The Effect of Collusion in Congestion Games

Extended Abstract

Ara Hayrapetyan*
Cornell University
Dept. of Computer Science
4106 Upson Hall
Ithaca, NY 14853, USA
ara@cs.cornell.edu

Éva Tardos†
Cornell University
Dept. of Computer Science
5153 Upson Hall
Ithaca, NY 14853, USA
eva@cs.cornell.edu

Tom Wexler‡
Cornell University
Dept. of Computer Science
5156 Upson Hall
Ithaca, NY 14853, USA
wexler@cs.cornell.edu

ABSTRACT

In this paper we initiate the study of how collusion alters the quality of solutions obtained in competitive games. The price of anarchy aims to measure the cost of the lack of coordination by comparing the quality of a Nash equilibrium to that of a centrally designed optimal solution. This notion assumes that players act not only selfishly, but also independently. We propose a framework for modeling groups of colluding players, in which members of a coalition cooperate so as to selfishly maximize their collective welfare. Clearly, such coalitions can improve the social welfare of the participants, but they can also harm the welfare of those outside the coalition. One might hope that the improvement for the coalition participants outweighs the negative effects on the others. This would imply that increased cooperation can only improved the overall solution quality of stable outcomes. However, increases in coordination can actually lead to significant decreases in total social welfare. In light of this, we propose the *price of collusion* as a measure of the possible negative effect of collusion, specifying the factor by which solution quality can deteriorate in the presence of coalitions. We give examples to show that the price of collusion can be arbitrarily high even in convex games. Our main results show that in the context of load-balancing games, the price of collusion depends upon the disparity in market power among the game participants. We show that in some symmetric nonatomic games (where all users have access to the same set of strategies) increased cooperation always improves the solution quality, and in the discrete analogs of such games, the price of collusion is bounded by two.

Categories and Subject Descriptors

F.0 [**Theory of Computation**]: General—*Game Theory*

*Supported in part by NSF ITR grant CCR-0325453.

†Supported in part by NSF grant CCR-0311333, ITR grant CCR-0325453, and ONR grant N00014-98-1-0589.

‡Supported in part by NSF ITR grants CCR-0325453 and CCR-0325556.

General Terms

Theory, Economics

Keywords

Game Theory, Coalitions, Congestion Games, Load-Balancing Games, Price of Anarchy, Price of Collusion.

1. INTRODUCTION

In recent years, game theory has been embraced as a powerful tool for understanding the many environments in which the lack of coordination is the primary obstacle to optimization (rather than lack of computational resources or information). Modeling various problems from routing, network design, and scheduling as games played by selfish agents has led to many interesting results. One of the most intriguing lines of research to emerge from this work concerns the price of anarchy [14, 20]. For any game, this ratio compares the cost of solutions reached through selfish behavior (Nash equilibria) to those that can be attained cooperatively. Thus it provides a standard measure by which to ask "How costly is decentralized behavior?"

Much of the work on the price of anarchy uses Nash equilibrium as the solution concept, i.e., defines Nash equilibrium as the outcome in a competitive game. This equilibrium concept is based on the assumption that every player acts independently, in a solely self-interested fashion. And yet, we know that in reality, this tends not to be the case. In many competitive situations, the participating agents *do* collude: contracts are made, businesses agree to cooperate, and corporations merge. Just as we must acknowledge that agents may not all cooperate for the greater good, we must also recognize that, in many cases, agents will form small coalitions to improve their collective well-being. In this paper, we start to explore the impact of coalitions in such environments.

We consider the question of how increased cooperation affects the overall solution quality of stable outcomes. Coalitions cooperate to improve the social welfare of their members, possibly at the expense of those who are not in the coalition. Our question is, does the improvement in social welfare experienced by members of a coalition outweigh the negative effects felt by on others? We might expect that forming coalitions should only help the quality of the resulting outcome: at least members of the coalition are cooperating for a better collective outcome. At the extreme, if all players were to form a single coalition, they could reach an optimal outcome that minimizes the total cost for all players. So clearly, coalitions hold the possibility of improvement. But is it possible that under certain settings, the equilibria reached by players in smaller coalitions are in fact more costly than those reached by players who do not cooperate at all? And if so, to what extent could this happen?

We investigate this question in the context of nonatomic and discrete congestion games. Congestion games are one of the most well understood and well studies class of games in the context of the price of anarchy. Nonatomic congestion games have been studied as a model of traffic routing since the fifties (see Beckmann, McGuire and Winsten [5]). Discrete congestion games were introduced by Rosenthal [22] as a broad class of games possessing pure equilibria, and were shown to be equivalent to the class of potential games by Monderer and Shapley [16]. The complexity of computing Nash equilibria in such games is considered by Fabrikant, Papadimitriou and Talwar [11]. Congestion games are also well understood in the context of the price of anarchy and stability in both the nonatomic case [25, 23] and the discrete case [2, 4, 7, 8].

We present a framework for modeling coalitions and show that in the context of convex load-balancing games, the price of collusion depends on the disparity in market power among the game participants. We show that collusion may adversely affect the price of anarchy if players have access to different strategies, but for symmetric nonatomic load-balancing games, coalitions can only improve solution quality. For symmetric discrete load-balancing games, collusion can increase the solution cost, but we show that this increase is bounded by a factor of two, and that this bound is tight. In contrast, the price of collusion in asymmetric load-balancing games can grow linearly with the number of players.

Our model of coalitions is closely related to the atomic splittable version of congestion games studied by [25, 24, 10] in the context of the price of anarchy. Recently, Cominetti, Correa and Stier-Moses [9] pointed out that the proofs in all these papers have a basic flaw, and many bounds are, in fact, not valid. They also provide corrected (weaker) bounds in some cases. In general congestion games (nonatomic or discrete), price of anarchy results can be used to derive bounds on the price of collusion. However, we believe that the bounds derived from price of anarchy bounds (such as [9] in the nonatomic case, and Awerbuch, Azar and Epstein [4] in the discrete case) are not tight in general, and we show that they can be improved in certain special cases.

Our Model for Collusion.

We consider nonatomic and discrete congestion games, where the delay suffered by each player depends on the congestion of the link(s) it uses. If $\mathcal{A}$ denotes a (pure) strategy for each player, then the social cost of $\mathcal{A}$ is simply the sum of all players' costs. For a game G, if NE is the most costly Nash equilibria and OPT is the solution that minimizes cost, then the price of anarchy of G is defined to be the ratio $c(NE)/c(OPT)$.

We now present a general framework for modeling static coalitions. We assume that members of a coalition display full cooperation, and aim to minimize their collective cost. For a game G, we model a set of coalitions as a partition $P = (P_1, \dots, P_k)$ of the set of players. In the case of nonatomic games all players are partitioned into a finite set of coalitions, whereas in discrete games, some coalitions may contain a single player (players who do no collude at all). We create a new game $G(P)$, in which each coalition P_i of players from G is modeled as a single player i, who selects a strategy for each of the original players in the set P_i. Note that there is a natural one-to-one correspondence between strategies in the collusion-free game G and those in the coalition game $G(P)$, and corresponding strategies have the same social cost.

For concreteness, consider the following two extreme cases. If P groups all players together, then $G(P)$ is a game containing a single player, whose cost function is the social cost function, so the solution chosen by this player will maximize social value. For discrete games, we can also consider the other extreme, where P partitions the players into n singleton sets. In this case, coalitions are the single players, and $G(P) = G$.

We assume that coalitions display full cooperation; that is, they aim to maximize the total social welfare of coalition's members. This is the natural objective for the coalition in cases where the personal objective is money, and we allow monetary transfer. In the context of load-balancing or routing, the personal objectives are to minimize delay rather than cost, and delay may not be so easy to transfer. However, if all coalition members have access to the same set of strategies, then optimizing the collective welfare can still be viewed as the shared objective of all coalition members, assuming the coalitions use randomization. To see this, assume that whenever a coalition splits its players between a number of different strategies, it does so using randomization, so that all players have in expectation the same delay, namely the average delay of the coalition strategy. Note that we assume here that coalitions use coordinated randomization, so the players are always split exactly according to the desired strategy.

With coalitions defined, we consider Nash equilibria in the game where players represent coalitions. At equilibrium, no coalition has an incentive to deviate. Note that coalitions have a much richer set of deviations available to them than are available to players in the collusion-free game, as a coalition considers changing strategies for multiple members simultaneously. Coalitions also evaluate moves differently, as they consider tradeoffs between the costs their members experience. It is not hard to see that a Nash equilibrium in G may not correspond to an equilibrium in $G(P)$, and vice-versa.

Our primary goal is to explore how coalitions effect the cost of equilibria. More precisely, we are interested in the following question: if G is a collusion-free game, and then coalitions P are formed, yielding the game $G(P)$, how much more expensive can the equilibria in $G(P)$ be compared to those in G? Let NE be the most costly Nash equilibrium of G. Let CE be the most costly equilibrium over all games $G(P)$ generated by adding any coalitions P to G. We define the *price of collusion* for G to be $\frac{c(CE)}{c(NE)}$, the factor by which the quality of equilibria can deteriorate when coalitions form. We will see that the difference in players' possible strategies plays an important role in the price of collusion. We call a game *symmetric* if all users have access to the same set of strategies. For example, a routing game in networks is symmetric if all users share the same source-sink pair.

There is a close connection between the price of anarchy and the price of collusion we defined. On one hand, we will see that price of anarchy bounds can often be used to prove analogous bounds for the price of collusion. On the other hand, if we can bound both the price of anarchy and the price of collusion, then we can chain the two to obtain bounds on the quality of Nash equilibria with arbitrary coalitions.

Notice that our game models *static* coalitions, and we do not model the process of forming coalitions. We assume that coalitions are formed exogenously, and not through game play. We note that while dynamic coalitions are clearly of interest as well, a suitable model has thus far proved elusive.

Our Results.

In Section 2 we consider collusion in nonatomic congestion games, which essentially become the splittable flow games considered by Roughgarden [24] and Correa, Schulz, and Stier-Moses [10]. Our main result is that for symmetric nonatomic load-balancing games with convex delays, collusion only improves the solution, so the price of collusion is 1.

In contrast, we show that the price of collusion in asymmetric games can be arbitrarily large, even for a load-balancing game on two links with convex delays. The price of anarchy results of Cominetti, Correa and Stier-Moses [9] for atomic splittable games can be used to derive bounds on the price of collusion for our games. We believe that these bounds are not tight, and for a network with two parallel links, we can show an upper bound of 13/12 for the price of collusion, improving the upper bound of 3/2 implied by the bound on the price of anarchy.

In Section 3 we consider discrete congestion games and observe that collusion can adversely effect the solution even in symmetric load-balancing games with convex delays. However, we show that in such games the price of collusion is bounded by 2. We also present a sequence of games with two links for which the price of collusion tends towards 2 as the number of players increases. Furthermore, we show that although these games are not exact potential games, they still possess pure equilibria. For general discrete congestion games we extend the price of anarchy result of [4] to derive bounds on the price of collusion.

In the case of discrete games, we also consider concave delays. While the coalition-free game always has a pure Nash equilibrium, after forming coalitions, the resulting game is not known to have pure equilibria when latencies are concave. We show that the price of collusion is again at most 2 for pure equilibria (if they exist). We also consider mixed equilibria, and show a somewhat weaker bound of 4 on the price of collusion. This proof compares any randomized Nash equilibrium with coalitions to a pure equilibrium of the basic game. We also present a weaker lower bound, showing that the price of collusion is at least $8/7$.

Related Work.

The price of anarchy was introduced by Koutsoupias and Papadimitriou in [14, 20] and first applied to bound the cost of selfish behavior in a simple load-balancing game. This game is very similar to the congestion games we consider in Section 3, with a few notable distinctions. Their game considers more general weighted jobs (we mostly consider unit jobs); assumes linear delays (we mostly consider convex delays) and analyzes the worst case delays (we consider the average of delay).

Roughgarden and Tardos [25, 23] considered the price of anarchy in a similar game in which players route flow through a general network and attempt to minimize their latency. In contrast to [14], they considered a nonatomic game, in which a continuum of players route flow, each controlling a negligible amount. We will consider such nonatomic games in Section 2. In this model there is no distinction between randomized and pure equilibria. In evaluating a solution [25, 23] consider social welfare (the sum of user delays), while [14] consider the egalitarian objective function of the worst delay. Suri, Tóth and Zhou [26, 13] were the first to focus on the social welfare and pure Nash equilibria in a discrete version of the game. Social welfare of Nash equilibria in discrete routing and load-balancing games were further explored by Christodoulou and Koutsoupias [7, 8] and Awerbuch, Azar and Epstein [4], who extended the results to weighted games and mixed equilibria.

The nonatomic routing game and the unit job versions of both the discrete load-balancing and routing games are all special cases of congestion games. In a congestion game, the cost of using a machine or an edge depends on its load. Congestion games were introduced by Rosenthal [22] as a broad class of games possessing pure equilibria. Rosenthal proved that any congestion game is an *exact potential game*, a game where selfish moves are local steps that decrease a potential function. For finite games, the existence of such a function guarantees a pure Nash equilibrium. Later, Monderer and Shapley [16] proved the converse; every potential game is equivalent to a congestion game.

Rosen [21], Orda, Rom and Shimkin [19], Catoni and Pallottino [6], and Altman et al [1] consider atomic routing games with splittable flow, in which each player control a tangible volume of flow, and may route it fractionally over multiple paths. This is identical to the nonatomic routing game with coalitions. Rosen [21] showed that such games do have pure Nash equilibria. For the case of load balancing, Orda, Rom, and Shimkin [19] show that the equilibria in these games are essentially unique. Altman et al [1] extended the uniqueness results to general networks for a restricted class of delay functions. Roughgarden and Tardos [25], Roughgarden [24] and Correa, Schulz and Stier-Moses [10] study this class of games in the context of the price of anarchy. Recently, Cominetti, Correa and Stier-Moses [9] pointed out that the proofs in all these papers have a basic flaw, and in fact many of the bounds provided are not valid. They also provide corrected (weaker) bounds in some cases.

The discrete version of the load balancing game with coalitions gives rise to a game that is related to weighted load-balancing games. In these games, each job has a weight, and the delay of a machine depends on the total weight of jobs assigned to it. Such weighted games do not have exact potential functions, yet they do have pure Nash equilibria. Milchtaich [15] considers a generalization of load-balancing games in which players have different payoff functions. As in the weighted load-balancing games, the result of this extension is not a potential game, and yet the author proves that such games do have pure Nash equilibria. This proof technique can be adapted to show that our discrete load balancing game with coalitions has pure Nash equilibria as well.

Catoni and Pallottino [6] study how nonatomic equilibria compare to atomic splittable equilibria. They show an instance of a network with two commodities, each of which chooses between a private and shared path, for which the value of the splittable atomic equilibrium is worse than the value of the nonatomic Nash equilibrium. Cominetti, Correa and Stier-Moses [9] give examples where the maximum price of anarchy is affected by collusion. For the case of linear delays, Roughgarden and Tardos [25] show that the price of anarchy is bounded by 4/3, while in [9], the authors show that the price of anarchy can exceed 4/3 in the same class of games with collusion (i.e., with atomic splittable flow), but it is at most 3/2. Our results imply that this degradation of the quality of Nash equilibria cannot happen on symmetric load balancing games. Cominetti, Correa, and Stier-Moses [9] also show that for the case of routing games where all players route between the same source-sink pair, and all coalitions have the same amount of flow to route, the social cost of the coalition equilibrium is no worse than the nonatomic equilibrium. While their results are for the more general case of routing between a source-sink pair, and we consider only load balancing, they assume a higher level of symmetry between players: not only do they assume that all players have access to the same set of strategies, but they also assume that all coalitions have the same size (and hence the same market power).

While there has been much work in the social science literature trying to understand the process of coalition formation, there is no generally accepted game theoretic model of such a process. There have been a number of solution concepts considered for equilibria where coalitions may act together, starting with the strong equilibrium concept of Aumann [3]. In a strong equilibrium, no coalition has a deviation that is improving for each of its players. This concepts is too strong in the sense that strong equilibria may not exists. There have been many other variants proposed, such as the notion of coalition-proof equilibria of Moreno and Wooders [17]. Coalitions have also been considered from the perspective of mechanism design. For example, a number of papers, such as Moulin and Shenker [18], have dealt with the problem of group-strategyproof cost sharing, while Goldberg and Hartline [12] consider group-strategyproof mechanisms in the context of auction design.

In contrast, our work studies the possible equilibria that result from the formation of coalitions, rather than finding solutions, such as strong equilibria, for which coalitions have no incentive to form. Our primary goal is to understand the potential negative effects that coalitions may have on the outcome of selfish interactions.

2. COLLUSION IN NONATOMIC GAMES

In this section we consider coalition equilibria for nonatomic games. A basic nonatomic routing game is given by a network $G = (V, E)$ with nonnegative, monotone increasing latency functions $\ell_e(x)$ on the edges $e \in E$, and source-sink pairs $s_i, t_i \in V$ with demands r_i that need to get routed from the source s_i to the sink t_i in G. Consider a flow f that ships r_i units from source s_i to sink t_i, and let f_e denote the flow on edge e. Given this flow f, each edge has delay $\ell_e(f_e)$. The flow f is said to be at Nash equilibrium, if for all source-sink pairs i, all r_i units of the flow travel on the minimum delay paths between s_i and t_i subject to the delays $\ell_e(f_e)$). It is well known that for any such network, flows at Nash equilibrium exist and the delays suffered by this flow are unique (see Beckmann, McGuire and Winsten [5]). We will further assume that the latencies $\ell_e(x)$ are convex (or at least that $x\ell_e(x)$ is convex). Roughgarden [23] calls this class of latency functions *standard*. Convex functions are often used to model delay or response time as a function of congestion or load.

Now consider a small quantity of flow that forms a coalition. We will assume that such a coalition aims to minimize the average delay of the flow belonging to the coalition, rather than minimizing the delay greedily for all units. In the case when all members of a coalition have the same source-sink pair, this shared objective function can be viewed as the expected delay of all members of the coalition, assuming that coalition members get randomly assigned to the different paths used by the coalition. If the flow r_i associated with each source-sink pair forms a coalition, then the notion of coalition flow is identical to the atomic splittable flow considered by Roughgarden [24], Correa, Schulz and Stier-Moses [10] and by Roughgarden and Tardos [25]. Rosen in [21] proves that Nash equilibria in such games exist. Orda, Rom, and Shimkin [19] show that the equilibria are essentially unique for the case of load-balancing games, and Altman, et al [1] extended the uniqueness results to general networks for a restricted class delay functions.

An example by Catoni and Pallottino [6] establishes that coalitions can actually hurt the overall quality of the resulting flow. They give an example of this on three links with two classes of asymmetric players, where each class has access to a private and a shared path. The resulting game has a tiny price of collusion. We give an example of an asymmetric load-balancing game with two links that has arbitrarily high price of collusion.

Our focus will be to show bounds on the *price of collusion*, and we will compare the cost of coalition equilibrium to that of nonatomic flow equilibrium. Price of anarchy bounds can be used to bound the price of collusion, but our results suggest that for many classes of games, the price of collusion is significantly smaller than the bounds implied by the price of anarchy.

In Section 2.1 we consider symmetric load-balancing games (in which all players can use all edges), and show that the price of collusion is 1 for any class of convex latencies. Load-balancing games are routing games where G is a graph with two nodes connected by parallel links. Recall, that the price of anarchy can be arbitrarily high for these games, depending on the class of latency functions. In contrast, we prove that in this setting there is no price to collusion, and the intuition that coordination is beneficial is correct.

We believe that for many classes of games, the price of collusion is significantly smaller than the price of anarchy in the corresponding games with collusion. In Section 2.2 we establish this in an asymmetric load-balancing games for graphs with two parallel links, and linear delays. Two parallel links provide the worst price of anarchy in routing games with nonatomic flow, which is 4/3 for linear delays. We consider such networks in Section 2.2 and show that the price of collusion is at most 13/12. We also give an example to show that the price of collusion is at least 16/15.

2.1 Symmetric Nonatomic Games

In this section we prove that the price of collusion for symmetric nonatomic load-balancing games with convex latencies is equal to 1. In other words, forming coalitions can only improve the overall solution quality.

In a symmetric nonatomic load-balancing game, there is a flow of r that wants to route from a source to a sink using a set of parallel links $\{1, \ldots, n\}$ (or wants to balance load on n machines) with associated convex latencies $\ell_1(.), \ldots, \ell_n(.)$. We assume that all links are available for all flow. At Nash equilibrium NE, the flow will choose to route through the minimum latency edges. As a result, at Nash equilibrium all flow will suffer a common delay L and the total cost (total delay) will be $c(NE) = L \cdot r$.

We now assume that this flow is organized into coalitions. More formally, we consider the formation of m coalitions, where coalition i controls a volume of r_i flow, and $\sum_i r_i = r$. Each coalition is interested in minimizing the total latency suffered by its flow. To analyze the Nash equilibrium of this new game, we need to understand how a coalition will distribute its flow among different links. We use CE to denote the resulting coalition equilibrium outcome and $c(CE)$ to denote its total cost. We will also use f_s to denote the amount of flow on link s under CE. Using this notation the cost of the solution $c(CE)$ can be expressed as $c(CE) = \sum_s f_s \cdot \ell_s(f_s)$, as f_s flow suffers $\ell_s(f_s)$ delay on link s, regardless of which coalitions that flow belongs to.

Before proving our main theorem, we state the convex optimality condition that flows in the coalition equilibrium must satisfy. Consider a link s that has f_s units of flow, and a coalition i with x_s flow on link s. The cost of this flow to coalition i is $x_s \cdot \ell_s(f_s)$. Increasing i's flow on link i increases the cost at the rate of $\ell_s(f_s) + x_s \cdot \ell'_s(f_s)$, as the added flow suffers delay $\ell_s(f_s)$ while the existing x_s units of flow of coalition i on link s will increase their delay at the rate of $\ell'_s(f_s)$.

Lemma 2.1 *A flow is a coalition equilibrium if, for all links s and t and for all coalitions i that have $x_s > 0$ flow on link s and x_t flow on link t, the following inequality holds.*

$$\ell_s(f_s) + x_s \cdot \ell'_s(f_s) \le \ell_t(f_t) + x_t \cdot \ell'_t(f_t).$$

Now we are ready to prove the main theorem of this section. Recall that at Nash equilibrium NE, all flow uses links with delay L. We call a link *overloaded* if in CE it carries non-zero flow and has delay more than L, and *underloaded* if it has delay less than L. The key idea is to show that in CE, moving a small amount of flow from any overloaded link to any underloaded one increases the social cost.

Lemma 2.2 *Assume link s is overloaded and link t is underloaded in a coalition Nash equilibrium CE. Then moving a small amount of flow from s to t in this equilibrium cannot decrease the social cost.*

The flows on links s and t contribute $f_s \cdot \ell_s(f_s) + f_t \cdot \ell_t(f_t)$ to the total cost of the coalition equilibrium. The lemma considers moving an (infinitesimally) small amount of flow from s to t. To show that this change does not decrease the overall cost, we need to consider the corresponding derivative: the rate of increase and decrease in cost as we decrease flow on link s and increase flow on link t. Recall that $[x \cdot \ell(x)]' = \ell(x) + x \cdot \ell'(x)$ by the chain rule. Hence the above lemma states that

$$\begin{aligned} [f_s \cdot \ell_s(f_s)]' &= \ell_s(f_s) + f_s \cdot \ell_s'(f_s) \\ &\leq \ell_t(f_t) + f_t \cdot \ell_t'(f_t) = [f_t \cdot \ell_t(f_t)]'. \end{aligned} \quad (1)$$

Before proving the lemma, we want to establish that the lemma immediately implies the main theorem.

Theorem 2.3 *For a nonatomic, symmetric load-balancing game with latencies $\ell_1(.), \ldots, \ell_n(.)$ such that $x \cdot \ell_s(x)$ is convex for all links s, and for any choice of coalitions, $c(CE) \leq c(NE)$.*

PROOF. Lemma 2.2 states that moving an infinitesimally small amount of flow from some overloaded link s that has flow f_s to an underloaded link t that has flow f_t does not decrease the social cost. Note that, for such pairs of links s and t, the lemma also implies that for any flow amounts $\hat{f}_s \leq f_s$ and $\hat{f}_t \geq f_t$ on links s and t respectively, moving an infinitesimally small amount of flow from link s to link t increases the cost as well. This immediately follows from (1), and the assumption that both $x \cdot \ell_s(x)$ and $x \cdot \ell_t(x)$ are convex functions, and hence have monotone increasing derivatives.

One can obtain the Nash equilibrium NE from the coalition equilibrium CE by moving flow from overloaded links to underloaded links. The observation above implies that this process will only increase the cost of the solution. Hence $c(CE) \leq c(NE)$. □

PROOF OF LEMMA 2.2. Recall that by Lemma 2.1 if a coalition i puts $x_s > 0$ flow on link s and x_t flow on link t, then the delays must satisfy $\ell_s(f_s) + x_s \cdot \ell_s'(f_s) \leq \ell_t(f_t) + x_t \cdot \ell_t'(f_t)$. If we sum this inequality over all coalitions (say k of them) that have flow on s, then on the left side we get $k \cdot \ell_s(f_s) + f_s \cdot \ell_s'(f_s)$. Furthermore, by adding $x \cdot \ell_t'(f_t) \geq 0$ to the right side, where x is the amount of flow on t that does not belong to any of these k coalitions, we get $k \cdot \ell_s(f_s) + f_s \cdot \ell_s'(f_s) \leq k \cdot \ell_t(f_t) + f_t \cdot \ell_t'(f_t)$. By assumption s is overloaded and t is underloaded, so we have $\ell_s(f_s) > \ell_t(f_t)$. Hence, $\ell_s(f_s) + f_s \cdot \ell_s'(f_s) \leq \ell_t(f_t) + f_t \cdot \ell_t'(f_t)$. □

2.2 Asymmetric Nonatomic Games

In this section we analyze the effect of coalitions in asymmetric nonatomic games. First, we provide an example to show that the price of collusion in asymmetric games can be arbitrarily high even with convex latencies. The example considers two links and two classes of players grouped into two coalitions as shown in Figure 1. Let $\alpha > 0$ be a large number. Coalition 1 needs to route $r_1 = \alpha + 1/\alpha$ units of flow, and has access only to link 1, which has latency $\ell_1(x) = \max(0, \alpha(x - \alpha))$. Note that r_1 flow on link 1 incurs a latency of 1. Coalition 2 needs to route $r_2 = \alpha + 1$ units of flow, and has access to both links, where the second link has latency $\ell_2(x) = \max(0, x - \alpha)$. At Nash equilibrium all of coalition 2's flow will be routed on link 2. Hence the delay on both links is 1, and the total cost of coalition-free equilibrium is $\Theta(\alpha)$. At coalition equilibrium, the second coalition will move 0.5 flow to link 1, thereby increasing the delay on link 1 to approximately $\alpha/2$. Thus, the total cost of flow on link 1 becomes $\approx \alpha^2$, which is an $\Omega(\alpha)$ factor larger than the coalition-free equilibrium cost.

On a positive note, a bound on the price of anarchy of routing games with atomic splittable flows immediately implies a bound on the price of collusion of nonatomic routing games. More generally,

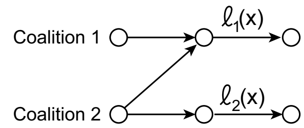

Figure 1: An asymmetric game with high price of collusion.

Theorem 2.4 *For any game G, the price of collusion of G is at most the maximum price of anarchy of $G(P)$ over all coalitions P.*

However, such bounds are not always tight. In particular, for the special case of linear delays (delays of the form $\ell_s(x) = a_s x + b_s$) and two parallel links, the price of anarchy is at least 4/3. In the extended version of this paper we prove that for this special case, the price of collusion is significantly smaller.

Theorem 2.5 *In nonatomic load-balancing games with linear delays and two parallel links, the price of collusion is at most 13/12.*

Finally, we consider games with pure linear delays, where $\ell_s(x) = a_s x$ with no constant terms. It is interesting to note that the Nash equilibrium of this game is actually optimal (the delay $a_s x$ and the derivative of the cost $[a_s x^2]' = 2a_s x$ differ exactly by a factor of 2), so it is easy to see that in such games collusion does have a price: Nash is optimal and unique, but coalition Nash often rearranges flow, and thus may not be optimal. In the full version of this paper, we prove the following upper bound for this case.

Theorem 2.6 *In nonatomic load-balancing games with pure linear delays and two parallel links, the price of collusion is at most 16/15.*

Finally, we give an example that achieves this 16/15 price of collusion on two parallel links with pure linear delays. We have latency functions $\ell_1(x) = x$ and $\ell_2(x) = \frac{7}{8}x$, with a total of 10 units of flow, 5 of which is fixed on the first link. We form two coalitions; one containing all 5 units of flow that is restricted to the first link, and the other with the remaining unrestricted 5 units. Then $c(NE) = 375/8$, while $c(CE) = 400/8$, yielding the desired lower bound.

3. COLLUSION IN DISCRETE GAMES

In this section we will consider a discrete version of the load-balancing or routing game from the previous section. A discrete congestion game consists of n players each with a non-negative weight $w_i \geq 0$ and a source-sink pair, together with the network and the delays as given in the nonatomic game. Player i wants to select a path connecting its terminals and route its w_i unit of flow on this single path. The delay $\ell_s(x)$ on each edge s is now a function of the total weight of the players using the edge. The goal of the players is still to select a path with minimal total latency.

We will be primarily concerned with the special case of load-balancing games, where the network G consists of a set of parallel links, and players have access to a different subsets of the links. We will focus on the uniform case when all weights are the same ($w_i = 1$ for all i). We say that the game is symmetric if all players have access to all links. The goal of this section is to see to what extent the results obtained for the nonatomic games in Section 2 carry over to discrete games. As in the nonatomic model, the lack of symmetricity can induce a large price of collusion even in the uniform case: one can design examples of asymmetric uniform load-balancing games on two links with convex latencies where the price of collusion is linear in the number of players.

One would hope that the main positive result of the previous section would also carry through - that the price of collusion for uniform symmetric load-balancing games with *convex* latencies would also be 1. Unfortunately, as we shall see below, this is not true. Our main result in this section, which is presented in Section 3.1, proves that the price of collusion is at most 2 for these games, and this bound is tight. In Section 3.2 we extend the work of [4] to give bounds for the price of anarchy (and hence the price of collusion) of general weighted congestion games with coalitions, where link latencies are restricted to degree-d polynomials.

For the case of discrete games we also consider uniform symmetric load-balancing games with *concave* latencies. In this case the existence of pure Nash equilibrium is not known to be guaranteed. In Section 3.3 we give bounds for the price of collusion for pure and mixed coalition equilibria for such games.

For conciseness, we will drop the adjectives *uniform* and *symmetric* in Sections 3.1 and 3.3, since they are assumed throughout.

3.1 Games with Convex Latencies

In this section we analyze load-balancing games with convex latencies. In Theorem 3.5 we prove that the price of collusion is at most 2, and that this bound is asymptotically tight. We then show in Theorem 3.9 that pure equilibria always exist in these games with coalitions, even though they are often not potential games. We begin with an example in which the price of collusion approaches 2 as the number of players grows.

Example : Consider an instance of the game with two links and $2k$ units. Define latencies as follows.

$$\ell_1(x) = \begin{cases} 0 & x \leq k-1 \\ 2 & x = k \\ \infty & x \geq k+1 \end{cases} \qquad \ell_2(x) = \begin{cases} 0 & x \leq k \\ 1 & x = k+1 \\ \infty & x \geq k+2 \end{cases}$$

The Nash equilibrium without coalitions places $k-1$ units on the first link and $k+1$ on the second, for a total cost of $k+1$. Now form k coalitions of size 2. If each coalition places one unit on both links, the resulting strategy is at equilibrium; moving a unit from the first link to the second does not yield an improvement in cost. The total cost of this solution is $2k$. Thus the price of collusion for this game is $\frac{2k}{k+1}$. ■

The Price of Collusion.

We now prove that 2 is in fact an upper bound on the price of collusion for atomic load-balancing games with convex latencies. For ease of presentation, we will assume that link latencies are distinct, i.e. $\ell_s(x) \neq \ell_t(y)$ for all x, y and $s \neq t$ (this assumption can be dropped with minor changes to the proofs). A useful consequence of this assumption is that there is now a unique Nash equilibrium NE in the collusion-free game.

We use the following notation. Let CE be any Nash equilibrium in the game with coalitions. Let n_s^i denote the number of units that coalition i puts on link s under CE. Let $n_s = \sum_i n_s^i$ denote the total number of units on link s under CE and let $n^i = \sum_s n_s^i$ denote the total number of units controlled by coalition i. Similarly, let n'_s denote the number of units on link s under NE. Finally, we denote the latency of link s under CE and NE as $\ell_s = \ell_s(n_s)$ and $\ell'_s = \ell_s(n'_s)$ respectively.

We partition the links into classes based on the loads they carry under NE and CE. We refer to $B = \{s \mid n_s = n'_s\}$ as the set of *balanced* links, and $U = \{s \mid n_s < n'_s\}$ as the *underloaded* links (i.e. those links for which CE has fewer units than NE). The links $L = \{s \mid n_s = n'_s + 1\}$ are called *lightly overloaded*, and the remaining links $H = \{s \mid n_s > n'_s + 1\}$ are called *heavily overloaded*. We call a link *overloaded* if it is in either L or H.

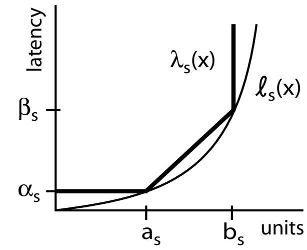

Figure 2: Simplifying the latency curves.

To make the calculations simpler, we begin by showing that we can simplify the latency functions without loss of generality. We modify the latencies of the links so as to preserve both the cost and the stability of NE and CE. For a link s, let $\alpha_s = \min(\ell_s, \ell'_s)$ and $\beta_s = \max(\ell_s, \ell'_s)$ denote the latencies of this link under NE and CE, and let $a_s = \min(n_s, n'_s)$ and $b_s = \max(n_s, n'_s)$ denote the corresponding loads. Define $\delta_s = \frac{\beta_s - \alpha_s}{b_s - a_s}$. We now define a new latency function

$$\lambda_s(x) = \begin{cases} \alpha_s & \text{if } x < a_s \\ \alpha_s + (x - a_s)\delta_s & \text{if } a_s \leq x \leq b_s \\ \infty & \text{if } x > b_s \end{cases}$$

See the example shown in Figure 2. Notice that for the loads given by NE and CE on link s, the new latency function $\lambda_s(x)$ matches the original function $\ell_s(x)$. Furthermore, for any other load, $\lambda_s(x)$ is at least as large $\ell_s(x)$. Hence, both NE and CE are still Nash equilibria for their respective games. Moreover, the social costs of the two equilibria are unchanged and thus, from now on, we will assume without loss of generality that all latency functions $\ell_s(x)$ are of this simplified form.

The main challenge will be to bound the cost paid by CE on the overloaded links, since on the remaining links, CE pays no more than NE. A key insight is that the heavily overloaded links behave analogously to overloaded links in the nonatomic version of this game: they realize a social benefit of collusion. The proof of the following lemma is essentially an adaptation of the the proof of the analogous Lemma 2.2 for nonatomic games to discrete load-balancing games.

Lemma 3.1 *Let $s \in H$ be a heavily overloaded link, let $t \in U$ be an underloaded link, and let $\mathcal{A}$ be the strategy that results from starting at CE and moving a single unit from s to t. Then $c(\mathcal{A}) \geq c(CE)$.*

PROOF. Consider an overloaded link s. Suppose there exists an underloaded link t, such that moving a single unit from s to t improves the cost of CE. It will suffice to show that $n_s = n'_s + 1$. In other words, we will argue that s was in fact only lightly overloaded.

Under CE, the total cost of the units on links s and t is $n_s\ell_s + n_t\ell_t$. For $\mathcal{A}$, the total cost of these links is $(n_s - 1)(\ell_s - \delta_s) + (n_t + 1)(\ell_t + \delta_t)$. These two solutions have the same cost for all other links. Hence, by our assumption, we have the following inequality:

$$n_s\ell_s + n_t\ell_t > (n_s - 1)(\ell_s - \delta_s) + (n_t + 1)(\ell_t + \delta_t). \quad (2)$$

On the other hand, we know that CE is an equilibrium for all coalitions, and thus no coalition has incentive to move one unit from s to t. In particular, for any coalition i who participates on s we must have

$$n_s^i \ell_s + n_t^i \ell_t \leq (n_s^i - 1)(\ell_s - \delta_s) + (n_t^i + 1)(\ell_t + \delta_t).$$

Summing this inequality over all coalitions who participate on s, yields

$$\sum_i n_s^i \ell_s + \sum_i n_t^i \ell_t \qquad (3)$$
$$\leq \sum_i (n_s^i - 1)(\ell_s - \delta_s) + \sum_i (n_t^i + 1)(\ell_t + \delta_t).$$

where the summation for i is running only over all coalitions that have a unit on link s. Now adding the trivial inequality $n_t^i \ell_t \leq n_t^i(\ell_t + \delta_t)$ for coalitions that participate in link t but not in link s, we get that the left hand side is equal to $n_s \ell_s + n_t \ell_t$, matching the upper bound of inequality (2). The right side is $(n_s - 1)(\ell_s - \delta_s) + (n_t + 1)(\ell_t + \delta_t) - (k_s - 1)(\ell_s - \delta_s) + (k_s - 1)(\ell_t + \delta_t)$, where k_s denotes the number of coalitions participating on link s. Notice that this latter quantity matches the right hand side of inequality (2) with the addition of the term $(k_s - 1)[(\ell_t + \delta_t) - (\ell_s - \delta_s)]$. Therefore for the inequalities 2 and 3 to hold simultaneously, it must be the case that $(k_s - 1)[(\ell_t + \delta_t) - (\ell_s - \delta_s)] > 0$. Thus $(\ell_t + \delta_t) > (\ell_s - \delta_s)$. Since t was underloaded, $\ell_t + \delta_t \leq \ell'_t$. But we also know that $\ell'_t \leq \ell'_s + \delta_s$, since NE is a stable configuration without coalitions and hence no unit can benefit by moving from t to s. Combining these last three inequalities yields

$$\ell'_s + \delta_s > \ell_s - \delta_s. \qquad (4)$$

Recall that ℓ_s and ℓ'_s were the latencies of this link under CE and NE respectively, and adding (subtracting) δ_s gives the latency with one more (less) unit. Therefore, inequality (4) implies that the difference between the number of units that CE and NE put on link j can not be more than one. □

Thus the heart of this proof lies in bounding the cost of the lightly overloaded links. To begin, we group links from L and U into clusters as follows. Each cluster R contains some number d (which may vary for each cluster) of lightly overloaded links and a single underloaded link t such that $n_t \leq n'_t - d$. In other words, the underloaded edge has sufficient space to take one unit from all of the lightly overloaded links in R without becoming overloaded itself. We can choose any such grouping, so long as we ensure that each lightly overloaded link appears in some cluster. We will use NE_R and CE_R to denote NE and CE restricted to the links in R.

For each cluster R, we construct an intermediate allocation $\mathcal{I}_R$ by starting with CE_R and moving a single unit from each lightly overloaded link to the underloaded one. Thus, $\mathcal{I}_R$ is identical to NE on the overloaded links, and puts exactly d units more than CE on the underloaded link t.

Let us assume without loss of generality that the cluster R consists of the first $d+1$ links, with the first d of these being the lightly overloaded links, and the $(d+1)$-st link being the underloaded one. We will use δ to denote δ_{d+1}, and κ to denote the number of coalitions participating in any of the first d links. The following lemma will be crucial in bounding the cost of the CE_R.

Lemma 3.2 *For all coalitions i participating on any of the links $1, \ldots, d$, $n_{d+1}^i \geq d$. In other words, under CE, any coalition that uses an overloaded link in R must have at least d units on the underloaded link.*

PROOF. For the sake of contradiction, assume that coalition i has at least one unit on some overloaded link $s \in \{1, \ldots, d\}$, and has at most $d-1$ units on the last link. As CE was a stable state for this coalition, it has no incentive to move a unit from link s to link $(d+1)$. The saving it gets from taking the unit off of s is at least ℓ_s, the current latency at of that link. The extra cost it incurs by adding that unit to link $(d+1)$ is at most $(d-1)\delta + (\ell_{d+1} + \delta) = d \cdot \delta + \ell_{d+1}$ (the first part accounts for the extra cost that all of its units on the last link will incur, and the second part accounts for the latency of the new unit). Hence $\ell_s \leq d \cdot \delta + \ell_{d+1}$.

On the other hand, $\ell_s > d \cdot \delta + \ell_{d+1}$, since otherwise a player in the collusion-free Nash equilibrium NE (which has at least d more units on link $(d+1)$ than CE) would have an incentive to move from link $(d+1)$ to link s and incur a lower cost [1]. Thus we have a contradiction. □

We can now bound the cost of CE_R in terms of $\mathcal{I}_R$.

Lemma 3.3 *For any cluster R, $c(CE_R) \leq 2c(\mathcal{I}_R)$*

PROOF. It is sufficient to prove that the cost difference $\Delta = c(CE_R) - c(\mathcal{I}_R)$ is at most $c(\mathcal{I}_R)$. To do this, we need a lower bound on $c(\mathcal{I}_R)$, and an upper bound on Δ. Recall that $\mathcal{I}_R$ puts exactly d units more on link $(d+1)$ than CE. These extra units increase the cost paid by the original n_{d+1} units on this link by at least δ (as $d \geq 1$), so the original n_{d+1} units on link $(d+1)$ pay at least $n_{d+1}(\ell_{d+1} + \delta)$. Counting only the cost of these n_{d+1} units we get $c(\mathcal{I}_R) \geq n_{d+1}(\ell_{d+1} + \delta)$. Lemma 3.2 implies that $n_{d+1} \geq \kappa \cdot d$, where κ is the number of coalitions participating in the first d links. So we get the lower bound we will use:

$$c(\mathcal{I}_R) \geq n_{d+1}(\ell_{d+1} + \delta) \geq \kappa \cdot d \cdot (\ell_{d+1} + \delta). \qquad (5)$$

Now consider the difference $\Delta = c(CE_R) - c(\mathcal{I}_R)$. First we express Δ in terms of n_s, ℓ_s and δ_s. Then we consider each of the κ coalitions participating on any of the first d links. None of these coalitions has an incentive to move one unit from an earlier link to link $(d+1)$. We use this to give an inequality regarding the delays on the links. Summing over all coalitions will allow us to bound the difference Δ.

We can express the difference Δ exactly in terms of the quantities n_s, ℓ_s and δ_s as follows. The cost of CE_R on any link $s \in \{1 \ldots d\}$ exceeds the costs of $\mathcal{I}_R$ on that link by exactly $(\ell_s - \delta_s) + n_s \cdot \delta_s$, since the two configurations differ by only a single unit. On link $(d+1)$, $\mathcal{I}_R$ puts d units more than CE_R, so it overpays by $(n_{d+1} + d)(\ell_{d+1} + d \cdot \delta) - n_{d+1} \cdot \ell_{d+1}$. Hence,

$$\Delta = \sum_{s \in [1 \ldots d]} [(\ell_s - \delta_s) + n_s \cdot \delta_s] \qquad (6)$$
$$- [(n_{d+1} + d)(\ell_{d+1} + d \cdot \delta) - n_{d+1} \cdot \ell_{d+1}].$$

Since CE is a coalition equilibrium, no coalition i has an incentive to move a unit from any link $s \in \{1, \ldots, d\}$ to link $(d+1)$. Hence for any coalition i participating on link s, we have $(\ell_s - \delta_s) + n_s^i \cdot \delta_s \leq (\ell_{d+1} + \delta) + n_{d+1}^i \cdot \delta$. Summing this inequality over all coalitions on link s yields

$$k_s(\ell_s - \delta_s) + n_s \cdot \delta_s \leq k_s(\ell_{d+1} + \delta) + \sum_{i \text{ on } s} n_{d+1}^i \cdot \delta$$
$$\leq \kappa(\ell_{d+1} + \delta) + n_{d+1} \cdot \delta, \qquad (7)$$

[1]This is the only place we use the fact that latencies are distinct. This assumption can be removed by altering the proof slightly.

where k_s denotes the number of coalitions participating on link s and recall that κ denotes the number of coalitions participating on the first d links. Since $k_s \geq 1$, we can apply inequality 7 to equation 6 yielding

$$\begin{aligned}\Delta &\leq \sum_{s\in[1\ldots d]} [\kappa(\ell_{d+1}+\delta) + n_{d+1}\cdot\delta] - \\ &\qquad [(n_{d+1}+d)(\ell_{d+1}+d\cdot\delta) - n_{d+1}\cdot\ell_{d+1}] \\ &= d\cdot\kappa\cdot\ell_{d+1} + d\cdot\kappa\cdot\delta + d\cdot n_{d+1}\cdot\delta - \\ &\qquad n_{d+1}\cdot d\cdot\delta - d\cdot\ell_{d+1} - d^2\cdot\delta \\ &\leq d\cdot\kappa(\ell_{d+1}+\delta). \qquad (8)\end{aligned}$$

Note that our upper bound on Δ matches our lower bound for $c(\mathcal{I}_R)$, completing the proof. □

We can now define an intermediate allocation $\mathcal{I}$ for all links by using the intermediate configurations $\mathcal{I}_R$ for each cluster R and using the allocation given by CE for all remaining links. Lemma 3.3 implies that $c(CE) \leq 2\cdot c(\mathcal{I})$.

Analogous to the proof of the main theorem 2.3 we can use Lemma 3.1 to show that the cost of the intermediate solution $\mathcal{I}$ does not exceed that of the collusion-free Nash equilibrium NE.

Lemma 3.4 $c(\mathcal{I}) \leq c(NE)$.

PROOF. Consider converting allocation $\mathcal{I}$ to NE by moving a unit at a time from an overloaded link to an underloaded one. Recall that the construction of $\mathcal{I}$ ensured that the only overloaded links are heavily overloaded. Furthermore, Lemma 3.1 implies that moving a unit away from a heavily overloaded link H to an underloaded link in U does not decrease social cost. Since in $\mathcal{I}$ the underloaded links have no fewer units than in CE, this property still holds. The convexity of latencies implies that if moving one unit between two links increases the social cost, then moving subsequent units will further increase the social cost. Hence in each step of the conversion from $\mathcal{I}$ to NE the social cost only increases, and thus $c(\mathcal{I}) \leq c(NE)$. □

Taken together, Lemmas 3.3 and 3.4 prove our main result.

Theorem 3.5 *The price of collusion for congestion games with convex latencies is at most 2.*

The Existence of Pure Equilibria.

In the absence of coalitions, congestion games are exact potential games, and thus possess pure equilibria. However, after adding coalitions, exact potential functions may no longer exist. In fact, there are potential games that do not have pure equilibria after coalitions are added. Nevertheless, as we now show, load balancing games with coalitions always have pure equilibria. The proof of the theorem is along the same lines as that of Milchtaich [15].

Consider a load-balancing game with coalitions $G(P)$. Let k be the number of coalitions, and let n^i be the number of units possessed by coalition i. We start by considering a variation of $G(P)$, in which each coalition controls only a single unit. This game has a pure equilibrium, since it is a potential game. We will then repeatedly pick a coalition i that does not yet control n^i units, and add a single unit to that coalition. For each addition, we inductively argue that the resulting game also has a pure equilibrium. In particular, we argue that if we start from a Nash equilibrium, and a single coalition is given an additional unit to place, natural game play will quickly converge to a stable solution. Thus, our result also suggests a natural method for efficiently computing pure equilibria.

Suppose we have a strategy $\mathcal{A}$ that is at equilibrium. Consider assigning an additional unit to some coalition i. Our first lemma characterizes coalition i's best response given that it has an extra unit to place.

Lemma 3.6 *Given that all other players assign their units as described by $\mathcal{A}$, a best response for coalition i is to add the new unit to one of the links without rearranging its existing allocation.*

PROOF. Let A_i denote coalition i's strategy prior to the introduction of the additional unit. A coalition's strategy can be naturally viewed as a vector of length m, with each term indicating how many units are placed on the corresponding link. Consider the best response A'_i for coalition i (with the additional unit) that has minimum L_1-distance to A_i. It suffices to show that A'_i places at least as many units as A_i on any given link. Suppose otherwise. Then there is some link s on which i places fewer units under A'_i than under A_i. Since A'_i assigns more units in total, there must be a link t where A'_i places more units than A_i.

Consider starting from A'_i and moving one of i's units from t to s, creating A''_i. Since A''_i has a smaller L_1-distance to A_i, it can't be a best response, and hence it must be beneficial to move that unit back from s to t. In particular, the marginal savings of removing a unit from s must strictly exceed the marginal cost of placing a unit on t. But now consider modifying A_i by moving a single unit from s to t. Since A_i has at most as many units on t as A''_i does, and latencies are convex, the marginal cost of placing a unit on t under A_i is no greater than under A''_i. Likewise, the marginal savings of removing a unit from s it at least as large under A_i as it is under A''_i. Thus there is a positive net savings incurred by moving a unit from s to t under A_i. This contradicts our assumption that we were at equilibrium before the arrival of the new unit. □

We now know how coalition i will deal with an additional unit. Unfortunately, after placing that unit on link s, other coalitions may have an incentive to deviate. Observe that only those coalitions using link s could now have improving moves. Let j be a coalition using link s with an improving move (if none exist, we have an equilibrium), and consider removing one unit controlled by that coalition from the game.

Lemma 3.7 *After removing one unit of coalition j from link s, the resulting strategy is at equilibrium.*

PROOF. From the perspective of all coalitions other than i or j, the resulting strategy is indistinguishable from $\mathcal{A}$, and hence they do not have incentives to move. The cost of i's strategy A'_i only improved when one unit was removed from s; hence if i has an improving move, it can only be to move an extra unit to link s, but then this move would have been even more beneficial in the original strategy $\mathcal{A}$. To see that coalition j has no improving move, note that, all of j's remaining units have the same cost as they did under strategy $\mathcal{A}$. If j has an improving move, it is to move a new unit to link s (where j now has fewer units). However, if coalition j has a such an improving move, then j would have been content before a unit was removed from s. □

Thus, ignoring the removed unit, all players are at equilibrium, and the number of units per link is the same as under $\mathcal{A}$. Hence the scenario is equivalent to the initial setting, only now, an extra unit is assigned to coalition j, rather than i. Coalition j might place this unit on some link t, causing some other coalition using t to remove a unit, which must be then placed on another link, and so forth. We now consider a sequence of such moves and argue that it has bounded length.

Lemma 3.8 *The best response dynamic described above terminates.*

PROOF. It suffices to show that if a coalition adds a unit to a link, it never removes a unit from that link. Suppose this is not the case for coalition i on link s. Over all possible sequences of best response steps, choose one of minimum length. Denote by $T_0, \ldots, T_f$ the time steps within this sequence during which coalition i deviates. Notice that i can not move units to or from s on any of the time steps $T_1, \ldots, T_{f-1}$.

Lemma 3.7 implies that during each of the time steps T_τ, coalition i removes a single unit from some link and moves it to another. We can decompose this process as follows: (a) Coalition i first removes the unit from a link, (b) the marginal cost of placing the removed unit on each link t is calculated (thus creating a list of numbers $\alpha_{\tau,t}$ for each time step T_τ), where $\alpha_{T_0,s}$ is the marginal cost when moving a unit of coalition i to link s and (c) the extra unit is placed on a link $t(\tau)$ minimizing $\alpha_{\tau,t(\tau)}$.

Consider two consecuive values $\alpha_{\tau,t(\tau)}$ and $\alpha_{\tau+1,t(\tau+1)}$. If $t(\tau) \neq t(\tau+1)$, then we argue that $\alpha_{\tau,t(\tau)} \leq \alpha_{\tau+1,t(\tau+1)}$ must hold, as the link $t(\tau+1)$ was not selected at time τ, and had the same number of units, hence the same marginal cost for coalition i. If $t(\tau) = t(\tau+1)$ the inequality holds as at time $\tau+1$ coalition i has more units on this link than at time τ.

Now consider the final step T_f when coalition i is moving a unit from link s to another link with marginal cost $\alpha_{T_f,t(T_f)}$. From the above sequence of inequalities we get that $\alpha_{T_0,s} = \alpha_{T_0,t(T_0)} \leq \alpha_{T_f,t(T_f)}$. However, the marginal cost of keeping the unit on link s is exactly $\alpha_{T_0,s}$, and hence coalition i will be content with leaving the unit at link s. $\square$

Combining the above lemmas yields the following result.

Theorem 3.9 *Load-balancing games with coalitions possess pure equilibria when latencies are convex.*

3.2 Weighted Congestion Games

In this section we analyze the price of collusion for general (not only load-balancing) weighted congestion games. While the price of collusion can be arbitrarily large with arbitrary convex latencies, we get improved bounds when link latencies are restricted to degree-d polynomials. In particular, we adapt of the techniques of Awerbuch, Azar and Epstein [4] to obtain the following upper bound on the price of anarchy for weighted congestion games with coalitions, which in turn upper bounds the price of collusion.

Theorem 3.10 *The price of anarchy of a weighted congestion game with coalitions is at most 2.618 when latencies are linear, and $O(2^d \cdot d^{d+1})$ when latencies are degree-d polynomials.*

PROOF. We only present the proof of the bound of 2.618 for linear latencies. The bound when latencies are degree-d polynomials is similar. Let $\ell_s(x) = a_s \cdot x + b_s$ denote the latency of link s. In the congestion game G without coalitions, player i's strategy involves choosing a subset of links A_i from a set of possible subsets Λ_i. A strategy $\mathcal{A} = (A_1, \ldots, A_n)$ denotes a strategy for each player. The cost for a player i under this strategy is $w_i \cdot \sum_{s \in A_i} \ell_s(w_s)$, where w_s is the total weight of players whose strategy includes the link s.

We now take any coalitions P, and consider the game G with coalitions, $G(P)$. We wish to argue that the price of anarchy for this new game is bounded by $\frac{3+\sqrt{5}}{2} \approx 2.618$.

Consider a coalition Nash equilibrium strategy CE and socially optimal allocation OPT. We will use I to denote a generic coalition and i to denote a generic member within this coalition. Also, we will use Q_I and Q_I^* (Q_i and Q_i^*) to denote the strategy for this coalition (member) under CE and OPT respectively. Since coalition I has no incentive to deviate from its strategy Q_I at equilibrium, it should in particular not gain from deviating to Q_I^*. Thus,

$$\sum_{i,s \mid i \in I, s \in Q_i} (a_s \cdot w_s + b_s) w_i \leq \sum_{i,s \mid i \in I, s \in Q_i^*} (a_s \cdot (w_s + w_{s,I}^*) + b_s) w_i,$$

where $w_{s,I}^*$ is the load that coalition I puts on link s in OPT. Summing this inequality over all coalitions yields the Nash equilibrium cost $c(NE) = \sum_s (a_s w_s + b_s) w_s$ on the left side, and on the right side we get

$$\begin{aligned}
&\sum_s \sum_I \sum_{i \in I \mid s \in Q_i^*} (a_s \cdot (w_s + w_{s,I}^*) + b_s) w_i \\
&= \sum_s \sum_I (a_s \cdot (w_s + w_{s,I}^*) + b_s) w_{s,I}^* \\
&= \sum_s \sum_I (a_s w_s w_{s,I}^* + a_s (w_{s,I}^*)^2 + b_s w_{s,I}^*) \\
&\leq \sum_s a_s w_s w_s^* + a_s {w_s^*}^2 + b_s w_s^* \\
&= \sum_s a_s w_s w_s^* + \sum_s (a_s w_s^* + b_s) w_s^*,
\end{aligned}$$

where $w_s^* = \sum_I w_{s,I}^*$ denotes the load of link s at OPT, and the third inequality holds, because $\sum_I (w_{s,I}^*)^2 \leq {w_s^*}^2$. Notice that the second part of the last formula is $c(OPT)$. Now, applying the Cauchy-Schwarz inequality to the term $\sum_s a_s w_s w_s^*$ we get

$$\sum_s (a_s w_s + b_s) w_s \leq \sqrt{\sum_s (a_s w_s + b_s) w_s \sum_s (a_s w_s^* + b_s) w_s^*} + \sum_s (a_s w_s^* + b_s) w_s^*.$$

Finally, if we use x to denote the square root of the price of anarchy, then the above inequality is equivalent to $x^2 \leq x + 1$, which yields $x \leq \frac{3+\sqrt{5}}{2}$, completing the proof. $\square$

3.3 Games with Concave Latencies

In this section we give a simple proof that the price of collusion in load-balancing games with concave latencies is at most 2. Unfortunately, in this setting the coalition game is not known to have a pure equilibrium. Hence we also consider mixed equilibria and show a weaker bound of 4 on the price of collusion. Lastly, we give an example of a game with concave latencies for which the price of collusion is non-trivial (strictly greater than 1), although unlike our convex example, this does not match our upper bound.

First, consider a collusion-free equilibrium NE. We start with a simple claim.

Claim 3.11 *At NE the largest latency L and the smallest latency ℓ differ by at most a factor of two.*

PROOF. Assume that $2\ell < L$ and consider the link with latency ℓ. Due to concavity, adding a unit to that link can not increase its latency beyond $2\ell < L$. Hence a unit on the link with latency L has incentive to switch, contradicting the assumption of NE. $\square$

This allows us to prove the following result.

Theorem 3.12 *For any pure coalition equilibrium CE, $c(CE) \leq 2c(NE)$.*

PROOF. Let ℓ and L denote respectively the smallest and the largest latencies in NE. Claim 3.11 implies that $L \leq 2 \cdot \ell$. Obviously, $c(NE) \geq n \cdot \ell$.

We call a link s *underloaded* if its load at CE is less than its load at NE, i.e. $n_s < n'_s$. Consider a particular coalition i with n^i units and imagine that coalition extracts its units from CE, possibly creating additional underloaded links. Now consider what happens when i places its units on underloaded links, such that the resulting load of those links does not exceed their loads at NE. Such a placement is clearly possible, since the total number of units of NE and CE are identical.

The cost of such a move is at most $n^i \cdot L$. Since CE is a coalition equilibrium, this is an upper bound on what coalition i pays in CE. Summing over all coalitions yields that the total cost $c(CE) \leq \sum_i n^i \cdot L = n \cdot L \leq n \cdot 2\ell \leq 2c(NE)$, completing the proof. □

We now consider mixed equilibria, which are known to always exist. In a mixed equilibrium, players select distributions over pure strategies so as to minimize the expected cost they incur given all other players' distributions. Note that for any pure strategy A_i, player i must evaluate the cost of A_i by measuring the expected cost given that other players act according to their distributions, and that i *deterministically* plays A_i. A randomized strategy for player i is a distribution over deterministic strategies, and hence if a player plays a randomized strategy, the expected load experienced by the player may exceed the expected load of the machines he is randomizing between. Because of this we lose a factor of 2 in our bound. The proof of the following theorem can be found in the full version of this paper.

Theorem 3.13 *For any mixed coalition equilibrium CE, $\mathbf{E}[c(CE)] \leq 4 \cdot c(NE)$.*

We conclude by giving an example where the price of collusion with concave latencies is 8/7.

Example : Consider an instance of the game with 2 links and 4 players. The latencies are defined as follows: The first link has latency $\ell_1(1) = 1/2$ and $\ell_1(x) = 1$ for all $x > 1$, while the second has latency $\ell_2(x) = 1 + (x-4)\epsilon$. NE will place a single unit on the first link and three units on the second, for a total cost of $\frac{1}{2} + 3(1-\epsilon) \sim 3.5$. If we now form two coalitions of size 2, we have a coalition equilibrium in which both coalitions put a single unit on each link. The resulting outcome has a cost of $2 + 2(1-2\epsilon) \sim 4$. Hence the price of collusion for this game is $8/7$. ■

4. REFERENCES

[1] E. Altman, T. Başar, T. Jiménez, and N. Shimkin. Competitive Routing in Networks with Polynomial Costs. *IEEE Transactions on Automatic Control*, 92–96, 2002.

[2] E. Anshelevich, A. Dasgupta, J. Kleinberg, É. Tardos, T. Wexler, and T. Roughgarden. The Price of Stability for Network Design with Fair Cost Allocation. *IEEE Symposium on Foundations of Computer Science*, 295–304, 2004.

[3] R. Aumann. Acceptable Points in General Cooperative n-Person Games. *Contributions to the Theory of Games IV* Princeton University Press, 1959.

[4] B. Awerbuch, Y. Azar, and A. Epstein. The Price of Routing Unsplittable Flow. *ACM Symposium on Theory of Computing*, 57–66, 2005.

[5] M. Beckmann, C. B. McGuire, and C. B. Winsten. Studies in the Economics of Transportation. *Yale University Press*, 1956.

[6] S. Catoni and S. Pallottino. Traffic Equilibrium Paradoxes. *Transportation Science*, 240–244, 1991.

[7] G. Christodoulou and E. Koutsoupias. The Price of Anarchy of Finite Congestion Games. *ACM Symposium on Theory of Computing*, 67–73, 2005.

[8] G. Christodoulou and E. Koutsoupias. On The Price of Anarchy and Stability of Correlated Equilibria of Linear Congestion Games. *European Symposium on Algorithms*, 59–70, 2005.

[9] R. Cominetti, J. R. Correa, and N. E. Stier-Moses. Network Games with Atomic Players. *Columbia Working Paper # DRO-2006-03*, 2006.

[10] J. R. Correa, A. S. Schulz, and N. E. Stier-Moses. On The Inefficiency of Equilibria in Congestion Games. *Conference on Integer Programming and Combinatorial Optimization*, 167–181, 2005.

[11] A. Fabrikant, C. H. Papadimitriou, and K. Talwar. The Complexity of Pure Nash Equilibria. *ACM Symposium on Theory of Computing*, 604–612, 2004.

[12] A. V. Goldberg and J. D. Hartline. Collusion-Resistant Mechanisms for Single-Parameter Agents. *ACM-SIAM Symposium on Discrete Algorithms*, 620–629, 2005.

[13] A. Kothari, S. Suri, C. D. Tóth, and Y. Zhou. Congestion Games, Load Balancing, and Price of Anarchy. *Workshop on Combinatorial and Algorithmic Aspects of Networking*, 2004.

[14] E. Koutsoupias and C. H. Papadimitriou. Worst-Case Equilibria. *International Symposium on Theoretical Aspects of Computer Science*, 404–413, 1999.

[15] I. Milchtaich. Congestion Games with Player-Specific Payoff Functions. *Games and Economic Behavior*, 111–124, 1996.

[16] D. Monderer and L. S. Shapley. Potential Games. *Games and Economic Behavior*, 124–143, 1996.

[17] D. Moreno and J. Wooders. Coalition-Proof Equilibrium. *Games and Economic Behavior*, 80–112, 1996.

[18] H. Moulin and S. Shenker. Strategyproof Sharing of Submodular Costs: Budget Balance Versus Efficiency. *Economic Theory*, 511–533, 2001.

[19] A. Orda, R. Rom, and N. Shimkin. Competitive Routing in Multiuser Communication Networks. *IEEE/ACM Transactions on Networking*, 510–521, 1993.

[20] C. H. Papadimitriou. Algorithms, Games and the Internet. *ACM Symposium on Theory of Computing*, 749–753, 2001.

[21] J. B. Rosen. Existence and Uniqueness of Equilibrium Points for Concave N-Person Games. *Econometrica*, 520–534, 1965.

[22] R. W. Rosenthal. A Class of Games Possessing Pure-Strategy Nash Equilibria. *International Journal of Game Theory*, 65–67, 1973.

[23] T. Roughgarden. The Price of Anarchy is Independent of the Network Topology. *ACM Symposium on Theory of Computing*, 428–437, 2002.

[24] T. Roughgarden. Selfish Routing with Atomic Players. *ACM-SIAM Symposium on Discrete Algorithms*, 1184–1185, 2005.

[25] T. Roughgarden and É. Tardos. How Bad Is Selfish Routing? *Journal of the ACM*, 236–259, 2002.

[26] S. Suri, C. D. Toth, and Y. Zhou. Selfish Load Balancing and Atomic Congestion Games. *ACM Symposium on Parallel Algorithms and Architectures*, 188–195, 2004.