ROUTING WITHOUT REGRET: ON CONVERGENCE TO NASH EQUILIBRIA OF REGRET-MINIMIZING ALGORITHMS IN ROUTING GAMES[†]

AVRIM BLUM[∗], EYAL EVEN-DAR[†], AND KATRINA LIGETT[‡]

Abstract. There has been substantial work developing simple, efficient *no-regret* 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 no-regret 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 (formally, on most days, the cost of the flow 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 [2]. Our nonatomic results also apply to a more general class of games known as congestion games.

1. Introduction. There has been substantial work in learning theory and game theory on adaptive *no-regret* algorithms for problems of repeated decision-making. These algorithms have the property that in any online, repeated game 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 [19], 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 $O(\log N)$ in more recent exponential-weighting algorithms for this problem [23, 6, 16] (also called the problem of "combining expert advice"). Most recently, a number of algorithms have been developed for achieving such guarantees *efficiently* in many settings where the number of choices N is exponential in the natural description-length of the problem [21, 30, 31].

One specific setting where these efficient algorithms apply 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 simultaneously an adversary (or nature) chooses a set of edge costs $\{c_e^t\}_{e \in E}$. The edge costs are then revealed and the player pays the cost of its path. Even though the number of possible paths can be exponential in the size of the graph, no-regret algorithms exist (e.g., [21, 31]) that achieve running time and convergence rates (to the cost of the best fixed path in hindsight) which are polynomial in the size of the graph and the maximum edge cost. Moreover, a number of extensions [1, 24] 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 [22] 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, where the cost of an edge is a function of the amount of traffic using that edge [7, 8, 22, 27, 11]. Such work implicitly assumes that selfish individual behavior results in Nash equilibria.

In this work we consider the question: if all players in a routing game use no-regret algorithms to choose their paths each day, what can we say about the overall behavior of the system? In particular, the no-regret property (also called Hannan Consistency) can be viewed as a natural *definition* of well-reasoned self-interested behavior over time. Thus, if all players are adapting their behavior in such a way, can we say that the system as a whole will approach Nash equilibrium? Our main result is that in the Wardrop setting of multicommodity flow and infinitesimal agents, the flows will approach equilibrium in the sense that a $1 - \epsilon$ 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 one new parameter—the dependence on slope—is necessary. In addition, we give stronger results for special cases such as the case of n parallel links and also consider the finite-size (non-infinitesimal) load-balancing model of Azar [2]. Our results for nonatomic players also hold for a more general class of games called congestion games, although efficient regret-minimizing algorithms need not exist for the most general of these games.

One way our result can be viewed is as follows. No-regret algorithms are very compelling from the point of view of individuals: if you use a no-regret 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 routing model, so long as edge latencies have bounded slope, we can view Nash equilibria as not just a stable steady-state or the result of adaptive procedures specifically designed to find them, but in fact as the inevitable 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 follows strategies that are no-regret in the worst-case, as long as their behavior satisfies the no-regret 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 no-regret 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 §5, the no-regret property can be used 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 no-regret 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, no-regret algorithms when played against each other will approach a minimax optimal solution [17]. However, it is known that even in small 2-player *general-sum* games, no-regret 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 [32].

1.2. Regret and Correlated equilibria. It is known that certain algorithms such as that of Hart and Mas-Colell [20], as well as any algorithms satisfying the stronger property of "no internal regret" [15], have the property that the empirical distribution of play approaches a *correlated* equilibrium. 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 §2.4).

1.3. Related work. Fischer and Vöcking [13] 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 [14]. Most recently, Fischer, Raecke, and Vöcking [12] give 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* algorithms satisfying the no-regret 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 no-regret, we show they achieve convergence. In addition, because efficient no-regret algorithms exist even in the bandit setting where each agent gets feedback only about its own actions [1, 24], 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.

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 [25, 9]. The work of Goldberg [18] 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 [10] 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. [4] 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 no-regret.

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 multiset 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* determines 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 $(\int_{A_i} (f_i)^1, \int_{A_i} (f_i)^2, \ldots, \int_{A_i} (f_i)^{|\mathcal{P}_i|})$. A flow thus induces a distribution over paths, which we write for a path P in $\mathcal{P}_i$ as $f_P = (f_i)^P$ for P of type i. 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, and the cost incurred by a player is simply the latency of the path she plays.

We define $|E| = m$ and write n for the number of edges in the largest path in $\mathcal{P}$. We will assume all edge latency functions have range $[0, 1]$, so the latency of a path is always between 0 and n. 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., $\hat{f}_e = \frac{1}{T} \sum_{t=1}^{T} f_e^t$.

REMARK 2.1. *Network 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 *Nash equilibrium* if no user would prefer to reroute her traffic, given the existing flow.

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 $f_{P_1} > 0$, $\ell_{P_1}(f) \leq \ell_{P_2}(f)$.*

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. In addition, given our assumption that all latency functions are continuous and non-decreasing, one can prove the existence of Nash equilibria:

PROPOSITION 2.3. *(Schmeidler [29], generalization of Beckman et al. [3]) Every nonatomic congestion game 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 [26], generalization of Beckman et al. [3]) Distinct equilibria for a nonatomic congestion game have equal social cost.*

2.3. No-Regret 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,*

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

An online algorithm for selecting paths at each time step t is no-regret if, for any sequence of flows, the expected regret (over internal randomness in the algorithm) goes to 0 as T goes to infinity.

Here and in the rest of this paper, excluding §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 no-regret 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*

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

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 number of time steps required to get* $R(T) \leq \epsilon$.

For example, for the case of a routing game consisting of only two nodes and m parallel edges, exponential-weighting algorithms [23, 6, 16] give $T_\epsilon = O(\frac{1}{\epsilon^2}\log m)$. For general graphs, results of Kalai and Vempala yield $T_\epsilon = O(\frac{mn \log n}{\epsilon^2})$ [21]. For general graphs where an agent can observe only its path cost, results of Awerbuch and Kleinberg yield $T_\epsilon = \tilde{O}(\frac{n^7 m}{\epsilon^3})$ [1].

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 to some true Nash 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 to a true Nash 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 no-regret 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 [1, 24]). 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* ϵ-*Nash equilibrium if the average cost under this flow is within* ϵ *of the minimum cost paths under this flow, i.e.* $C(f) - \sum_{i=1}^{k} a_i \min_{P \in \mathcal{P}_i} \sum_{e \in P} \ell_e(f_e) \leq \epsilon$.

Note that Definition 2.8 implies that at most a $\sqrt{\epsilon}$ fraction of traffic can have more than a $\sqrt{\epsilon}$ incentive to deviate from their path, and as a result is very similar to the definition of (ϵ, δ)-Nash equilibria in [12].

We also are able to show that one can apply price-of-anarchy results to ϵ-Nash flows; we discuss this in §6.

We will begin by focusing on the *time-average* flow $\hat{f}$, showing that for no-regret 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 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 ϵ-Nash. To achieve bounds of this form, which we show in § 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 ϵ-Nash for *all* (sufficiently large) time-steps t, because no-regret 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. To see why this is necessary, 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 Nash 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 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 no-regret 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 $\frac{1}{3}$-Nash, because on each day, the cost of the flow f is at least $1/3$.

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. $\ell_e(\hat{f}_e) = \frac{1}{T} \sum_{t=1}^{T} \ell_e(f_e^t)$ for all e.

THEOREM 3.1. *Suppose the latency functions are linear. Then for $T \geq T_\epsilon$, the average flow $\hat{f}$ is ϵ-Nash, i.e.*

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

Proof. From the linearity of the latency functions, we have for all e, $\ell_e(\hat{f}_e) = \frac{1}{T} \sum_{t=1}^{T} \ell_e(f_e^t)$. 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

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

□

COROLLARY 3.2. *Assume that all latency functions are linear. In general routing games, if all agents use the Kalai-Vempala algorithm [21], the average flow converges*

to an ϵ-Nash equilibrium at $T_\epsilon = O(\frac{mn \log n}{\epsilon^2})$. On networks consisting of two nodes and m parallel links, if all agents use optimized "combining expert advice"-style algorithms (with each edge an expert), the average flow converges to an ϵ-Nash equilibrium at $T_\epsilon = O(\frac{\log m}{\epsilon^2})$.

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

4. Infinitesimal Users: General Latency Functions. The case of general latency functions is more complicated because the first and third transitions in the proof above do not apply. Here, the additive term depends on the maximum slope of any latency function.

THEOREM 4.1. *Let $\epsilon' = \epsilon + 2\sqrt{s\epsilon n}$. Then for general functions with maximum slope s, for $T \geq T_\epsilon$, the time-average flow is ϵ'-Nash, that is,*

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

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

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

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

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

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

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

Our goal in proving Theorem 4.1 is to show that (4.1) is not too much greater than (4.5). We will prove this as follows. We know that $(4.3) \leq \epsilon + (4.4)$ by the no-regret property and that $(4.2) \leq (4.3)$ by convexity. So, what remains to show is that (4.4) is not much greater than (4.5) and that (4.1) is not much greater than (4.2). We prove these in Lemmas 4.2 and 4.3 below.

LEMMA 4.2. *For general latency functions with maximum slope s, $(4.4) \leq \sqrt{s\epsilon n} + (4.5)$.*

Proof. First, observe that, because our latency functions are non-decreasing, the average latency of an edge must be less than or equal to the latency of that edge as seen by a random user on a random day. That is, for all e,

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

Define $\epsilon_e = \frac{1}{T}\sum_{t=1}^{T}\ell_e(f_e^t)f_e^t - \frac{1}{T}\hat{f}_e\sum_{t=1}^{T}\ell_e(f_e^t)$ to be the gap between the above two terms. Now, notice that the right-hand side of the above inequality, summed over all

edges, is precisely quantity (4.3). By the no-regret 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:

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

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

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,

$$\sum_{e \in E}\epsilon_e \leq \epsilon.$$

We now use this fact, together with the assumption of bounded slope, to show that edge latencies cannot be varying wildly over time. Specifically, we can rewrite the definition of ϵ_e as:

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

where we are using the fact that $\hat{f}_e = \frac{1}{T}\sum_{t=1}^{T}f_e^t$ and so $\frac{1}{T}\sum_{t=1}^{T}\ell_e(\hat{f}_e)(f_e^t - \hat{f}_e) = 0$.

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

$$|\ell_e(f_e^t) - \ell_e(\hat{f}_e)| \leq \sqrt{s\left(\ell_e(f_e^t) - \ell_e(\hat{f}_e)\right)\left(f_e^t - \hat{f}_e\right)}$$

for all e.

We then get

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

Using equation (4.6) above, this yields

$$(4.7) \qquad \frac{1}{T}\sum_{t=1}^{T}(\ell_e(f_e^t) - \ell_e(\hat{f}_e)) \leq \sqrt{s\epsilon_e}.$$

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 the fact that $\sum_i a_i \sum_{e \in P_i^*}\sqrt{s\epsilon_e} \leq \sqrt{s\epsilon n}$, we have

$$(4.5) + \sqrt{s\epsilon n} \;\geq\; \sum_{e \in P^*}\frac{1}{T}\sum_{t=1}^{T}\ell_e(f_e^t) \;\geq\; (4.4),$$

as desired.   □

LEMMA 4.3. *For general latency functions with maximum slope s, $(4.1) \leq \sqrt{s\epsilon n} + (4.2)$.*

Proof. Equation (4.7) above directly gives us

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

We then use the fact that $\hat{f}_e \leq 1$ for all e to obtain the desired result. $\square$

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

Proof. [of Theorem 4.1]

Since $(4.3) \leq \epsilon + (4.4)$ by Assumption 2.7, and $(4.2) \leq (4.3)$ by convexity, we get

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

as desired. $\square$

COROLLARY 4.4. *Let $\epsilon' = \epsilon + 2\sqrt{s\epsilon n}$. Assume that all latency functions are positive, non-decreasing, and continuous, with maximum slope s. In general routing games, if all agents use the Kalai-Vempala algorithm [21], the average flow converges to an ϵ'-Nash equilibrium at $T_\epsilon = O(\frac{mn\log n}{\epsilon^2}) = O(\frac{mn^3 s^2 \log n}{\epsilon'^4})$. On networks consisting of two nodes and m parallel links, if all agents use optimized "combining expert advice"-style algorithms, the average flow converges to an ϵ'-Nash equilibrium at $T_\epsilon = O(\frac{\log m}{\epsilon^2}) = O(\frac{n^2 s^2 \log m}{\epsilon'^4})$.*

Once again we remark that not only have we proved that the average flow approaches 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 ϵ-Nash equilibrium.

THEOREM 5.1. *In general routing games with general latency functions with maximum slope s, for all but a $(ms^{1/4}\epsilon^{1/4})$ fraction of time steps up to time T_ϵ, f^t is a $(\epsilon + 2\sqrt{s\epsilon n} + 2m^{3/4}s^{1/4}\epsilon^{1/4})$-Nash flow. We can rewrite this as: for all but an ϵ' fraction of time steps up to T_ϵ, f^t is an ϵ'-Nash flow for $\epsilon = \Omega\left(\frac{\epsilon'^4}{sm^4 + s^2 n^2}\right)$.*

Proof. Based on equation (4.6),

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

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

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

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

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

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 $m^{3/4}s^{1/4}\epsilon^{1/4}$. Also on most time steps, the cost incurred by flow f^t differs from the cost incurred by flow $\hat{f}$ by at most $m^{3/4}s^{1/4}\epsilon^{1/4}$. Thus since $\hat{f}$ is an $(\epsilon + 2\sqrt{s\epsilon n})$-Nash equilibrium, f^t is an $(\epsilon + 2\sqrt{s\epsilon n} + 2m^{3/4}s^{1/4}\epsilon^{1/4})$-Nash equilibrium on all but a $ms^{1/4}\epsilon^{1/4}$ fraction of time steps. $\square$

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

This demonstrates that no-regret algorithms are a reasonable, stable response in a network setting: if a player knows that all other players are using no-regret algorithms, there is no strategy that will significantly improve her expected cost on more than a small fraction of days. By using a no-regret 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.

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 $m^{3/4}s^{1/4}\epsilon^{1/4}$ of the cost of the best path given $\hat{f}$, which is at most $2\sqrt{s\epsilon n}$ worse than the cost of the best path in hindsight. Combining this with the no-regret property achieved by each user gives the desired result. $\square$

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 show that any ϵ-Nash equilibrium in a congestion game is closely related to a true Nash equilibrium in a related congestion game. This allows us to apply Price of Anarchy results for the congestion game to the regret-minimizing players in the original game. In our second result in this section, we give an argument paralleling that of Roughgarden and Tardos [28] that directly relates the costs of regret-minimizing users to the cost of the social optimum.

THEOREM 6.1. *If f is an ϵ-Nash equilibrium flow for a nonatomic congestion game Γ, then $C(f) \leq \frac{\rho}{1-\sqrt{\epsilon}}(C(OPT) + s\sqrt{\epsilon}n + \sqrt{\epsilon} + \epsilon)$, where OPT is the min cost flow and ρ is the price of anarchy in a related congestion game Γ' with the same class of latency functions as Γ but with additive offsets.*

For example, Theorem 6.1 implies that for linear latency functions, an ϵ-Nash flow f will have cost at most $\frac{4/3}{1-\sqrt{\epsilon}}(C(OPT) + \sqrt{\epsilon}(n + 1) + \epsilon)$. Note that for regret minimizing players, Theorem 6.3 below improves this to $\frac{4}{3}C(OPT) + \epsilon$.

The proof idea for this theorem is as follows: For every nonatomic congestion game Γ and flow f at 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 a Nash 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 ϵ-Nash equilibrium on Γ, then at most a $\sqrt{\epsilon}$ fraction of users are experiencing costs more than $\sqrt{\epsilon}$ 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. Call the remaining $(1 - \delta)$ users f_2.

Then $C(f \text{ on } \Gamma) \leq C(f_2 \text{ on } \Gamma_2) + \sqrt{\epsilon} C(f \text{ on } \Gamma) + \epsilon$. We can rewrite this as

$$C(f \text{ on } \Gamma) \leq \frac{1}{1 - \sqrt{\epsilon}} \left(C(f_2 \text{ on } \Gamma_2) + \epsilon \right).$$

We now construct an alternate congestion game Γ_3 (not necessarily a routing game, even if the original game was a routing game) such that f_2 interpreted on Γ_3 is a Nash equilibrium. To do this, we create a new edge for each commodity, and include that edge in every allowable path for that 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_2 on Γ_2. The maximum cost we need to assign to any entry edge in order to achieve this is $\sqrt{\epsilon}$, since we already removed all users paying more than $\sqrt{\epsilon}$ plus the cost of the best path available to them. Thus $C(f_2 \text{ on } \Gamma_2) \leq C(f_2 \text{ interpreted on } \Gamma_3)$, so we have

$$C(f \text{ on } \Gamma) \leq \frac{1}{1 - \sqrt{\epsilon}} \left(C(f_2 \text{ interpreted on } \Gamma_3) + \epsilon \right).$$

Define ρ to be the price of anarchy of the new congestion game Γ_3 when played with one unit of flow. The price of anarchy when played with less than one unit of flow can only be lower. Thus, defining $OPT_\alpha(H)$ to be the min-cost flow of size α in game H, we have

$$C(f \text{ on } \Gamma) \leq \frac{\rho}{1 - \sqrt{\epsilon}} \left(C(OPT_{1-\delta}(\Gamma_3)) + \epsilon \right).$$

Since we added at most $\sqrt{\epsilon}$ to the cost of any solution in going from Γ_2 to Γ_3, this gives

$$C(f \text{ on } \Gamma) \leq \frac{\rho}{1 - \sqrt{\epsilon}} \left(C(OPT_{1-\delta}(\Gamma_3) \text{ interpreted on } \Gamma_2) + \sqrt{\epsilon} + \epsilon \right),$$

and since $OPT_{1-\delta}(\Gamma_2)$ is the min-cost flow of size $(1 - \delta)$ on Γ_2,

$$C(f \text{ on } \Gamma) \leq \frac{\rho}{1 - \sqrt{\epsilon}} \left(C(OPT_{1-\delta}(\Gamma_2)) + \sqrt{\epsilon} + \epsilon \right),$$

We now must quantify the amount by which the cost of $OPT_{1-\delta}$ on Γ_2 could exceed the cost of OPT_1 on Γ. Since the cost of any edge in Γ_2 is at most $s\sqrt{\epsilon}$ more than the cost of that edge in Γ, this gives

$$C(f \text{ on } \Gamma) \leq \frac{\rho}{1 - \sqrt{\epsilon}} \left(C(OPT) + s\sqrt{\epsilon}n + \sqrt{\epsilon} + \epsilon \right).$$

□

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 $4/3$ [28].

We can also directly characterize the costs incurred by regret-minimizing players without going through the intermediate step of analyzing ϵ-Nash flows by arguing from scratch paralleling the Price of Anarchy proofs of Roughgarden and Tardos [28].

DEFINITION 6.2. *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 the marginal social cost $\ell_e^*(\xi) = \ell_e(\xi) + \xi \cdot \ell_e'(\xi)$, $\lambda \in [0,1]$ 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 6.3. *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 asymptotically at most $\alpha(\mathcal{L})$ (which is the Price of Anarchy bound given by Roughgarden and Tardos [28]).*

Proof. Let f^* be an optimal action distribution and $f_1, \ldots, f_T$ 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 minimizing property,

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

and thus

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

which gives us

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

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

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

as desired.   □

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 $\ell_e(\hat{f}_e \setminus i) = \frac{1}{T} \sum_{t=1}^{T} \ell_e(f_e^t \setminus i)$, 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 best expert algorithm. Then for all users, for all $T \geq O(\frac{\log m}{\epsilon^2})$,*

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

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 $\frac{1 - w_i}{m}$. Since user i is using an optimized best expert algorithm and the maximal latency is 1, we have that

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

$$\leq \frac{1 - w_i}{m} + \sqrt{\frac{\log m}{T}}$$

$$\leq \frac{1 - w_i}{m} + \epsilon$$

where the last inequality holds for $T \geq O(\frac{\log m}{\epsilon^2})$.   □

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

8. Conclusions. In this paper, we consider the question: if each player in a routing game (or more general congestion game) uses a no-regret strategy, will behavior converge to a Nash equilibrium, and under what conditions and in what sense? Our main result is that in the setting of multicommodity flow and infinitesimal agents, a $1 - \epsilon$ fraction of the daily flows are at 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 [21], we get a bound of $O(\frac{mn \log n}{\epsilon^2})$ on the number of time steps needed for the time-average flow to reach an ϵ-Nash equilibrium in the linear case, but $O(\frac{mn^3 \log n}{\epsilon^4})$ 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 Nash equilibria can also be applied to the results of regret-minimizing behavior. Recent work [5] 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 Nash equilibria may be hard to find.

REFERENCES

[1] B. AWERBUCH AND R. KLEINBERG, *Adaptive routing with end-to-end feedback: Distributed learning and geometric approaches*, in Proceedings of the 36th ACM Symposium on Theory of Computing, 2004.

[2] Y. AZAR, *On-line Load Balancing Online Algorithms - The State of the Art*, Springer, 1998, ch. 8, pp. 178–195.

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

[4] P. BERENBRINK, T. FRIEDETZKY, L. A. GOLDBERG, P. GOLDBERG, Z. HU, AND R. MARTIN, *Distributed selfish load balancing*, in Proceedings of the 17th Annual ACM-SIAM Symposium on Discrete Algorithms(SODA), 2006.

[5] A. BLUM, M. T. HAJIAGHAYI, K. LIGETT, AND A. ROTH, *Regret minimization and the price of total anarchy*. Unpublished manuscript.

[6] N. CESA-BIANCHI, Y. FREUND, D. HELMBOLD, D. HAUSSLER, R. SCHAPIRE, AND M. WARMUTH, *How to use expert advice*, Journal of the ACM, 44 (1997), pp. 427–485.

[7] A. CZUMAJ, P. KRYSTA, AND B. VÖCKING, *Selfish traffic allocation for server farms*, in Proceedings of the 34th Symposium on Theory of Computing, 2002, pp. 287–296.

[8] A. CZUMAJ AND B. VÖCKING, *Tight bounds on worse case equilibria*, in Proceedings of the Thirteenth Annual ACM-SIAM Symposium on Discrete Algorithms, 2002, pp. 413–420.

[9] E. EVEN-DAR, A. KESSELMAN, AND Y. MANSOUR, *Convergence time to nash equilibria*, in 30th International Conference on Automata, Languages and Programming (ICALP), 2003, pp. 502–513.

[10] E. EVEN-DAR AND Y. MANSOUR, *Fast convergence of selfish rerouting*, in Proceedings of the Sixteenth Annual ACM-SIAM Symposium on Discrete Algorithms, 2005, pp. 772–781.

[11] A. FABRIKANT, A. LUTHRA, E. MANEVA, C. H. PAPADIMITRIOU, AND S. SHENKER, *On a network creation game*, in Proceedings of the twenty-second annual symposium on Principles of distributed computing (PODC), ACM Press, 2003, pp. 347–351.

[12] S. FISCHER, H. RAECKE, AND B. VÖCKING, *Fast convergence to wardrop equilibria by adaptive sampling methods*, in Proceedings of 38th ACM Symposium on Theory of Computing (STOC), 2006.

[13] S. FISCHER AND B. VÖCKING, *On the evolution of selfish routing*, in Proceedings of the 12th European Symposium on Algorithms (ESA), 2004, pp. 323–334.

[14] ———, *Adaptive routing with stale information*, in Proceedings of the 24th Annual ACM SIGACT-SIGOPS Symposium on Principles of Distributed Computing (PODC), 2005.

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

[16] Y. Freund and R. Schapire, *A decision-theoretic generalization of on-line learning and an application to boosting*, J. Comput. System Sci., 55 (1997), pp. 119–139.

[17] Y. Freund and R. Schapire, *Adaptive game playing using multiplicative weights*, Games and Economic Behavior, 29 (1999), pp. 79–103.

[18] P. Goldberg, *Bounds for the convergence rate of randomized local search in multiplayer games, uniform resource sharing game*, in Proceedings of the Twenty-Third PODC, 2004, pp. 131–144.

[19] J. Hannan, *Approximation to Bayes risk in repeated play*, in Contributions to the Theory of Games, M. Dresher, A. Tucker, and P. Wolfe, eds., vol. III, Princeton University Press, 1957, pp. 97–139.

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

[21] A. Kalai and S. Vempala, *Efficient algorithms for on-line optimization*, in Proceedings of the The 16th Annual Conference on Learning Theory, 2003, pp. 26–40.

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

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

[24] H. McMahan and A. Blum, *Online geometric optimization in the bandit setting against an adaptive adversary*, in Proceedings of the 17th Annual Conference on Learning Theory (COLT), 2004, pp. 109–123.

[25] I. Milchtaich, *Congestion games with player-specific payoff functions*, Games and Economic Behavior, 13 (1996), pp. 111–124.

[26] I. Milchtaich, *Generic uniqueness of equilibrium in large crowding games*, Mathematics of Operations Research, 25 (2000), pp. 349–364.

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

[28] T. Roughgarden and E. Tardos, *Bounding the inefficiency of equilibria in nonatomic congestion games*, Games and Economic Behavior, 47 (2004), pp. 389–403.

[29] D. Schmeidler, *Equilibrium points of nonatomic games*, Journal of Statistical Physics, 7 (1973), pp. 295–300.

[30] E. Takimoto and M. K. Warmuth, *Path kernels and multiplicative updates*, in Proceedings of the 15th Annual Conference on Computational Learning Theory, Lecture Notes in Artificial Intelligence, Springer, 2002.

[31] M. Zinkevich, *Online convex programming and generalized infinitesimal gradient ascent*, in Proceedings of the 20th International Conference on Machine Learning, 2003, pp. 928–936.

[32] ———, *Theoretical guarantees for algorithms in multi-agent settings*, Tech. Rep. CMU-CS-04-161, Carnegie Mellon University, 2004.