Multiplicative Updates Outperform Generic No-Regret Learning in Congestion Games

Extended Abstract

Robert Kleinberg*
Cornell University
rdk@cs.cornell.edu

Georgios Piliouras†
Cornell University
gpil@cs.cornell.edu

Éva Tardos‡
Cornell University
eva@cs.cornell.edu

ABSTRACT

We study the outcome of natural learning algorithms in atomic congestion games. Atomic congestion games have a wide variety of equilibria often with vastly differing social costs. We show that in almost all such games, the well-known multiplicative-weights learning algorithm results in convergence to pure equilibria. Our results show that natural learning behavior can avoid bad outcomes predicted by the price of anarchy in atomic congestion games such as the load-balancing game introduced by Koutsoupias and Papadimitriou, which has super-constant price of anarchy and has correlated equilibria that are exponentially worse than any mixed Nash equilibrium.

Our results identify a set of mixed Nash equilibria that we call weakly stable equilibria. Our notion of weakly stable is defined game-theoretically, but we show that this property holds whenever a stability criterion from the theory of dynamical systems is satisfied. This allows us to show that in every congestion game, the distribution of play converges to the set of weakly stable equilibria. Pure Nash equilibria are weakly stable, and we show using techniques from algebraic geometry that the converse is true with probability 1 when congestion costs are selected at random independently on each edge (from any monotonically parametrized distribution). We further extend our results to show that players can use algorithms with different (sufficiently small) learning rates, i.e. they can trade off convergence speed and long term average regret differently.

*Supported by NSF grants CCF-0643934 and CCF-0729102, an Alfred P. Sloan Foundation Fellowship, and a Microsoft Research New Faculty Fellowship.

†Supported by NSF grant CCF-0325453.

‡Supported by NSF grants CCF-0325453, CCF-0729006, ONR grant N00014-98-1-0589, and a Yahoo! Research Alliance Grant.

STOC'09, May 31–June 2, 2009, Bethesda, Maryland, USA.

Categories and Subject Descriptors

F.2 and J.4 [**Analysis of Algorithms and Social and Behavior Sciences**]:

General Terms

Algorithms, Economics, Theory

Keywords

Learning Theory, Game Theory, Price of Anarchy

1. INTRODUCTION

Congestion games have been studied extensively in computer science, often from the standpoint of analyzing the price of anarchy: the ratio of solution quality achieved by the worst-case Nash equilibrium versus the optimal solution. Koutsoupias and Papadimitriou [21] introduced the price of anarchy in the context of a load-balancing game studying the makespan objective function. Congestion games both with makespan and with social welfare objective function are well understood; see the surveys by Vöcking [38] and Roughgarden [31]. Analyzing the inefficiency of Nash equilibria provides useful information about the solution quality achieved by selfish players once they reach an equilibrium, but does not provide a model of how selfish players behave, and it says little about whether selfish players will coordinate on an equilibrium, nor which equilibria they are likely to coordinate on if the game has more than one.

Learning has been suggested as a natural model of players' behavior in games. No-regret learning algorithms suggest simple and plausible adaptive procedures where players do not have regrets about their past behavior in some precise sense, which makes them natural candidates to model selfish play. The theory of learning in games has studied the limit of repeated play when all players use such no-regret learning strategies. The resulting equilibrium concepts (variants of correlated equilibrium) typically have worst-case equilibria that fall short of the solution quality achieved by Nash equilibrium. Thus, although researchers studying certain classes of games have proven that the outcome of no-regret learning matches the price-of-anarchy bound [3, 4, 5, 32], there are broad classes of games in which there is a large gap between the predictions arising from analysis of Nash equilibria versus analysis of learning processes (correlated equilibria).

To illustrate this point, consider the load balancing game introduced by Koutsoupias and Papadimitriou [21]. In this game, there are n balls and n bins. Each ball chooses a bin and experiences a cost equal to the number of balls that choose the same bin. If the objective function is the makespan (i.e. maximum load in a bin) then an optimal solution places each ball in a separate bin. These solutions coincide with the pure Nash equilibria of the game. However, there are many other mixed Nash equilibria, including the fully mixed equilibrium in which each ball chooses a bin uniformly at random and the expected makespan is $\Theta(\log n/\log\log n)$. The same game with a non-linear congestion cost results in an arbitrarily high price of anarchy both in the makespan and in the average congestion cost models: the same symmetric fully mixed equilibria are expected to have some bin with congestion $\Theta(\log n/\log\log n)$, which can have arbitrarily high congestion cost. Worse yet, the game can have correlated equilibria that are exponentially worse than the worst mixed Nash equilibrium. In the simple case of linear edge costs the expected makespan is $\Theta(\sqrt{n})$. The ratio is even worse if the congestion cost is decreasing, as in the cost-sharing games where a bin with x players costs $1/x$ to each player. As before, pure equilibria coincide with the social optimal solution, which in this case has a total cost of 1, while the fully mixed Nash equilibrium is expected to use a constant fraction of all bins, and hence to cost $\Theta(n)$.

Our question.

We focus on understanding the quality of outcomes reached by players using "realistic" learning algorithms. Restricting attention to realistic learning algorithms is consistent with our goal of modeling realistic player behavior, and it is also necessary because within the class of *all* no-regret learning algorithms one can find contrived algorithms whose distribution of play converges to an arbitrary (e.g. worst-case) correlated equilibrium of any game[1], as well as contrived algorithms[2] whose distribution of play converges into the set of Nash equilibria of any game.

Our results.

We consider a class of learning dynamics, called the aggregate monotonic selection (AMS) dynamics, that extends the multiplicative weights learning algorithm [2, 22] (also known as Hedge [13]) to players whose learning rates may differ and may vary over the strategy space. We show that if players use AMS dynamics to adjust their strategies, then game play converges to a subset of Nash equilibria, which we call *weakly stable equilibria*. These are mixed Nash equilibria $(\sigma_1, \ldots, \sigma_n)$ with the additional property that each player i remains indifferent[3] between the strategies in the support of σ_i whenever *any other single player* j modifies its mixed strategy to any pure strategy in the support of σ_j. Pure Nash equilibria are weakly stable, and we show that the converse is true with probability 1 when congestion costs are selected at random independently on each edge (e.g. from any monotonically parameterized distribution). Thus, our results imply that in congestion games, learning via AMS dynamics surpasses the Price of Total Anarchy (as defined by Blum et al. [4]) and also the Price of Anarchy for mixed Nash equilibria.

Intuitively, players using this learning algorithm are able to steer clear of undesirable mixed Nash equilibria because of symmetry-breaking properties resulting from the inherent randomness in the algorithm. In a load-balancing game, for instance, when one player randomly chooses a machine it causes others to reduce their probability of choosing that machine in the future, and this asymmetry is self-reinforcing. We justify this intuition by showing that the symmetry-breaking is implied by spectral properties of a matrix that is defined at each mixed equilibrium of the game, and that these spectral properties in turn imply weak stability in the sense defined earlier.

We show that a discrete version of the process with a small amount of added noise at each step follows the solution of the differential equation closely enough to converge to the set of ν-stable equilibria, a generalization of weakly stable equilibria in which the Jacobian is allowed to have eigenvalues whose real part is at most ν, an arbitrarily small positive real number.

Our techniques.

Our technique is based on analyzing a differential equation expressing a continuum limit of the multiplicative-weights update process, as the multiplicative factor approaches 1 and time is renormalized accordingly. For the case of the Hedge algorithm this differential equation turns out to be identical to the asymmetric replicator dynamic studied in evolutionary game theory. More generally, the differential equation is the extension of the replicator dynamic called aggregate monotonic selection (AMS) dynamics introduced by Samuelson and Zhang [33].

As a first step in analyzing the dynamical system, we show that every flow line of the differential equation converges to the set of fixed points. As in prior work on replicator dynamics in potential games (e.g. [1]) we do this by proving that the potential function [26] associated with the congestion game is a *Lyapunov function* for any AMS dynamics; that is, it is non-increasing along flow lines of the differential equation and is strictly decreasing except at fixed points.

The set of fixed points of the differential equation includes all the mixed Nash equilibria (not just the weakly stable or pure ones) as well as some mixed strategy profiles that are not Nash equilibria at all. To see which fixed points arise as limit points of the flow starting from a generic initial condition, we need to distinguish between stable and unstable fixed points. For a fixed point p that is not a Nash equilibrium, it is not hard to argue that p is unstable. For Nash equilibria, we prove that the dynamical-systems notion of stability — a fixed point where the Jacobian matrix has no complex eigenvalues in the open right half-plane — implies our game-theoretic notion of weakly stable equilibria. To do this, we prove that when the Jacobian matrix is restricted to the subspace of strategies played with positive probability, this submatrix $\mathfrak{J}$ is nilpotent, i.e. its only eigenvalue is 0. It is easy to see that $\mathrm{Tr}(\mathfrak{J}) = 0$. The difficult part here lies in showing that we also have $\mathrm{Tr}(\mathfrak{J}^2) \geq 0$, which uses Steele's [37] non-symmetric version of the Efron-Stein inequality (see Theorem 3.6). The fact that $\mathrm{Tr}(\mathfrak{J}) = 0$ and $\mathrm{Tr}(\mathfrak{J}^2) \geq 0$, together with the absence of complex eigenvalues in the right

[1] We provide a proof for this statement in the full version of the paper. A well-known result of a similar flavor, but using calibrated learning rather than no-regret play, is due to Foster and Vohra [10].

[2] For examples, see the discussion of related work below.

[3] Note that the definition does not require each of the strategies in the support of σ_i to remain a best response after player j modifies σ_j. This modification may cause player i to prefer a strategy lying outside the support of σ_i.

half-plane, implies that all eigenvalues are 0. This in turn entails a linear relation on "two-player marginal cost terms" that implies our game-theoretic notion of weak stability.

Clearly all pure equilibria are weakly stable. To show the opposite is true with probability 1 when congestion costs are selected at random independently on each edge, we in fact prove a stronger statement — the existence of a non-pure weakly stable equilibrium implies the vanishing of a non-zero polynomial function of the edge costs — using techniques from algebraic geometry. To illustrate the idea, consider the special case of load balancing games with monotonic cost functions, where weakly stable Nash equilibria are "almost pure" in the sense that each machine has at most one randomizing player using it. (For example, in a load-balancing game with one player and 2 identical machines, any mixed strategy of the one player is stable.) When congestion costs are selected at random, such almost-pure mixed Nash equilibria cannot exist: the probability that two machines have the same cost is 0. To extend this reasoning to general congestion games we need to use more sophisticated techniques, as weakly stable Nash equilibria need not be almost-pure. However, a weakly stable mixed Nash equilibrium must satisfy many polynomial constraints (Nash constraints, insensitivity to one player's change). Using an algebraic-geometric version of Sard's Theorem, we show that congestion costs that have stable mixed Nash equilibria satisfy a nontrivial polynomial equation. In the case of load-balancing games this is a linear relation (two machines having equal cost), but in general it will be a higher-degree polynomial.

Related work.

No-regret learning algorithms have long been studied in the context of adaptive game playing. There are a number of simple and natural families of no-regret learning algorithms such as the regret matching of Hart and Mas-Colell [18] and the multiplicative weights or Hedge algorithm, introduced by Freund and Schapire [13], which generalizes the well-known weighted majority algorithm of Littlestone and Warmuth [22]. In general games these algorithms converge into the set of coarse correlated equilibria, but not necessarily into the set of Nash equilibria as we prove here for congestion games and the Hedge algorithm. Our decision to analyze the Hedge algorithm in this work was motivated by the algorithm's ubiquitousness in learning theory and other areas of theoretical computer science [2]; the issue of whether similar results can be obtained for other algorithms such as regret matching [18] is an interesting open question.

There are a number of other learning-like processes, such as fictitious play [27], calibrated forecasting [11, 9], regret testing [12, 15], and others (see [19]) whose play is known to converge to Nash equilibria in some games. Some of these results use simple stochastic processes that are tantamount to stochastic search for a pure Nash equilibrium (e.g. [19]) but are not regret-minimizing, while others use complicated adaptive procedures satisfying calibration properties that are closely related to the no-regret property (e.g. [9]). Unlike these works, our goal is not to discover uncoupled adaptive procedures for finding a Nash equilibrium but to analyze the behavior of a particular simple, realistic, and well-known adaptive procedure.

The dynamics of repeated play in congestion games has been studied in non-learning-theoretic models such as sink equilibria and selfish rerouting [6, 16, 25]. For *nonatomic* congestion games (i.e., games with infinitesimal players) Fischer, Räcke and Vöcking [8, 7] and Blum, Even-Dar, and Ligett [3] considered learning dynamics in the setting of multicommodity flow (or more generally congestion games) and showed that the dynamics converges to a Wardrop equilibrium. Fischer, Räcke and Vöcking [8, 7] consider a replication-exploration protocol, while Blum, Even-Dar, and Ligett [3] show that in this setting, if each player uses any no-regret strategy the behavior will approach the Wardrop equilibrium. The setting of atomic congestion games studied here is much more intricate because the game typically has many Nash equilibria forming a disconnected set with many components, and we need to distinguish between the stable and unstable ones.

Blum, Hajiaghayi, Ligett, and Roth [4] defined the price of total anarchy and showed that in a number of games, the known bounds on the price of anarchy extend also to the price of total anarchy: the worst case bound on the quality of coarse correlated equilibria to the optimum outcome is already achieved over pure Nash equilibria. Roughgarden [32] extends this to a wider class of games. Our results complement these by showing that if one assumes players use a specific, standard no-regret algorithm (namely, Hedge, or AMS) rather than arbitrarily bad no-regret algorithms, one obtains much stronger guarantees about the distribution of play in atomic congestion games: it converges to a weakly stable Nash equilibrium.

The replicator dynamic and other differential equations are studied in evolutionary game theory [23, 35], which also considers associated notions of stability such as evolutionarily stable states (ESS) and neutrally stable states (NSS). The book of Fudenberg and Levine [14] provides an excellent survey on these topics.For each of the three main steps in our analysis, we identify here the most closely related work in evolutionary game theory. **In congestion games, the replicator dynamic converges to its fixed point set.** Proofs of this theorem and generalizations, using the game's potential function as a Lyapunov function, have appeared in many prior works, e.g. [1, 20, 27, 34]. Our short proof conveniently provides a quantitative bound on the rate of decrease of the potential. **A weakly stable fixed point of the replicator dynamic is a weakly stable equilibrium.** Our notion of weakly stable equilibrium is similar to, but weaker than, the notion of neutrally stable states. It is known that neutrally stable states of a game are Lyapunov stable points of the replicator dynamic (hence weakly stable), but our line of attack requires the converse, which is not true in general [39]. Here, we are able to deduce this converse by introducing a weaker game-theoretic notion of stability. **In almost every game, the weakly stable equilibria coincide with the pure Nash equilibria.** The most closely-related result in evolutionary game theory is Ritzberger and Weibull's theorem [29] that every *asymptotically stable* fixed point set of the replicator dynamics is not contained in the relative interior of any face of the mixed strategy polytope, and hence contains a pure equilibrium. Their result applies to general games and not just congestion games. It is weaker than ours in two key respects: it assumes a stronger stability property (that need not hold at the weakly stable fixed points considered in our analysis) and derives a weaker conclusion (the asymptotically stable set contains a pure Nash equilibrium but may also contain other equilibria).

2. MODEL AND PRELIMINARIES

Congestion games.

We assume the reader is familiar with the notions of a strategic-form game, pure and mixed strategies, and pure and mixed Nash equilibria.

Congestion games [30] are non-cooperative games in which the utility of each player depends only on the player's strategy and the number of other players that either choose the same strategy, or some strategy that "overlaps" with it. Formally, a *congestion game* is defined by a tuple of the form $(N; E; (\mathcal{S}_i)_{i\in N}; (c_e)_{e\in E})$ where N is the set of *players*, E is a set of *facilities* (also known as *edges* or *bins*), and each player i has a set $\mathcal{S}_i$ of subsets of E ($\mathcal{S}_i \subseteq 2^E$). Each pure strategy $S_i \in \mathcal{S}_i$ is a set of edges (a *path*), and c_e is a cost (negative utility) function associated with facility e. Given a pure strategy profile $S = (S_1, S_2, \ldots, S_N)$, the cost of player i is given by $c_i(S) = \sum_{e\in S_i} c_e(k_e(S))$, where $k_e(S)$ is the number of players using e in S. Congestion games admit a *potential function* $\Phi(S) = \sum_{e\in E}\sum_{j=1}^{k_e(S)} c_e(j)$, and pure Nash equilibria are local minima of $\Phi(S)$ [30].

Multiplicative update (Hedge) and the AMS dynamic.

An online learning algorithm is an algorithm for choosing a sequence of elements of some fixed set of *actions*, in response to an observed sequence of cost functions mapping actions to real numbers. The t^{th} action chosen by the algorithm may depend on the first $t-1$ observations but not on any later observations. The regret of an online learning algorithm is defined as the maximum over all input instances of the expected difference in payoff between the algorithm's actions and the best action. If this difference is guaranteed to grow sublinearly with time, we say it is a *no-regret learning algorithm* or that it is *Hannan consistent* [17].

The family of regret minimizing algorithms that we study in this paper are called *aggregate monotonic selection (AMS) dynamics* [29, 33], and they generalize the weighted majority algorithm introduced by Littlestone and Warmuth [22] and the Hedge algorithm of Freund and Schapire [13]. These algorithms maintain a vector of n probabilities for the n available actions, and at each round they sample an action according to this distribution. (Initially, the probabilities are all equal.) After each round, AMS dynamics updates the weights multiplicatively, favoring actions that exhibit low cost. The update is governed by a global parameter $\varepsilon > 0$, and a state-dependent parameter $\beta = \beta(i,p)$ which is determined by the player, i, and by the profile of mixed strategies, p, representing the current state of *every* player's learning algorithm. We can interpret the parameter $\varepsilon\beta(i,p)$ as the learning rate of player i. If the cost of action a at time t is $c^a(t)$ then the weights are updated by multiplying with $(1-\varepsilon\beta)^{c^a(t)}$ and then renormalizing:

$$p_a^{t+1} = \frac{p_a^t(1-\varepsilon\beta)^{c^a(t)}}{\sum_j p_j^t(1-\varepsilon\beta)^{c^j(t)}} \tag{1}$$

We will assume throughout that $\forall i,p \;\; 0 < \beta(i,p) \le 1$. For example, the algorithm Hedge(ε) is obtained by setting $\beta(i,p) = 1$ for all i,p. Hedge(ϵ) is not a no-regret algorithm, but it is an ε-regret algorithm, and well known tricks (halving ε as the algorithm proceeds) can be employed so that it becomes Hannan consistent.

3. THE CONTINUOUS-TIME PROCESS

In this section we define and analyze the continuous-time version of the Hedge (multiplicative weights) algorithm. We consider the algorithm's update rule for probabilities, and derive the limit as $\varepsilon \to 0$, a first-order differential equation (ODE) known as the replicator dynamic.

Setting up the differential equation.

The continuous-time process we want to analyze in this section arises as the limit of the update rule (1) as $\varepsilon \to 0$. The cost of a strategy S_i for a player i is a random variable $C^i(S_i) = \sum_{e\in S_i} c_e(1+K_i(e))$, where $K_i(e)$ denotes $|\{j : j \neq i, e \in S_j\}|$, the number of players other than i that use edge e. Taking the derivative of the above update with respect to ε, and substituting $\varepsilon = 0$, we get a differential equation using the random variables $C^i(S_i)$. Taking expectation, and using the notation $c^i(e) = \mathbb{E}(c_e(1+K_i(e))$ for the expected cost of edge e for player i, we get that the expected cost is $c^i(R) = \mathbb{E}(C^i(R)) = \sum_{e\in R} c^i(e)$ and the expected update in the probabilities is the following differential equation, $\dot{p} = \xi(P)$, where:

$$\xi_{iR} \;\; = \;\; \beta(i,p)p_{iR}\left(\sum_{\bar{R}} p_{i\bar{R}}(c^i(\bar{R}) - c^i(R))\right). \tag{2}$$

Fixed points and convergence to fixed points.

By inspecting (2), one can see that the fixed points of the ODE are the distributions where players mix between options with equal (but not necessarily minimum) cost.

THEOREM 3.1. *Probability distributions p_i for $i \in N$ form a fixed point of (2) if and only if $\forall\, i$, all strategies $R \in \mathcal{S}_i$ with $p_{iR} > 0$ have the same expected cost $c^i(R)$ for player i.*

Before we study fixed points further, we want to establish that the solutions of the differential equation converge to fixed points (e.g., do not cycle). To do this we will consider the standard potential function Φ of the congestion game. This is analogous to the proof of Monderer and Shapley [27] who show that in games when players have identical interest (and also in potential games that are strategically equivalent to such games for fictitious play and also for the AMS dynamic), repeated play with fictitious play converges to the set of Nash equilibria. Recall that the standard potential function of a congestion game is defined as $\sum_e \sum_{k=1}^{K_e} c_e(k)$, where we use the notation $K_e = |\{j : e \in S_j\}|$ for a pure strategy profile $S = (S_1, \ldots, S_n)$. We will denote the expected value of this function by Ψ, and we will use it as our potential function.

THEOREM 3.2. *The time derivative $\dot{\Psi}$ of the potential function is 0 at fixed points, and negative at all other points. In fact, $-\dot{\Psi} \ge \frac{||\xi||_1^2}{2\beta_{1..i}(p)}$, where $\beta_{1..i}(p)$ denotes $\sum_{i=1}^n \beta(i,p)$.*

PROOF. Let Φ_{-i} denote the potential function of the game without player i, i.e. $\Phi_{-i}(S_{-i}) = \sum_{e\in E}\sum_{j=1}^{k_e(S_{-i})} c_e(j)$, where $k_e(S_{-i})$ denotes the number of players using edge e in strategy profile S_{-i}. It is well known that the actual potential function Φ satisfies $\Phi(S) = \Phi_{-i}(S_{-i}) + C^i(S_i)$ for all strategy profiles S. Note that

$$\Psi = \mathbb{E}[\Phi_{-i}(S_{-i})] + \mathbb{E}[C^i(S_i)] = \mathbb{E}[\Phi_{-i}(S_{-i})] + \sum_R p_{iR}c^i(R).$$

The terms $\mathbb{E}[\Phi_{-i}(S_{-i})]$ and $c^i(R)$ don't depend on player i's mixed strategy, so $\partial\Psi/\partial p_{iR} = c^i(R)$. Now,

$$\begin{aligned}\dot{\Psi} &= \sum_{i,R} \left(\frac{\partial \Psi}{\partial p_{iR}}\right) \dot{p}_{iR} \\ &= \sum_i \beta(i,p) \sum_{R,\bar{R}} p_{iR} p_{i\bar{R}} \left(c^i(R)c^i(\bar{R}) - c^i(R)^2\right) \\ &= -\sum_i \beta(i,p) \sum_{R<\bar{R}} p_{iR} p_{i\bar{R}} \left(c^i(R) - c^i(\bar{R})\right)^2.\end{aligned}$$

The second line is produced by substituting $\frac{\partial\Psi}{\partial p_{iR}}$ and $\dot{p}_{iR}$ and rearranging the resulting terms. In deriving the last line, we have assumed that the strategy set of player i is totally ordered, and we have paired the terms on the preceding line corresponding to $R, \bar{R}$ and $\bar{R}, R$. Finally, to bound $-\dot{\Psi}$ from below in terms of $\|\xi\|_1$ (which is needed for the discrete-time analysis) we use the Cauchy Schwartz inequality:

$$(-2\beta_{1..i}(p))\dot{\Psi} = \beta_{1..i}(p) \sum_{i,R,\bar{R}} \beta(i,p) p_{iR} p_{i\bar{R}} (c^i(R) - c^i(\bar{R}))^2$$

$$= \left[\sum_{i,R,\bar{R}} \beta(i,p) p_{iR} p_{i\bar{R}}\right] \left[\sum_{i,R,\bar{R}} \beta(i,p) p_{iR} p_{i\bar{R}} (c^i(R) - c^i(\bar{R}))^2\right]$$

$$\geq \left[\sum_{i,R,\bar{R}} \beta(i,p) p_{iR} p_{i\bar{R}} \left|c^i(R) - c^i(\bar{R})\right|\right]^2 \geq \|\xi\|_1^2 \quad \square$$

Unstable fixed points and the Jacobian.

We will use the notion of stability from dynamical systems. In the neighborhood of a fixed point p_0 the ODE can be approximated by $\dot{p} \approx J(p - p_0)$, where J is the matrix of partial derivatives, the Jacobian. A fixed point of a dynamical system is said to be unstable ([28]) if the Jacobian matrix has an eigenvalue with positive real part.

For an ODE represented by a vector field $\xi(x)$ the entry J_{ij} of the Jacobian is the partial derivative of the i-th coordinate $\xi_i(x)$ in the direction of x_j. Our ODE has coordinates p_{iR} corresponding to a player i and a strategy R, and our vector field ξ is defined in (2). Observe in (2) that ξ_{iR} is the product of $\beta(i,p)$ with a term that vanishes at a fixed point. When we take the partial derivative in any direction, using the product rule we find that the derivative of $\beta(i,p)$ vanishes, as it is multiplied by 0 at a fixed point. Now let us examine the entries of the Jacobian matrix case by case. For directions corresponding to the same player we get

$$\begin{aligned}\frac{\partial \xi_{iR}}{\partial p_{iR}} &= \beta(i,p) \sum_{\bar{R}} p_{i\bar{R}} (c^i(\bar{R}) - c^i(R)) \\ \frac{\partial \xi_{iR}}{\partial p_{i\bar{R}}} &= \beta(i,p) p_{iR} (c^i(\bar{R}) - c^i(R))\end{aligned}$$

where the sum in the partial derivative should be for $\bar{R} \neq R$, but can equally well be understood to be for all $\bar{R}$. The second expression is for $\bar{R} \neq R$.

Finally, taking a derivative in the direction p_{jQ} for $j \neq i$ involves understanding how the cost $c^i(R)$ depends on the probability p_{jQ}. We get that $mc_e^{ij} := \mathbb{E}(c_e(2 + K_{ij}(e))) - \mathbb{E}(c_e(1 + K_{ij}(e)))$ is the coefficient of p_{jQ} in the cost $c^i(e)$. Using the notation $mc^{ij}(A) = \sum_{e \in A} mc_e^{ij}$ we get that:

$$\frac{\partial \xi_{iR}}{\partial p_{jQ}} = \beta(i,p) p_{iR} \sum_{\bar{R}} p_{i\bar{R}} (mc^{ij}(\bar{R} \cap Q) - mc^{ij}(R \cap Q))$$

Note that the marginal cost depends on the probability distributions of players other than i and j, but does not depend i and j. In particular, we have that $mc_e^{ij} = mc_e^{ji}$ for every edge e.

THEOREM 3.3. *The Jacobian matrix is expressed by the equations above.*

LEMMA 3.4. *A fixed point p that is stable corresponds to a Nash equilibrium.*

PROOF. Consider a fixed point p of the ODE that is not a Nash equilibrium. If a fixed point is not a Nash equilibrium, than there is a player i and a strategy R with $p_{iR} = 0$ that has $\beta(i,p) \sum_{\bar{R}} p_{i\bar{R}} (c^i(\bar{R}) - c^i(R)) = \lambda > 0$. The unit vector w^{iR} with a 1 in the (i,R) coordinate is a left eigenvector of J with $w^{iR} J = \lambda w^{iR}$, hence J has a positive eigenvalue. $\square$

At a fixed point p that is a Nash equilibrium, let $\mathfrak{J}$ be the submatrix of the Jacobian restricted to the subset of strategies played with positive probability ($p_{iR} > 0$).

LEMMA 3.5. *For any right eigenvector w of the matrix $\mathfrak{J}$ that satisfies $\mathfrak{J}w = \lambda w$, the vector w° extending w with 0 values to the remaining coordinates is an eigenvector of the full Jacobian, J, with eigenvalue λ.*

First note that the trace of $\mathfrak{J}$ is 0, which follows directly from the definition of fixed point, and definition of $\mathfrak{J}$, as all diagonal entries are 0. To help establish that the submatrix $\mathfrak{J}$ (as well as J) has an eigenvalue with positive real part we will prove that the trace of $\mathfrak{J}^2$ is nonnegative.

THEOREM 3.6. *Consider a fixed point of the ODE, and let $\mathfrak{J}$ be the submatrix of the Jacobian defined above. Then* $\mathrm{Tr}(\mathfrak{J}^2) \geq 0$, *and in fact it is equal to*

$$\sum_{i,j} \beta(i,p)\beta(j,p) \sum_{R<\bar{R}, Q<\bar{Q}} p_{iR} p_{i\bar{R}} p_{jQ} p_{j\bar{Q}} (M_{i,j}^{R,\bar{R},Q,\bar{Q}})^2,$$

where $M_{i,j}^{R,\bar{R},Q,\bar{Q}}$ is defined to be $mc^{ij}(R \cap Q) - mc^{ij}(R \cap \bar{Q}) - mc^{ij}(\bar{R} \cap Q) + mc^{ij}(\bar{R} \cap \bar{Q})$.

To prove this simply check that each term of the form $mc^{ij}(R \cap Q) mc^{ij}(\bar{R} \cap \bar{Q})$ occurs in the two expressions with the same multiplier. We are ready to show that if a fixed point is stable, then $\mathfrak{J}$ must have 0 as its only eigenvalue.

THEOREM 3.7. *For a stable fixed point p, all eigenvalues of the submatrix $\mathfrak{J}$ of the Jacobian corresponding to the coordinates with $p_{iR} > 0$ are zero. Also, for all players i, j and all strategies $R, \bar{R}, Q, \bar{Q}$ played with positive probability by the players i and j respectively, we must have $mc_{ij}(R \cap Q) - mc_{ij}(R \cap \bar{Q}) - mc_{ij}(\bar{R} \cap Q) + mc_{ij}(\bar{R} \cap \bar{Q}) = 0$.*

PROOF. For any fixed point the sum of the eigenvalues, $\mathrm{Tr}(\mathfrak{J})$, is zero, hence if $\mathfrak{J}$ has no eigenvalues with positive real part then all eigenvalues must be pure imaginary. But in this case $\mathrm{Tr}(\mathfrak{J}^2)$ is nonpositive, as it is the sum of squares of pure imaginary numbers. We know that $\mathrm{Tr}(\mathfrak{J}^2) \geq 0$ and hence it must equal zero. Hence all eigenvalues of $\mathfrak{J}$ must equal zero, as claimed. $\square$

Using the condition derived in above for stable fixed points of the ODE, we can connect the notion of stable for the dynamical system to our game theoretic notion of weakly stable, in the sense defined in the introduction, that each player i remains indifferent between the strategies in the support of σ_i whenever *any other single player* j modifies its mixed strategy to any pure strategy in the support of σ_j.

THEOREM 3.8. *If a Nash equilibrium is stable for the dynamical system then it is a weakly stable Nash equilibrium.*

PROOF. We have proved that if a Nash equilibrium is stable for the dynamical system, then for all players i, j and all strategies $R, \bar{R}, Q, \bar{Q}$ played with positive probability by the players i and j, respectively, $mc_{ij}(R \cap Q) - mc_{ij}(\bar{R} \cap Q) = mc_{ij}(R \cap \bar{Q}) - mc_{ij}(\bar{R} \cap \bar{Q})$. Using j's mixed strategy σ_j to take a weighted average over all $\bar{Q}$, we get the claim that if i is indifferent between its strategies R, R' when j randomizes, he remains indifferent when j plays only strategy Q. □

We have seen the solution of the ODE converges, and that fixed points of the dynamic system that are not weakly stable Nash equilibria are unstable for the dynamic system. Using the theory of differential equations [28], we can conclude that starting from a generic initial condition, the ODE converges to weakly stable Nash equilibria.

THEOREM 3.9. *From all but a measure 0 set of starting points, the solution of the ODE* (2) *converges to weakly stable Nash equilibria.*

4. WEAKLY STABLE EQUILIBRIA

If one fixes the set of players and facilities, and the strategy sets of each player of a congestion game — which we collectively denote as the game's "combinatorial structure" — the game itself is determined by the vector of edge costs $\vec{c}$, i.e., the vector whose components are the numbers $c_e(k)$ for every edge e and every possible load value k on that edge. One can thus identify the set of congestion games having a fixed combinatorial structure with the vector space $\mathbb{R}^N$ where N is equal to the number of pairs (e, k) for which $c_e(k)$ is defined. If one imposes other constraints such as non-negativity and monotonicity on the edge costs, then the set of games is identified with a convex subset of $\mathbb{R}^N$ rather than $\mathbb{R}^N$ itself.

Our goal in this section is to prove the following theorem.

THEOREM 4.1. *For almost every congestion game, every weakly stable equilibrium is a pure Nash equilibrium. In other words, the set of congestion games having non-pure weakly stable equilibria is a measure-zero subset of* $\mathbb{R}^N$.

In fact, we will prove the following stronger version of Theorem 4.1.

THEOREM 4.2. *There is a non-zero multivariate polynomial* W*, defined on* $\mathbb{R}^N$*, such that for every game with a non-pure weakly stable equilibrium, its edge costs satisfy the equation* $W(\vec{c}) = 0$.

This strengthening implies, for example, that among all the congestion games with a fixed combinatorial structure and with cost functions taking integer values between 0 and B, the fraction of such games having a non-pure weakly stable equilibrium tends to zero as $B \to \infty$.

Define an equilibrium to be *fully mixed* if it satisfies $p_{iR} > 0$ for every player i and every strategy R in that player's strategy set. Every mixed equilibrium $\vec{p}$ of a game is a fully mixed equilibrium of the subgame obtained by deleting the strategies that satisfy $p_{iR} = 0$. Since there are only finitely many such subgames, we can establish Theorems 4.1 and 4.2 by proving the corresponding statements about fully mixed weakly stable equilibria. Theorem 4.1 then follows because a finite union of measure-zero sets has measure zero, and Theorem 4.2 follows because the union of the zero-sets of polynomials $W_1, W_2, \ldots, W_k$ is the zero-set of their product $W_1 W_2 \ldots W_k$.

Let X be the set of pairs $(\vec{p}, \vec{c})$ such that $\vec{p}$ is a fully mixed weakly stable equilibrium of the game with edge costs $\vec{c}$, let $f : X \to \mathbb{R}^N$ be the function that projects such a pair $(\vec{p}, \vec{c})$ to its second component, $\vec{c}$, and let $Y \subseteq \mathbb{R}^N$ be the set $f(X)$, i.e. the set of games having a fully mixed weakly stable equilibrium. To prove that Y has measure zero and is contained in the zero-set of a nontrivial polynomial, we will first prove a "local, linearized version" of the same statement. Lemma 4.3 below asserts, roughly[4], that for every point $x \in X$, with tangent space $T_x X$, the projection of $T_x X$ to $\mathbb{R}^N$ has dimension strictly less than N. (And thus, the image of $T_x X$ in $\mathbb{R}^N$ has measure zero and is contained in the zero-set of a nontrivial linear function.) Theorems 4.1 and 4.2 are then obtained using general theorems that allow global conclusions to be deduced from these local criteria. To obtain Theorem 4.1 we work in the category of differentiable manifolds and apply Sard's Theorem [24]: the set of critical values of a differentiable function has measure zero. To obtain Theorem 4.2 we work in the category of algebraic varieties and apply an "algebraic geometry version" of Sard's Theorem ([36], Lemma II.6.2.2): if $f : X \to Y$ is a regular map of varieties defined over a field of characteristic 0, and f is surjective, then there exists a nonempty open set $V \subseteq X$ such that the differential $d_x f$ is surjective for all $x \in V$.[5] As an aid to the reader unfamiliar with algebraic geometry we summarize some standard definitions in Section 6.

Linearized version of Theorems 4.1 and 4.2.

Each of the expressions $c^i(R), mc^{ij}(R \cap Q)$ used in Section 3 actually refers to a polynomial — in fact, a multilinear polynomial — in the variables p_{**} and $c_*(*)$, because the probability of any given pure strategy profile being sampled is a multilinear polynomial in the p_{**} variables, and the cost of any edge, in any given pure strategy profile, is one of the variables $c_*(*)$. Let the polynomial equation $A_i^{R,R'} = 0$ express the fact that player i is indifferent between strategies R and R', i.e. $c^i(R) - c^i(R') = 0$. By definition of a weakly stable equilibrium we must have $mc^{ij}(R \cap Q) - mc^{ij}(R' \cap Q) = \sum_{\bar{Q}} p_{j\bar{Q}}(mc^{ij}(R \cap \bar{Q}) - mc^{ij}(R' \cap \bar{Q}))$ for any Q, hence we get $M_{i,j}^{R,R',Q,Q'} = mc^{ij}(R \cap Q) - mc^{ij}(R' \cap Q) - mc^{ij}(R \cap Q') +$

[4]The actual statement is more complicated because X may have singularities, so the tangent space $T_x X$ may be ill-defined. To deal with this, what we actually show is that X can be partitioned into finitely many nonsingular subsets $X_1, X_2, \ldots, X_k$ — possibly of different dimensions — such that for every $x \in X_i$ $(1 \leq i \leq k)$, the projection of $T_x X_i$ to $\mathbb{R}^N$ has dimension strictly less than N.

[5]Actually the Lemma as stated in [36] requires the field to be algebraically closed, and it requires the variety X to be nonsingular. We describe how to work around these technical difficulties in the proof of Theorem 4.2.

$mc^{ij}(R' \cap Q') = 0$ for all R, R', Q, Q'. Finally let $P_i = 0$ encode $\sum_{R \in \mathcal{S}_i} p_{iR} = 1$. In earlier sections of this paper, we have seen that all of these equations must hold when $\vec{p}$ is a fully mixed weakly stable equilibrium of the game with edge costs $\vec{c}$. In other words, if I denotes the polynomial ideal generated by $\{A_i^{R,R'}\} \cup \{M_{i,j}^{R,R',Q,Q'}\} \cup \{P_i\}$, then the set of fully mixed weakly stable equilibria, X, is contained in the algebraic variety $V(I)$ defined by the vanishing of all the polynomials in I.

LEMMA 4.3. *The ideal I contains a polynomial $F \in \mathbb{R}[\vec{p}, \vec{c}]$ that satisfies:*

1. *$\partial F / \partial p_{iR} \in I$ for all variables p_{iR}.*
2. *$\partial F / \partial c_e(k) \notin I$ for at least one variable $c_e(k)$. In fact, there exists an edge e such that the sum of the partial derivatives in directions $c_e(k)$, for $k = 1, \ldots, n$, is in $1 + I$.*

PROOF. Fix any player i, and fix any two strategies R, R' for that player. For all players $j \neq i$ fix a strategy $Q_j^0 \in \mathcal{S}_j$. Consider the polynomial

$$F = A_i^{R,R'} + \sum_{j \neq i} \left[mc^{ij}(R' \cap Q_j^0) - mc^{ij}(R \cap Q_j^0) \right] P_j.$$

We can show that it satisfies $\partial F / \partial p_{jQ} = M_{i,j}^{R,R',Q,Q_j^0}$ if $j \neq i$ and $Q \neq Q_j^0$, and otherwise $\partial F / \partial p_{jQ} = 0$. This confirms property (1). Property (2) follows from the formula

$$\forall e \in R \setminus R' \quad \sum_{k=1}^{n} \frac{\partial F}{\partial c_e(k)} = \prod_{j \neq i} (P_j + 1) +$$

$$+ \sum_{j \neq i} \left[\sum_{k=1}^{n} \left(\frac{\partial mc^{ij}(R \cap Q_j^0)}{\partial c_e(k)} - \frac{\partial mc^{ij}(R' \cap Q_j^0)}{\partial c_e(k)} \right) \right] P_j$$

The detailed derivations will be included in the full version of the paper. □

Measure-theoretic and algebraic conclusions.

In order to apply Sard's Theorem, we need to work with a smooth manifold, whereas the set X of fully mixed weakly stable equilibria may have singularities. However, we know that X is contained in the affine algebraic variety $V(I)$, so we may use the fact that X decomposes into a union of finitely many nonsingular varieties. (A precise statement and proof of this decomposition theorem are contained in Lemma 6.1.)

PROOF OF THEOREM 4.1. By Lemma 6.1, the set X of fully mixed weakly stable equilibria can be covered by finitely many smooth manifolds $X_1, \ldots, X_m$, so it suffices to prove that each of them projects to a measure-zero subset of $\mathbb{R}^N$. If F is the polynomial defined in Lemma 4.3, then for every $x \in X_j$ $(1 \leq j \leq m)$ and every tangent vector $v \in T_x X_j$, we have $\nabla F(x) \cdot v = 0$, by Lemma 6.1. Recalling that X is a subset of the vector space $\mathbb{R}^{M+N}$, where M is the combined number of strategies in all players' strategy sets, and N is the combined number of pairs (e, k) such that the edge cost $c_e(k)$ is well-defined, then we may write $v = (v_p, v_c)$, where v_p denotes the first M components of v (corresponding to the "probability coordinates") and v_c denotes the last N components of v (corresponding to the "edge cost coordinates"). Recalling that every polynomial in I vanishes at x, properties (1)-(2) of Lemma 4.3 imply that the first M coordinates of $\nabla F(x)$ vanish whereas the last N coordinates do not. Thus, the equation $\nabla F(x) \cdot v = 0$ imposes a nontrivial linear constraint on the vector v_c. This implies that the tangent space $T_x X_j$ projects to a proper linear subspace of $\mathbb{R}^N$. Since x was arbitrary, we have proven that the differential of the projection map $X_j \to \mathbb{R}^N$ has rank less than N at every point of X_j. By Sard's Theorem, the image of X_j in $\mathbb{R}^N$ has measure zero. □

In fact, the technique used to prove Theorem 4.1 actually allows us to establish a stronger theorem, in which we consider congestion games whose edge cost functions are drawn from a specified class of cost functions. Let us define a *smooth, monotonically parameterized class of cost functions* to be a collection of functions $c_\gamma : \{1, 2, \ldots, n\} \to \mathbb{R}$ parameterized by a vector of real numbers $\gamma = (\gamma_0, \gamma_1, \ldots, \gamma_d) \in U$ for some open subset $U \subseteq \mathbb{R}^{d+1}$, satisfying the following two properties: (1) for all $k \in \{1, 2, \ldots, n\}$, the function $\gamma \mapsto c_\gamma(k)$ is a smooth function, which we will denote by $h_k(\gamma)$; (2) for all $k \in \{1, 2, \ldots, n\}$ and all $\gamma \in U$, the partial derivative $\partial c_\gamma(k) / \partial \gamma_0$ is strictly positive. For example, edge costs that are specified by degree-d polynomials $c_e(k) = \gamma_0 + \gamma_1 k + \ldots + \gamma_d k^d$ constitute a smooth, monotonically parameterized class of cost functions. The following generalization of Theorem 4.1 can be proven using essentially the same technique.

THEOREM 4.4. *Suppose that for each edge e of a congestion game we are given a smooth, monotonically parameterized class of cost functions with parameter $\gamma^e \in U_e$. The set of congestion games having non-pure weakly stable equilibria is a measure-zero subset of $\prod_e U_e$.*

Turning now from measure-theoretic statements to algebraic ones, we present the proof of Theorem 4.2.

PROOF OF THEOREM 4.2. We work over the field $\mathbb{C}$ of complex numbers in order to apply theorems about varieties over an algebraically closed field; in the last sentence of the proof we will translate the result back to the field $\mathbb{R}$. Let I be the ideal defined in Section 4, and let $X_{\mathbb{C}} = V(I)$, the zero-set of I over $\mathbb{C}$. By Lemma 6.1, $X_{\mathbb{C}}$ can be covered by finitely many nonsingular varieties $X_1, \ldots, X_m$. We will prove that each of them projects to a proper closed subset of the affine space $\mathbb{C}^N$. If F is the polynomial defined in Lemma 4.3, then as in the proof of Theorem 4.1, the equation $\nabla F(x) \cdot v = 0$ implies[6] that $\mathrm{rank}(d_x f) < N$, where f denotes the projection map $X_j \to \mathbb{C}^N$. However, if the set $f(X_j)$ were dense in $\mathbb{C}^N$ then by ([36], Lemma II.6.2.2) we would have $\mathrm{rank}(d_x f) = N$ for all x in a nonempty open subset of X_j. It follows that $f(X_j)$ is contained in a proper closed subset of $\mathbb{C}^N$ for all j, hence $f(X_{\mathbb{C}}) = \cup_j f(X_j)$ is also contained in a proper closed subset of $\mathbb{C}^N$. This means there is a nonzero polynomial $P \in \mathbb{C}[c_1, c_2, \ldots, c_N]$ that vanishes on $f(X_{\mathbb{C}})$. If $\bar{P}$ denotes the complex conjugate of P, then $W = P\bar{P}$ is a polynomial in $\mathbb{R}[c_1, c_2, \ldots, c_N]$ that vanishes on $f(X)$. □

5. THE DISCRETE-TIME PROCESS

In this section we give an approximate version of our main result using the discrete-time learning process with a small

[6]See Lemma 6.2 in section 6 for a rigorous derivation of this step using the standard algebraic-geometry definitions of tangent spaces and differentials.

amount of added noise. The added noise is a useful artifact for proving general convergence guarantees in the discrete time setting.

Let $p'(t)$ denote the mixed strategy profile at time t before we introduce the noise and let $p(t)$ express the resulting distribution. The original strategy distribution $p'(t)$ is scaled down by a factor of $1-\varepsilon^2$ and then shifted by mixing it with the uniform distribution. The updated vector has the form $(1-\varepsilon^2)p'(t)+\varepsilon^2\vec{1}$. Next, a player is chosen at random (i.e. player i) who in turns chooses uniformly and independently two of her strategies R and $\bar{R}$. If both $p_{iR}, p_{i\bar{R}}$ are greater than ε then the player proceeds to move $\varepsilon/2$ from the first strategy to the second. If that is not the case, then she proceeds to move merely $\varepsilon^2/2$. This update is always possible because of our first mixing step.

Approximately Stable Equilibria.

We need to extend the definitions of stable and unstable fixed points to approximate stability. For the dynamical systems definition, we define a *ν-stable fixed point*, for any real number $\nu \geq 0$, to be a fixed point at which the Jacobian has no eigenvalues whose real part is greater than ν.

For simplicity, we assume in this section that $\beta(i,p) = 1$ for all players i and all probability distributions p. This section outlines a proof of the following result.

THEOREM 5.1. *The discrete-time learning process with a small amount of added noise satisfies the following guarantee for all congestion games with and non-decreasing cost functions.*

> *For all $\nu > 0$ there exists an $\epsilon > 0$, and for all $\epsilon < \epsilon_0$ there exists T_0, so that when all players are using $Hedge(\epsilon)$ to optimize their strategies, then for all times $T > T_0$ with probability at least $1-\nu$, the mixed strategy profile is a ν-stable equilibrium point in all but νT of the first T steps of the history of play.*

To use this theorem in the context of games, we need to relate our game-theoretic notion of weak stability to this notion of ν-stable fixed points. We say that a Nash equilibrium p is $\bar{\nu}$-weakly stable, if the following holds for each pair of players i and j. Suppose we randomly sample a strategy of j with probability distribution p_j, and assume that j plays this sampled strategy Q and all other players $k \neq i, j$ play with the given probability distribution p_k. Now sampling two strategies R and $\bar{R}$ of i with probability distribution p_i, the expected difference of payoffs of player i between strategies R and $\bar{R}$ is at most $\bar{\nu}$. Note that 0-weakly stable is exactly our definition of weakly stable. We will show that the learning process spends almost all the time near a ν-stable equilibrium point. To conclude the same for $\bar{\nu}$-weakly stable Nash equilibrium, we need to show that ν-stable implies $\bar{\nu}$-weakly stable.

THEOREM 5.2. *For every $\bar{\nu} > 0$ there is a $\nu > 0$ so that if a Nash equilibrium p is ν-stable for the dynamical system, then it is a $\bar{\nu}$-weakly stable Nash equilibrium.*

PROOF. Assume p is ν-stable, that is, for all eigenvalues λ of the Jacobian J we have $\Re(\lambda) \leq \nu$. If there are N players and a total of M strategies, this implies $\mathrm{Tr}(\mathfrak{J}^2) \leq \nu^2 M$. We know from Theorem 3.6 that this trace is equal to

$$\sum_{i,j} \sum_{R<\bar{R},Q<\bar{Q}} p_{iR}p_{i\bar{R}}p_{jQ}p_{j\bar{Q}}(M_{i,j}^{R,\bar{R},Q,\bar{Q}})^2 \tag{3}$$

where $M_{i,j}^{R,\bar{R},Q,\bar{Q}} = mc_{ij}(R\cap Q) - mc_{ij}(R\cap\bar{Q}) - mc_{ij}(\bar{R}\cap Q) + mc_{ij}(\bar{R}\cap\bar{Q})$, and recall that we are assuming here that $\beta(i,p) = 1$ for all i and p for simplicity.

Now consider the change in the cost of strategy R for player i when player j selects a strategy Q. This change is $mc_{ij}(R\cap Q) - \sum_{\bar{Q}} p_{j\bar{Q}} mc_{ij}(R\cap\bar{Q})$. The difference in cost between strategies R and $\bar{R}$ is then

$$\sum_{\bar{Q}} p_{j\bar{Q}}(mc_{ij}(R\cap Q) - mc_{ij}(R\cap\bar{Q}) - $$
$$- mc_{ij}(\bar{R}\cap Q) + mc_{ij}(\bar{R}\cap\bar{Q})) = \sum_{\bar{Q}} p_{j\bar{Q}} M_{i,j}^{R,\bar{R},Q,\bar{Q}}.$$

Taking the expectation defining the $\bar{\nu}$-weakly stable Nash equilibrium, the sum that we need to bound is

$$\sum_{R,\bar{R},Q,\bar{Q}} p_{iR}p_{i\bar{R}}p_{jQ}p_{j\bar{Q}} M_{i,j}^{R,\bar{R},Q,\bar{Q}}. \tag{4}$$

Restricting the sum of squares in (3) to only the (i,j) pair and using the Cauchy-Schwarz inequality we can bound the sum in (4) by $2\nu\sqrt{M}$, which establishes the theorem with $\bar{\nu} = 2\nu\sqrt{M}$. $\square$

Overview of the Discrete-Time Analysis.

In deriving Theorem 5.1 from the foregoing continuous-time analysis, we must address several sources of error: the noise introduced by the players' random sampling, the evolution of mixed strategies in discrete jumps rather than along continuous flow lines, and the approximation error resulting from treating ε as infinitesimally small when estimating the coefficients of the vector field ξ. Resolving these issues requires careful manipulations of Taylor series, the details of which can be found in the full version of the paper, but it is also possible to distinguish a few main ideas which constitute a road map for this stage of the proof.

Amortizing unstable steps.
Our analysis of the stochastic process $p(t)$ distinguishes three types of time steps: *potential-diminishing* steps in which the expected decrease in Ψ is at least $\Omega(\varepsilon^2)$, *stable* steps in which $p(t)$ is near a ν-stable equilibrium point, and *unstable* steps in which $p(t)$ is near an equilibrium point which is not ν-stable. To deal with unstable steps, we show that a sufficiently long time window starting at an unstable step t will contain (with high probability) many more potential-diminishing steps than unstable steps. Amortizing the change in Ψ over the entire time window, we can show that the expected potential decrease is $\Omega(\varepsilon^2)$. Thus, the entire time history $t = 1, 2, \ldots$ can be broken up into *good stages* consisting of a single stable step, and *bad stages* of bounded length such that the expected potential decrease during a bad stage is $\Omega(\varepsilon^2)$. Since Ψ can only decrease by a bounded amount over the entire history of play, we may easily conclude that with high probability, the good stages vastly outnumber the bad stages in any sufficiently long time history.

Balancing error terms by completing the square.
The most involved step in the preceding outline is the proof

that every sufficiently long time window which begins with an unstable step is likely to contain many more potential-diminishing steps than unstable steps. The difficulty is that the rate at which an unstable fixed point p^0 repels points at distance ρ from p^0 (along an unstable direction) is $O(\rho^2)$. This second-order effect is offset by a second-order correction term arising from our approximation of the multiplicative-update rule by the vector field ξ. To compare these two effects we use an analogue of "completing the square": instead of basing a Taylor expansion at the fixed point p^0 we choose a nearby basepoint p^1, resulting in a Taylor expansion whose leading-order term has an unambiguous sign reflecting the system's tendency to move away from p^1 in the repelling direction.

6. ALGEBRAIC GEOMETRY REVIEW

In this section, we prove some standard facts from algebraic geometry that are needed in Section 4.

LEMMA 6.1. *If $I \subseteq \mathbb{R}[x_1, x_2, \ldots, x_n]$ then the variety $Z = V(I)$ is the union of finitely many subsets $Z_1, Z_2, \ldots, Z_m$, each of which is a nonsingular quasi-affine algebraic variety[7], and therefore also a smooth manifold. For any polynomial $P \in I$ and any point $z \in Z_j$ $(1 \le j \le m)$, the gradient vector ∇P at z is orthogonal to the entire tangent space $T_z Z_j$ of the manifold Z_j.*

PROOF. To prove the existence of the decomposition into nonsingular subsets, we induct on the dimension of Z. Every algebraic variety is the union of finitely many irreducible varieties ([36], Theorem I.3.1.1), so it suffices to prove the statement when Z is irreducible. A zero-dimensional irreducible variety is a point, so the base case is trivial. Assuming now that the theorem holds for all varieties of dimension less than d, let Z be an arbitrary irreducible variety of dimension d. The nonsingular points of Z form an open, hence quasi-affine, subset Z_1 ([36], Section II.1.4) and the singular points of Z form a closed proper subvariety ([36], Section II.1.4), whose dimension is strictly smaller than d because Z is irreducible ([36], Theorem I.6.1.1). By the induction hypothesis, the set of singular points of Z is the union of finitely many nonsingular quasi-affine algebraic varieties $Z_2, \ldots, Z_m$. This completes the induction.

A nonsingular quasi-affine algebraic variety over $\mathbb{R}$ is a smooth manifold ([36], Section II.2.3). For any $z \in Z_j$ $(1 \le j \le m)$ and any tangent vector $v \in T_z Z_j$, let $\gamma : [-1, 1] \to Z_j$ be any smooth parameterized curve in Z_j such that $\gamma(0) = z, \gamma'(0) = v$. Letting $h(t) = P(\gamma(t))$, we have

$$h'(0) = \nabla P(\gamma(0)) \cdot \gamma'(0) = \nabla P(z) \cdot v.$$

But $h(t) = 0$ for all t because P vanishes on Z_j. Thus $h'(0) = 0$, which establishes that $\nabla P(z) \cdot v = 0$ as claimed. □

We now present some definitions about tangent spaces and differentials, leading up to Lemma 6.2 below, which rigorously establishes a key step in the proof of Theorem 4.2.

If k is a field and $I \subset k[x_1, \ldots, x_n]$ is an ideal, then the affine algebraic variety $X = V(I)$ is the set of points $x \in k^n$ such that $P(x) = 0$ for all $P \in I$. (In algebraic geometry it is customary to denote k^n by $\mathbb{A}_k^n$ and to call if *affine n-space over k*.) The *coordinate ring* $k[X]$ is the quotient ring $k[x_1, \ldots, x_n]/I$. Note that every $P \in k[X]$ determines a well-defined function on X, because if $P \equiv Q \pmod{I}$ then $P(x) = Q(x)$ for all $x \in X$. Functions from X to k defined in this way are called *regular functions* on X.

An m-tuple of n-variate polynomials $f_1, \ldots, f_m$ collectively determine a *regular map* $f : \mathbb{A}_k^n \to \mathbb{A}_k^m$ that sends every point $(x_1, \ldots, x_n)$ to the point $(f_1(\vec{x}), \ldots, f_m(\vec{x}))$. Note that a regular map uniquely determines a homomorphism from $k[y_1, \ldots, y_m]$ to $k[x_1, \ldots, x_n]$ by mapping the polynomial y_i to f_i for each i. If $X = V(I), Y = V(J)$ are two affine algebraic varieties in $\mathbb{A}_k^n, \mathbb{A}_k^m$, respectively, a *regular map* from X to Y is a mapping from the points of X to the points of Y obtained by restricting a regular map $f : \mathbb{A}_k^n \to \mathbb{A}_k^m$ to X. A regular map from X to Y uniquely determines a homomorphism from $k[Y]$ to $k[X]$.

If $p \in X = V(I) \subseteq \mathbb{A}_k^n$, then the degree-1 polynomials $x_1 - p_1, \ldots, x_n - p_n$ generate a maximal ideal of $k[X]$ denoted by $\mathfrak{m}_p$. Every function in $k[X]$ that vanishes at p belongs to $\mathfrak{m}_p$. If $\mathfrak{m}_p^2$ denotes the ideal generated by all products of pairs of elements of $\mathfrak{m}_p$ (i.e., all regular functions on X that vanish to second order at p) then the quotient $\mathfrak{m}_p/\mathfrak{m}_p^2$ is a k-vector-space called the *cotangent space of X at p*. The dual vector space $T_p X = (\mathfrak{m}_p/\mathfrak{m}_p^2)^*$ is the *tangent space of X at p*.

If $f : X \to Y$ is a regular map and $f(p) = q$, then the induced homomorphism $h : k[Y] \to k[X]$ sends $\mathfrak{m}_q$ to a subset of $\mathfrak{m}_p$, hence it induces a well-defined linear transformation from $\mathfrak{m}_q/\mathfrak{m}_q^2$ to $\mathfrak{m}_p/\mathfrak{m}_p^2$. The dual of this linear transformation is denoted by $d_p f : T_p X \to T_q Y$.

It's useful to consider how these ideas play out in the case of the projection map that sends $\mathbb{A}_k^{r+s}$ to $\mathbb{A}_k^s$ by mapping an $(r+s)$-tuple to its last s coordinates. Denoting the coordinate rings of $\mathbb{A}_k^{r+s}$ and $\mathbb{A}_k^s$ by $k[x_1, \ldots, x_r, y_1, \ldots, y_s] = k[\vec{x}, \vec{y}]$ and $k[y_1, \ldots, y_s] = k[\vec{y}]$, respectively, then the projection map induces the homorphism $k[\vec{y}] \to k[\vec{x}, \vec{y}]$ that simply includes $k[\vec{y}]$ as a subring of $k[\vec{x}, \vec{y}]$. If $X = V(I) \subseteq \mathbb{A}_k^{r+s}$ is an affine algebraic variety, then the regular map f defined by the function composition $X \hookrightarrow \mathbb{A}_k^{r+s} \to \mathbb{A}_k^s$ induces the ring homomorphism $k[\vec{y}] \to k[\vec{x}, \vec{y}]/I$ that maps a polynomial in the variables $y_1, \ldots, y_s$ to its equivalence class modulo I.

If $a = (b, c)$ is a point of $\mathbb{A}_k^{r+s}$ then a polynomial $P \in k[\vec{x}, \vec{y}]$ belongs to $\mathfrak{m}_a$ if and only if it vanishes at a, and P belongs to $\mathfrak{m}_a^2$ if and only if its Taylor expansion at a has vanishing degree-0 and degree-1 terms, i.e. $P(a) = 0$ and $\nabla P(a) = 0$. If $X = V(I) \subseteq \mathbb{A}_k^{r+s}$ and $a = (b, c)$ is a point of X, then we can also consider $\mathfrak{m}_a$ and $\mathfrak{m}_a^2$ as ideals in $k[X]$. (The notation doesn't distinguish between these two meanings.) In the case of $\mathfrak{m}_a$ the distinction is insignificant: an element of $\mathfrak{m}_a$ in $k[X]$ is an equivalence class (modulo I) of polynomials that vanish at a. But with $\mathfrak{m}_a^2$ we need to be more careful: a polynomial P represents an element of $\mathfrak{m}_a^2$ in $k[X]$ if and only if it is equivalent (modulo I) to another polynomial Q that satisfies $Q(a) = 0$ and $\nabla Q(a) = 0$.

So, for example, suppose that the gradient $\nabla F(a)$ of a polynomial $F \in I$ at a point $a = (b, c) \in X$ is a vector of the form $(0, 0, \ldots, 0, w_1, \ldots, w_s)$ whose first r components are 0 and whose last s components constitute a nonzero vector w. Let P denote the degree-1 polynomial $\sum_{j=1}^s w_j (y_j - c_j) \in k[y_1, \ldots, y_s]$. Then $P \in \mathfrak{m}_a^2$ because P is congruent, modulo I, to the polynomial $Q = P - F$, and Q satisfies $Q(a) = 0$ and $\nabla Q(a) = \nabla P(a) - \nabla F(a) = (0, \vec{w}) - (0, \vec{w}) = 0$. Thus, P is a nonzero element of $\mathfrak{m}_c/\mathfrak{m}_c^2$ that maps to zero in $\mathfrak{m}_a/\mathfrak{m}_a^2$; i.e. the induced map of cotangent spaces $\mathfrak{m}_c/\mathfrak{m}_c^2 \to \mathfrak{m}_a/\mathfrak{m}_a^2$ is not one-to-one. Dualizing this statement, we conclude that

[7]A quasi-affine algebraic variety is any variety isomorphic to an open subset of an affine algebraic variety

the differential $d_a f$ is not surjective, i.e. its rank is strictly less than s. Thus we have established the following lemma.

LEMMA 6.2. *Suppose $X = V(I) \subseteq \mathbb{A}_k^{r+s}$ is an affine algebraic variety and $f : X \to \mathbb{A}_k^s$ is the composition of the inclusion and projection maps $X \hookrightarrow \mathbb{A}_k^{r+s} \to \mathbb{A}_k^s$. If $F \in k[x_1, \ldots, x_r, y_1, \ldots, y_s]$ is a polynomial in I whose gradient $\nabla F(a)$ at a is a vector of the form $(0, \vec{w})$ for some nonzero $w \in k^s$, then the differential $d_a f : T_a X \to T_c \mathbb{A}_k^s$ is a linear transformation of rank strictly less than s.*

7. REFERENCES

[1] E. Akin. *The Geometry of Population Genetics*. Springer-Verlag, 1979.

[2] S. Arora, E. Hazan, and S. Kale. The multiplicative weights update method: A meta-algorithm and applications. Preprint.

[3] A. Blum, E. Even-Dar, and K. Ligett. Routing without regret: on convergence to Nash equilibria of regret-minimizing algorithms in routing games. *In Proc. of PODC*, 45-52, 2006.

[4] A. Blum, M. Hajiaghayi, K. Ligett, and A. Roth. Regret Minimization and the Price of Total anarchy. *In Proc. of STOC*, 373-382, 2008.

[5] A. Blum and Y. Mansour. Learning, Regret Minimization and Equilibria, in *Algorithmic Game Theory* (eds. N. Nisan, T. Roughgarden, É. Tardos and V. Vazirani), Cambridge University Press, 2007.

[6] E. Even-Dar and Y. Mansour. Fast convergence of selfish rerouting. *SODA*, 2005.

[7] S. Fischer, H. Racke, and B. Vocking. Fast convergence to wardrop equilibria by adaptive sampling methods. *STOC*, 2006.

[8] S. Fischer, B. Vocking. On the evolution of selfish routing. *In Proc. of the 12th European Symposium on Algorithms (ESA '04)*, pages 323-334. Springer, 2004.

[9] D. Foster and S. M. Kakade. Deterministic calibration and Nash equilibrium. *In Proc. of COLT*, pp. 33-48, 2004.

[10] D. Foster, and R. Vohra. Calibrated learning and correlated equilibrium. *Games and Economic Behavior* 21:40–55, 1997.

[11] D. Foster and H. P. Young. Learning, hypothesis testing, and Nash equilibrium. *Games and Economic Behavior*, 45:73, 96, 2003.

[12] D. Foster and H. P. Young. Regret testing: Learning to play Nash equilibrium without knowing you have an opponent. *Theoretical Economics*, 1:341-367, 2006.

[13] Y. Freund, and R. Schapire. Adaptive game playing using multiplicative weights. *Games and Economic Behavior*, 29:79-103, 1999.

[14] D. Fudenberg and D. Levine. The theory of learning in games. MIT Press, 1998.

[15] F. Germano and G. Lugosi. Global Nash convergence of Foster and Young's regret testing. *Games and Economic Behavior*, 60:135-154, 2007.

[16] M. Goemans, V. Mirrokni, and A. Vetta, Sink Equilibria and Convergence, *IEEE Symposium on Foundations of Computing*, 2005.

[17] J. Hannan, Approximation to Bayes risk in repeated plays. In M. Dresher, A. Tucker, and P. Wolfe eds, *Contributions to the Theory of Games*, vol 3, pp 79-139, Princeton University Press.

[18] S. Hart, and A. Mas-Colell. A simple adaptive procedure leading to correlated equilibrium. *Econometrica*, 68:1127-1150, 2000.

[19] S. Hart and A. Mas-Colell. Stochastic uncoupled dynamics and Nash equilibrium. *Games and Economic Behavior*, 57:286-303, 2006.

[20] J. Hofbauer and K. Sigmund. *Evolutionary Games and Population Dynamics.* Cambridge Univ, Press, 1998.

[21] E. Koutsoupias and C. H. Papadimitriou. Worst-case equilibria. *In Proceedings of the 16th Annual Symposium on Theoretical Aspects of Computer Science*, 404-413, 1999.

[22] Nick Littlestone and Manfred K. Warmuth. The weighted majority algorithm. *Information and Computation*, 108(2):212–260, 1994.

[23] J. Maynard Smith. *Evolution and the Theory of Games.* Cambridge University Press, 1982.

[24] J. W. Milnor. *Topology from the differentiable viewpoint.* Princeton Landmarks in Mathematics. Princeton University Press, 1997.

[25] V. Mirrokni and A. Vetta, Convergence Issues in Competitive Games, *RANDOM-APPROX*, 2004.

[26] D. Monderer and L. S. Shapley. Potential games. *Games and Economic Behavior* 14: 124-143, 1996.

[27] D. Monderer and L. S. Shapley. Fictitious play property for games with identical interests. *Journal of Economic Theory* 68:258-265, 1996.

[28] L. Perko. *Differential Equations and Dynamical Systems*. Springer, 1991.

[29] K. Ritzberger and J. Weibull. Evolutionary selection in normal-form games. *Econometrica* 63, 1371-1399, 1995.

[30] R. W. Rosenthal. A class of games possessing pure-strategy Nash equilibria. *International Journal of Game Theory* 2:65-67, 1973.

[31] T. Roughgarden. Routing Games, in *Algorithmc Game Theory* (eds. N. Nisan, T. Roughgarden, É. Tardos, V. Vazirani). Cambridge University Press. 2007.

[32] T. Roughgarden. Intrinsic Robustness of the Price of Anarchy. *In Proceedings of the 41th Annual ACM Symposium on Theory of Computing (STOC)*, 2009.

[33] L. Samuelson and J. Zhang. Evolutionary stability in asymmetric games. *Journal of Economic Theory* 57:363-391, 1992.

[34] W. H. Sandholm. Potential games with continuous player sets. *J. of Economic Theory* 97:81-108, 2001.

[35] W. H. Sandholm. *Population Games and Evolutionary Dynamics.* Manuscript, October 2008.

[36] I. R. Shafarevich. *Basic Algebraic Geometry*, volume 1. Miles Reid, translator. Springer-Verlag, 1994.

[37] J. M. Steele. An Efron-Stein inequality for nonsymmetric statistics. *Annals of Statistics* 14(2):753-758, 1986.

[38] B. Vöcking. Selfish Load Balancing. In *Algorithmc Game Theory* (eds. N. Nisan, T. Roughgarden, É. Tardos, V. Vazirani). Cambridge Univ. Press. 2007.

[39] J. Weibull. *Evolutionary Game Theory.* MIT Press, 1997.