equilibria and social cost

A flow f is at *Wardrop equilibrium* if no user would prefer to reroute her traffic, given the existing flow. This is essentially a continuous (infinitesimal, nonatomic) analogue of the Nash equilibrium.

Definition 2.2. A flow f on game $(E, \ell, \mathcal{P}, A)$ is at equilibrium if and only if for every player type i, and paths $P_1, P_2 \in \mathcal{P}_i$ with $f_{P_1} > 0$, $\ell_{P_1}(f) \le \ell_{P_2}(f)$.

It is useful to note that in this domain, the flows at equilibrium are those for which all flow-carrying paths for a particular player type have the same latency, and this latency is minimal among all paths for players of that type. In addition, given our assumption that all latency functions are continuous and non-decreasing, one can prove the existence of equilibria:

Proposition 2.3 (Schmeidler [37], generalization of Beckman et al. [5])**.** *Every nonatomic congestion game with continuous, non-decreasing latency functions admits a flow at equilibrium.*

We define the social cost of a flow to be the average cost incurred by the players:

Definition 2.4. Define the *cost* $C(f)$ of a flow f to be $C(f) = \sum_{e\in E} \ell_e(f_e) f_e$.

In addition, for any nonatomic congestion game, there is a unique equilibrium cost:

Proposition 2.5 (Milchtaich [31], generalization of Beckman et al. [5])**.** *Distinct equilibria for a nonatomic congestion game have equal social cost.*

2.3 Regret-minimizing algorithms

Definition 2.6. Consider a series of flows $f^1, f^2, \ldots, f^T$ and a user who has experienced latencies $c^1, c^2, \ldots, c^T$ over these flows. The per-time-step *regret* of the user is the difference between her average latency and the latency of the best fixed path in hindsight for players of her type i, that is,

$$\frac{1}{T}\sum_{t=1}^{T} c^t - \min_{P\in\mathcal{P}_i} \frac{1}{T}\sum_{t=1}^{T}\sum_{e\in P} \ell_e(f_e^t)\,.$$

The concept which we refer to here as *regret* is known more generally as *external regret*, to distinguish it from other regret notions that compare with a different set of reference outcomes. An online algorithm for selecting paths at each time step t is *regret-minimizing* (has the *regret-minimizing property*) if, for any sequence of flows, the expected per-time step regret (over internal randomness in the algorithm) goes to 0 as T goes to infinity.

Here and in the rest of this paper, excluding Section 7, we consider infinitesimal users using a finite number of different algorithms; in this setting, we can get rid of the expectation. In particular, if each user is running a regret-minimizing algorithm, then the average regret over users also approaches 0. Thus, since all players have bounded per-timestep cost, applying the strong law of large numbers, we can make the following assumption:

Assumption 2.7. The series of flows $f^1, f^2, \ldots$ satisfies

$$\frac{1}{T}\sum_{t=1}^{T}\sum_{e\in E}\ell_e(f_e^t)f_e^t \le R(T) + \frac{1}{T}\sum_{i=1}^{k} a_i \min_{P\in\mathcal{P}_i}\sum_{t=1}^{T}\sum_{e\in P}\ell_e(f_e^t)$$

where $R(T) \to 0$ as $T \to \infty$. The function $R(T)$ may depend on the size of the network and its maximum possible latency. We then define T_ε as the value of T required for a particular algorithm to get $R(T) \le \varepsilon$ (this is well defined for any algorithm with non-increasing expected regret).

For example, for the case of a routing game consisting of only two nodes and m parallel edges, exponential-weighting algorithms [28, 8, 20] give $T_\varepsilon = O((\log m)/\varepsilon^2)$. For general graphs, an alternative algorithm and analysis of Kalai and Vempala yields $T_\varepsilon = O(mn(\log n)/\varepsilon^2)$ [25]. For general graphs where an agent can observe only its path cost, results of Abernathy et al. [1] achieve $T_\varepsilon = O(m^3/\varepsilon^2)$.

2.4 Approaching Nash equilibria

We now need to specify in what sense flow will be approaching a Nash equilibrium. The first notion one might consider is the L_1 distance from some true equilibrium flow. However, if some edges have nearly-flat latency functions, it is possible for a flow to have regret near 0 and yet still be far in L_1 distance from a true equilibrium flow. A second natural notion would be to say that the flow f has the property that no user has cost much more than the cheapest path given f. However, notice that the regret-minimization property allows users to occasionally take long paths, so long as they perform well on average (and in fact algorithms for the bandit problem will have exploration steps that do just that [3, 29]). So, one cannot expect that on any time step *all* users are taking cheap paths.

Instead, we require that *most* users be taking a nearly-cheapest path given f. Specifically,

Definition 2.8. A flow f is at average-ε-Nash equilibrium if the average cost under this flow is within ε of the minimum cost paths under this flow, i. e., $C(f) - \sum_{i=1}^{k} a_i \min_{P\in\mathcal{P}_i}\sum_{e\in P}\ell_e(f_e) \le \varepsilon$.

Note that Definition 2.8 implies that at most a $\sqrt{\varepsilon}$ fraction of traffic can have more than a $\sqrt{\varepsilon}$ incentive to deviate from their path, and as a result is very similar to the definition of (ε, δ)-Nash equilibria in [16] (similarly, an $r < \varepsilon$ fraction of traffic could have an ε/r incentive to deviate). We also are able to show that one can apply Price of Anarchy results to average-ε-Nash flows; we discuss this in Section 6.

We will begin by focusing on the *time-average* flow $\hat{f}$, showing that for regret-minimizing algorithms, this flow is approaching equilibrium. That is, for a given T_ε we will give bounds on the number of time steps before $\hat{f}$ is average-ε-Nash. After analyzing $\hat{f}$, we then extend our analysis to show that in fact for *most* time steps t, the flow f^t itself is average-ε-Nash. To achieve bounds of this form, which we show in Section 5, we will however need to lose an additional factor polynomial in the size of the graph. Again, we cannot hope to say that f^t is average-ε-Nash for *all* (sufficiently large) time-steps t, because regret-minimizing algorithms may occasionally take long paths, and an "adversarial" set of such algorithms may occasionally all take long paths at the same time.

2.5 Dependence on slope

Our convergence rates will depend on the maximum slope s allowed for any latency function. For example, consider the case of a routing game with two parallel links, where one edge has latency 0 up to a load of $1/3$ and then rises immediately to 1, and the other edge has latency 0 up to a load of $2/3$ and then rises directly to 1. In this case the equilibrium cost is 0, and moreover for *any* flow f' we have $\min_{P\in\mathcal{P}}\sum_{e\in P}\ell_e(f'_e)=0$. Thus, the only way f' can be average-ε-Nash is for it to actually have low cost, which means the algorithm must precisely be at a $1/3$-$2/3$ split. If players use regret-minimizing algorithms, traffic will instead oscillate, each edge having cost 1 on about half the days and each player incurring cost 1 on not much more than half the days (and thus not having much regret). However, none of the daily flows will be better than average-$(1/3)$-Nash, because on each day, the cost of the flow f is at least 1/3.

This example demonstrates that some dependence on the shape of the latency functions is necessary. However, it is possible that one could obtain stronger results using a different parameter than the slope.

3 Infinitesimal users: Linear latency functions

We begin as a warm-up with the easiest case, infinitesimal users and linear latency functions, which simplifies many of the arguments. In particular, for linear latency functions, the latency of any edge under the time-average flow $\hat{f}$ is guaranteed to be equal to the average latency of that edge over time, i. e., $\ell_e(\hat{f}_e)=(1/T)\sum_{t=1}^{T}\ell_e(f_e^t)$ for all e.

Theorem 3.1. *Suppose the latency functions are linear. Then for flows satisfying the regret-minimizing assumption* (2.7)*, for $T\geq T_\varepsilon$, the time-average flow $\hat{f}$ is average-ε-Nash, i. e.,*

$$C(\hat{f})\leq\varepsilon+\sum_i a_i\min_{P\in\mathcal{P}_i}\sum_{e\in P}\ell_e(\hat{f}_e)\,.$$

Proof. From the linearity of the latency functions, we have for all e, $\ell_e(\hat{f}_e)=(1/T)\sum_{t=1}^{T}\ell_e(f_e^t)$. Since $\ell_e(f_e^t)f_e^t$ is a convex function of the flow, this implies

$$\ell_e(\hat{f}_e)\hat{f}_e\leq\frac{1}{T}\sum_{t=1}^{T}\ell_e(f_e^t)f_e^t\,.$$

Summing over all e, we have

$$\begin{array}{rcll} C(\hat{f}) & \leq & \frac{1}{T}\sum_{t=1}^{T} C(f^t) & \\ & \leq & \varepsilon + \sum_i a_i \min_{P\in\mathcal{P}_i} \frac{1}{T}\sum_{t=1}^{T}\sum_{e\in P}\ell_e(f_e^t) & \text{(by Assumption 2.7)} \\ & = & \varepsilon + \sum_i a_i \min_{P\in\mathcal{P}_i}\sum_{e\in P}\ell_e(\hat{f}_e) & \text{(by linearity.)} \end{array}$$

□

Corollary 3.2. *Assume that all latency functions are linear. In network routing games, if all agents use, for example, the algorithm of Kalai and Vempala [25], the time-averaged flow converges to an average-ε-Nash equilibrium at $T_\varepsilon = O(mn(\log n)/\varepsilon^2)$. On networks consisting of two nodes and m parallel links, if all agents use optimized regret-minimizing algorithms (in particular, "combining expert advice"-style algorithms with each edge an expert [28, 8, 20]), the time-averaged flow converges to an average-ε-Nash equilibrium at $T_\varepsilon = O((\log m)/\varepsilon^2)$.*

Note that we not only proved that the average flow approaches an average-ε-Nash equilibrium, but by connecting the first and third steps of the proof, we also showed that the *actual* average cost incurred by a user of type i is at most ε worse than the best path in $\mathcal{P}_i$ in the average flow.

4 Infinitesimal users: General latency functions

The case of general latency functions is more complicated because the proof above used a convexity assumption in its first step and a linearity (concavity) assumption in the third step; the same argument does not apply to general latency functions. In the case of general functions, the additive term in our bounds depends on the maximum slope of any latency function.

Theorem 4.1. *Let $\varepsilon' = \varepsilon + 2\sqrt{\varepsilon L}$. (Recall, $L = sn$.) Then for flows satisfying the regret-minimizing assumption (2.7), for general functions with maximum slope s, for $T \geq T_\varepsilon$, the time-average flow $\hat{f}$ is average-ε'-Nash, that is,*

$$\sum_{e\in E}\ell_e(\hat{f}_e)\hat{f}_e \leq \varepsilon + 2\sqrt{\varepsilon L} + \sum_i a_i \min_{P\in\mathcal{P}_i}\sum_{e\in P}\ell_e(\hat{f}_e)\,.$$

Before giving the proof, we list several quantities we will need to relate:

$$\sum_{e\in E}\ell_e(\hat{f}_e)\hat{f}_e \qquad \text{(cost of } \hat{f}\text{)} \tag{4.1}$$

$$\frac{1}{T}\sum_{t=1}^{T}\sum_{e\in E}\ell_e(f_e^t)\hat{f}_e \qquad \text{("cost of } \hat{f} \text{ in hindsight")} \tag{4.2}$$

$$\frac{1}{T}\sum_{t=1}^{T}\sum_{e\in E}\ell_e(f_e^t)f_e^t \qquad \text{(avg cost of flows up to time } T\text{)} \tag{4.3}$$

$$\sum_i a_i \min_{P\in\mathcal{P}_i}\sum_{e\in P}\frac{1}{T}\sum_{t=1}^{T}\ell_e(f_e^t) \qquad \text{(cost of best path in hindsight)} \tag{4.4}$$

$$\sum_i a_i \min_{P\in\mathcal{P}_i}\sum_{e\in P}\ell_e(\hat{f}_e) \qquad \text{(cost of best path given } \hat{f}\text{)} \tag{4.5}$$

Our goal in proving Theorem 4.1 is to show that quantity (4.1) is not too much greater than quantity (4.5). We will prove this through a chain of inequalities in the following lemmas.

Lemma 4.2. *For general latency functions,*

$$\frac{1}{T}\sum_{t=1}^{T}\sum_{e\in E}\ell_e(f_e^t)\hat{f}_e \le \frac{1}{T}\sum_{t=1}^{T}\sum_{e\in E}\ell_e(f_e^t)f_e^t\,.$$

Proof. This is a direct application of Chebyshev's sum inequality. □

Lemma 4.3. *For general latency functions with maximum slope s, for flows satisfying the regret-minimizing assumption* (2.7)*,*

$$\min_{P\in\mathcal{P}_i}\sum_{e\in P}\frac{1}{T}\sum_{t=1}^{T}\ell_e(f_e^t) \le \min_{P\in\mathcal{P}_i}\sum_{e\in P}\ell_e(\hat{f}_e)+\sqrt{\varepsilon L}\,.$$

Proof. Since our latency functions are non-decreasing, we can apply Chebyshev's sum inequality to obtain for every e that,

$$\frac{1}{T}\hat{f}_e\sum_{t=1}^{T}\ell_e(f_e^t) \le \frac{1}{T}\sum_{t=1}^{T}\ell_e(f_e^t)f_e^t\,.$$

Now, notice that the right-hand side of the above inequality, summed over all edges, is precisely quantity (4.3). By the regret-minimization property, this is at most ε larger than the time-average cost of the best paths in hindsight, which in turn is clearly at most the time-average cost of $\hat{f}$. Therefore, we have:

$$\begin{aligned}
\frac{1}{T}\sum_{t=1}^{T}\sum_{e\in E}\hat{f}_e\ell_e(f_e^t) &\le \frac{1}{T}\sum_{t=1}^{T}\sum_{e\in E}\ell_e(f_e^t)f_e^t \\
&\le \varepsilon+\sum_i a_i\min_{P\in\mathcal{P}_i}\sum_{e\in P}\frac{1}{T}\sum_{t=1}^{T}\ell_e(f_e^t) \\
&\le \varepsilon+\frac{1}{T}\sum_{t=1}^{T}\sum_{e\in E}\hat{f}_e\ell_e(f_e^t).
\end{aligned}$$

That is, we have "sandwiched" the flow-average latency between the time-average latency and the time-average latency plus ε. This implies that for every edge e, its time-average cost must be close to its flow-average cost, namely,

$$\sum_{e\in E}\varepsilon_e \le \varepsilon\,.$$

We now use this fact, together with the assumption of bounded slope, to show that edge latencies cannot be varying wildly over time. Define

$$\varepsilon_e=\frac{1}{T}\sum_{t=1}^{T}\ell_e(f_e^t)f_e^t-\frac{1}{T}\hat{f}_e\sum_{t=1}^{T}\ell_e(f_e^t)\,.$$

We can then use the fact that $\hat{f}_e = (1/T)\sum_{t=1}^{T} f_e^t$ and so $(1/T)\sum_{t=1}^{T} \ell_e(\hat{f}_e)(f_e^t - \hat{f}_e) = 0$ to rewrite this definition of ε_e as:

$$\varepsilon_e = \frac{1}{T}\sum_{t=1}^{T}(\ell_e(f_e^t) - \ell_e(\hat{f}_e))(f_e^t - \hat{f}_e) \geq 0\,. \tag{4.6}$$

From the bound on the maximum slope of any latency function, we know that $|\ell_e(f_e^t) - \ell_e(\hat{f}_e)| \leq s|f_e^t - \hat{f}_e|$ and thus,

$$\begin{aligned} |\ell_e(f_e^t) - \ell_e(\hat{f}_e)| &= \sqrt{(\ell_e(f_e^t) - \ell_e(\hat{f}_e))^2} \\ &\leq \sqrt{s\left(\ell_e(f_e^t) - \ell_e(\hat{f}_e)\right)\left(f_e^t - \hat{f}_e\right)} \end{aligned}$$

for all e.

We then get

$$\frac{1}{T}\sum_{t=1}^{T}\left|\ell_e(f_e^t) - \ell_e(\hat{f}_e)\right| \leq \frac{\sqrt{s}}{T}\sum_{t=1}^{T}\sqrt{(\ell_e(f_e^t) - \ell_e(\hat{f}_e))(f_e^t - \hat{f}_e)}\,.$$

Using equation (4.6) above and the fact the square root is a concave function, an application of the Jensen inequality yields

$$\frac{1}{T}\sum_{t=1}^{T}\left|\ell_e(f_e^t) - \ell_e(\hat{f}_e)\right| \leq \sqrt{s\varepsilon_e}\,. \tag{4.7}$$

Finally, let P_i^* be the best path of type i given $\hat{f}$. Summing equation (4.7) over the edges in P_i^*, and using an application of the Cauchy-Schwartz inequality to get $\sum_{e\in P_i^*}\sqrt{s\varepsilon_e} \leq \sqrt{\varepsilon L}$, we have

$$\min_{P\in\mathcal{P}_i}\sum_{e\in P}\frac{1}{T}\sum_{t=1}^{T}\ell_e(f_e^t) \;\leq\; \sum_{e\in P_i^*}\frac{1}{T}\sum_{t=1}^{T}\ell_e(f_e^t) \;\leq\; \min_{P\in\mathcal{P}_i}\sum_{e\in P}\ell_e(\hat{f}_e) + \sqrt{\varepsilon L},$$

as desired. □

Lemma 4.4. *For general latency functions with maximum slope s, for flows satisfying the regret-minimizing assumption* (2.7)*,*

$$\sum_{e\in E}\ell_e(\hat{f}_e)\hat{f}_e \leq \frac{1}{T}\sum_{t=1}^{T}\sum_{e\in E}\ell_e(f_e^t)\hat{f}_e + \sqrt{\varepsilon L}\,.$$

Proof. Equation (4.7) above directly gives us

$$(4.1) \leq \sum_{e\in E}\sqrt{s\varepsilon_e}\hat{f}_e + (4.2)\,.$$

An application of the Cauchy-Schwartz inequality then gives us

$$\left(\sum_{e\in E}\sqrt{s\varepsilon_e}\hat{f}_e\right)^2 \leq \left(\sum_{e\in E}s\varepsilon_e\right)\left(\sum_{e\in E}\hat{f}_e^2\right).$$

Since $\hat{f}_e \leq 1$ for all e, this is at most $\left(\sum_{e\in E} s\varepsilon_e\right)\left(\sum_{e\in E} \hat{f}_e\right)$. Since $\sum_{e\in E} \hat{f}_e \leq n$, this is at most εL, and thus

$$\sum_{e\in E} \sqrt{s\varepsilon_e}\hat{f}_e \leq \sqrt{\varepsilon L},$$

which gives the desired result. □

Given the above lemmas we now present the proof of Theorem 4.1.

Proof of Theorem 4.1. We can now piece together the following chain of inequalities:

$$\begin{aligned}
(4.1) &\leq \sqrt{s\varepsilon n} + (4.2)\\
&\leq \sqrt{\varepsilon L} + (4.3)\\
&\leq \varepsilon + \sqrt{\varepsilon L} + (4.4)\\
&\leq \varepsilon + 2\sqrt{\varepsilon L} + (4.5)\,.
\end{aligned}$$

The first line is from Lemma 4.4 and the second from Lemma 4.2. The third line is a consequence of the regret-minimization property. Finally, the fourth line follows from Lemma 4.3. □

Corollary 4.5. *Let $\varepsilon' = \varepsilon + 2\sqrt{\varepsilon L}$. Assume that all latency functions have maximum slope s. In general routing games, if all agents use, for example, the algorithm of Kalai and Vempala [25], the average flow converges to an average-ε'-Nash equilibrium at*

$$T_\varepsilon = O\left(\frac{mn\log n}{\varepsilon^2}\right) = O\left(\frac{mn^3 s^2 \log n}{\varepsilon'^4}\right).$$

On networks consisting of two nodes and m parallel links, if all agents use optimized regret-minimizing algorithms (in particular "combining expert advice"-style algorithms [28, 8, 20]), the average flow converges to an average-ε'-Nash equilibrium at

$$T_\varepsilon = O\left(\frac{\log m}{\varepsilon^2}\right) = O\left(\frac{n^2 s^2 \log m}{\varepsilon'^4}\right).$$

Once again we remark that not only have we proved that the average flow approaches average-ε'-Nash equilibrium, but as an intermediate step in our proof we showed that *actual* average cost obtained by the users is at most ε' worse than the best path in the average flow.

5 Infinitesimal users: Bounds on most timesteps

Here we present results applicable to general graphs and general functions showing that on *most* time steps t, the flow f^t will be at average-ε-Nash equilibrium.

Theorem 5.1. *In routing games with general latency functions with maximum slope s, for flows satisfying the regret-minimizing assumption (2.7), for all but a $(ms^{1/4}\varepsilon^{1/4})$ fraction of time steps up to time T for $T \geq T_\varepsilon$, f^t is a average-$(\varepsilon + 2\sqrt{\varepsilon L} + 2m^{3/4}s^{1/4}\varepsilon^{1/4})$-Nash flow. We can rewrite this as: for all but an ε' fraction of time steps up to T for $T \geq T_\varepsilon$, f^t is an average-ε'-Nash flow for $\varepsilon = \Omega\left(\varepsilon'^4/(sm^4 + s^2n^2)\right)$.*

Proof. Based on equation (4.6),

$$\sqrt{s\varepsilon_e} \geq \frac{1}{T}\sum_{t=1}^{T} |\ell_e(f_e^t) - \ell_e(\hat{f}_e)|$$

for all edges. Thus, for all edges, for all but $s^{1/4}\varepsilon_e^{1/4}$ of the time steps,

$$s^{1/4}\varepsilon_e^{1/4} \geq |\ell_e(f_e^t) - \ell_e(\hat{f}_e)|\,.$$

Using a union bound over edges, this implies that on all but a $ms^{1/4}\varepsilon^{1/4}$ fraction of the time steps, *all* edges have

$$s^{1/4}\varepsilon_e^{1/4} \geq |\ell_e(f_e^t) - \ell_e(\hat{f}_e)|\,.$$

From this, it follows directly that on most time steps, the cost of the best path given f^t differs from the cost of the best path given $\hat{f}$ by at most $n^{3/4}s^{1/4}\varepsilon^{1/4}$. Also on most time steps, the cost incurred by flow f^t differs from the cost incurred by flow $\hat{f}$ by at most $m^{3/4}s^{1/4}\varepsilon^{1/4}$. Thus since $\hat{f}$ is an average-$(\varepsilon+2\sqrt{\varepsilon L})$-Nash equilibrium, f^t is an average-$(\varepsilon+2\sqrt{\varepsilon L}+2m^{3/4}s^{1/4}\varepsilon^{1/4})$-Nash equilibrium on all but a $ms^{1/4}\varepsilon^{1/4}$ fraction of time steps. □

Corollary 5.2. *In general routing games with general latency functions with maximum slope s, for all but a* $(ms^{1/4}\varepsilon^{1/4})$ *fraction of time steps up to time* $T = T_\varepsilon$*, the expected average cost* $(1/T)\sum_{t=1}^{T} c^t$ *incurred by any user is at most* $(\varepsilon+2\sqrt{\varepsilon L}+m^{3/4}s^{1/4}\varepsilon^{1/4})$ *worse than the cost of the best path on that time step.*

Proof. From the proof of Theorem 5.1 we see that on most days, the cost of the best path given the flow for that day is within $m^{3/4}s^{1/4}\varepsilon^{1/4}$ of the cost of the best path given $\hat{f}$, which is at most $2\sqrt{\varepsilon L}$ worse than the cost of the best path in hindsight. Combining this with the regret-minimization property achieved by each user gives the desired result. □

This demonstrates that regret-minimizing algorithms are a reasonable, stable response in a network setting: if a player knows that all other players are using regret-minimizing algorithms, there is no strategy that will significantly improve her expected cost on more than a small fraction of days. By using a regret-minimizing algorithm, she gets the guarantee that on most time steps her expected cost is within some epsilon of the cost of the best path given the flow for that day.

6 Regret minimization and the price of anarchy

In this section, we relate the costs incurred by regret-minimizing players in a congestion game to the cost of the social optimum. We approach this problem in two ways: First, we give an argument paralleling that of Roughgarden and Tardos [35] that directly relates the costs of regret-minimizing users to the cost of the social optimum. In our second result in this section, we show that any average-ε-Nash equilibrium in a congestion game is closely related to a true equilibrium in a related congestion game. This is an interesting property of approximate equilibria in their own right, and further allows us to apply Price of Anarchy results for the congestion game to the regret-minimizing players in the original game.

We can directly characterize the costs incurred by regret-minimizing players by an argument that parallels the Price of Anarchy proofs of Roughgarden and Tardos [35].

Definition 6.1. Let $\mathcal{L}$ be the set of cost functions used by a nonatomic congestion game, with all $\ell(\xi)\xi$ convex on $[0,\infty)$. For a nonzero cost function $\ell \in \mathcal{L}$, we define $\alpha(\ell)$ by

$$\alpha(\ell) = \sup_{n>0:\ell(n)>0} [\lambda\mu + (1-\lambda)]^{-1}$$

where $\ell_e^*(\xi) = \ell_e(\xi) + \xi \cdot \ell_e'(\xi)$, $\lambda \in [0,1]$, (the marginal social cost) satisfies $\ell^*(\lambda n) = \ell(n)$, and $\mu = \ell(\lambda n)/\ell(n) \in [0,1]$. We define $\alpha(\mathcal{L})$ by

$$\alpha(\mathcal{L}) = \sup_{0\neq\ell\in\mathcal{L}} \alpha(\ell)\,.$$

Theorem 6.2. *If Γ is a nonatomic congestion game with cost functions $\mathcal{L}$ with all $\ell(\xi)\xi$ convex on $[0,\infty)$, then the ratio of the costs incurred by regret-minimizing players to the cost of the global optimum flow is at most $\alpha(\mathcal{L})$ (which is the Price of Anarchy bound given by Roughgarden and Tardos [35]), once the players achieve the regret-minimizing assumption.*

Proof. Let f^* be an optimal action distribution and $f_1,\dots,f_T$ be a sequence of action distributions obtained by regret-minimizing players. We can lower bound the optimum social cost using a linear approximation of the function $\ell_e(\xi)\xi$ at the point $\lambda_e^t f_e^t$, where $\lambda_e^t \in [0,1]$ solves $\ell_e^*(\lambda_e^t f_e^t) = \ell_e(f_e^t)$:

$$\begin{aligned}\ell_e(f_e^*)f_e^* &= \ell_e(\lambda_e^t f_e^t)\lambda_e^t f_e^t + \int_{\lambda_e^t f_e^t}^{f_e^*} \ell_e^*(f)\,dx\\ &\geq \ell_e(\lambda_e^t f_e^t)\lambda_e^t f_e^t + (f_e^* - \lambda_e^t f_e^t)\ell_e^*(\lambda_e^t f_e^t)\\ &= \ell_e(\lambda_e^t f_e^t)\lambda_e^t f_e^t + (f_e^* - \lambda_e^t f_e^t)\ell_e(f_e^t)\end{aligned}$$

for all edges and time steps, and thus

$$C(f^*) \geq \frac{1}{T}\sum_{t=1}^{T}\sum_{e\in E}[\ell_e(\lambda_e^t f_e^t)\lambda_e^t f_e^t + (f_e^* - \lambda_e^t f_e^t)\ell_e(f_e^t)]\,.$$

We can rewrite this as

$$C(f^*) \geq \frac{1}{T}\sum_{t=1}^{T}\sum_{e\in E}[\mu_e^t\lambda_e^t f_e^t + (1-\lambda_e^t)f_e^t]\ell_e(f_e^t) + \sum_{e\in E}[f_e^* - f_e^t]\ell_e(f_e^t)\,,$$

where $\mu_e^t = \ell_e(\lambda_e^t f_e^t)/\ell_e(f_e^t)$. By the regret minimization property,

$$\frac{1}{T}\sum_{t=1}^{T}\sum_{e\in E} f_e^t\ell_e(f_e^t) \leq \varepsilon + \sum_i a_i \min_{P\in\mathcal{P}_i}\frac{1}{T}\sum_{t=1}^{T}\sum_{e\in E}\ell_e(f_e^t)$$

and thus

$$\frac{1}{T}\sum_{t=1}^{T}\sum_{e\in E} f_e^t\ell_e(f_e^t) \leq \varepsilon + \frac{1}{T}\sum_{t=1}^{T}\sum_{e\in E} f_e^*\ell_e(f_e^t)\,,$$

which gives us

$$C(f^*) + \varepsilon \geq \frac{1}{T}\sum_{t=1}^{T}\sum_{e\in E}[\mu_e^t\lambda_e^t f_e^t + (1-\lambda_e^t)f_e^t]\ell_e(f_e^t)\,.$$

By definition, $\mu_e^t \lambda_e^t + (1-\lambda_e^t) \geq 1/\alpha(\mathcal{L})$ for each e and t, so $\mu_e^t \lambda_e^t f_e^t + (1-\lambda_e^t) f_e^t] \ell_e(f_e^t)$ and $\ell_e(f_e^t) f_e^t$ differ by at most a multiplicative $\alpha(\mathcal{L})$ factor for every e and t. This gives us

$$C(x^*) + \varepsilon \geq \frac{1}{\alpha(\mathcal{L})} \frac{1}{T} \sum_{t=1}^{T} \sum_{e \in E} \ell_e(f_e^t) f_e^t = \frac{C(x)}{\alpha(\mathcal{L})},$$

as desired. □

Subsequent work of Christodoulou et al. [9] also studies the Price of Anarchy of approximate equilibria in congestion games. Their definition of approximate equilibrium is stronger than ours (though with a multiplicative rather than additive definition of the players' incentive to deviate), but their results apply only to *linear* congestion games.

We can also analyze the performance of regret-minimizing players indirectly, by gaining better understanding of flows at average-ε-Nash equilibrium. In the proof of this theorem, we need to consider costs of different flows on a few different congestion games; we will use the notation $C(f \text{ on } \Gamma)$ to denote the cost $C(f)$ of a flow (or partial, non-unit-sized flow) f in the congestion game Γ.

Theorem 6.3. *If f is an average-ε-Nash equilibrium flow for a nonatomic congestion game Γ, then*

$$C(f \text{ on } \Gamma) \leq \rho \left(C(OPT) + (L+1)\sqrt{\varepsilon} \right) + n\sqrt{\varepsilon},$$

where $L = sn$, OPT is the minimum cost flow, and ρ is the Price of Anarchy in a related congestion game with the same class of latency functions as Γ but with additive offsets.

For example, Theorem 6.3 implies that for linear latency functions of slope less than or equal to one, an average-ε-Nash flow f will have cost at most $4/3(C(OPT) + \sqrt{\varepsilon}(n+1)) + n\sqrt{\varepsilon}$. Note that for regret minimizing players, Theorem 6.2 improves this to $(4/3)C(OPT) + \varepsilon$.

The proof idea for this theorem is as follows: For every nonatomic congestion game Γ and flow f at average-ε-Nash equilibrium on Γ, there exists a nonatomic congestion game Γ' that approximates Γ and a flow f' that approximates f such that: (a) f' is an equilibrium flow on Γ', (b) the cost of f' on Γ' is close to the cost of f on Γ, and (c) the cost of the optimal flow on Γ' is close to the cost of the optimal flow on Γ. These approximations allow one to apply Price of Anarchy results from f' and Γ' to f and Γ.

Proof. Note that since f is at average-ε-Nash equilibrium on Γ, then at most a $\sqrt{\varepsilon}$ fraction of users are experiencing costs more than $\sqrt{\varepsilon}$ worse than the cost of their best path given f. We can modify Γ to Γ_2 to embed the costs associated with these "meandering" users such that the costs experienced by the remaining users do not change. For example, a latency function $\ell(x)$ associated with an edge with a flow of δ meandering users can be modified to the function $\ell'(x) = \ell(x+\delta)$ for $x \leq 1-\delta$ and $\ell'(x) = \ell(1)$ for $x > 1-\delta$. Call the remaining $\eta \geq (1-\sqrt{\varepsilon})$ flow of non-meandering users f', all of whom experience costs at most $\sqrt{\varepsilon}$ worse than the cost of their best path given f. Then $C(f \text{ on } \Gamma) \leq C(f' \text{ on } \Gamma_2) + \sqrt{\varepsilon}n$, since the flow of the meandering users is at most $\sqrt{\varepsilon}$ and the cost of each path is at most n.

We now construct an alternate congestion game Γ' (not necessarily a routing game, even if the original game was a routing game) such that f' interpreted on Γ' is at equilibrium. To do this, we create a new edge for every allowable path for each commodity. We can now assign costs to these new "entry edges" to cause the minimum cost of any available path for each commodity to be equal to the cost of the worst flow-carrying path for that commodity in f' on Γ_2. The maximum cost we need to assign to any entry edge in order to achieve this is $\sqrt{\varepsilon}$, since we already removed all users paying more than $\sqrt{\varepsilon}$ plus the cost of the best path available to them. Thus $C(f' \text{ on } \Gamma_2) \leq C(f' \text{ interpreted on } \Gamma')$, so we have

$$C(f \text{ on } \Gamma) \leq C(f' \text{ interpreted on } \Gamma') + n\sqrt{\varepsilon}\,.$$

Define ρ to be the Price of Anarchy of the new congestion game Γ' when played with up to one unit of flow. Thus, defining $OPT_\alpha(H)$ to be the min-cost flow of size α in game H, we have

$$C(f \text{ on } \Gamma) \leq \rho\left(C(OPT_\eta(\Gamma'))\right) + n\sqrt{\varepsilon}\,.$$

Since we added at most $\sqrt{\varepsilon}$ to the cost of any solution in going from Γ_2 to Γ' and $OPT_\eta(\Gamma_2)$ is the min-cost flow of size η on Γ_2,

$$C(f \text{ on } \Gamma) \leq \rho\left(C(OPT_\eta(\Gamma_2)) + \sqrt{\varepsilon}\right) + n\sqrt{\varepsilon}\,,$$

We now must quantify the amount by which the cost of OPT_η on Γ_2 could exceed the cost of OPT_1 on Γ. Since the cost of any edge in Γ_2 is at most $s\sqrt{\varepsilon}$ more than the cost of that edge in Γ, this gives

$$C(f \text{ on } \Gamma) \leq \rho\left(C(OPT) + (sn+1)\sqrt{\varepsilon}\right) + n\sqrt{\varepsilon}\,.$$

□

In particular, when all latency functions are linear, we can apply results of Roughgarden and Tardos bounding the price of anarchy in a congestion game with linear latency functions by $4/3$ [35].

7 Discrete users: Parallel paths

In contrast with the previous sections, we now consider discrete users, where we denote the ith user weight as w_i. Without loss of generality, we assume that the weights are normalized such that $\sum_{i=1}^n w_i = 1$. We limit ourselves in this section to the single-commodity version of the parallel paths routing game model and to functions with latency equal to the load, that is, for a path e we have $\ell_e = f_e$. For each user i, we let the latency excluding her own path e at time t be $\ell_e(f_e^t \setminus i)$ and her average latency on path e be $\ell_e(\hat{f}_e \setminus i) = (1/T)\sum_{t=1}^T \ell_e(f_e^t \setminus i)$, where $f_e^t \setminus i = f_e^t$ if user i is not routing on path e and $f_e^t \setminus i = f_e^t - w_i$ otherwise. We always exclude the ith player from the latency function, since the ith player always pays for its weight.

Next we observe that at time t, there always exists a path with load at most the average load.

Observation 7.1. At any time step t, for every user i, there exists a path e such that

$$\ell_e(\hat{f}_e \setminus i) \leq \frac{1-w_i}{m}\,.$$

The following theorem differs from other theorems in the paper in the sense that it is an expectation result and holds for every user.

Theorem 7.2. *Consider the parallel paths model, with latency functions such that the latency equals the load. Assume that each discrete user i uses an optimized "combining expert advice"-style algorithm with each edge an expert [28, 8, 20]. Then for all users, for all* $T \geq O((\log m)/\varepsilon^2)$,

$$\frac{1}{T}\sum_{t=1}^{T} E_{e\sim q_t}[\ell_e(f_e^t \setminus i)] \leq \frac{1-w_i}{m} + \varepsilon\,,$$

where q_t *is the distribution over the m paths output by the best expert algorithm at time t.*

Proof. By Observation 7.1 we have that there exists a path with average cost at most $(1-w_i)/m$. Since user i is using an optimized "combining expert advice"-style algorithm and the maximal latency is 1, we have that

$$\begin{aligned}\frac{1}{T}\sum_{t=1}^{T} \mathbf{E}_{e\sim q_t}[\ell_e(f_e^t \setminus i)] &\leq \min_{e\in E} \ell_e(\hat{f}_e \setminus i) + \sqrt{\frac{\log m}{T}} \\ &\leq \frac{1-w_i}{m} + \sqrt{\frac{\log m}{T}} \\ &\leq \frac{1-w_i}{m} + \varepsilon\end{aligned}$$

where the last inequality holds for $T \geq O((\log m)/\varepsilon^2)$. □

Consider an instance of this model where every user plays uniformly at random. The resulting flow is clearly at equilibrium, and the expected latency for the ith player is $(1-w_i)/m$ excluding its own weight. We thus have shown that the expected latency experienced by each user i is at most ε worse than this equilibrium latency.

8 Conclusions

In this paper, we consider the question: if each player in a routing game (or more general congestion game) uses a regret-minimizing strategy, will behavior converge to an equilibrium, and under what conditions and in what sense? Our main result is that in the setting of multicommodity flow and infinitesimal agents, a $1-\varepsilon$ fraction of the daily flows are at average-ε-Nash equilibrium for ε approaching 0 at a rate that depends polynomially on the players' regret bounds and the maximum slope of any latency function. Moreover, we show the dependence on slope is necessary.

Even for the case of reasonable (bounded) slopes, however, our bounds for general nonlinear latencies are substantially worse than our bounds for the linear case. For instance if agents are running the Kalai-Vempala algorithm [25], we get a bound of $O(mn(\log n)/\varepsilon^2)$ on the number of time steps needed for the time-average flow to reach an average-ε-Nash equilibrium in the linear case, but $O(mn^3(\log n)/\varepsilon^4)$ for general latencies. We do not know if these bounds in the general case can be improved. In addition, our bounds on the daily flows lose additional polynomial factors which we suspect are not tight.

We also show that Price of Anarchy results can be applied to regret-minimizing players in routing games, that is, that existing results analyzing the quality of equilibria can also be applied to the results of regret-minimizing behavior. Recent work [7] shows that in fact Price of Anarchy results can be extended to cover regret-minimizing behavior in a wide variety of games, including many for which this behavior may not approach equilibria and where equilibria may be hard to find.

References

[1] JACOB ABERNETHY, ELAD HAZAN, AND ALEXANDER RAKHLIN: Competing in the dark: An efficient algorithm for bandit linear optimization. In *Proc. 21st Ann. Conf. Learning Theory (COLT)*. Springer, 2008. http://colt2008.cs.helsinki.fi/papers/123-Abernethy.pdf. 185

[2] BARUCH AWERBUCH, YOSSI AZAR, AMIR EPSTEIN, VAHAB SEYED MIRROKNI, AND ALEXANDER SKOPALIK: Fast convergence to nearly optimal solutions in potential games. In *Proc. 9th ACM Conf. Electronic Commerce (EC)*, pp. 264–273. ACM Press, 2008. [doi:10.1145/1386790.1386832]. 183

[3] BARUCH AWERBUCH AND ROBERT D. KLEINBERG: Adaptive routing with end-to-end feedback: Distributed learning and geometric approaches. In *Proc. 36th STOC*, pp. 45–53. ACM Press, 2004. [doi:10.1145/1007352.1007367]. 180, 182, 185

[4] YOSSI AZAR: *Online Algorithms: The State of the Art*, chapter "On-line load balancing", pp. 178–195. Springer, 1998. 181

[5] M. BECKMANN, C. B. MCGUIRE, AND C. B. WINSTEN: *Studies in the Economics of Transportation*. Yale University Press, 1956. 184

[6] PETRA BERENBRINK, TOM FRIEDETZKY, LESLIE ANN GOLDBERG, PAUL W. GOLDBERG, ZENGJIAN HU, AND RUSSELL MARTIN: Distributed selfish load balancing. *SIAM J. Comput.*, 37(4):1163–1181, 2007. [doi:10.1137/060660345]. 183

[7] AVRIM BLUM, MOHAMMADTAGHI HAJIAGHAYI, KATRINA LIGETT, AND AARON ROTH: Regret minimization and the price of total anarchy. In *Proc. 40th STOC*, pp. 373–382. ACM Press, 2008. [doi:10.1145/1374376.1374430]. 183, 196

[8] NICOLÒ CESA-BIANCHI, YOAV FREUND, DAVID HAUSSLER, DAVID P. HELMBOLD, ROBERT E. SCHAPIRE, AND MANFRED K. WARMUTH: How to use expert advice. *J. ACM*, 44(3):427–485, 1997. [doi:10.1145/258128.258179]. 180, 185, 187, 190, 195

[9] GEORGE CHRISTODOULOU, ELIAS KOUTSOUPIAS, AND PAUL G. SPIRAKIS: On the performance of approximate equilibria in congestion games. In *Proc. 17th European Symp. Algorithms (ESA'09)*, pp. 251–262. Springer, 2009. [doi:10.1007/978-3-642-04128-0_22]. 193

[10] ARTUR CZUMAJ, PIOTR KRYSTA, AND BERTHOLD VÖCKING: Selfish traffic allocation for server farms. *SIAM J. Comput.*, 39:1957–1987, 2010. [doi:10.1137/070693862]. 180

[11] ARTUR CZUMAJ AND BERTHOLD VÖCKING: Tight bounds for worst-case equilibria. *ACM Trans. Algorithms*, 3(1):4, 2007. [doi:10.1145/1186810.1186814]. 180

[12] EYAL EVEN-DAR, ALEX KESSELMAN, AND YISHAY MANSOUR: Convergence time to Nash equilibria. In *Proc. 30th Intern. Colloq. Autom. Lang. Program. (ICALP)*, number 2719 in LNCS, pp. 193–193. Springer, 2003. [doi:10.1007/3-540-45061-0_41]. 183

[13] EYAL EVEN-DAR AND YISHAY MANSOUR: Fast convergence of selfish rerouting. In *Proc. 16th Ann. ACM-SIAM Symp. Discrete Algorithms (SODA)*, pp. 772–781. ACM Press, 2005. [ACM:1070432.1070541]. 183

[14] EYAL EVEN-DAR, YISHAY MANSOUR, AND URI NADAV: On the convergence of regret minimization dynamics in concave games. In *Proc. 41st STOC*, pp. 523–532. ACM Press, 2009. [doi:10.1145/1536414.1536486]. 183

[15] ALEX FABRIKANT, ANKUR LUTHRA, ELITZA MANEVA, CHRISTOS H. PAPADIMITRIOU, AND SCOTT SHENKER: On a network creation game. In *Proc. 22nd Symp. Principles of Distrib. Comput. (PODC)*, pp. 347–351. ACM Press, 2003. [doi:10.1145/872035.872088]. 180

[16] SIMON FISCHER, HARALD RÄCKE, AND BERTHOLD VÖCKING: Fast convergence to Wardrop equilibria by adaptive sampling methods. In *Proc. 38th STOC*, pp. 653–662. ACM Press, 2006. [doi:10.1145/1132516.1132608]. 182, 185

[17] SIMON FISCHER AND BERTHOLD VÖCKING: On the evolution of selfish routing. In *Proc. 12th European Symp. Algorithms (ESA)*, number 3221 in LNCS, pp. 323–334. Springer, 2004. [doi:10.1007/978-3-540-30140-0_30]. 182

[18] SIMON FISCHER AND BERTHOLD VÖCKING: Adaptive routing with stale information. *Theoretical Computer Science*, 410(36):3357–3371, 2009. [doi:10.1016/j.tcs.2008.01.055]. 182

[19] DEAN P. FOSTER AND RAKESH V. VOHRA: Calibrated learning and correlated equilibrium. *Games Econom. Behav.*, 21(1–2):40–55, 1997. [doi:10.1006/game.1997.0595]. 182

[20] YOAV FREUND AND ROBERT E. SCHAPIRE: A decision-theoretic generalization of on-line learning and an application to boosting. *J. Comput. System Sci.*, 55(1):119–139, 1997. [doi:10.1006/jcss.1997.1504]. 180, 185, 187, 190, 195

[21] YOAV FREUND AND ROBERT E. SCHAPIRE: Adaptive game playing using multiplicative weights. *Games Econom. Behav.*, 29:79–103, 1999. [doi:10.1006/game.1999.0738]. 182

[22] PAUL W. GOLDBERG: Bounds for the convergence rate of randomized local search in a multiplayer load-balancing game. In *Proc. 23rd Ann. ACM Symp. Principles of Distrib. Comput. (PODC)*, pp. 131–140. ACM Press, 2004. [doi:10.1145/1011767.1011787]. 183

[23] J.F. HANNAN: Approximation to Bayes risk in repeated play. In M. DRESHER, A.W. TUCKER, AND P. WOLFE, editors, *Contributions to the Theory of Games*, volume III, pp. 97–139. Princeton University Press, 1957. 180

[24] SERGIU HART AND ANDREU MAS-COLELL: A simple adaptive procedure leading to correlated equilibrium. *Econometrica*, 68(5):1127–1150, 2000. [doi:10.1111/1468-0262.00153]. 182

[25] ADAM KALAI AND SANTOSH VEMPALA: Efficient algorithms for online decision problems. *J. Comput. System Sci.*, 71(3):291–307, 2005. [doi:10.1016/j.jcss.2004.10.016]. 180, 185, 187, 190, 195

[26] ROBERT KLEINBERG, GEORGIOS PILLOURAS, AND ÉVA TARDOS: Multiplicative updates outperform generic no-regret learning in congestion games. In *Proc. 41st STOC*. ACM Press, 2009. [doi:10.1145/1536414.1536487]. 183

[27] ELIAS KOUTSOUPIAS AND CHRISTOS PAPADIMITRIOU: Worst-case equilibria. *Comput. Sci. Rev.*, 3(2):65–69, 2009. [doi:10.1016/j.cosrev.2009.04.003]. 180

[28] N. LITTLESTONE AND M. K. WARMUTH: The weighted majority algorithm. *Inform. and Comput.*, 108(2):212–261, 1994. [doi:10.1006/inco.1994.1009]. 180, 185, 187, 190, 195

[29] H. BRENDAN MCMAHAN AND AVRIM BLUM: Online geometric optimization in the bandit setting against an adaptive adversary. In *Proc. 17th Ann. Conf. Learning Theory (COLT)*, number 3120 in LNCS, pp. 109–123. Springer, 2004. [Springer:38p3f4h61cgy7gg7]. 180, 182, 185

[30] IGAL MILCHTAICH: Congestion games with player-specific payoff functions. *Games Econom. Behav.*, 13(1):111–124, 1996. [doi:10.1006/game.1996.0027]. 183

[31] IGAL MILCHTAICH: Generic uniqueness of equilibrium in large crowding games. *Math. Oper. Res.*, 25(3):349–364, 2000. [doi:10.1287/moor.25.3.349.12220]. 184

[32] ABRAHAM NEYMAN: Correlated equilibrium and potential games. *Internat. J. Game Theory*, 26(2):223–227, 1997. [doi:10.1007/BF01295851]. 182

[33] TIM ROUGHGARDEN: Intrinsic robustness of the price of anarchy. In *Proc. 41st STOC*, pp. 513–522. ACM Press, 2009. [doi:10.1145/1536414.1536485]. 183

[34] TIM ROUGHGARDEN AND ÉVA TARDOS: How bad is selfish routing? *J. ACM*, 49(2):236–259, 2002. [doi:10.1145/506147.506153]. 180

[35] TIM ROUGHGARDEN AND ÉVA TARDOS: Bounding the inefficiency of equilibria in nonatomic congestion games. *Games Econom. Behav.*, 47(2):389–403, 2004. [doi:10.1016/j.geb.2003.06.004]. 191, 192, 194

[36] WILLIAM H. SANDHOLM: Potential games with continuous player sets. *J. Econom. Theory*, 97(1):81–108, 2001. [doi:10.1006/jeth.2000.2696]. 182

[37] DAVID SCHMEIDLER: Equilibrium points of nonatomic games. *J. Stat. Phys.*, 7(4):295–300, 1973. [doi:10.1007/BF01014905]. 184

[38] EIJI TAKIMOTO AND MANFRED K. WARMUTH: Path kernels and multiplicative updates. *J. Mach. Learn. Res.*, 4:818, 2003. http://mlr.csail.mit.edu/papers/volume4/takimoto03a/takimoto03a.pdf. 180

[39] MARTIN ZINKEVICH: Online convex programming and generalized infinitesimal gradient ascent. In *Proc. 20th Intern. Conf. Mach. Learn. (ICML)*, pp. 928–936. AAAI Press, 2003. 180

[40] MARTIN ZINKEVICH: Theoretical guarantees for algorithms in multi-agent settings. Technical Report CMU-CS-04-161, Carnegie Mellon University, 2004. 182

AUTHORS

Avrim Blum
Professor
Department of Computer Science
Carnegie Mellon University
Pittsburgh, PA
avrim@cs.cmu.edu
http://www.cs.cmu.edu/~avrim

Eyal Even-Dar
Google Research
76 9th Avenue, New York, NY
evendar@google.com
http://sites.google.com/site/evendareyal/

Katrina Ligett
Postdoctoral Associate
Department of Computer Science
Cornell University
Ithaca, NY
katrina@cs.cornell.edu
http://www.cs.cornell.edu/~katrina

ABOUT THE AUTHORS

AVRIM BLUM received his Ph. D. at MIT in 1991 under Ron Rivest. His main research interests are machine learning theory, approximation algorithms, on-line algorithms, algorithmic game theory, and mechanism design.

EYAL EVEN-DAR graduated from Tel Aviv University in 2005 under the supervision of Yishay Mansour. After two years of postdoctoral research at the University of Pennsylvania under the supervision of Michael Kearns, he joined Google Research NY for three years. His research interests include learning theory and algorithmic game theory.

KATRINA LIGETT graduated from Carnegie Mellon in 2009; her advisor was Avrim Blum. Her interests include online learning, algorithmic game theory, and data privacy.