Theory Comput Syst (2011) 48: 781–802
DOI 10.1007/s00224-010-9269-4

Stackelberg Strategies and Collusion in Network Games with Splittable Flow

Tobias Harks

Published online: 22 May 2010

Abstract We study the impact of collusion in network games with splittable flow and focus on the well established price of anarchy as a measure of this impact. We first investigate symmetric load balancing games and show that the price of anarchy is at most m, where m denotes the number of coalitions. For general networks, we present an instance showing that the price of anarchy is unbounded, even in the case of two coalitions. If latencies are restricted to polynomials with nonnegative coefficients and bounded degree, we prove upper bounds on the price of anarchy for general networks, which improve upon the current best ones except for affine latencies.

In light of the negative results even for two coalitions, we analyze the effectiveness of Stackelberg strategies as a means to improve the quality of Nash equilibria. In this setting, an α fraction of the entire demand is first routed centrally by a *Stackelberg leader* according to a predefined *Stackelberg strategy* and the remaining demand is then routed selfishly by the coalitions (*followers*).

For a single coalitional follower and parallel arcs, we develop an efficiently computable Stackelberg strategy that reduces the price of anarchy to one. For general networks and a single coalitional follower, we show that a simple strategy, called SCALE, reduces the price of anarchy to $1 + \alpha$. Finally, we investigate SCALE for multiple coalitional followers, general networks, and affine latencies. We present the first known upper bound on the price of anarchy in this case. Our bound smoothly varies between 1.5 for $\alpha = 0$ and full efficiency for $\alpha = 1$.

Keywords Atomic splittable flow games · Coalitions · Stackelberg routing · Price of anarchy

A preliminary version of this paper appeared in the *Proceedings of the 6th International Workshop on Approximation and Online Algorithms (WAOA)*, 2008, see [18].
Research supported by the Federal Ministry of Education and Research (BMBF grant 03MOPAI1).

T. Harks (✉)
Institute of Mathematics, Technical University Berlin, Berlin, Germany
e-mail: harks@math.tu-berlin.de

1 Introduction

Over the past years, the impact of the behavior of selfish, uncoordinated users in congested networks has been investigated intensively in the theoretical computer science and operations research literature. In this context, *network routing games* have proved to be a reasonable means of modeling selfish behavior in networks. The basic idea is to model the interaction of selfish network users as a *noncooperative game*. We are given a directed graph with latency functions on the arcs and a set of origin-destination pairs, called *commodities*. Every commodity is associated with a *demand*, which specifies the rate of flow that needs to be sent from the respective origin to the destination. In the nonatomic variant, every demand represents a continuum of agents, each controlling an infinitesimal amount of flow. The latency that an agent experiences to traverse an arc is given by a (non-decreasing) function of the total flow on that arc. Agents are assumed to act selfishly and route their flow along a minimum-latency path from their origin to the destination; a solution in which no agent can switch to a path with smaller travel time corresponds to a Wardrop equilibrium [15, 37].

Koutsoupias and Papadimitriou [24] initiated the investigation of the efficiency loss caused by selfish behavior. They introduced a measure to quantify the inefficiency of Nash equilibria which they termed the *price of anarchy*. The price of anarchy is defined as the worst-case ratio of the cost of a Nash equilibrium over the cost of a system optimum. In a seminal work, Roughgarden and Tardos [34] showed that the price of anarchy for network routing games with nonatomic players and linear latency functions is $4/3$; in particular, this bound holds independently of the underlying network topology. The case of more general families of latency functions has been studied by Roughgarden [29] and Correa et al. [13]. (For an overview of these results, we refer to the book by Roughgarden [31].) Despite these bounds for specific classes of latency functions, it is known that the price of anarchy for general latency functions is unbounded even on simple parallel-arc networks [34].

In this paper, we study nonatomic network games in which the agents are partitioned into a (in)finite number of sets, which we interprete (and term) as coalitions of agents. We allow that agents of different commodities may belong to the same coalition and further assume that every coalition aims at minimizing the *average delay* experienced by this coalition. In this setting, we study the worst case efficiency (price of anarchy) of Nash equilibria: stable points, where no coalition can unilaterally improve its cost by rerouting its flow. While the model under consideration (also known as *atomic splittable flow games*) has been studied by many researchers, see among others Cominetti et al. [11], Hayrapetyan et al. [20], Korilis et al. [23], and Roughgarden and Tardos [34], several intriguing open questions still persist.

Cominetti et al. [11] discovered that the price of anarchy in these games may exceed that of corresponding nonatomic games without coalitions. More precisely, Cominetti et al. [11] presented an instance showing that for polynomial latency functions of degree d, the price of anarchy grows as $\Omega(d)$. On the positive side, they presented upper bounds of 1.5, 2.56, and 7.83, for polynomial latency functions of degree $d = 1, 2, 3$, respectively. For polynomials of larger degree, the previously best known upper bound is $O(2^d d^{d+1})$, which is due to Hayrapetyan et al. [20].

1.1 Our Results

We investigate nonatomic network routing games with coalitions. Our contribution in this setting is the following:

1. First, we consider symmetric load balancing games, that is, we are given parallel arcs that connect a common source and a common sink. For this setting, we show that the price of anarchy is at most m, where m denotes the number of coalitions. This result holds for *arbitrary* convex latencies and is related to a previous result of Cominetti et al. [11], who showed that for single-commodity network games with m coalitions each of which controlling the same amount of flow is at most m. Our result is a generalization in the sense that we do not require that the flow is evenly distributed among coalitions. On the other hand, our result is more restrictive as it only holds for parallel arcs.
2. We then investigate the efficiency of Nash equilibria for general networks. We show that the price of anarchy in such games is unbounded, even for two coalitions. For semi-convex latency functions, we derive a generic upper bound on the price of anarchy using a variational inequality approach. We further show that if the class of allowable latencies is restricted to polynomials with nonnegative coefficients and maximum degree d, the price of anarchy is at most $d^{\sqrt{d}}$ for $d \geq 4$. Our bounds improve upon all previous known bounds, except for affine latencies, i.e., $d = 1$. For an overview of our bounds, we refer to Table 1.

Due to the large efficiency loss of Nash equilibria, researchers have proposed different approaches to reduce the price of anarchy in network routing games. One of the most promising approaches is the use of *Stackelberg routing*, see [23, 30]. In this setting, it is assumed that a fraction $\alpha \in [0, 1]$ of the entire demand is controlled by a central authority, termed *Stackelberg leader*, while the remaining demand is controlled by the selfish players, also called the *followers*. In a *Stackelberg game*, the Stackelberg leader first routes the centrally controlled flow according to a predetermined policy, called the *Stackelberg strategy*, and then the remaining demand is routed by the selfish followers. The aim is to devise Stackelberg strategies so as to minimize the price of anarchy of the resulting combined flow.

3. In light of the negative results that hold even for only two coalitions, we investigate Stackelberg strategies as a way to improve the quality of Nash equilibria. Recently, Bonifaci et al. [6] showed that for nonatomic followers and single commodity networks, no Stackelberg strategy can reduce the price of anarchy to a constant. This result, however, does not rule out the existence of a Stackelberg strategy inducing a constant price of anarchy, when the number of coalitional followers is small. For a single coalitional follower, parallel arcs and semi-convex latencies, we develop an efficiently computable Stackelberg strategy (called SFS) that reduces the price of anarchy to one. For general networks, semi-convex latencies and a single coalitional follower, we prove that the SCALE strategy (see Roughgarden [30]) reduces the price of anarchy to $1 + \alpha$. This result holds for convex latencies and general networks.
4. Finally, we consider general networks and multiple coalitional followers. For affine linear latencies, we prove that the SCALE strategy yields an upper bound

on the price of anarchy, which smoothly varies between the best known bound on the price of anarchy of 1.5 when $\alpha = 0$ and full efficiency when $\alpha = 1$.

1.2 Applications

There are numerous applications that can be interpreted as a network routing game with coalitions. Here, we focus on highlighting only a few (as we find) particularly interesting ones.

In recent years, the number of traffic participants that use a navigation device has increased significantly. Already nowadays, navigation systems feature bidirectional data communication which, among other services, opens the possibility to transmit the current location of a customer to a central server of the service provider (see, e.g., [26]). This way, the current traffic situation can be monitored accurately in real-time (given that a sufficient number of traffic participants are using this technology). Based on this data, the service provider can provide a better route guidance, e.g., in the case of traffic congestion, by centrally computing routes for their customers which are then communicated back to the respective navigation devices. A natural objective that the service provider might want to achieve in order to provide a good quality of service is to minimize the average travel time of their customers. This scenario can be modeled as nonatomic network game with coalitions, where the members of coalition are the customers of a specific service-provider.

One important application of Stackelberg routing is the routing of Internet traffic within the domain of an Internet service provider, see also Sharma and Williamson [35]. Here, the Internet service provider centrally controls a fraction of the overall traffic traversing its domain, while the remaining traffic is controlled by other service providers. In this setting, a natural goal for a service provider is to devise routes to the centrally controlled flow so as to minimize the overall delay in its domain. Our results for the Stackelberg strategies SCALE and SFS provides the Internet service provider with efficient algorithms to compute routes for the centrally controlled traffic.

1.3 Related Work

Awerbuch et al. [3], Christodoulou and Koutsoupias [10], and Aland et al. [1] derived tight bounds on the price of anarchy for weighted and unweighted congestion games with polynomial latency functions. These works, however, did not study the impact of coalitions on the price of anarchy.

Closer to our work are the papers by Hayrapetyan et al. [20] and Cominetti et al. [11]. The former presented a general framework for studying congestion games with colluding players. Their goal is to investigate the price of collusion: the factor by which the quality of Nash equilibria can deteriorate when coalitions form. Their results imply that for symmetric nonatomic load balancing games with coalitions the price of anarchy does not exceed that of the game without coalitions. For weighted congestion games with coalitions and polynomial latencies they proved upper bounds of $O(2^d d^{d+1})$, where d denotes the degree of the considered polynomials. They also presented examples showing that in discrete (atomic) network games, the price of collusion may be strictly larger than 1, i.e., coalitions may strictly increase the social cost.

Cominetti et al. [11] studied the atomic splittable selfish routing model in which the flow of every commodity forms a coalition (atomic player). Thus, this model can be incorporated as a special case of nonatomic network congestion games with arbitrary coalitions. They observed that the price of anarchy of this game may exceed that of the standard nonatomic selfish routing game without coalitions. Based on the work of Catoni and Pallotino [9], they presented an instance with affine latency functions in which the price of anarchy is 1.34. Using a variational inequality approach, they presented bounds on the price of anarchy for linear and polynomial latency functions of degree two and three of 1.5, 2.56, and 7.83, respectively. As noted by Cominetti et al., these positive bounds directly carry over to the case of nonatomic network congestion games with arbitrary coalitions (considered in this paper), since the variational inequalities are still valid in this more general model. For polynomials of larger degree, their approach does not yield bounds. For single commodity networks with symmetric demands (every coalition controls the same amount of flow), Cominetti et al. [11] proved an upper bound of m on the price of anarchy.

Altman et al. [2] proved for monomial latency functions and single commodity networks that there is a Nash flow, which is optimal. They also derived conditions under which Nash equilibria are unique. Uniqueness of Nash equilibria has been further studied by Fleischer et al. [4] and Orda et al. [27].

Haurie and Marcotte [19] presented a general framework for studying atomic splittable network games with *elastic* demands. They characterized the relationship between nonatomic and atomic splittable network games. Haurie and Marcotte, however, do not study the efficiency of Nash equilibria with respect to an optimal solution.

Fotakis et al. [17] studied algorithmic issues in the setting of atomic congestion games with coalitions and unsplittable flows. They proved upper bounds on the price of anarchy, where the cost of a coalition is defined as the maximum latency, see also the KP-model [24].

The idea of using Stackelberg strategies to improve the performance of a system was first proposed by Korilis et al. [23]. The authors identified necessary and sufficient conditions for the existence of Stackelberg strategies that induce a system optimum; their model also considers atomic splittable followers. In particular, they showed that for a single atomic splittable follower, parallel arcs, and $M/M/1$ latencies, there exists an optimal Stackelberg strategy that reduces the price of anarchy to one.

Roughgarden [30] proposed some natural Stackelberg strategies, e.g., SCALE and Largest-Latency-First (LLF). For parallel-arc networks he showed that the price of anarchy for LLF is bounded by $4/(3+\alpha)$ and $1/\alpha$ for linear and arbitrary latency functions, respectively. Both bounds are best possible. Moreover, he also proved that it is NP-hard to compute the best Stackelberg strategy. Kumar and Marathe [25] gave a PTAS to compute the best Stackelberg strategy for the case of parallel-arc networks. Karakostas and Kolliopoulos [22] proved upper bounds on the price of anarchy for SCALE and LLF. Their bounds hold for arbitrary multi-commodity networks and linear latency functions. Swamy [36] obtained upper bounds on the price of anarchy for SCALE and LLF for polynomial latency functions. He also proved a bound of $1 + 1/\alpha$ for single-commodity, series-parallel networks with arbitrary latency functions. Bonifaci et al. [6] proved that even for single-commodity networks no Stackelberg strategy can induce a bounded price of anarchy for any $\alpha \in (0, 1)$. On the positive side, they proved that LLF induces an upper bound on the price of anarchy, which only depends on the size of the network (number of vertices, arcs and commodities). They also derived almost tight bounds for SCALE and polynomial latencies. Correa and Stier-Moses [12] proved, besides some other results, that strategies in which the Stackelberg leader sends no more flow on every edge than the system optimum, does not increase the price of anarchy. Sharma and Williamson [35] considered the problem of determining the smallest value of α such that the price of anarchy can be improved. They obtained results for parallel-arc networks and linear latency functions. Kaporis and Spirakis [21] studied a related question of finding the minimum demand that the Stackelberg leader needs to control in order to enforce an optimal flow. Given that the Stackelberg leader controls a sufficiently large fraction of the overall demand, they also showed that one can efficiently compute the optimal Stackelberg strategy. Finally, Fotakis [16] studied Stackelberg routing with unsplittable flows and proved (among other results) that the $1/\alpha$ bound for parallel links still holds.

2 The Model

In a network routing game, we are given a directed network $G = (V, A)$ and k origin-destination pairs $(s_1, t_1), \dots, (s_k, t_k)$ called *commodities*. We will use the shorthand $[k] := \{1, 2, \dots, k\}$. For every commodity $i \in [k]$, a demand $r_i > 0$ is given that specifies the amount of flow with origin s_i and destination t_i. Let $\mathcal{P}_i$ be the set of all paths from s_i to t_i in G and let $\mathcal{P} = \bigcup_i \mathcal{P}_i$. A *flow* is a function $f : \mathcal{P} \to \mathbb{R}_+$, and we denote by $f_P = f(P)$ the amount of flow that is send along path P. The flow f is *feasible* (with respect to r) if for all i, $\sum_{P \in \mathcal{P}_i} f_P = r_i$.

For a given flow f, we define the flow on an arc $a \in A$ as $f_a = \sum_{P \ni a} f_P$. Moreover, each arc $a \in A$ has an associated load-dependent *latency* denoted by $\ell_a(\cdot)$. For each $a \in A$, the latency function ℓ_a is assumed to be nonnegative, nondecreasing and differentiable. We also assume that ℓ_a is defined on $[0, \infty)$ and that $x\ell_a(x)$ is a convex function of x. Such functions are called *semi-convex* or *standard* [29]. The latency of a path P with respect to a flow f is defined as the sum of the latencies of the arcs in the path, denoted by $\ell_P(f) = \sum_{a \in P} \ell_a(f_a)$. The *total cost* of a flow f is $C(f) = \sum_{a \in A} f_a \ell_a(f_a)$. The feasible flow of minimum total cost is called *optimal*. We will denote the optimal flow by o.

In a nonatomic network game, infinitely many agents are carrying the flow rate and each agent controls only an infinitesimal fraction of the demand. The continuum of agents of type j (traveling from s_j to t_j) is represented by the interval $[0, r_j]$. It is well known that for this setting Nash flows exist and their total cost is unique, see [31]. Furthermore, the price of anarchy, which measures the worst case ratio of the total cost of any Nash flow and that of an optimal flow is well understood, see Correa et al. [13, 14], Perakis [28], Roughgarden [31], and Roughgarden and Tardos [34].

In this paper, we study nonatomic network games in which the agents are partitioned into a (in)finite set of coalitions. In our model, we allow that agents of different commodities, i.e., agents traveling from different sources to different destinations, may belong to the same coalition. We assume that the partition of agents into coalitions is *fixed* and given a priori.

Let $[m]=\{1,\dots,m\}$ denote a set of coalitions. To this end, we represent every agent of commodity i as a real number in $[0,r_i]$. Then, the distribution of agents among the coalitions is modeled by a collection of Lebesgue-measurable functions $c^i:[0,r_i]\to[m]$, $i\in[k]$, which map an agent of type $i\in[k]$ to coalition $j\in[m]$. The continuum of agents of type i belonging to coalition j is defined as the Lebesgue-measure of $\{\xi\in[0,r_i]:c^i(\xi)=j\}$ and denoted by $c^{i,j}$. Using this notation, we define by f^j the flow for coalition j and say that f^j is *feasible for coalition* j if f^j satisfies the demands $c^{i,j}, i\in[k]$ in the usual sense. The amount of flow of coalition j on arc a is defined as $f_a^j=\sum_{P\in\mathcal{P}:P\ni a}f_P^j$, where f_P^j denotes the flow of coalition j along path P.

We assume that every coalition aims at minimizing the *average delay* or *total travel time* experienced by this coalition, see also [11]. Thus, the cost for coalition j is defined as $C^j(f^j;f^{-j}):=\sum_{a\in A}\ell_a(f_a)f_a^j$, where f^{-j} denotes the flow of all other coalitions.

The tuple $I=(G,r,\ell,c,m)$ is called an *instance* of the nonatomic network game with coalitions. Our model is similar to the one proposed by Hayrapetyan et al. [20] and it includes the special case, where we have exactly k coalitions each of which controlling the flow for commodity k.

Definition 1 A feasible flow f is a Nash equilibrium if and only if for all $j\in[m]$: $C^j(f^j;f^{-j})\le C^j(x^j;f^{-j})$ for all feasible flows x^j for coalition $j\in[m]$.

In a Nash equilibrium, every coalition routes its flow so as to minimize $C^j(f^j;f^{-j})$ with the understanding that coalition j optimizes over f^j while the flow f^{-j} of all other coalitions is fixed.

Definition 2 Let $\mathcal{L}$ be a class of latency functions. Let $\mathcal{I}_m(\mathcal{L})$ be the set of all instances with at most m coalitions and latency functions in $\mathcal{L}$. For $I\in\mathcal{I}_m(\mathcal{L})$, let o_I be an optimal profile and let Θ_I be the set of Nash equilibria, respectively. Then, the *price of anarchy* is defined by

$$\sup_{I\in\mathcal{I}_m(\mathcal{L})}\sup_{f\in\Theta_I}\frac{C(f)}{C(o_I)}.$$

Note that this definition of the price of anarchy is slightly different from the standard nonatomic selfish routing model [31], since there may be qualitatively different equilibria, see [4].

If latencies are restricted to be standard, minimizing $C^j(f^j;f^{-j})$ is a convex optimization problem. The following necessary and sufficient optimality conditions characterize Nash flows for a nonatomic network game with coalitions. This characterization can also be found in Haurie and Marcotte [19] (Theorem 2.3) and in Cominetti et al. [11].

Lemma 1 *A feasible flow $(f^1, \dots, f^m)$ is a Nash equilibrium for a nonatomic network game with m coalitions if and only if for every $j \in [m]$ the following inequality is satisfied*:

$$\sum_{a \in A} (\ell_a(f_a) + \ell'_a(f_a) f_a^j)(f_a^j - x_a^j) \le 0 \quad \text{for all feasible flows } x^j. \tag{1}$$

Proof A flow f is a Nash equilibrium if and only if every f^j, $j \in [m]$, is a global minimizer of $C^j(f^j; f^{-j})$. Since the feasible region of all feasible flows for coalition j forms a convex and compact set, and the objective $C^j(f^j; f^{-j})$ is nondecreasing, differentiable and convex, the variational inequality (1) constitutes a first order necessary and sufficient optimality condition for the global minimum of $C^j(\cdot; f^{-j})$ at f^j, see the book by Boyd and Vandenberghe [7]. This condition expresses that at the optimum f^j, there is no feasible gradient descent direction. □

3 Nonatomic Network Games with Coalitions

In the subsequent sections, we will investigate the price of anarchy for specific network topologies and classes of latency functions.

3.1 Symmetric Load Balancing Games

A symmetric load balancing game is a network game, where the underlying digraph simply connects two distinguished nodes with parallel links.

Theorem 1 *For symmetric load balancing games with m coalitions and nondecreasing, differentiable, and standard latency functions, the price of anarchy is at most m.*

Proof As usual, let f denote a Nash flow and o an optimal flow. We bound the cost of each coalition individually. Assume the flow for coalition j carries α_j units of flow. We claim that there exists a feasible flow g^j such that $g_a^j + f_a^{-j} \le o_a$ for all $a \in A$ with $g_a^j > 0$. To see this, we define the flow $\bar{g} = [o - f^{-j}]^+$, where the positive projection is applied component wise, that is, for arc a we have $[\bar{g}_a]^+ = \bar{g}_a$, if $\bar{g}_a \ge 0$, and 0 otherwise. It is straight-forward to verify that $\bar{g}$ is a feasible flow for $\beta \ge \alpha_j$ units of flow. Hence, the flow $g = \frac{\alpha_j}{\beta} \bar{g}$ is feasible for coalition j. The cost of coalition j when applying strategy g can be bounded by $C^j(g; f^{-j}) = \sum_{a \in A} \ell_a(g_a + f_a^{-j}) g_a \le \sum_{a \in A} \ell_a(o_a) g_a \le \sum_{a \in A} \ell_a(o_a) o_a$. The first inequality is valid since for arcs a with $g_a > 0$, we have $\frac{\alpha_j}{\beta}[o_a - f_a^{-j}]^+ + f_a^{-j} \le o_a$, because $o_a \ge f_a^{-j}$ and $\frac{\alpha_j}{\beta} \le 1$. The second inequality follows since g is by definition opt-restricted, that is, $g_a \le o_a$ for all $a \in A$. Using that coalition j plays a best response in equilibrium, we have $C^j(f^j; f^{-j}) \le C^j(g; f^{-j}) \le C(o)$. We apply the same argument for every coalition, thus, $C(f) = \sum_{j \in [m]} C^j(f^j; f^{-j}) \le mC(o)$. □

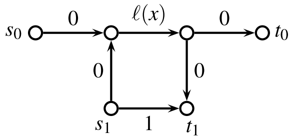

Fig. 1 The graph G, used in the proof of Proposition 1

3.2 Multi-commodity Networks

We present the following negative result.

Proposition 1 *Let $M > 0$. There is a multi-commodity instance $I = (G, r, \ell, c, m)$ with $m = 2$ such that for a Nash flow f, and an optimal flow o, $C(f) \geq \Omega(M) \cdot C(o)$.*

Proof Consider the construction in Fig. 1. We have two players, where one player has a demand of size M from s_0 to t_0. The second player has a demand of size 1 from s_1 from t_1. All latencies are constant (1 or 0 as indicated in Fig. 1) except for the latency function $\ell(x)$, which is defined as $\ell(x) = \max\{0, x - M\}$. In a Nash equilibrium, the second player will route 1/2 of its flow along the upper path. Indeed, in this case the marginal latency evaluates to $\ell(1/2 + M) + \ell'(1/2 + M)1/2 = 1$. The total cost of the combined flow f evaluates to $C(f) = 1/2(M + 1/2) + 1/2 = \Omega(M)$. A feasible flow can always be constructed by routing the flow of the two commodities along the direct path. Thus, we obtain $C(o) \leq 1$, proving the proposition. □

Note that the function $\ell(x)$ used in the above proposition is not differentiable in $x = M$. But this can be removed by defining a different function $\bar{\ell}(x)$, which smoothly interpolates between $\ell(M) = 0$ and $\ell(1/2 + M)$ and satisfies $\ell(1/2 + M) + \ell'(1/2 + M)1/2 = 1$.

3.3 Bounding the Price of Anarchy via the λ-Approach

The previous example showed that for multi-commodity networks, the price of anarchy is unbounded even for two coalitions. In the following, we will therefore restrict the class of allowable latency functions in order to obtain upper bounds on the price of anarchy.

For a latency function ℓ and nonnegative parameter λ we define the following nonnegative value:

$$\omega(\ell; m, \lambda) := \sup_{f,x \geq 0} \frac{(\ell(f) - \lambda\ell(x))x + \ell'(f)(\sum_{j\in[m]}(f^j x^j - (f^j)^2))}{\ell(f)f}. \quad (2)$$

Here, we slightly abuse notation and denote by f (under the supremum) the vector $f = (f^1, \ldots, f^m)$ and also the sum $f = \sum_{j=1}^m f^j$.

We assume $0/0 = 0$ by convention. For a given class of latency functions $\mathcal{L}$, we define $\omega_m(\mathcal{L}; \lambda) := \sup_{\ell\in\mathcal{L}} \omega(\ell; m, \lambda)$ and $\Lambda_m(\mathcal{L}) := \{\lambda \in \mathbb{R}^+ | (1 - \omega_m(\mathcal{L}; \lambda)) > 0\}$.

Theorem 2 *Consider a family of instances $\mathcal{I}_m(\mathcal{L})$, where $\mathcal{L}$ is a class of nondecreasing, differentiable, and standard latency functions. Then, the price of anarchy is at most* $\inf_{\lambda\in\Lambda_m(\mathcal{L})}[\lambda(1 - \omega_m(\mathcal{L}; \lambda))^{-1}]$.

Proof Let f be a Nash flow, and x be any feasible flow. Then,

$$C(f) \le \sum_{a\in A}\Big(\ell_a(f_a)f_a + \sum_{j\in[m]}(\ell_a(f_a)+\ell'_a(f_a)f_a^j)(x_a^j - f_a^j)\Big) \tag{3}$$

$$= \sum_{a\in A}\Big(\lambda\ell_a(x_a)x_a + (\ell_a(f_a)-\lambda\ell_a(x_a))x_a + \sum_{j\in[m]}\ell'_a(f_a)f_a^j(x_a^j - f_a^j)\Big)$$

$$\le \lambda C(x) + \omega_m(\mathcal{L};\lambda)C(f). \tag{4}$$

Here, (3) follows from the variational inequality stated in Lemma 1. The last inequality (4) follows from the definition of $\omega_m(\mathcal{L};\lambda)$. Taking x as the optimal flow the claim is proven. □

Note that whenever $\Lambda_m(\mathcal{L}) = \emptyset$ or $\Lambda_m(\mathcal{L}) = \{\infty\}$, the approach does not yield a finite price of anarchy. Our definition of $\omega_m(\mathcal{L};\lambda)$ is closely related to the parameter $\beta^m(\mathcal{L})$ in Cominetti et al. [11] and $\alpha^m(\mathcal{L})$ in Roughgarden [32] for the atomic splittable selfish routing model. For $\lambda = 1$, we have $\beta^m(\mathcal{L}) = \omega_m(\mathcal{L};1)$ and $\alpha^m(\mathcal{L}) = (1-\omega_m(\mathcal{L};1))^{-1}$. As we show in the next section, the generalized value $\omega_m(\mathcal{L};\lambda)$ implies improved bounds for a large class of latency functions, e.g., polynomial latency functions. The previous approaches with $\beta^m(\mathcal{L})$ (or $\alpha^m(\mathcal{L})$) failed for instance to generate upper bounds for polynomials of degree $d \ge 4$ because this value exceeds 1 (or is infinite). The advantage of Theorem 2 is that we can tune the parameter λ and, hence, $\omega_m(\mathcal{L};\lambda)$ so as to minimize the price of anarchy given by $\lambda/(1-\omega_m(\mathcal{L};\lambda))$.

We make use of a result of Cominetti et al. [11].

Theorem 3 (Cominetti et al. [11]) *The value* $\beta^m(\ell) = \omega(\ell;m,1)$ *is at most*

$$\sup_{x,f\ge 0}\frac{(\ell(f)-\ell(x))x + \ell'(f)[(x)^2/4 - (f-x/2)^2/m]}{\ell(f)f}.$$

Since the necessary calculations to prove the above claim only affect the last term in (2), which is the same for $\omega(\ell;m,\lambda)$ and $\beta^m(\ell)$, this bound carries over for arbitrary nonnegative values of λ.

Corollary 1 *If* $\lambda \ge 0$, *the value* $\omega(\ell;m,\lambda)$ *is at most*

$$\sup_{x,f\ge 0}\frac{(\ell(f)-\lambda\ell(x))x + \ell'(f)[x^2/4 - (f-x/2)^2/m]}{\ell(f)f}.$$

3.4 Linear and Affine Linear Latency Functions

Cominetti et al. [11] proved an upper bound of 1.5 for affine latencies. In the following, we present a stronger result for linear latencies. We also show that for affine latencies the best bound can be achieved by setting $\lambda = 1$. In this case, we have $\beta^m(\mathcal{L}) = \omega_m(\mathcal{L};1)$.

Theorem 4 *Consider linear latency functions in* $\mathcal{L}_1^* = \{a_1 z : a_1 \geq 0\}$ *and* $m \geq 2$ *coalitions. Then, the price of anarchy is at most*

$$P(m) = \frac{(2m + \sqrt{2}\sqrt{m(m+1)})(m + 1 + \sqrt{2}\sqrt{m(m+1)})\sqrt{2}}{8\sqrt{m(m+1)}(m+1)}.$$

Furthermore, $\lim_{m\to\infty} P(m) = \frac{3}{4} + \frac{1}{2}\sqrt{2} \approx 1.46$.

Proof For proving the first claim, we start with the bound on $\omega(\ell; m, \lambda)$ given in Corollary 1. We define $\mu := \frac{x}{f}$ for $f > 0$ and 0, otherwise, and replace $x = \mu f$. This yields $\omega(\ell; m, \lambda) \leq \max_{\mu \geq 0}(\mu^2(\frac{m-1-\lambda 4m}{4m}) + \mu(\frac{m+1}{m}) - \frac{1}{m})$. For $\lambda > \frac{m-1}{4m}$ this is a strictly convex program with a unique solution given by $\mu^* = \frac{-2(m+1)}{m-1-\lambda 4m}$. Inserting the solution, yields $\omega(\ell; m, \lambda) \leq \frac{m+3-4\lambda}{4\lambda m+1-m}$. The condition $\lambda \in \Lambda_m(\mathcal{L}_1^*)$ is equivalent to $\lambda > \max\{\frac{m-1}{4m}, \frac{m-2}{2m-2}\}$. We define the value $\lambda = \frac{1}{2} + \frac{1}{4}\sqrt{2(m+1)/m}$, which is contained in $\Lambda_m(\mathcal{L}_1^*)$. Applying Theorem 2 with this value proves the claim. □

The proof for affine latencies is similar and leads to $C(f) \leq \min_{\lambda \geq 1} \frac{4\lambda^2 - \lambda}{4\lambda - 2} C(x)$ showing that the best bound can be achieved by setting $\lambda = 1$.

3.5 Polynomial Latency Functions

To facilitate the result of Theorem 2 for polynomial latency functions, one needs to bound $\omega_m(\mathcal{L}_d; \lambda)$ for the class $\mathcal{L}_d$ of polynomials with nonnegative coefficients and degree at most $d \in \mathbb{N}$:

$$\mathcal{L}_d = \{c_d x^d + \cdots + c_1 x + c_0 : c_s \geq 0,\ s = 0, \ldots, d\}.$$

Note that polynomials in $\mathcal{L}_d$ are nonnegative for nonnegative arguments, continuous, nondecreasing, and convex.

We focus in the following on the general case $m \in \mathbb{N} \cup \{\infty\}$. Therefore, we define

$$\omega(\ell; \infty, \lambda) := \sup_{x, f \geq 0} \frac{(\ell(f) - \lambda \ell(x))x + \ell'(f)(x)^2/4}{\ell(f) f}. \tag{5}$$

Then, it follows from Theorem 3 that $\omega(\ell; m, \lambda) \leq \omega(\ell; \infty, \lambda)$, since the square is nonnegative and $\lim_{m\to\infty}(f - x/2)^2/m = 0$.

We now observe that the total cost function $C(f)$ is linear in each of the latency functions $\ell(\cdot)$. We can therefore reduce the analysis to monomial latency functions. For this we subdivide each arc a into d arcs $a_1, \ldots, a_d$ with monomial latency functions $\ell_{a_s}(x) = c_s x^s$ for $s = 1, \ldots, d$.

Lemma 2 *Consider the class* $\mathcal{M}_s := \{c_s x^s : c_s \geq 0\}$ *for* $s \in \mathbb{N}$. *Then,* $\omega_\infty(\mathcal{M}_s; \lambda) \leq \max_{0 \leq \mu} \mu(1 - \lambda \mu^s + s\mu/4)$.

Proof Let $\ell \in \mathcal{M}_s$. Then, by (5) we get

$$\omega(\ell; \infty, \lambda) \le \sup_{x,f \ge 0} \frac{(f^s - \lambda x^s)x + sf^{s-1}x^2/4}{f^{s+1}}.$$

Substituting $x = \mu f$, $\mu \ge 0$, we obtain

$$\omega(\ell; \infty, \lambda) \le \max_{0\le\mu} \mu(1 - \lambda\mu^s + s\mu/4).$$ □

The next lemma states that $\omega_\infty(\mathcal{M}_s; \lambda)$ is monotonically increasing in s for $\lambda \ge 1$.

Lemma 3 *For* $\lambda \ge 1$, $\omega_\infty(\mathcal{M}_s; \lambda) \le \omega_\infty(\mathcal{M}_d; \lambda)$ *for all* $s \le d$, $s \in \mathbb{N}$, $d \in \mathbb{N}$.

Proof Let $\lambda \ge 1$. By Lemma 2, we have for $\ell \in \mathcal{M}_s$

$$\omega(\ell; \infty, \lambda) \le \max_{0\le\mu} \mu(1 - \lambda\mu^s + s\mu/4).$$

It is enough to prove that the argument maximum satisfies $\mu^* \le 1$. We define $T_s(\mu) := \mu(1 - \lambda\mu^s + s\mu/4)$ and show that $T_s'(\mu) \le 0$ for all $\mu \ge 1$. To this end, we obtain

$$\begin{aligned} T_s'(\mu) &= 1 - (s+1)\lambda\mu^s + (s\mu)/2 = 1 - \mu((s+1)\lambda\mu^{s-1} - s/2) \\ &\le 1 - \mu((s+1)\lambda - s/2) \le 1 - \mu(s/2+1) \le 0, \end{aligned}$$

where the first inequality follows from $\mu \ge 1$, while the second inequality follows from $\lambda \ge 1$. □

The next theorem presents an upper bound on the price of anarchy for latencies in $\mathcal{L}_d$.

Theorem 5 *Consider latency functions in* $\mathcal{L}_d$, $d \ge 2$. *Then, the price of anarchy is at most* $(\frac{1}{2}\sqrt{d} + \frac{1}{2})^d \frac{(d^2+1-\sqrt{d}-d^{\frac{3}{2}})}{(\sqrt{d}-1)(d-1)}$.

Proof We define $\lambda(d) := (\frac{1}{2}\sqrt{d} + \frac{1}{2})^d \frac{(d^2+1-\sqrt{d}-d^{\frac{3}{2}})}{(\sqrt{d}-1)(d^2-1)}$.

The proof proceeds by proving a claim, which yields a bound on $\omega(\mathcal{L}_d; \infty, \lambda(d))$.

Claim $\max_{0\le\mu\le1}[T(\mu) := \mu(1 - \lambda(d)\mu^d + d\frac{\mu}{4})] = d/(d+1)$, for all $d \ge 2$.

Proof To prove the claim it is convenient to write $\lambda(d)$ as

$$\lambda(d) = \frac{d^2 + 1 - \sqrt{d} - d^{\frac{3}{2}}}{\mu_1(d)^d(\sqrt{d} - 1)(d^2 - 1)},$$

where $\mu_1(d) := 2/(\sqrt{d} + 1)$.

Then, the claim is proven by verifying the following facts:

1. $T'(\mu_1(d)) = 0$, $T''(\mu_1(d)) < 0$ and $T''(\mu)$ has at most one zero in $(0, 1)$
2. $T(0) = 0$, $T(1) \leq d/(d+1)$ and $T(\mu_1(d)) = d/(d+1)$.

Before we prove these facts, we show how they imply the claim. The first fact implies that $\mu_1(d)$ is the *only* local maximum of $T(\mu)$ in the open interval $(0, 1)$. Then, by comparing $T(\mu_1(d))$ to the boundary values $T(0)$ and $T(1)$ it follows that $T(\mu_1(d)) = d/(d+1)$ is the global maximum.

We start by proving the first fact. The expression $T(\mu_1(d))$ evaluates to:

$$\begin{aligned}
T'(\mu_1(d)) &= 1 - (d+1)\lambda(d)\mu_1(d)^d + d\mu_1(d)/2 \\
&= 1 - (d+1)\frac{d^2 + 1 - \sqrt{d} - d^{\frac{3}{2}}}{(\sqrt{d}-1)(d^2-1)} + d\mu_1(d)/2 \\
&= 1 - \frac{d^2 + 1 - \sqrt{d} - d^{\frac{3}{2}}}{(\sqrt{d}-1)(d-1)} + \frac{d}{\sqrt{d}+1} \\
&= 0.
\end{aligned}$$

We now prove $T''(\mu_1(d)) < 0$. First, we simplify as follows

$$\begin{aligned}
T''(\mu_1(d)) &= -d(d+1)\lambda(d)\mu_1(d)^{d-1} + d/2 \\
&= -\frac{d(d^2 + 1 - \sqrt{d} - d^{\frac{3}{2}})}{2(\sqrt{d}-1)^2} + d/2.
\end{aligned}$$

Then, $T''(\mu_1(d)) < 0$ if and only if

$$\frac{d(d^2 + 1 - \sqrt{d} - d^{\frac{3}{2}})}{2(\sqrt{d}-1)^2} > 1/2 \quad \Leftrightarrow \quad d^2 + \sqrt{d} - d^{\frac{3}{2}} - d > 0.$$

The last inequality is fulfilled for all $d \geq 1$.

To verify that $T''(\mu)$ has at most one zero in $(0, 1)$, use for example Descartes' rule of signs. The second fact follows by simple calculations. □

The claim implies $\omega_\infty(\mathcal{L}_d; \lambda(d)) \leq d/(d+1)$, hence, $\lambda(d) \in \Lambda_\infty(\mathcal{L}_d)$ so we can use Theorem 2 to obtain the claimed bound of $(d+1)\lambda(d)$. □

In the following we analyze the growth of the derived upper bound for large d, $(d \geq 4)$. The proof consists of standard calculus and is omitted.

Corollary 2 $(\frac{1}{2}\sqrt{d} + \frac{1}{2})^d \frac{(d^2+1-\sqrt{d}-d^{\frac{3}{2}})}{(\sqrt{d}-1)(d-1)} \leq \sqrt{d}^d$ *for* $d \geq 4$.

In Table 1, we present an overview about achievable upper bounds on the price of anarchy when numerically optimizing over $\lambda \in \Lambda_m(\mathcal{L}_d)$ so as to calculate the minimum in Theorem 2.

Table 1 Overview of upper bounds on the price of anarchy for polynomials with nonnegative coefficients and maximum degree d. The result in the first column marked with (*) is with respect to linear latencies $\{a_1 x : a_1 \geq 0\}$. The result of the second column (affine latencies) is due to [11]

$d=1^*$	$d=1$	$d=2$	$d=3$	$d=4$	$d=5$	$d=6$	$d=7$	$d=8$	$d=9$	$d=10$
1.46	**1.5**	**2.55**	**5.06**	**11.09**	**26.32**	**66.89**	**180.27**	**512**	**1,524**	**4,734**

4 Stackelberg Strategies with Coalitional Followers

Since the price of anarchy in network games with only two coalitions is already unbounded (Proposition 1), we investigate coordination mechanisms as a means to improve the quality of Nash equilibria. One of the most prominent coordination mechanisms in the context of network routing games is the use of *Stackelberg routing*, see Korilis et al. [23] and Roughgarden [30].

In this setting, it is assumed that a fraction $\alpha \in [0,1]$ of the entire demand is controlled by a central authority, termed *Stackelberg leader*, while the remaining demand is controlled by selfish *followers* which in our case are the selfish coalitions. In a *Stackelberg game*, the Stackelberg leader first routes the centrally controlled flow according to a predetermined policy, called the *Stackelberg strategy*, and then the remaining demand is routed by the selfish followers. The aim is to devise Stackelberg strategies so as to minimize the price of anarchy of the resulting combined flow with respect to an optimal solution for the entire demand.

An instance of a Stackelberg routing game with coalitional followers is characterized by a tuple $I(\alpha) = (G, r, \ell, c, m, \alpha)$, where in addition to G, r, ℓ, c and m, a parameter $\alpha \in (0,1)$ is given that specifies the fraction of the demand controlled by the Stackelberg leader.

A (*strong*) *Stackelberg strategy* is a flow g feasible with respect to the demand vector $r' = (\alpha_1 r_1, \ldots, \alpha_k r_k)$, for some $\alpha_1, \ldots, \alpha_k \in [0,1]$ such that $\sum_{i=1}^k \alpha_i r_i = \alpha \sum_{i=1}^k r_i$. If $\alpha_i = \alpha$ for all i, g is called a *weak Stackelberg strategy*. Thus, both strong and weak strategies route a fraction α of the overall traffic, but a strong strategy can choose how much flow of each commodity is centrally controlled. For single-commodity networks the two definitions coincide. A Stackelberg strategy g is called *opt-restricted* if $g_a \leq o_a$ for all $a \in A$.

Given a Stackelberg strategy g, let $\tilde{\ell}_a(x) = \ell_a(g_a + x)$ for all $a \in A$ and let $\tilde{r} = r - r'$. We assume that the Stackelberg leader may choose arbitrarily which amount of flow (up to αr) of a commodity *and* coalition it controls. Thus, the remaining set and demands of the coalitional followers denoted by $\tilde{m}$ and $\tilde{c}$, respectively, is obtained by reducing every $c^{i,j}$ by the amount of demand that the Stackelberg leader wishes to control from coalition j and commodity i.

We say that a flow h is *induced by* g if it is a Nash flow for the instance $(G, \tilde{r}, \tilde{\ell}, \tilde{c}, \tilde{m})$.

A Nash flow h can be characterized by the following *variational inequality* (see Lemma 1): h is a Nash flow induced by g if and only if for all flows x feasible with respect to $\tilde{r}$,

$$\sum_{j \in [m]} \sum_{a \in A} (\ell_a(g_a + h_a) + \ell'_a(g_a + h_a) h_a^j)(x_a^j - h_a^j) \geq 0. \tag{6}$$

We will mainly be concerned with the total cost of the combined induced flow $g+h$, given by $C(g+h)=\sum_{a\in A}(g_a+h_a)\ell_a(g_a+h_a)$. In particular, we are interested in bounding the *price of anarchy*, that is, the worst case ratio of $C(g+h)/C(o)$. It will be convenient to separate the total cost $C(g+h)$ in $C_1(g;h):=\sum_{a\in A}\ell_a(g_a+h_a)g_a$ and $C_2(h;g):=\sum_{a\in A}\ell_a(g_a+h_a)h_a$.

4.1 Symmetric Load Balancing Games

We consider symmetric load balancing games in which the underlying digraph simply connects two distinguished nodes with parallel links. Let g be a flow according to the Largest-Latency-First (LLF) strategy introduced by Roughgarden [30]. LLF simply calculates an optimal flow o and saturates the arcs with largest latencies first. On the one hand, Roughgarden showed that for Stackelberg routing games with nonatomic followers (without coalitions), LLF reduces the price of anarchy to $1/\alpha$. On the other hand, Hayrapetyan et al. [20] showed that for symmetric load balancing games colluding nonatomic players only decrease the total cost. Combining these two results (Hayrapetyan et al. [20] (Theorem 2.3) and Roughgarden [30] (Theorem 4.2)), it follows that the LLF strategy induces a flow of total cost of at most $1/\alpha C(o)$. Thus, the LLF strategy reduces the price of anarchy to $1/\alpha$ even in Stackelberg routing games with coalitional followers.

4.2 Symmetric Load Balancing Games with a Single Follower

We now consider the case of a single follower. This setting has been previously studied by Korilis et al. [23]. The authors showed that for a single coalitional follower, parallel arcs, and $M/M/1$ latencies, there exists an efficiently computable Stackelberg strategy that reduces the price of anarchy to one. Our main result in this section is a generalization of their result to arbitrary semi-convex latencies. We are given an instance $I(\alpha)$ of a Stackelberg game on parallel arcs and a single coalitional follower.

We define a Stackelberg strategy g that we call *single-follower-support* (SFS) strategy as in Algorithm 1.

We prove that this algorithm computes an optimal Stackelberg strategy.

Theorem 6 *Consider an instance $I(\alpha)$ of a Stackelberg game on parallel arcs with a single coalitional follower. Let g be according to the SFS strategy and let h be an induced Nash flow. Then, the combined flow $g+h$ is optimal.*

Proof First, we consider the case $\sum_{a\in A_1} g_a^1=\alpha$, which implies $g=g^1$.

Since g is opt-restricted, it suffices to prove that the flow $h_a=o_a-g_a$ is a feasible Nash flow for $1-\alpha$. More precisely, we have to verify that

$$\ell_a(o_a)+\ell_a'(o_a)(o_a-g_a)\le \ell_{\hat a}(o_{\hat a})+\ell_{\hat a}'(o_{\hat a})(o_{\hat a}-g_{\hat a}),$$

for all $a,\hat a\in A$ with $o_a-g_a>0$. These inequalities are satisfied since $\ell_a'(o_a)=0$ for all $a\in A_1$.

Algorithm 1 Single-follower-support

Input: $I(\alpha)$

Output: Stackelberg strategy g

1: compute a system optimal flow o

2: $A_1 := \{a \in A : o_a > 0 \text{ and } \ell'_a(o_a) = 0\}$, $A_2 := \{a \in A : o_a > 0 \text{ and } \ell'_a(o_a) > 0\}$

3:

$$x^* := \arg \max_{0 \le x_a \le o_a, a \in A_1} \sum_{a \in A_1} x_a, \quad \text{s.t. : } \sum_{a \in A_1} x_a \le \alpha$$

4: $g^1 := x^*$, $g^2 = (g^2_a)_{a \in A_2} := 0$

5: **if** $\sum_{a \in A_1} g^1_a < \alpha$ **then**

6:

$$g^2_a := \frac{\alpha - \sum_{a \in A_1} g^1_a}{\ell'_a(o_a)(\sum_{\bar{a} \in A_2} \frac{1}{\ell'_{\bar{a}}(o_{\bar{a}})})} \quad \text{for all } a \in A_2.$$

7: **end if**

8: **Return** $g = g^1 + g^2$

Now we consider the case $\sum_{a \in A_1} g^1_a < \alpha$. Notice that in this case $o_a - g_a = 0$ for all $a \in A_1$. Thus, we have to show that

$$\ell_a(o_a) + \ell'_a(o_a)(o_a - g_a) = C \quad \text{for some } C \ge 0 \text{ and all } a \in A_2, \tag{7}$$

$$C \le \ell_{\hat{a}}(o_{\hat{a}}) + \ell'_{\hat{a}}(o_{\hat{a}})(o_{\hat{a}} - g_{\hat{a}}) \quad \text{for all } \hat{a} \in A. \tag{8}$$

We now use that the system optimal flow o satisfies

$$\ell_a(o_a) + \ell'_a(o_a)(o_a) = \bar{C} \quad \text{for some } \bar{C} \ge 0 \text{ and all } a \in A_2,$$

$$\bar{C} \le \ell_{\hat{a}}(o_{\hat{a}}) + \ell'_{\hat{a}}(o_{\hat{a}})(o_{\hat{a}}) \quad \text{for all } \hat{a} \in A.$$

Hence, the conditions (7) and (8) are equivalent to

$$\ell'_a(o_a) g_a = D \quad \text{for some } D \ge 0 \text{ and all } a \in A_2.$$

Defining $D = (\alpha - \sum_{a \in A_1} g^1_a)/(\sum_{\bar{a} \in A_2} \frac{1}{\ell'_{\bar{a}}(o_{\bar{a}})})$ together with $g_a = D/\ell'_a(o_a)$ proves the result. □

4.3 General Networks with a Single Follower

In the following section, we will analyze a simple and easy-to-implement Stackelberg strategy termed *SCALE*. According to the SCALE strategy, a flow g is obtained by computing an optimal flow o and scaling this flow by α, i.e., $g = \alpha o$.

We show that SCALE achieves a bound of $(1 + \alpha)$ on the price of anarchy that even holds for general networks and latency functions.

Theorem 7 *Consider a family of instances $\mathcal{I}(\alpha)$ of Stackelberg games with a single coalitional follower and let g be according to the SCALE strategy. Then, the price of anarchy of the equilibrium flow $g+h$ is at most $1+\alpha$.*

Proof We bound the cost $C_1(g;h)$ and $C_2(h;g)$ separately. For the follower, we know that $\bar{h}=(1-\alpha)o_a$ is a feasible flow. Since the follower plays a best response in equilibrium, we have $C_2(h;\alpha o)\le C_2((1-\alpha)o;\alpha o)=\sum_{a\in A}\ell_a(o_a)(1-\alpha)o_a\le (1-\alpha)C(o)$. Now we bound the cost of the leader. Let h denote the best response of the follower. We consider the following cases. (i) $0\le h_a\le(1-\alpha)o_a$. In this case it follows that $\ell_a(\alpha o_a+h_a)\alpha o_a\le\alpha\ell_a(o_a)o_a$. (ii) $h_a>(1-\alpha)o_a$. This case implies $o_a<\frac{1}{1-\alpha}h_a$ and we get $\ell_a(\alpha o_a+h_a)\alpha o_a\le\frac{\alpha}{1-\alpha}\ell_a(\alpha o_a+h_a)h_a$. Using both cases, we have $C_1(\alpha o;h)\le\alpha C(o)+\frac{\alpha}{1-\alpha}C_2(h;\alpha o)\le 2\alpha C(o)$, where the last inequality follows because $C_2(h;\alpha o)\le(1-\alpha)C(o)$. Summing both inequalities for C_1 and C_2 proves the claim. □

Based on a simple single-commodity Braess instance [8], one can show that no Stackelberg strategy can induce a price of anarchy of one, even if there is only a single coalitional follower.

4.4 General Networks with Multiple Followers

In this section, we study SCALE for general networks and multiple coalitional followers.

Lemma 4 *Consider an instance $I(\alpha)$ of a Stackelberg game and let g be according to the SCALE strategy. Then, the following inequality holds*:

$$\sum_{k\in[m]}\sum_{a\in A}(\ell_a(\alpha o_a+h_a)+\ell'_a(\alpha o_a+h_a)h_a^k)(x_a^k-h_a^k)\ge 0,$$

where h is the flow of the followers and x is any feasible flow for the demand $(1-\alpha)r$.

Proof The lemma follows directly from (6). Taking $x_a:=(1-\alpha)o_a$, which is a feasible flow for the remaining $(1-\alpha)r$ demand, we get

$$\sum_{k\in[m]}\sum_{a\in A}(\ell_a(\alpha o_a+h_a)+\ell'_a(\alpha o_a+h_a)h_a^k)((1-\alpha)o_a^k-h_a^k)\ge 0.$$

□

For a latency function ℓ and a nonnegative number λ_1, we define the nonnegative value:

$$\omega_1(\ell;\alpha,\lambda_1):=\sup_{o,x\ge 0}\frac{\ell(\alpha o+h)\alpha o-\lambda_1\ell(o)o}{\ell(\alpha o+h)(\alpha o+h)}. \tag{9}$$

We assume by convention $0/0=0$.

For a given class $\mathcal{L}$, we further define $\omega_1(\mathcal{L};\alpha,\lambda_1) := \sup_{\ell\in\mathcal{L}} \omega_1(\ell;\alpha,\lambda_1)$. Similarly,

$$\omega_2(\ell;\alpha,m,\lambda_2) := \sup_{o,h\geq 0} \frac{((1-\alpha)\ell(\alpha o+h)-\lambda_2\ell(o))o+z(f,h)}{\ell(\alpha o+h)(\alpha o+h)},$$

with $z(f,h) := \ell'(\alpha o+h)(\sum_{k\in[m]}[(1-\alpha)h^k o^k - (h^k)^2])$. Note that the value $z(f,h)$ is at most $\ell'(\alpha o+h)\frac{(1-\alpha)(o)^2}{4}$. We define $\omega_2(\mathcal{L};\alpha,m,\lambda_2) := \sup_{\ell\in\mathcal{L}} \omega_2(\ell;\alpha, m,\lambda_2)$.

Proposition 2 *Consider an instance $I(\alpha)$ of a Stackelberg game and let g be according to the SCALE strategy. Then,*

$$C_1(g;h) \leq \lambda_1 C(o) + \omega_1(\mathcal{L};\alpha,\lambda_1)C(g+h),$$
$$C_2(h;g) \leq \lambda_2 C(o) + \omega_2(\mathcal{L};\alpha,m,\lambda_2)C(g+h).$$

The proof simply uses Lemma 4 and the definitions of ω_1 and ω_2.

Before we state the main theorem, we define

$$\Lambda_m(\mathcal{L};\alpha) := \{(\lambda_1,\lambda_2)\in\mathbb{R}^2_+ \,|\, (1-(\omega_1(\mathcal{L};\alpha,\lambda_1)+\omega_2(\mathcal{L};\alpha,m,\lambda_2)))>0\}.$$

Note that the set $\Lambda(\mathcal{L};\alpha)$ may be empty.

Theorem 8 *Consider a family of instances $\mathcal{I}(\alpha)$ of Stackelberg games, where g is defined according to the SCALE strategy. Then, the price of anarchy is at most* $\inf_{(\lambda_1,\lambda_2)\in\Lambda_m(\mathcal{L};\alpha)}[\frac{\lambda_1+\lambda_2}{1-(\omega_1(\mathcal{L};\alpha,\lambda_1)+\omega_2(\mathcal{L};\alpha,m,\lambda_2))}]$.

The proof uses the previous proposition.

Affine Latency Functions We will use Theorem 8 to prove upper bounds on the price of anarchy for affine latencies. First, we need two technical lemmas.

Lemma 5 *For $\lambda_1\in\mathbb{R}_+$, $\omega_1(\mathcal{L}_1;\alpha,\lambda_1)\leq\max\{\frac{\alpha-\lambda_1}{\alpha},\frac{\alpha^2}{4\lambda_1}\}$.*

Proof We start with constant latency functions $\ell(z)=c_0$. By definition of $\omega_1(\mathcal{L};\alpha,\lambda_1)$ we get

$$\omega_1(\mathcal{L};\alpha,\lambda_1) = \sup_{o,h\geq 0}\frac{\alpha o c_0 - \lambda_1 o c_0}{(\alpha o+h)c_0} \leq \max\left\{\frac{\alpha-\lambda_1}{\alpha},0\right\}.$$

For linear latency functions $\ell(z)=c_1 z$, we get

$$\omega_1(\mathcal{L};\alpha,\lambda_1) = \sup_{o,h\geq 0}\frac{c_1(\alpha o+h)\alpha o-\lambda_1 c_1 o^2}{c_1(\alpha o+h)^2} = \sup_{o,h\geq 0}\frac{(\alpha o+h)\alpha o-\lambda_1 o^2}{(\alpha o+h)^2}.$$

We define $\mu := \frac{h}{o}$ if $o > 0$ and zero otherwise. This yields

$$\omega_1(\mathcal{L}; \alpha, \lambda_1) \le \max_{\mu \ge 0} \frac{\alpha^2 + \alpha\mu - \lambda_1}{(\alpha + \mu)^2} \le \frac{\alpha^2}{4\lambda_1}.$$

Since $\frac{\alpha^2}{4\lambda_1} \ge 0$, we get the claim. $\square$

Lemma 6 *For* $\lambda_2 \ge \frac{1+\alpha-2\alpha^2}{4}$, $\omega_2(\mathcal{L}_1; \alpha, m, \lambda_2) \le \max\{\frac{1-\alpha-\lambda_2}{\alpha}, \frac{(1-\alpha)^2}{4\lambda_2+\alpha-1}\}$.

Proof We start with constant latency functions $\ell(z) = c_0$. By definition of $\omega_2(\mathcal{L}; \alpha, \lambda_2)$ and since $h \ge 0$ we get

$$\omega_2(\mathcal{L}; \alpha, m, \lambda_2) = \sup_{o,h \ge 0} \frac{(1-\alpha)oc_0 - \lambda_2 oc_0}{(\alpha o + h)c_0} \le \frac{1 - \alpha - \lambda_2}{\alpha}.$$

For linear latency functions $\ell(z) = c_1 z$, we get

$$\begin{aligned}\omega_2(\mathcal{L}; \alpha, m, \lambda_2) &\le \sup_{o,h \ge 0} \frac{c_1(\alpha o + h)(1-\alpha)o - \lambda_2 c_1 o^2 + c_1 \frac{1-\alpha}{4} o^2}{c_1(\alpha o + h)^2} \\ &= \sup_{o,h \ge 0} \frac{(\alpha o + h)(1-\alpha)o - \lambda_2 o^2 + \frac{1-\alpha}{4} o^2}{(\alpha o + h)^2}.\end{aligned}$$

We define $\mu := \frac{h}{o}$ if $o > 0$ and zero otherwise. This yields

$$\omega_2(\mathcal{L}; \alpha, m, \lambda_2) \le \max_{\mu \ge 0} \frac{(1-\alpha)(\alpha + \mu) - \lambda_2 + \frac{1-\alpha}{4}}{(\alpha + \mu)^2} \le \frac{(\alpha - 1)^2}{\alpha + 4\lambda_2 - 1},$$

where $\mu^* = \frac{2\alpha^2 + 4\lambda_2 - 1 - \alpha}{2(1-\alpha)}$ is the optimal solution to the above convex program. Using $\lambda_2 \ge \frac{1+\alpha-2\alpha^2}{4}$ we have $\mu^* \ge 0$, which proves the claim. $\square$

Theorem 9 *Consider a family of instances $\mathcal{I}(\alpha)$ of Stackelberg games such that latency functions are affine. Then, the price of anarchy for the SCALE strategy and m coalitional followers is at most*

$$\frac{(1 + 2\sqrt{1-\alpha})(1 + \sqrt{1-\alpha})^2}{4 + 4\sqrt{1-\alpha} - 3\alpha} \quad \textit{for } \alpha \in \left[0, \frac{1}{2}\sqrt{3}\right]$$

and

$$\frac{(-3\alpha - 2\alpha\sqrt{1-\alpha} - 1 + 2\alpha^2)(1 + \sqrt{1-\alpha})\alpha}{2(-3\alpha - 3\alpha\sqrt{1-\alpha} + 1 + \sqrt{1-\alpha} + \alpha^2)} \quad \textit{for } \alpha \in \left[\frac{1}{2}\sqrt{3}, 1\right].$$

Proof We define for $\alpha \in [0, \frac{1}{2}\sqrt{3}]$

$$\lambda_1 = \frac{1}{2}(1 + \sqrt{1-\alpha})\alpha, \qquad \lambda_2 = \frac{1}{2}(1 + \sqrt{1-\alpha})(1-\alpha).$$

This choice satisfies the conditions:

$$\frac{\alpha-\lambda_1}{\alpha}=\frac{\alpha^2}{4\lambda_1},\qquad \frac{1-\alpha-\lambda_2}{\alpha}\le\frac{(1-\alpha)^2}{4\lambda_2}.$$

Note that for $\alpha\in[0,\frac{1}{2}\sqrt{3}]$ we have $\lambda_2\ge\frac{1+\alpha-2\alpha^2}{4}$ as required in Lemma 6. From Lemmas 5 and 6, we thus obtain

$$\begin{aligned}
&\omega_1(\mathcal{L}_1;\alpha,\lambda_1)+\omega_2(\mathcal{L}_1;\alpha,m,\lambda_2)\\
&\quad=\frac{1-\lambda_1}{\alpha}+\frac{(1-\alpha)^2}{4\lambda_2+\alpha-1}\\
&\quad=\frac{2+2\sqrt{1-\alpha}-\alpha}{2(1+\sqrt{1-\alpha})(1+2\sqrt{1-\alpha})}\\
&\quad=\frac{1}{(1+2\sqrt{1-\alpha})}-\frac{\alpha}{2(1+\sqrt{1-\alpha})(1+2\sqrt{1-\alpha})}<1.
\end{aligned}$$

Thus $(\lambda_1,\lambda_2)\in\Lambda_m(\mathcal{L}_1;\alpha)$ and applying Theorem 8 proves the first claim.

For $\alpha\in[\frac{1}{2}\sqrt{3},1]$ we define

$$\lambda_2=\frac{1+\alpha-2\alpha^2}{4}.$$

It is easy to prove that for $\alpha\in[\frac{1}{2}\sqrt{3},1]$ we have

$$\frac{1-\alpha-\lambda_2}{\alpha}\le\frac{(1-\alpha)^2}{4\lambda_2}.$$

Then, it also straightforward to check that $(\lambda_1,\lambda_2)\in\Lambda_m(\mathcal{L}_1;\alpha)$. Hence, applying Theorem 8 proves the second claim. □

5 Conclusions and Final Remarks

In the first part of this paper, we investigated the price of anarchy in nonatomic network games with coalitions. On the positive side, we derived an upper bound on the price of anarchy for restricted topologies (load balancing games). For general topologies and (semi-convex) latency functions, we developed a generic upper bound on the price of anarchy which depends on the specific class of allowable latency functions. We note that this bound actually holds for the larger class of congestion games with fractional demand assignments, because the proof technique does not use the network structure, but only uses variational inequalities which remain valid in this more general setting.

After the publication of a preliminary version of this article [18], there has been some work extending our results. Bhaskar et al. [5] showed that the upper bound of m on the price of anarchy for load balancing games (see Theorem 1) continues to hold for series-parallel networks. Roughgarden and Schoppmann [33] proved that the generic upper bound of Theorem 2 is in fact tight. They also give an exact closed-form expression for the price of anarchy for polynomial latency functions with nonnegative coefficients and bounded degree.

In the second part of this paper, we investigated Stackelberg routing as a means to improve the quality of Nash equilibria. In this setting, we investigated and designed Stackelberg strategies and derived bounds on the price of anarchy for restricted network topologies, number of followers, and classes of latency functions, respectively. Perhaps, the most intriguing open question in this setting is whether there exists a Stackelberg strategy that induces a constant price of anarchy (depending on α) for a finite number of following coalitions. So far, we only understand the extreme cases: for one follower, the answer is yes (Theorem 7), while for infinitely many followers, the answer is no, see [6].

References

1. Aland, S., Dumrauf, D., Gairing, M., Monien, B., Schoppmann, F.: Exact price of anarchy for polynomial congestion games. In: Proc. of the 23th Annual Symposium on Theoretical Aspects of Computer Science (STACS). Lecture Notes in Computer Science, vol. 3884, pp. 218–229. Springer, Berlin (2006)
2. Altman, E., Basar, T., Jimnez, T., Shimkin, N.: Competitive routing in networks with polynomial costs. IEEE Trans. Autom. Control **47**, 92–96 (2002)
3. Awerbuch, B., Azar, Y., Epstein, A.: The price of routing unsplittable flow. In: Proc. of the 37th Annual ACM Symposium on Theory of Computing, pp. 57–66. ACM, New York (2005)
4. Bhaskar, U., Fleischer, L., Hoy, D., Huang, C.: Equilibria of atomic flow games are not unique. In: Proc. of the 20th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA), pp. 748–757. SIAM, Philadelphia (2009)
5. Bhaskar, U., Fleischer, L., Huang, C.: The price of collusion in series-parallel networks. In Proc. of the 14th International IPCO Conference Integer Programming and Combinatorial Optimization. Springer (2010, to appear)
6. Bonifaci, V., Harks, T., Schäfer, G.: Stackelberg routing in arbitrary networks. Math. Oper. Res. (2010). doi:10.1287/moor.1100.0442
7. Boyd, S., Vandenberghe, L.: Convex Optimization. Cambridge University Press, Cambridge (2004)
8. Braess, D.: Über ein Paradoxon aus der Verkehrsplanung. Unternehmensforschung **12**, 258–268 (1969)
9. Catoni, S., Pallotino, S.: Traffic equilibrium paradoxes. Transp. Sci. **25**, 240–244 (1991)
10. Christodoulou, G., Koutsoupias, E.: The price of anarchy of finite congestion games. In: Proc. of the 37th Annual ACM Symposium on Theory of Computing (STOC), pp. 67–73. ACM, New York (2005)
11. Cominetti, R., Correa, J.R., Stier-Moses, N.E.: The impact of oligopolistic competition in networks. Oper. Res. **57**(6), 1421–1437 (2009)
12. Correa, J.R., Stier-Moses, N.E.: Stackelberg routing in atomic network games. Technical report, Columbia Business School, February 2007
13. Correa, J.R., Schulz, A.S., Stier-Moses, N.E.: Selfish routing in capacitated networks. Math. Oper. Res. **29**, 961–976 (2004)
14. Correa, J.R., Schulz, A.S., Stier-Moses, N.E.: A geometric approach to the price of anarchy in nonatomic congestion games. Games Econ. Behav. **64**(2), 457–469 (2008)
15. Dafermos, S.C., Sparrow, F.T.: The traffic assignment problem for a general network. J. Res. Natl. Bur. Stand. B **73**, 91–118 (1969)
16. Fotakis, D.: Stackelberg strategies for atomic congestion games. In: Proc. of the 15th Annual European Symposium on Algorithms (ESA). Lecture Notes in Computer Science, vol. 4698, pp. 299–310. Springer, Berlin (2007)
17. Fotakis, D., Kontogiannis, S., Spirakis, P.G.: Atomic congestion games among coalitions. ACM Trans. Algorithms **4**(4), 1–27 (2008)

18. Harks, T.: Stackelberg strategies and collusion in network games with splittable flow. In: Proc. of the 6th Workshop on Approximation and Online Algorithms (WAOA). Lecture Notes in Computer Science, vol. 5426. Springer, Berlin (2008)
19. Haurie, A., Marcotte, P.: On the relationship between Nash-Cournot and Wardrop equilibria. Networks **15**, 295–308 (1985)
20. Hayrapetyan, A., Tardos, E., Wexler, T.: The effect of collusion in congestion games. In: Proc. of the 38th Annual ACM Symposium on Theory of Computing (STOC), pp. 89–98. ACM, New York (2006)
21. Kaporis, A.C., Spirakis, P.G.: The price of optimum in Stackelberg games on arbitrary single commodity networks and latency functions. Theor. Comput. Sci. **410**(8–10), 745–755 (2009)
22. Karakostas, G., Kolliopoulos, S.G.: Stackelberg strategies for selfish routing in general multicommodity networks. Algorithmica **53**(1), 132–153 (2009)
23. Korilis, Y.A., Lazar, A.A., Orda, A.: Achieving network optima using Stackelberg routing strategies. IEEE/ACM Trans. Netw. **5**(1), 161–173 (1997)
24. Koutsoupias, E., Papadimitriou, C.H.: Worst-case equilibria. In: Proc. of the 16th Annual Symposium on Theoretical Aspects of Computer Science (STACS). Lecture Notes in Computer Science, vol. 1563, pp. 404–413. Springer, Berlin (1999)
25. Kumar, V.S.A., Marathe, M.V.: Improved results for Stackelberg scheduling strategies. In: Proc. of the 33rd International Colloquium of Automata, Languages and Programming (ICALP), pp. 776–787. Springer, Berlin (2002)
26. Mossberg, W.S.: Dash's car navigation gives smart directions, if others participate. Wall Str. J., March 27, 2008. http://ptech.allthingsd.com/20080327/dashs-car-navigator-gives-smart-directions-if-others-participate
27. Orda, A., Rom, R., Shimkin, N.: Competitive routing in multiuser communication networks. IEEE/ACM Trans. Netw. **1**(5), 510–521 (1993)
28. Perakis, G.: The price of anarchy when costs are non-separable and asymmetric. In: Proc. of 10th International Conference on Integer Programming and Combinatorial Optimization, (IPCO), New York. Lecture Notes in Computer Science, vol. 3064, pp. 46–58. Springer, Berlin (2004)
29. Roughgarden, T.: The price of anarchy is independent of the network topology. J. Comput. Syst. Sci. **67**, 341–364 (2002)
30. Roughgarden, T.: Stackelberg scheduling strategies. SIAM J. Comput. **33**(2), 332–350 (2004)
31. Roughgarden, T.: Selfish Routing and the Price of Anarchy. MIT Press, Cambridge (2005)
32. Roughgarden, T.: Selfish routing with atomic players. In: Proc. of the 16th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA), pp. 973–974. SIAM, Philadelphia (2005)
33. Roughgarden, T., Schoppmann, F.: The price of anarchy in atomic splittable congestion games (2010, unpublished manuscript)
34. Roughgarden, T., Tardos, E.: How bad is selfish routing? J. ACM **49**(2), 236–259 (2002)
35. Sharma, Y., Williamson, D.: Stackelberg thresholds in network routing games or the value of altruism. In: Proc. of the 8th ACM Conference on Electronic Commerce (EC), pp. 93–102. ACM, New York (2007)
36. Swamy, C.: The effectiveness of Stackelberg strategies and tolls for network congestion games. In: Proc. of the 18th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA), pp. 1133–1142. SIAM, Philadelphia (2007)
37. Wardrop, J.G.: Some theoretical aspects of road traffic research. Proc. Inst. Civ. Eng. **1**(Part II), 325–378 (1952)