The Price of Stochastic Anarchy

Christine Chung[1], Katrina Ligett[2⋆], Kirk Pruhs[1⋆⋆], and Aaron Roth[2]

[1] Department of Computer Science
University of Pittsburgh
{chung,kirk}@cs.pitt.edu
[2] Department of Computer Science
Carnegie Mellon University
{katrina,alroth}@cs.cmu.edu

Abstract. We consider the solution concept of stochastic stability, and propose the *price of stochastic anarchy* as an alternative to the *price of (Nash) anarchy* for quantifying the cost of selfishness and lack of coordination in games. As a solution concept, the Nash equilibrium has disadvantages that the set of stochastically stable states of a game avoid: unlike Nash equilibria, stochastically stable states are the result of natural dynamics of computationally bounded and decentralized agents, and are resilient to small perturbations from ideal play. The price of stochastic anarchy can be viewed as a smoothed analysis of the price of anarchy, distinguishing equilibria that are resilient to noise from those that are not. To illustrate the utility of stochastic stability, we study the load balancing game on unrelated machines. This game has an unboundedly large price of Nash anarchy even when restricted to two players and two machines. We show that in the two player case, the price of stochastic anarchy is 2, and that even in the general case, the price of stochastic anarchy is bounded. We conjecture that the price of stochastic anarchy is $O(m)$, matching the price of strong Nash anarchy without requiring player coordination. We expect that stochastic stability will be useful in understanding the relative stability of Nash equilibria in other games where the worst equilibria seem to be inherently brittle.

⋆ Partially supported by an AT&T Labs Graduate Fellowship and an NSF Graduate Research Fellowship.

⋆⋆ Supported in part by NSF grants CNS-0325353, CCF-0448196, CCF-0514058 and IIS-0534531.

1 Introduction

Quantifying the *price of (Nash) anarchy* is one of the major lines of research in algorithmic game theory. Indeed, one fourth of the authoritative algorithmic game theory text edited by Nisan et al. [20] is wholly dedicated to this topic. But the Nash equilibrium solution concept has been widely criticized [15, 4, 9, 10]. First, it is a solution characterization without a road map for how players might arrive at such a solution. Second, at Nash equilibria, players are unrealistically assumed to be perfectly rational, fully informed, and infallible. Third, computing Nash equilibria is PPAD-hard for even 2-player, n-action games [6], and it is therefore considered very unlikely that there exists a polynomial time algorithm to compute a Nash equilibrium even in a centralized manner. Thus, it is unrealistic to assume that selfish agents in general games will converge precisely to the Nash equilibria of the game, or that they will necessarily *converge* to anything at all. In addition, the price of Nash anarchy metric comes with its own weaknesses; it blindly uses the worst case over all Nash equilibria, despite the fact that some equilibria are more resilient than others to perturbations in play.

Considering these drawbacks, computer scientists have paid relatively little attention to if or how Nash equilibria will in fact be reached, and even less to the question of which Nash equilibria are more likely to be played in the event players do converge to Nash equilibria. To address these issues, we employ the stochastic stability framework from evolutionary game theory to study simple dynamics of computationally efficient, imperfect agents. Rather than defining a-priori states such as Nash equilibria, which might not be reachable by natural dynamics, the stochastic stability framework allows us to define a natural dynamic, and from it derive the stable states. We define the *price of stochastic anarchy* to be the ratio of the worst stochastically stable solution to the optimal solution. The stochastically stable states of a game may, but do not necessarily, contain all Nash equilibria of the game, and so the price of stochastic anarchy may be strictly better than the price of Nash anarchy. In games for which the stochastically stable states are a subset of the Nash equilibria, studying the ratio of the worst stochastically stable state to the optimal state can be viewed as a smoothed analysis of the price of anarchy, distinguishing Nash equilibria that are brittle to small perturbations in perfect play from those that are resilient to noise.

The evolutionary game theory literature on *stochastic stability* studies n-player games that are played repeatedly. In each round, each player observes her action and its outcome, and then uses simple rules to select her action for the next round based only on her size-restricted memory of the past rounds. In any round, players have a small probability of deviating from their prescribed decision rules. The state of the game is the contents of the memories of all the players. The *stochastically stable states* in such a game are the states with non-zero probability in the limit of this random process, as the probability of error approaches zero. The play dynamics we employ in this paper are the imitation dynamics studied by Josephson and Matros [16]. Under these dynamics, each player imitates the strategy that was most successful for her in recent memory.

To illustrate the utility of stochastic stability, we study the price of stochastic anarchy of the unrelated load balancing game [2, 1, 11]. To our knowledge, we are the first to quantify the loss of efficiency in any system when the players are in stochastically stable equilibria. In the load balancing game on unrelated machines, even with only two players and two machines, there are Nash equilibria with arbitrarily high cost, and so the price of Nash anarchy is unbounded. We show that these equilibria are inherently brittle, and that for two players and two machines, the price of stochastic anarchy is 2. This result matches the strong price of anarchy [1] without requiring coordination (at strong Nash equilibria, players have the ability to coordinate by forming coalitions). We further show that in the general n-player, m-machine game, the price of stochastic anarchy is bounded. More precisely the price of stochastic anarchy is upper bounded by the nmth n-step Fibonacci number. We also show that the price of stochastic anarchy is at least $m + 1$.

Our work provides new insight into the equilibria of the load balancing game. Unlike some previous work on dynamics for games, our work does not seek to propose practical dynamics with fast convergence; rather, we use simple dynamics as a tool for understanding the inherent relative stability of equilibria. Instead of relying on player coordination to avoid the Nash equilibria with unbounded cost (as is done in the study of strong equilibria), we show that these bad equilibria are inherently unstable in the face of occasional uncoordinated mistakes. We conjecture that the price of stochastic anarchy is closer to the linear lower bound, paralleling the price of strong anarchy.

In light of our results, we believe the techniques in this paper will be useful for understanding the relative stability of Nash equilibria in other games for which the worst equilibria are brittle. Indeed, for a variety of games in the price of anarchy literature, the worst Nash equilibria of the lower bound instances are not stochastically stable.

1.1 Related Work

We give a brief survey of related work in three areas: alternatives to Nash equilibria as a solution concept, stochastic stability, and the unrelated load balancing game.

Recently, several papers have noted that the Nash equilibrium is not always a suitable solution concept for computationally bounded agents playing in a repeated game, and have proposed alternatives. Goemans et al. [15] study players who sequentially play myopic best responses, and quantify the *price of sinking* that results from such play. Fabrikant and Papadimitriou [9] propose a model in which agents play restricted finite automata. Blum et al. [4, 3] assume only that players' action histories satisfy a property called *no regret*, and show that for many games, the resulting social costs are no worse than those guaranteed by price of anarchy results.

Although we believe this to be the first work studying stochastic stability in the computer science literature, computer scientists have recently employed other tools from evolutionary game theory. Fisher and Vöcking [13] show that under replicator dynamics in the routing game studied by Roughgarden and Tardos [22], players converge to Nash. Fisher et al. [12] went on to show that using a simultaneous adaptive sampling method, play converges quickly to a Nash equilibrium. For a thorough survey of algorithmic results that have employed or studied other evolutionary game theory techniques and concepts, see Suri [23].

Stochastic stability and its adaptive learning model as studied in this paper were first defined by Foster and Young [14], and differ from the standard game theory solution concept of evolutionarily stable strategies (ESS). ESS are a refinement of Nash equilibria, and so do not always exist, and are not necessarily associated with a natural play dynamic. In contrast, a game always has stochastically stable states that result (by construction) from natural dynamics. In addition, ESS are resilient only to single shocks, whereas stochastically stable states are resilient to persistent noise.

Stochastic stability has been widely studied in the economics literature (see, for example, [24, 17, 19, 5, 7, 21, 16]). We discuss in Sect. 2 concepts from this body of literature that are relevant to our results. We recommend Young [25] for an informative and readable introduction to stochastic stability, its adaptive learning model, and some related results. Our work differs from prior work in stochastic stability in that it is the first to quantify the social utility of stochastically stable states, the *price of stochastic anarchy*.

We also note a connection between the stochastically stable states of the game and the sinks of a game, recently introduced by Goemans et al. as another way of studying the dynamics of computationally bounded agents. In particular, the stochastically stable states of a game under the play dynamics we consider correspond to a subset of the sink equilibria, and so provide a framework for identifying the stable sink equilibria. In potential games, the stochastically stable states of the play dynamics we consider correspond to a subset of the Nash equilibria, thus providing a method for identifying which of these equilibria are stable.

In this paper, we study the price of stochastic anarchy in load balancing. Even-Dar et al. [8] show that when playing the load balancing game on unrelated machines, any turn-taking improvement dynamics converge to Nash. Andelman et al. [1] observe that the price of Nash anarchy in this game is unbounded and they show that the strong price of anarchy is linear in the number of machines. Fiat et al. [11] tighten their upper bound to match their lower bound at a strong price of anarchy of exactly m.

2 Model and Background

We now formalize (from Young [24]) the adaptive play model and the definition of stochastic stability. We then formalize the play dynamics that we consider. We also provide in this section the results from the stochastic stability literature that we will later use for our results.

2.1 Adaptive Play and Stochastic Stability

Let $G = (X, \pi)$ be a game with n players, where $X = \prod_{j=1}^{n} X_i$ represents the strategy sets X_i for each player i, and $\pi = \prod_{j=1}^{n} \pi_i$ represents the payoff functions $\pi_i : X \to \mathbb{R}$ for each player. G is played repeatedly for successive time periods $t = 1, 2, \ldots$, and at each time step t, player i plays some action $s_i^t \in X_i$. The collection of all players' actions at time t defines a play profile $S^t = (S_1^t, S_2^t, \ldots, S_n^t)$. We wish to model computationally efficient agents, and so we imagine that each agent has some finite memory of size z, and that after time step t, all players remember a history consisting of a sequence of play profiles $h^t = (S^{t-z+1}, S^{t-z+2}, \ldots, S^t) \in (X)^z$.

We assume that each player i has some efficiently computable function $p_i : (X)^z \times X_i \to \mathbb{R}$ that, given a particular history, induces a sampleable probability distribution over actions (for all players i and histories h, $\sum_{a \in X_i} p_i(h, a) = 1$). We write p for $\prod_i p_i$. We wish to model imperfect agents who make mistakes, and so we imagine that at time t each player i plays according to p_i with probability $1 - \epsilon$, and with probability ϵ plays some action in X_i uniformly at random.[3] That is, for all players i, for all actions

[3] The mistake probabilities need not be uniform random—all that we require is that the distribution has support on all actions in X_i.

$a \in X_i$, $\Pr[s_i^t = a] = (1-\epsilon)p_i(h^t, a) + \frac{\epsilon}{|X_i|}$. The dynamics we have described define a Markov process $P^{G,p,\epsilon}$ with finite state space $H = (X)^z$ corresponding to the finite histories. For notational simplicity, we will write the Markov process as P^ϵ when there is no ambiguity.

The potential successors of a history can be obtained by observing a new play profile, and "forgetting" the least recent play profile in the current history.

Definition 2.1. *For any $S' \in X$, A history $h' = (S^{t-z+2}, S^{t-z+3}, \ldots, S^t, S')$ is a* successor *of history $h^t = (S^{t-z+1}, S^{t-z+2}, \ldots, S^t)$.*

The Markov process P^ϵ has transition probability $p^\epsilon_{h,h'}$ of moving from state $h = (S^1, \ldots, S^z)$ to state $h' = (T^1, \ldots, T^z)$:

$$p^\epsilon_{h,h'} = \begin{cases} \prod_{i=1}^n (1-\epsilon)\, p_i(h, T_i^z) + \frac{\epsilon}{|X_i|} & \text{if } h' \text{ is a successor of h;} \\ 0 & \text{otherwise.} \end{cases}$$

We will refer to P^0 as the unperturbed Markov process. Note that for $\epsilon > 0$, $p^\epsilon_{h,h'} > 0$ for every history h and successor h', and that for any two histories h and $\hat{h}$ not necessarily a successor of h, there is a series of z histories $h_1, \ldots, h_z$ such that $h_1 = h$, $h_z = \hat{h}$, and for all $1 < i \le z$, h_i is a successor of h_{i-1}. Thus there is positive probability of moving between any h and any $\hat{h}$ in z steps, and so P^ϵ is irreducible. Similarly, there is a positive probability of moving between any h and any $\hat{h}$ in $z+1$ steps, and so P^ϵ is aperiodic. Therefore, P^ϵ has a unique stationary distribution μ^ϵ.

The stochastically stable states of a particular game and player dynamics are the states with nonzero probability in the limit of the stationary distribution.

Definition 2.2 (Foster and Young [14]). *A state h is* stochastically stable *relative to P^ϵ if $\lim_{\epsilon \to 0} \mu^\epsilon(h) > 0$.*

Intuitively, we should expect a process P^ϵ to spend almost all of its time at its stochastically stable states when ϵ is small.

When a player i plays at random rather than according to p_i, we call this a mistake.

Definition 2.3 (Young [24]). *Suppose $h' = (S^{t-z+1}, \ldots, S^t)$ is a successor of h. A* mistake *in the transition between h and h' is any element S_i^t such that $p_i(h, S_i^t) = 0$. Note that mistakes occur with probability $\le \epsilon$.*

We can characterize the number of mistakes required to get from one history to another.

Definition 2.4 (Young [24]). *For any two states h, h', the* resistance *$r(h, h')$ is the minimum total number of mistakes involved in the transition $h \to h'$ if h' is a successor of h. If h' is not a successor of h, then $r(h, h') = \infty$.*

Note that the transitions of zero resistance are exactly those that occur with positive probability in the unperturbed Markov process P^0.

Definition 2.5. *We refer to the sinks of P^0 as* recurrent classes. *In other words, a* recurrent class *of P^0 is a set of states $C \subseteq H$ such that any state in C is reachable from any other state in C and no state outside C is accessible from any state inside C.*

We may view the state space H as the vertex set of a directed graph, with an edge from h to h' if h' is a successor of h, with edge weight $r(h, h')$.

Observation 2.6. *We observe that the recurrent classes $H_1, H_2, \ldots$, where each $H_i \subseteq H$, have the following properties:*

1. *From every vertex $h \in H$, there is a path of cost 0 to one of the recurrent classes.*
2. *For each H_i and for every pair of vertices $h, h' \in H_i$, there is a path of cost 0 between h and h'.*
3. *For each H_i, every edge (h, h') with $h \in H_i, h' \notin H_i$ has positive cost.*

Let $r_{i,j}$ denote the cost of the shortest path between H_i and H_j in the graph described above. We now consider the complete directed graph $\mathcal{G}$ with vertex set $\{H_1, H_2, \ldots\}$ in which the edge (H_i, H_j) has weight $r_{i,j}$. Let T_i be a directed minimum-weight spanning in-tree of $\mathcal{G}$ rooted at vertex H_i. (An in-tree is a directed tree where each edge is oriented toward the root.) The *stochastic potential* of H_i is defined to be the sum of the edge weights in T_i.

Young proves the following theorem characterizing stochastically stable states:

Theorem 2.7 (Young [24]). *In any n-player game G with finite strategy sets and any set of action distributions p, the stochastically stable states of $P^{G,p,\epsilon}$ are the recurrent classes of minimum stochastic potential.*

2.2 Imitation Dynamics

In this paper, we study agents who behave according to a slight modification of the imitation dynamics introduced by Josephson and Matros [16]. (We note that this modification is of no consequence to the results of Josephson and Matros [16] that we present below.) Player i using imitation dynamics parameterized by $\sigma \in \mathbb{N}$ chooses his action at time $t+1$ according to the following mechanism:

1. Player i selects a set Y of σ play profiles uniformly at random from the z profiles in history h_t.
2. For each play profile $S \in Y$, i recalls the payoff $\pi_i(S)$ he obtained from playing action S_i.
3. Player i plays the action among these that corresponds to his highest payoff; that is, he plays the i^{th} component of $\mathrm{argmax}_{S \in Y}\, \pi_i(S)$. In the case of ties, he plays a highest-payoff action at random.

The value σ is a parameter of the dynamics that is taken to be $n \le \sigma \le z/2$. These dynamics can be interpreted as modeling a situation in which at each time step, players are chosen at random from a pool of identical players, who each played in a subset of the last z rounds. The players are computationally simple, and so do not counterspeculate the actions of their opponents, instead playing the action that has worked the best for them in recent memory.

We will say that a history h is *monomorphic* if the same action profile S has been repeated for the last z rounds: $h = (S, S, \ldots, S)$. Josephson and Matros [16] prove the following useful fact:

Proposition 2.8. *A set of states is a recurrent class of the imitation dynamics if and only if it is a singleton set consisting of a monomorphic state.*

Since the stochastically stable states are a subset of the recurrent classes, we can associate with each stochastically stable state $h = (S, \ldots, S)$ the unique action profile S it contains. This allows us to now define the price of stochastic anarchy with respect to imitation dynamics. For brevity, we will refer to this throughout the paper as simply the price of stochastic anarchy.

Definition 2.9. *Given a game $G = (X, \pi)$ with a social cost function $\gamma : X \to \mathbb{R}$, the* price of stochastic anarchy *of G is equal to $\max \frac{\gamma(S)}{\gamma(\mathbf{OPT})}$, where $\mathbf{OPT}$ is the play profile that minimizes γ and the $\max$ is taken over all play profiles S such that $h = (S, \ldots, S)$ is stochastically stable.*

Given a game G, we define the *better response graph* of G: The set of vertices corresponds to the set of action profiles of G, and there is an edge between two action profiles S and S' if and only if there exists a player i such that S' differs from S only in player i's action, and player i does not decrease his utility by unilaterally deviating from S_i to S'_i. Josephson and Matros [16] prove the following relationship between this better response graph and the stochastically stable states of a game:

Theorem 2.10. *If $\mathbb{V}$ is the set of stochastically stable states under imitation dynamics, then $V = \{S : (S, \ldots, S) \in \mathbb{V}\}$ is either a strongly connected component of the better response graph of G, or a union of strongly connected components.*

Goemans et al. [15] introduce the notion of sink equilibria and a corresponding notion of the "price of sinking", which is the ratio of the social welfare of the worst sink equilibrium to that of the social optimum. We note that the strongly connected components of the better response graph of G correspond to the sink equilibria (under sequential better-response play) of G, and so Theorem 2.10 implies that the stochastically stable states under imitation dynamics correspond to a subset of the sinks of the better response graph of G, and we get the following corollary:

Corollary 2.11. *The price of stochastic anarchy of a game G under imitation dynamics is at most the price of sinking of G.*

3 Load Balancing: Game Definition and Price of Nash Anarchy

The load balancing game on unrelated machines models a set of agents who wish to schedule computing jobs on a set of machines. The machines have different strengths and weaknesses (for example, they may have different types of processors or differing amounts of memory), and so each job will take a different amount of time to run on each machine. Jobs on a single machine are executed in parallel such that all jobs on any given machine finish at the same time. Thus, each agent who schedules his job on machine M_i endures the *load* on machine M_i, where the load is defined to be the sum of the running times of all jobs scheduled on M_i. Agents wish to minimize the completion time for their jobs, and social cost is defined to be the *makespan*: the maximum load on any machine.

Formally, an instance of the load balancing game on unrelated machines is defined by a set of n players and m machines $M = \{M_1, \ldots, M_m\}$. The action space for each player is $X_i = M$. Each player i has some cost $c_{i,j}$ on machine j. Denote the cost of machine M_j for action profile S by $C_j(S) = \sum_{i \text{ s.t. } S_i=j} c_{i,j}$. Each player i has utility function $\pi_i(S) = -C_{s_i}(S)$. The social cost of an action profile S is $\gamma(S) = \max_{j\in M} C_j(S)$. We define $\mathbf{OPT}$ to be the action profile that minimizes social cost: $\mathbf{OPT} = \operatorname{argmin}_{S\in X} \gamma(S)$. Without loss of generality, we will always normalize so that $\gamma(\mathbf{OPT}) = 1$.

The coordination ratio of a game (also known as the price of anarchy) was introduced by Koutsoupias and Papadimitriou [18], and is intended to quantify the loss of efficiency due to selfishness and the lack of coordination among rational agents. Given a game G and a social cost function γ, it is simple to quantify the $\mathbf{OPT}$ game state S: $\mathbf{OPT} = \operatorname{argmin}\gamma(S)$. It is less clear how to model rational selfish agents. In most prior work it has been assumed that selfish agents play according to a Nash equilibrium, and the price of anarchy has been defined as the ratio of the cost of the worst (pure strategy) Nash state to $\mathbf{OPT}$. In this paper, we refer to this measure as the price of Nash anarchy, to distinguish it from the price of stochastic anarchy, which we defined in Sect. 2.2.

Definition 3.1. *For a game G with a set of Nash equilibrium states $\mathcal{E}$, the* price of (Nash) anarchy *is* $\max_{S\in\mathcal{E}} \frac{\gamma(S)}{\gamma(\mathbf{OPT})}$.

We show here that even with only two players and two machines, the load balancing game on unrelated machines has a price of Nash anarchy that is unbounded by any function of m and n. Consider the two-player, two-machine game with $c_{1,1} = c_{2,2} = 1$ and $c_{1,2} = c_{2,1} = 1/\delta$, for some $0 < \delta < 1$. Then the play profile $\mathbf{OPT} = (M_1, M_2)$ is a Nash equilibrium with cost 1. However, observe that the profile $S^* = (M_2, M_1)$ is also a Nash equilibrium, with cost $1/\delta$ (since by deviating, players can only increase their cost from $1/\delta$ to $1/\delta + 1$). The price of anarchy of the load balancing game is therefore $1/\delta$, which can be unboundedly large, although $m = n = 2$.

4 Upper Bound on Price of Stochastic Anarchy

The load balancing game is an ordinal potential game [8], and so the sinks of the better-response graph correspond to the pure strategy Nash equilibria. We therefore have by Corollary 2.11 that the stochastically stable states are a subset of the pure strategy Nash equilibria of the game, and the price of stochastic anarchy is at most the price of anarchy. We have noted that even in the two-person, two-machine load balancing game, the price of anarchy is unbounded (even for pure strategy equilibria). Therefore, as a warmup, we bound the price of stochastic anarchy of the two-player, two-machine case.

4.1 Two Players, Two Machines

Theorem 4.1. *In the two-player, two-machine load balancing game on unrelated machines, the price of stochastic anarchy is 2.*

Note that the two-player, two-machine load balancing game can have at most two strict pure strategy Nash equilibria. (For brevity we consider the case of strict equilibria. The argument for weak equilibria is similar). Note also that either there is a unique Nash equilibrium at (M_1, M_1) or (M_2, M_2), or there are two at $N_1 = (M_1, M_2)$ and $N_2 = (M_2, M_1)$.

An action profile N *Pareto dominates* N' if for each player i, $C_{N_i}(N) \leq C_{N'_i}(N')$.

Lemma 4.2. *If there are two Nash equilibria, and N_1 Pareto dominates N_2, then only N_1 is stochastically stable (and vice versa).*

Proof. Note that if N_1 Pareto dominates N_2, then it also Pareto dominates (M_1, M_1) and (M_2, M_2), since each is a unilateral deviation from a Nash equilibrium for both players. Consider the monomorphic state $(N_2, \ldots, N_2)$. If both players make simultaneous mistakes at time t to N_1, then by assumption, N_1 will be the action profile in $h_{t+1} = (N_2, \ldots, N_2, N_1)$ with lowest cost for both players. Therefore, with positive probability, both players will draw samples of their histories containing the action profile N_1, and therefore play it, until $h_{t+z} = (N_1, \ldots, N_1)$. Therefore, there is an edge in $\mathcal{G}$ from $h = \{N_2, \ldots, N_2\}$ to $h' = \{N_1, \ldots, N_1\}$ of resistance 2. However, there is no edge from h' to any other state in $\mathcal{G}$ with resistance $< \sigma$. Recall our initial observation that in fact, N_1 Pareto dominates all other action profiles. Therefore, no set of mistakes will yield an action profile with higher payoff than N_1 for either player, and so to leave state h' will require at least σ mistakes (so that some player may draw a sample from their history that contains no instance of action profile h). Therefore, given any minimum spanning tree of $\mathcal{G}$ rooted at h, we may add an edge (h, h') of weight 2, and remove the outgoing edge from h', which we have shown must have cost $\geq \sigma$. This is a minimum spanning tree rooted at h' with strictly lower cost. We have therefore shown that h' has strictly lower stochastic potential than h, and so by Theorem 2.7, h is not stochastically stable. Since at least one Nash equilibrium must be stochastically stable, $h' = (N_1, \ldots, N_1)$ is the unique stochastically stable state. $\square$

Proof (of Theorem 4.1). If there is only one Nash equilibrium (M_1, M_1) or (M_2, M_2), then it must be the only stochastically stable state (since in potential games these are a nonempty subset of the pure strategy Nash equilibria), and must also be $\mathbf{OPT}$. In this case, the price of anarchy is equal to the price of stochastic anarchy, and is 1. Therefore, we may assume that there are two Nash equilibria, N_1 and N_2. If N_1 Pareto dominates N_2, then N_1 must be $\mathbf{OPT}$ (since load balancing is a potential game), and by Lemma 4.2, N_1 is the only stochastically stable state. In this case, the price of stochastic anarchy is 1 (strictly less than the (possibly unbounded) price of anarchy). A similar argument holds if N_2 Pareto dominates N_1. Therefore, we may assume that neither N_1 nor N_2 Pareto dominate the other.

Without loss of generality, assume that N_1 is $\mathbf{OPT}$, and that in $N_1 = (M_1, M_2)$, M_2 is the maximally loaded machine. Suppose that M_2 is also the maximally loaded machine in N_2. (The other case is similar.) Together with the fact that N_1 does not Pareto dominate N_2, this gives us the following:

$$c_{1,1} \leq c_{2,2}$$
$$c_{2,1} \leq c_{2,2}$$
$$c_{1,2} \geq c_{2,2}$$

From the fact that both N_1 and N_2 are Nash equilibria, we get:

$$c_{1,1} + c_{2,1} \geq c_{2,2}$$
$$c_{1,1} + c_{2,1} \geq c_{1,2}$$

In this case, the price of anarchy among pure strategy Nash equilibria is:

$$\frac{c_{1,2}}{c_{2,2}} \leq \frac{c_{1,1} + c_{2,1}}{c_{2,2}} \leq \frac{c_{1,1} + c_{2,1}}{c_{1,1}} = 1 + \frac{c_{2,1}}{c_{1,1}}$$

Similarly, we have:

$$\frac{c_{1,2}}{c_{2,2}} \leq \frac{c_{1,1} + c_{2,1}}{c_{2,2}} \leq \frac{c_{1,1} + c_{2,1}}{c_{2,1}} = 1 + \frac{c_{1,1}}{c_{2,1}}$$

Combining these two inequalities, we get that the price of Nash anarchy is at most $1 + \min(c_{1,1}/c_{2,1}, c_{2,1}/c_{1,1}) \leq 2$. Since the price of stochastic anarchy is at most the price of anarchy over pure strategies, this completes the proof. □

4.2 General Case: n Players, m Machines

Theorem 4.3. *The general load balancing game on unrelated machines has price of stochastic anarchy bounded by a function Ψ depending only on n and m, and*

$$\Psi(n,m) \leq m \cdot F_{(n)}(nm+1),$$

where $F_{(n)}(i)$ denotes the i^{th} n-step Fibonacci number.[4]

To prove this upper bound, we show that any solution worse than our upper bound cannot be stochastically stable. To show this impossibility, we take any arbitrary solution worse than our upper bound and show that there must always be a minimum cost in-tree in $\mathcal{G}$ rooted at a different solution that has strictly less cost than the minimum cost in-tree rooted at that solution. We then apply Proposition 2.8 and Theorem 2.7. The proof proceeds by a series of lemmas.

Definition 4.4. *For any monomorphic Nash state $h = (S, \ldots, S)$, let the* Nash Graph *of h be a directed graph with vertex set M and directed edges (M_i, M_j) if there is some player i with $S_i = M_i$ and $\mathbf{OPT}_i = M_j$. Let the* closure *$\bar{M}_i$ of machine M_i, be the set of states reachable from M_i by following 0 or more edges of the Nash graph.*

Lemma 4.5. *In any monomorphic Nash state $h = (S, \ldots, S)$, if there is a machine M_i such that $C_i(S) > m$, then every machine $M_j \in \bar{M}_i$ has cost $C_j(S) > 1$.*

Proof. Suppose this were not the case, and there exists an $M_j \in \bar{M}_i$ with $C_j(S) \leq 1$. Since $M_j \in \bar{M}_i$, there exists a simple path $(M_i = M_1, M_2, \ldots, M_k = M_j)$ with $k \leq m$. Since S is a Nash equilibrium, it must be the case that $C_{k-1}(S) \leq 2$ because by the definition of the Nash graph, the directed edge from M_{k-1} to M_k implies that there is some player i with $S_i = M_{k-1}$, but $\mathbf{OPT}_i = M_k$. Since $1 = \gamma(\mathbf{OPT}) \geq C_k(\mathbf{OPT}) \geq c_{i,k}$, if player i deviated from his action in Nash profile S to $S'_i = M_k$, he would experience cost $C_k(S) + c_{i,k} \leq 1 + 1 = 2$. Since he cannot benefit from deviating (by definition of Nash), it must be that his cost in S, $C_{k-1}(S) \leq 2$. By the same argument, it must be that $C_{k-2}(S) \leq 3$, and by induction, $C_1(S) \leq k \leq m$. □

Lemma 4.6. *For any monomorphic Nash state $h = (S, \ldots, S) \in \mathcal{G}$ with $\gamma(S) > m$, there is an edge from h to some $g = (T, \ldots, T)$ where $\gamma(T) \leq m$ with edge cost $\leq n$ in $\mathcal{G}$.*

[4] $F_{(n)}(i) = \begin{cases} 1 & \text{if } i \leq n; \\ \sum_{j=i-n}^{i} F_{(n)}(j) & \text{otherwise.} \end{cases}$ $F_{(n)}(i) \in o(2^i)$ for any fixed n.

Proof. Let $D = \{M_j : C_i(S) \geq m\}$, and define the closure of D, $\bar{D} = \bigcup_{M_i \in D} \bar{M}_i$. Consider the successor state h' of h that results when every player i such that $S_i^t \in \bar{D}$ makes a mistake and plays on their OPT machine $S_i^{t+1} = \mathbf{OPT}_i$, and all other players do not make a mistake and continue to play $S_i^{t+1} = S_i^t$. Note that by the definition of $\bar{D}$, for $M_j \in \bar{D}$, for all players i playing machine j in S, $\mathbf{OPT}_i \in \bar{D}$. Let $T = S^{t+1}$. Then for all j such that $M_j \in \bar{D}$, $C_j(T) \leq 1$, since $C_j(T) \leq C_j(\mathbf{OPT}) \leq 1$. To see this, note that for every player i such that $S_i^t = M_j \in \bar{D}$, $S_i^{t+1} = M_j$ if and only if $\mathbf{OPT}_i = M_j$. Similarly, for every player i such that $S_i^{t+1} = M_j \in \bar{D}$ but $S_i^t \neq M_j$, $\mathbf{OPT}_i = M_j$, and so for each machine $M_j \in \bar{D}$, the agents playing on M_j in T are a subset of those playing on M_j at $\mathbf{OPT}$. Note that by Lemma 4.5, for all $M_j \in \bar{D}$, $C_j(S) > 1$. Therefore, for every agent i with $S_i^t \in \bar{D}$, $\pi_i(T) > \pi_i(S)$, and so for $h'' = (S, \ldots, S, T, T)$ a successor of h', $r(h', h'') = 0$. Reasoning in this way, there is a path of zero resistance from h' to $g = (T, \ldots, T)$. We have therefore exhibited a path between h and g that involves only $|\{i : S_i^t \in \bar{D}\}| \leq n$ mistakes. Finally, we observe that if $M_j \in \bar{D}$ then $C_j(T) \leq 1$, and by construction, if $M_j \notin \bar{D}$, then $C_j(T) = C_j(S) < m$, since as noted above $M_j \notin \bar{D}$ implies that the players playing M_j in S are the same set playing M_j in T. Thus, we have $\gamma(T) \leq m$, which completes the proof. □

Lemma 4.7. *Let $h = (S, \ldots, S) \in \mathcal{G}$ be any monomorphic state with $\gamma(S) \leq m$. Any path in $\mathcal{G}$ from h to a monomorphic state $h' = (S', \ldots, S') \in \mathcal{G}$ where $\gamma(h') > m \cdot F_{(n)}(mn+1)$ must contain an edge with cost $\geq \sigma$, where $F_{(n)}(i)$ denotes the i^{th} n-step Fibonacci number.*

Proof. Suppose there were some directed path $\mathcal{P}$ in $\mathcal{G}$ $(h = h_1, h_2, \ldots, h_l = h')$ such that all edge costs were less than σ. We will imagine assigning costs to players on machines adversarially: for a player i on machine M_j, we will consider $c_{i,j}$ to be undefined until play reaches a monomorphic state h_k in which he occupies machine j, at which point we will assign $c_{i,j}$ to be the highest value consistent with his path from h_{k-1} to h_k. Note that since initially $\gamma(S) \leq m$, we must have for all $i \in N$, $c_{i,S_i} \leq m = mF_{(n)}(n)$.

There are mn costs $c_{i,j}$ that we may assign, and we have observed that our first n assignments have taken values $\leq mF_{(n)}(n) = mF_{(n)}(1)$. We will assume inductively that our k^{th} assignment takes value at most $mF_{(n)}(k)$. Let $h_k = (T, \ldots, T)$ be the last monomorphic state in $\mathcal{P}$ such that only k cost assignments have been made, and $h_{k+1} = (T', \ldots, T')$ be the monomorphic state at which the $k+1^{st}$ cost assignment is made for some player i on machine M_j. Since by assumption, fewer than σ mistakes are made in the transition $h_k \to h_{k+1}$, it must be that $c_{i,j} \leq C_{T_i}(T)$; that is, $c_{i,j}$ can be no more than player i's experienced cost in state T. If this were not so, player i would not have continued playing on machine j in T' without additional mistakes, since with fewer than σ mistakes, any sample of size σ would have contained an instance of T which would have yielded higher payoff than playing on machine j. Note however that the cost of any machine M_j in T is at most:

$$C_j(T) \leq \sum_{i:c_{i,j} \neq \text{ undefined}} c_{i,j} \leq \sum_{i=0}^{n-1} mF_{(n)}(k-i) = mF_{(n)}(k+1)$$

where the inequality follows by our inductive assumption. We have therefore shown that the k^{th} cost assigned is at most $mF_{(n)}(k)$, and so the claim follows since there are at most nm costs $c_{i,j}$ that may be assigned, and the cost on any machine in S' is at most the sum of the n highest costs. □

Proof (of Theorem 4.3). Given any state $h = (S, \ldots, S) \in \mathcal{G}$ where $\gamma(S) > m \cdot F_{(n)}(mn+1)$, we can exhibit a state $f = (U, U, \ldots, U)$ with lower stochastic potential than h such that $\gamma(U) \leq m \cdot F_{(n)}(nm+1)$ as follows.

Consider the minimum weight spanning in-tree T_h of $\mathcal{G}$ rooted at h. We will use it to construct a spanning in-tree T_f rooted at a state f as follows: We add an edge of cost at most n from h to some state $g = (T, \ldots, T)$ such that $\gamma(T) \leq m$ (such an edge is guaranteed to exist by Lemma 4.6). This induces a cycle through h and g. To correct this, we remove an edge on the path from g to h in T_h of cost $\geq \sigma$ (such an edge is guaranteed to exist by Lemma 4.7). Since this breaks the newly induced cycle, we now have a spanning in-tree T_f with root $f = (U, U, \ldots, U)$ such that $\gamma(U) \leq m \cdot F_{(n)}(mn+1)$. Since the added edge has lower cost than the removed edge, T_f has lower cost than T_h, and so f has lower stochastic potential than h.

Since the stochastically stable states are those with minimum stochastic potential by Theorem 2.7 and Proposition 2.8, we have proven that h is not stochastically stable. □

5 Lower Bound on Price of Stochastic Anarchy

In this section, we show that the price of stochastic anarchy for load balancing is at least m, the price of strong anarchy. We show this by exhibiting an instance for which the worst stochastically stable solution costs m times the optimal solution. Our proof that this bad solution is stochastically stable uses the following lemma to show that the min cost in-tree rooted at that solution in $\mathcal{G}$ has cost as low as the min cost in-tree rooted at any other solution. We then simply apply Theorem 2.7 and Proposition 2.8.

Lemma 5.1. *For two monomorphic states h and h' corresponding to play profiles S and S', if S' is a unilateral better response deviation from S by some player i, then the resistance $r(h, h') = 1$.*

Proof. Suppose player i makes the mistake of playing S'_i instead of S_i. Since this is a better-response move, he experiences lower cost, and so long as he samples an instance of S', he will continue to play S'_i. No other player will deviate without a mistake, and so play will reach monomorphic state h' after z turns. □

	M_1	M_2	M_3	M_4
1	1	$1-\delta$	∞	∞
2	$2-2\delta$	1	$2-3\delta$	∞
3	$3-4\delta$	∞	1	$3-5\delta$
4	$4-6\delta$	∞	∞	1

Fig. 1. A load-balancing game with price of stochastic anarchy m for $m = 4$. The entry corresponding to player i and machine M_j represents the cost $c_{i,j}$. The δs represent some sufficiently small positive value and the ∞s can be any sufficiently large value. The optimal solution is (M_1, M_2, M_3, M_4) and costs 1, but (M_2, M_3, M_4, M_1) is also stochastically stable and costs $4 - 6\delta$. This example can be easily generalized to arbitrary m.

Theorem 5.2. *The price of stochastic anarchy of the load balancing game on unrelated machines is at least m.*

Proof. To aid in the illustration of this proof, refer to the instance of the load balancing game pictured in Fig. 1. Consider the instance of the load balancing game on m unrelated machines where $n = m$ and the costs are as follows. For each player i from 1 to n, let $c_{i,i} = 1$. For each player i from 2 to n, let $c_{i,1} = i - 2(i-1)\delta$, where δ is a diminishingly small positive integer. Finally, for each player i from 1 to $n-1$, let $c_{i,i+1} = i-(2i-1)\delta$. Let all other costs be ∞ or some sufficiently large positive value.

Note that in this instance the optimal solution is achieved when each player i plays on machine M_i and thus $\gamma(\mathbf{OPT}) = 1$. Also note that the only pure-strategy Nash states in this instance are the profiles $N_1 = (M_1, M_2, \ldots, M_m)$, $N_2 = (M_2, M_1, M_3, M_4, \ldots, M_m), N_3 = (M_2, M_3, M_1, M_4, \ldots, M_m), \ldots, N_{m-1} = (M_2, M_3, M_4, \ldots, M_{m-1}, M_1, M_m),\ N_m = (M_2, M_3, M_4, \ldots, M_m, M_1)$. We observe that $\gamma(N_m) = m - 2(m-1)\delta \approx m$, and the monomorphic state corresponding to N_m is stochastically stable:

Note that for the monomorphic state corresponding to each Nash profile N_i, there is an edge of resistance 2 to any monomorphic state $(S_i, \ldots, S_i)$ where S_i is on a better-response path to Nash profile N_{i+1}. This transition can occur with two simultaneous mistakes as follows: At the same time step t, player i plays on machine M_{i+1}, and player $i+1$ plays on machine M_i. Since for this turn, player i plays on machine M_{i+1} alone, he experiences cost that is δ less than his best previous cost. Player $i+1$ experiences higher cost. Therefore, player $i+1$ returns to machine M_{i+1} and continues to play it (since N_i continues to be the play profile in his history for which he experienced lowest cost). Player i continues to sample the play profile from time step t for the next σ rounds, and so continues to play on M_{i+1} without further mistakes (even though player $i+1$ has now returned). In this way, play proceeds in z timesteps to a new monomorphic state S_i without any further mistakes. Note that in S_i, players i and $i+1$ both occupy machine M_{i+1}, and so S_i is one better-response move, and hence one mistake, away from N_{i+1} (by moving to machine M_1, player $i+1$ can experience δ less cost).

Finally, we construct a minimum spanning in-tree T_{N_m} from the graph $\mathcal{G}$ rooted at N_m. For the monomorphic state corresponding to the Nash profile N_i, $1 \leq i \leq m-1$, we include the resistance 2 edge to S_i. All other monomorphic states correspond to non-Nash profiles, and so are on better-response paths to some Nash state (since this is a potential game). When a state is on a better-response path to two Nash states N_i and N_j, we consider only the state N_i such that $i > j$. For each non-Nash monomorphic state, we insert the edge corresponding to the first step in the better-response path to N_i, which by Lemma 5.1 has cost 1. Since non-Nash monomorphic states are part of shortest-path in-trees to Nash monomorphic states, which have edges to Nash states of higher index, this process produces no cycles, and so forms a spanning in-tree rooted at N_m. Moreover, no spanning tree of $\mathcal{G}$ can have lower cost, since every edge in T_{N_m} is of minimal cost: the only edges in T_{N_m} that have cost > 1 are those leaving strict Nash states, but *any* edge leaving a strict Nash state must have cost ≥ 2. Therefore, by definition of stochastic potential, Theorem 2.7, and Proposition 2.8, the monomorphic state corresponding to N_m is stochastically stable. $\square$

Remark 5.3. More complicated examples than the one we provide here show that the price of stochastic anarchy is greater than m, and so our lower bound is not tight. For an example, see Figure 2.

	M_1	M_2	M_3	M_4
1	1	1	∞	$4-3\delta$
2	$2-\delta$	1	$2-\delta$	∞
3	$3-2\delta$	$3-2\delta$	1	$3-2\delta$
4	$4-3\delta$	$5-4\delta$	∞	1

Fig. 2. The optimal solution here is (M_1, M_2, M_3, M_4) and costs 1, but by similar reasoning as in the proof of Theorem 5.2, (M_4, M_3, M_1, M_2) is also stochastically stable and costs $5-4\delta$. This example can be easily generalized to arbitrary values of m.

We note the exponential separation between our upper and lower bounds. We conjecture, however, that the true value of the price of stochastic anarchy falls closer to our lower bound:

Conjecture 5.4. The price of stochastic anarchy in the load balancing game with unrelated machines is $O(m)$.

If this conjecture is correct, then the $O(m)$ bound from the strong price of anarchy [1] can be achieved without coordination.

6 Conclusion and Open Questions

In this paper, we propose the evolutionary game theory solution concept of stochastic stability as a tool for quantifying the relative stability of equilibria. We show that in the load balancing game on unrelated machines, for which the price of Nash anarchy is unbounded, the "bad" Nash equilibria are not stochastically stable, and so the price of stochastic anarchy is bounded. We conjecture that the upper bound given in this paper is not tight and the cost of stochastic stability for load balancing is $O(m)$. If this conjecture is correct, it implies that the fragility of the "bad" equilibria in this game is attributable to their instability, not only in the face of player coordination, but also to minor uncoordinated perturbations in play. We expect that the techniques used in this paper will also be useful in understanding the relative stability of Nash equilibria in other games for which the worst equilibria are brittle. This promise is evidenced by the fact that the worst Nash in the worst-case instances in many games (for example, the Roughgarden and Tardos [22] lower bound showing an unbounded price of anarchy for routing unsplittable flow) are not stochastically stable.

7 Acknowledgments

We would like to thank Yishay Mansour for bringing the unbounded price of anarchy in the load balancing game to our attention, Avrim Blum and Tim Roughgarden for useful discussion, and Alexander Matros for his early guidance. We also are grateful to Mallesh Pai and Sid Suri for helpful discussions about evolutionary game theory.

References

1. Nir Andelman, Michal Feldman, and Yishay Mansour. Strong price of anarchy. In *SODA '07*.

2. Baruch Awerbuch, Yossi Azar, Yossi Richter, and Dekel Tsur. Tradeoffs in worst-case equilibria. *Theor. Comput. Sci.*, 361(2):200–209, 2006.
3. Avrim Blum, Eyal Even-Dar, and Katrina Ligett. Routing without regret: On convergence to Nash equilibria of regret-minimizing algorithms in routing games. In *PODC '06*.
4. Avrim Blum, MohammadTaghi Hajiaghayi, Katrina Ligett, and Aaron Roth. Regret minimization and the price of total anarchy. In *STOC '08*.
5. Lawrence E. Blume. The statistical mechanics of best-response strategy revision. *Games and Economic Behavior*, 11(2):111–145, November 1995.
6. Xi Chen and Xiaotie Deng. Settling the complexity of 2-player Nash-equilibrium. In *FOCS '06*.
7. Glenn Ellison. Basins of attraction, long-run stochastic stability, and the speed of step-by-step evolution. *Review of Economic Studies*, 67(1):17–45, January 2000.
8. Eyal Even-Dar, Alexander Kesselman, and Yishay Mansour. Convergence time to Nash equilibria. In *ICALP '03*.
9. Alex Fabrikant and Christos Papadimitriou. The complexity of game dynamics: Bgp oscillations, sink equilibria, and beyond. In *SODA '08*.
10. Alex Fabrikant, Christos Papadimitriou, and Kunal Talwar. The complexity of pure nash equilibria. In *STOC '04*.
11. Amos Fiat, Haim Kaplan, Meital Levy, and Svetlana Olonetsky. Strong price of anarchy for machine load balancing. In *ICALP '07*.
12. Simon Fischer, Harald Räcke, and Berthold Vöcking. Fast convergence to wardrop equilibria by adaptive sampling methods. In *STOC '06*.
13. Simon Fischer and Berthold Vöcking. On the evolution of selfish routing. In *ESA '04*.
14. D. Foster and P. Young. Stochastic evolutionary game dynamics. *Theoret. Population Biol.*, 38:229–232, 1990.
15. Michel Goemans, Vahab Mirrokni, and Adrian Vetta. Sink equilibria and convergence. In *FOCS '05*.
16. Jens Josephson and Alexander Matros. Stochastic imitation in finite games. *Games and Economic Behavior*, 49(2):244–259, November 2004.
17. Michihiro Kandori, George J Mailath, and Rafael Rob. Learning, mutation, and long run equilibria in games. *Econometrica*, 61(1):29–56, January 1993.
18. E. Koutsoupias and C. Papadimitriou. Worst-case equilibria. In *16th Annual Symposium on Theoretical Aspects of Computer Science*, pages 404–413, Trier, Germany, 4–6 March 1999.
19. Samuelson Larry. Stochastic stability in games with alternative best replies. *Journal of Economic Theory*, 64(1):35–65, October 1994.
20. Noam Nisan, Tim Roughgarden, Eva Tardos, and Vijay V. Vazirani, editors. *Algorithmic Game Theory*. Cambridge University Press, 2007.
21. Arthur J. Robson and Fernando Vega-Redondo. Efficient equilibrium selection in evolutionary games with random matching. *Journal of Economic Theory*, 70(1):65–92, July 1996.
22. Tim Roughgarden and Éva Tardos. How bad is selfish routing? *J. ACM*, 49(2):236–259, 2002. Also appeared in FOCS 2000.
23. Siddarth Suri. Computational evolutionary game theory. In Noam Nisan, Tim Roughgarden, Éva Tardos, and Vijay V. Vazirani, editors, *Algorithmic Game Theory*. Cambridge University Press, 2007.
24. H Peyton Young. The evolution of conventions. *Econometrica*, 61(1):57–84, January 1993.
25. H Peyton Young. *Individual Strategy and Social Structure*. Princeton University Press, Princeton, NJ, 1998.