Stackelberg Thresholds in Network Routing Games or The Value of Altruism

Yogeshwer Sharma[*]
Department of Computer Science
Cornell University, Ithaca NY 14853.
yogi@cs.cornell.edu

David P. Williamson[†]
School of OR&IE
Cornell University, Ithaca NY 14853.
dpw@cs.cornell.edu

ABSTRACT

Noncooperative network routing games are a natural model of users trying to *selfishly* route flow through a network in order to minimize their own delays. It is well known that the solution resulting from this selfish routing (called the Nash equilibrium) can have social cost strictly higher than the cost of the optimum solution. One way to improve the quality of the resulting solution is to *centrally control* a fraction of the flow. A natural problem for the network administrator then is to route the centrally controlled flow in such a way that the overall cost of the solution is minimized after the remaining fraction has routed itself selfishly.

This problem falls in the class of well-studied Stackelberg routing games. We consider the scenario where the network administrator wants the final solution to be (strictly) better than the Nash equilibrium. In other words, she wants to control enough flow such that the cost of the resulting solution is strictly less than the cost of the Nash equilibrium.

We call the minimum fraction of users that must be centrally routed to improve the quality of the resulting solution the *Stackelberg threshold*. We give a closed form expression for the Stackelberg threshold for parallel links networks with linear latency functions. The expression is in terms of Nash equilibrium flows and optimum flows. It turns out that the Stackelberg threshold is the minimum of Nash flows on links which have more optimum flow than Nash flow.

Using our approach to characterize the Stackelberg thresholds, we are able to give a simpler proof of an earlier result which finds the minimum fraction required to be centrally controlled to induce an optimum solution.

Categories and Subject Descriptors: F.2.2 [Analysis of Algorithms and Problem Complexity]: Nonnumerical Algorithms and Problems—*Routing and layout*

General Terms: Theory.

Keywords: Game theory, Nash equilibrium, Stackelberg equilibrium, centrally controlled flow, altruistic flow, Stackelberg threshold, price of anarchy.

1. Introduction and related work

Noncooperative network routing games are a nice model of the behavior of selfish users trying to optimize their own benefit. In such a game, each player intends to send a fixed amount of flow from its source to its sink using a shortest delay path through the given network in a noncooperative manner.

The solution reached by players selfishly routing their flow is called the Nash equilibrium or Nash flow. Since players choose their paths to minimize their own delay alone, the quality of the resulting Nash equilibrium in general may be worse than the quality of the optimum way to route flow through the network so as to minimize the total overall latency of all users, which may be thought of as the social cost of the routing. A classic example of Pigou [13] shows that this can indeed be the case. The ratio of the cost of the Nash equilibrium to the optimum solution is called the *Price of Anarchy* [9]. The idea of bounding the price of anarchy in network routing games has become well-studied after the groundbreaking work of Roughgarden and Tardos [19]. Roughgarden and Tardos show that for general latency functions, the price of anarchy can be arbitrarily large. For the class of networks with linear latency functions, however, they prove that the price of anarchy is bounded by 4/3.

The Nash equilibrium is an attractive concept from the point of view of the study of stable equilibria since no player has any incentive to unilaterally change his/her strategy. But its inefficiency (that is, its potentially large cost compared to the social optimum) has always been a concern. There has been substantial work on ways to address this issue. Some such methods are: (i) *Mechanism design*, in which the rules of the game are established to help ensure that the quality of the resulting Nash compatible with the rules is good compared to the social optimum, (see, for example, [12, 11]), (ii) *Taxes and tolls* on network links to discourage users from using some links which lead to inefficient equilibria, (see, for example, [1, 2, 5, 4]), (iii) *Designing the network* in such a way that the network has good Nash to optimum ratio to start with (see, for example, [14, 8]), and (iv) *Capacity augmentation*, such that the cost of Nash equilibrium in augmented network is good compared to the cost of optimum in the original network (see, for example, [19]).

---

[*] Supported by NSF grant CCF-0514628.

[†] Supported by NSF grant CCF-0514628.

EC'07, June 13–16, 2007, San Diego, California, USA.

Also see Roughgarden's survey [18] for a discussion about coping with inefficiency of Nash equilibria. All these methods necessitate either a change in the way game is played in the existing network or a change in the network itself.

Another way to improve the quality of the Nash equilibrium is to consider situations in which not all flow is routed selfishly. The motivation comes from considering networks where there is a mix of selfish and centrally controlled players. An example of such a network mentioned in Roughgarden's thesis [15, Chapter 6] is that of a network where there may be two different prices. Clients paying the *premium* price get to choose their own route through the network and those paying the *bargain* price do not get a choice of routes—they are controlled centrally by the network administrator. Roughgarden [16] considers the problem of routing a β fraction of flow centrally in such a way that if the remaining $1-\beta$ fraction chooses their own paths selfishly then the cost of the resulting solution is minimized. He calls the routing of the centrally controlled flow a *Stackelberg strategy* and the resulting equilibrium the *equilibrium induced by the strategy* with fraction β; we will refer to the latter as simply the *Stackelberg equilibrium.* He addresses the question of finding a Stackelberg strategy such that the cost of the resulting Stackelberg equilibrium is close to the social optimum. For a network of parallel links and centrally controlled $0 \le \beta \le 1$ fraction of flow, he gives a Stackelberg strategy such that the resulting Stackelberg equilibrium comes within a $1/\beta$ factor of the social optimum for arbitrary latencies and within a $4/(3+\beta)$ factor for linear latencies.

To be more specific about the problem considered by Roughgarden, we let G be a network with two nodes $\{s, t\}$, a source s and a sink t, and k directed parallel links $\{e_1, e_2, \ldots, e_k\}$ from s to t. Each edge e_i is equipped with a latency function $l_i(x) : \mathbb{R}_{\ge 0} \to \mathbb{R}_{\ge 0}$ which is nonnegative, continuous, and nondecreasing. A total flow of amount r is to be routed from s to t such that the total latency experienced by the whole flow is minimized. In other words, the socially optimum flow $f = (f_1, f_2, \ldots, f_k)$ is such that $\sum_{i=1}^{k} f_i = r$ and $\sum_{i=1}^{k} f_i \cdot l_i(f_i)$ is minimized. A Stackelberg strategy $\bar{h}$ for a given fraction β is a flow $\bar{h} = (\bar{h}_1, \ldots, \bar{h}_k)$ such that $\sum_{i=1}^{k} \bar{h}_i = \beta r$. Define $\tilde{l}_i(x) = l_i(x + \bar{h}_i)$ for $i = 1, \ldots, k$. Then the Stackelberg equilibrium *induced by* $\bar{h}$ is a Nash equilibrium routing $(1-\beta)r$ flow in the graph G with latencies $\tilde{l}$.

For $0 \le \beta \le 1$, let $c(G, r, l, \beta, \bar{h})$ be the cost of the Stackelberg equilibrium induced by $\bar{h}$, and let $c(G, r, l, \beta) = \min_{\bar{h}} c(G, r, l, \beta, \bar{h})$ be the cost of the optimum Stackelberg equilibrium with centrally controlled flow fraction β. Then $c(G, r, l, 1)$ is the social optimum cost, and $c(G, r, l, 0)$ is the social cost of the Nash flow. Note that finding $c(G, r, l, \beta)$ for an arbitrary network and an arbitrary β is weakly NP-complete as proved in [15, Chapter 6]. Roughgarden [16] has shown that $c(G, r, l, \beta) \le \frac{1}{\beta} c(G, r, l, 1)$ for arbitrary latencies, and when the latency functions l_i are linear, then $c(G, r, l, \beta) \le \frac{4}{3+\beta} c(G, r, l, 1)$.

There has been a fair amount of followup work on finding good Stackelberg strategies. For parallel links networks equipped with latency functions represented as polynomials with non-negative coefficients, Kumar and Marathe [10] give a polynomial-time algorithm for finding a Stackelberg strategy $\bar{h}$ such that $c(G, r, l, \beta, \bar{h}) \le (1+\varepsilon) c(G, r, l, \beta)$ for any given $\varepsilon > 0$. Swamy [20] extends the results for Roughgarden's Largest Latency First strategy [16] to incorporate various topologies and arbitrary latency functions. For series-parallel graphs (with arbitrary latency functions), he bounds the price of anarchy by $1 + 1/\beta$ and for the parallel links graphs (with latency functions from a class $\mathcal{L}$), by $\beta + (1-\beta)\rho(\mathcal{L})$ where $\rho(\mathcal{L})$ is the price of anarchy for networks with latency functions from class $\mathcal{L}$. For general graphs, he obtains latency class specific bounds on the price of anarchy which give a continuous tradeoff between the fraction of flow controlled and the price of anarchy. Correa and Stier-Moses independently obtained the bound of $1 + \frac{1}{\beta}$ for the series-parallel graphs [3]. For general topology networks equipped with linear delay functions, and multicommodity users, Karakostas and Kolliopoulos [7] show that the cost of a particular Stackelberg equilibrium (corresponding to the SCALE strategy as suggested in [16]) with β fraction of centrally controlled flow is at most $(4-X)/3$ times the cost of the optimum solution where $X = \frac{(1-\sqrt{1-\beta})(3\sqrt{1-\beta}+1)}{2\sqrt{1-\beta}+1}$.

In this paper, we study a simple but interesting question regarding Stackelberg equilibria in this setting: what fraction β of flow needs to be centrally controlled for there to be *any* improvement in the social cost whatsoever? We call this amount the *Stackelberg threshold* and denote it by $\sigma(G, r, l)$. To be more precise, $\sigma(G, r, l)$ is the minimum value of β such that $c(G, r, l, \beta + \varepsilon) < c(G, r, l, 0)$ for any $\varepsilon > 0$. In the network setting of Roughgarden, the Stackelberg threshold is the minimum fraction of bargain price users that the network administrator must get in order for the overall routing cost to be less than the cost of the Nash equilibrium in which everyone is able to route their own flows selfishly.

At first glance, it might appear that the threshold is trivially 0: that is, $c(G, r, l, \varepsilon) < c(G, r, l, 0)$ for any network G of parallel links. However, if the latency functions are such that $c(G, r, l, 0) = c(G, r, l, 1)$—that is, the Nash equilibrium happens to have optimum social cost—this is clearly false.

As this example points out, the threshold depends on the price of anarchy of the instance. This is also implied by Roughgarden's result. For linear latency functions, Roughgarden and Tardos [19] show that $c(G, r, l, 0)/c(G, r, l, 1) \le 4/3$ (in any network, not necessarily parallel links). Let us denote the price of anarchy by $\rho(G, r, l) \equiv c(G, r, l, 0)/c(G, r, l, 1)$. Then by Roughgarden's result we have that

$$c(G, r, l, \beta) \le \frac{4}{3+\beta} c(G, r, l, 1) = \frac{4}{3+\beta} \frac{c(G, r, l, 0)}{\rho(G, r, l)}.$$

Then a sufficient condition for $c(G, r, l, \beta) < c(G, r, l, 0)$ is $\frac{4}{3+\beta} \frac{1}{\rho(G,r,l)} < 1$, or $\beta > \frac{4}{\rho(G,r,l)} - 3$. If the price of anarchy is as bad as it can be, i.e., $\rho(G, r, l) = 4/3$, then $c(G, r, l, \beta) < c(G, r, l, 0)$ whenever $\beta > 0$ so that the Stackelberg threshold is 0 for these instances.

Our central result is to give a precise characterization of the Stackelberg threshold for the case of parallel link graphs with linear delay functions $l(x) = ax + b$ with $a, b \ge 0$. If f is a Nash flow, g is an optimum flow, and $c(G, r, l, 0) > c(G, r, l, 1)$ then we show that $\sigma(G, r, l) = \min_{i : f_i < g_i} f_i$. If $c(G, r, l, 0) = c(G, r, l, 1)$ then clearly no improvement in the cost of the Nash flow is possible.

Kaporis and Spirakis [6] study a closely related problem of determining the minimum fraction β_M of flow in a graph such that controlling that fraction through some Stackelberg strategy gives rise to a Stackelberg equilibrium of cost equal to the cost of the optimum solution; that is, they find the minimum β_M such that $c(G, r, l, \beta_M) = c(G, r, l, 1)$. They call this the *Price of Optimum* and give an algorithm to compute the price of optimum for single commodity networks equipped with continuous, differentiable, and strictly increasing latency functions. Using some insights from the proof of our main result, we are able to give a short proof of this result as well for the case of parallel links networks.

Before turning to the proof of our results, we conclude with a brief reflection on our motivation for considering Stackelberg thresholds. Stackelberg network routing games are usually discussed in the context of the central control of flow. This creates images of technocrats coercing routings for the benefit of society, either directly by controlling users or indirectly via taxes. An alternate image (though equivalent mathematically) is that of small coalitions of users behaving altruistically; that is, deciding not to behave selfishly, but in ways that improve the overall social welfare. Our reason for studying Stackelberg thresholds is to ask: how big do such coalitions have to be in order to make a difference? Part of the answer given by Roughgarden's work is: it depends on how bad things are. When things are at their worst, even infinitesimally small coalitions make a difference. Studies of the price of anarchy ask how bad off we are if everyone behaves selfishly; part of our motivation is to flip the question and ask how much better off we can be if some small fraction of users do not. Hence we ask not what is the price of anarchy, but what is the value of altruism? How much of it is required to be useful? This perspective suggests an interesting research agenda to which this paper is a modest contribution.

The paper is structured as follows. We begin in Section 2 with some introductory notation, definitions, and lemmas. Then in Section 3 we state the lemmas we will need to prove, and show how they imply our main theorem. Section 5 gives a way to transform a Nash equilibrium into a Stackelberg equilibrium, which is used to prove the main theorem. Sections 6 and 7 prove main technical lemmas. Our proof of correctness of the Kaporis-Spirakis algorithm to compute the Price of Optimum is in Section 8.

2. Some introductory definitions and lemmas

Let G be a graph with two nodes, a source s and a sink t, and with k parallel links from s to t. We require r units of flow to be sent from s to t. Let the *latency* on link i be $l_i(x_i) = a_i x_i + b_i$ with $a_i, b_i \geq 0$; we will sometimes refer to this as the *delay* of the link. We assume without loss of generality that there is exactly one link whose latency function is constant. If there are more than one such links, then we can remove all but the one with the minimum latency. If there is none, we can add one with a large enough latency, say $\max_{i\in[k]} l_i(r)$, without affecting anything.

Let f be a Nash flow sending flow f_i on link i and g be an optimum flow sending flow g_i on link i. By feasibility of flows, $\sum_{i=1}^k f_i = \sum_{i=1}^k g_i = r$. The goal is to determine for this specific network G the minimum value of $\beta \in [0, 1]$ such that $c(G, r, l, \beta + \varepsilon) < c(G, r, l, 0)$ for any $\varepsilon > 0$, or equivalently the minimum value of β such that getting central control of infinitesimally more fraction of flow than β allows a Stackelberg equilibrium of cost strictly less than the cost of the Nash equilibrium.

Throughout this paper, we will let $\bar{h}$ denote a Stackelberg strategy and h denote an induced Stackelberg equilibrium. It is worth noting that the flow h does not include $\bar{h}$ in it; that is $\sum_j h_j = r - \sum_j \bar{h}_j$ (and not equal to r). We call the flow controlled by a Stackelberg strategy *centrally controlled flow* or *altruistic flow*, and the equilibrium flow h the *selfish flow*.

Before stating (and proving) our main theorem in the next section, we recall the following well-known lemmas and definitions specialized to the case of parallel links. Most of the omitted proofs can be found in [17].

LEMMA 1. *A flow f is at Nash equilibrium if and only if for every i with $f_i > 0$ and every $j \neq i$, $l_i(f_i) \leq l_j(f_j)$. For a Stackelberg strategy $\bar{h}$, h is an induced Stackelberg equilibrium if and only if for all i with $h_i > 0$ and $j \neq i$, $l_i(\bar{h}_i + h_i) \leq l_j(\bar{h}_j + h_j)$.*

DEFINITION 2. *Let $l_i^*(x) = \frac{d}{dx}(x \cdot l_i(x)) = l_i(x) + x \cdot \frac{d}{dx}(l_i(x))$. We call this the* marginal latency *of link i. We also denote $\frac{d}{dx}(l_i(x))$ by $l_i'(x)$.*

LEMMA 3. *A flow g is optimal if and only if for every i with $g_i > 0$ and every $j \neq i$, $l_i^*(g_i) \leq l_j^*(g_j)$.*

By Lemma 1, each link j with $f_j > 0$ must have the same latency in f; we denote this common latency by L. Similarly, the common marginal latency of links with $g_j > 0$ in the optimum solution is denoted by L^* and the common latency of all links with $h_j > 0$ in a Stackelberg equilibrium is denoted by L_h. For the Nash flow f, marginal latencies $l_j^*(f_j)$ may not all be same. We use ε_j^* to denote their deviation from L^*; that is for all j, we let ε_j^* be such that $L^* + \varepsilon_j^* = l_j^*(f_j)$.

OBSERVATION 4. *The latency (marginal latency) of any link carrying positive flow in Nash equilibrium (optimum flow) cannot be larger than the latency (marginal latency) of the constant link. That is, $f_j > 0$ implies $L = a_j f_j + b_j \leq b_z$ and $g_j > 0$ implies $L^* = 2a_j g_j + b_j \leq b_z$. In particular, $\varepsilon_z^* \geq 0$.*

We review some results about the uniqueness of Nash equilibrium and Stackelberg equilibrium.

LEMMA 5. *If f and f' are flows at Nash equilibrium for the instance (G, r, l), then (i) the cost of f is equal to the cost of f', (ii) for all $j \in [k]$, $l_j(f_j) = l_j(f_j')$, and (iii) $L_f = L_{f'}$ where L_f and $L_{f'}$ denote the common latencies experienced by selfish flow in f and f' respectively.*

LEMMA 6. *Let h and h' be two Stackelberg equilibria induced by Stackelberg strategy $\bar{h}$. Then the costs of these two equilibria are equal.*

We partition all links into two sets, the set of *good* links, and the set of *bad* links. We define them next.

DEFINITION 7. *Let f be a Nash equilibrium and g be an optimum solution. A link j is called a* good *link if $f_j < g_j$, otherwise it is called a* bad *link.*

3. The proof of the main theorem

We present in this section our main theorem and some lemmas which will be helpful in the proof of the theorem. Recall that a good link is the one with more optimum flow than the Nash flow ($f_j < g_j$) and a bad link is the one with at least as much Nash flow as the optimum flow ($f_j \geq g_j$). At a high level, we show that a Stackelberg strategy routing more than the Nash flow on a good link gives rise to a Stackelberg equilibrium of improved social cost, while a Stackelberg strategy routing more than the Nash flow on bad links gives rise to Stackelberg equilibria with worse social cost (subject to some caveats). We let f_* denote the minimum of Nash flows on good links, and i_* the corresponding link. Formally,

$$f_* \overset{\text{def}}{=} \min_{i:f_i<g_i} f_i, \qquad i_* \overset{\text{def}}{=} \operatorname*{argmin}_{i:f_i<g_i} f_i. \tag{1}$$

If there are many indices for which the flow is equal to f_*, define i_* to be an arbitrary such link, say the lowest indexed one. We have $f_* = f_{i_*}$. Our main theorem in essence states that the Stackelberg threshold for the network G is f_*/r.

We need some notation to introduce our main theorem.

Definition 8. *For Nash flow f and a Stackelberg strategy $\bar{h}$, let $S(\bar{h}) = \{j : f_j < \bar{h}_j\}$ and $U(\bar{h}) = \{j : 0 < \bar{h}_j \leq f_j\}$. Let $\beta(\bar{h})$ denote the fraction of flow centrally controlled by Stackelberg strategy $\bar{h}$, that is $\beta(\bar{h}) = \left(\sum_j \bar{h}_j\right)/r$.*

The set of all other links is $[k] - S(\bar{h}) - U(\bar{h})$. We now state two lemmas which are central to the proof of our main theorem. Their proofs appear in subsequent sections. The first one provides the upper bound of f^*/r on the Stackelberg threshold, and the second one provides a lower bound of f^*/r on the Stackelberg threshold.

Lemma 9. *Suppose $c(G, r, l, 0) > c(G, r, l, 1)$ and $\beta > \frac{f_*}{r}$. Then there exists a Stackelberg strategy $\bar{h}$ with $\beta(\bar{h}) = \beta$ and $c(G, r, l, \beta, \bar{h}) < c(G, r, l, 0)$ (which implies that $c(G, r, l, \beta) < c(G, r, l, 0)$). In other words, for the amount of centrally controlled flow strictly more than f_*, there exists a Stackelberg strategy controlling that amount of flow and having cost strictly less than the cost of the Nash equilibrium.*

We comment on the Stackelberg strategy (and equilibrium) alluded to above. If the altruistic flow amount is more than f_*, then the Stackelberg strategy referred to above routes $f_* + \varepsilon$ flow on link i_* for a small enough ε and remaining flow *appropriately* on links other than i. For this Stackelberg strategy, $S(\bar{h}) = \{i_*\}$ and $\bar{h}_{i_*} = f_* + \varepsilon$ for small ε.

Lemma 10. *Suppose $c(G, r, l, 0) > c(G, r, l, 1)$ and $\beta \leq \frac{f_*}{r}$. Also assume that there is no link with zero Nash flow and positive optimum flow, that is $\{j : f_j = 0, g_j > 0\} = \emptyset$. Then for any Stackelberg strategy $\bar{h}$ with $\beta(\bar{h}) = \beta$, $c(G, r, l, \beta, \bar{h}) \geq c(G, r, l, 0)$ (or equivalently $c(G, r, l, \beta) \geq c(G, r, l, 0)$). In other words, any Stackelberg strategy controlling at most f_* amount of flow induces Stackelberg equilibria of cost at least as much as that of the Nash equilibrium.*

The condition $\beta \leq \frac{f_*}{r}$ implies that $S(\bar{h})$ does not contain any good link (otherwise, the amount of altruistic flow will be more than f_*). The above lemma in essence states that if the Stackelberg strategy does not have enough flow to influence flow on a good link, then it cannot induce a Stackelberg equilibrium of cost less than the cost of the Nash equilibrium.

We are now ready to state and prove our main theorem.

Theorem 11. *If $c(G, r, l, 0) > c(G, r, l, 1)$, then for a parallel link network G with k links and delay functions of the form $l(x) = ax + b$ with $a, b \geq 0$, Nash flow f, and optimum flow g,*

$$\sigma(G, r, l) = f_* \,/\, r = \min_{i:f_i<g_i} f_i \Big/ r. \tag{2}$$

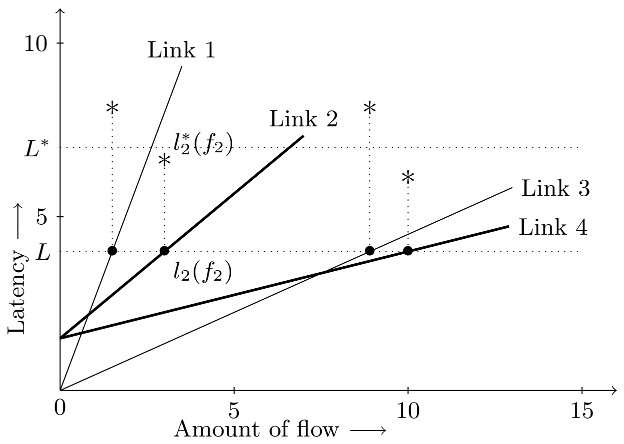

Figure 1: A figure of a network depicting latencies (and modified latencies). There are four links, curves $l_i(x)$ are drawn for all links. Here $l_1(x) = \frac{8}{3}x$, $l_2(x) = \frac{5}{6}x + \frac{3}{2}$, $l_3(x) = \frac{40}{89}x$, and $l_4(x) = \frac{1}{2}x + \frac{3}{2}$. Flow on link i is f_i; bullets denote $l_i(f_i)$ values and asterisks denote $l_i^*(f_i)$ values. Here $f_1 = \frac{3}{2}$, $f_2 = 3$, $f_3 = \frac{89}{10}$, and $f_4 = 10$. Good links (links 2 and 4, drawn in thick lines) have $l_i^*(f_i) < L^*$ (equivalently asterisk below L^* line) and bad links (links 1 and 3, drawn in thin lines) have $l_i^*(f_i) \geq L^*$ (equivalently asterisk on or above L^* line).

Proof. If there is a link with zero Nash flow and positive optimum flow, then $f_* = 0$. Lemma 9 states that controlling $\varepsilon > 0$ amount of flow gives rise to a Stackelberg equilibrium of cost less than the cost of the Nash equilibrium. It should also be clear that a Stackelberg strategy controlling zero amount of flow cannot have Stackelberg equilibrium of cost less than the cost of the Nash equilibrium. So, $\sigma(G, r, l) = 0 = f_*/r$ in this case.

We can now assume that there is no link with zero Nash flow and positive optimum flow. Lemma 9 states that $\sigma(G, r, l) \leq f_*/r$ since it gives a Stackelberg strategy controlling $f_* + \varepsilon$ flow which gives rise to a Stackelberg equilibrium with strictly smaller cost than the cost of the Nash equilibrium. Lemma 10 states that $\sigma(G, r, l) \geq f_*/r$ since any Stackelberg strategy controlling at most f_* flow cannot improve the cost to better than the cost of the Nash equilibrium. It follows that $\sigma(G, r, l) = f_*/r$. ☐

We comment here on the definition of f_*. The quantity f_* is defined with respect to a Nash equilibrium and an optimum solution (which may not be unique). But the condition $c(G, r, l, 0) > c(G, r, l, 1)$ together with Lemma 14 below guarantees that the Nash equilibrium is unique, as is the optimum solution, so this quantity is well defined.

We next give some intuition behind the proofs.

3.1 An informal explanation of the proof

We represent the network as a collection of links, and each link by a curve (corresponding to its latency) on two dimensional coordinate axes, where the x-axis corresponds to the amount of flow and the y-axis corresponds to the latency experienced.

If we look at the network in terms of Figure 1, controlling flow on link i has a natural interpretation. Say, the Stackelberg strategy pushes more than f_1 flow on link 1, this gives rise to a latency of more than L on link 1, and the flow on other links decreases so that all latencies on those links remain the same (the latency level comes down). We can think of latencies as the water-level which remains equal on all uncontrolled links. In terms of the water-level, the optimum solution has a water-level for marginal latency (in the same way as the Nash equilibrium has water-level for latency), called L^* in Figure 1. The marginal latencies of links in Nash equilibrium in general are not the same as L^*, they have a deviation from L^*; link 1 has a deviation of $(l_1^*(f_1) - L^*)$, link 2 has a deviation of $(l_2^*(f_2) - L^*)$, and so on. The intuition is that if a strategy tries to decrease the absolute value of this deviation, then the quality of resulting solution is better than the quality of the Nash equilibrium, if it tries to increase the absolute value of the deviations, then the quality is worse.

The proof of Lemma 9 basically requires us to prove that if we push more flow than the Nash flow on a good link such as link 2 (see Figure 1, link 2 is a good link) decreasing the absolute value of its deviation from L^*, then the overall cost of the solution decreases, the decrease coming from the decrease of flow on all links other than 2.

The proof of the second key lemma, Lemma 10, requires us to prove (modulo some caveats) that if we do not have enough flow to influence flow on a good link, then the quality of the Stackelberg equilibrium cannot be better than the Nash equilibrium. A strategy that controls less flow than the Nash flow on a good link can control flow on various bad links in whatever way it wishes: say it controls flow on links 1 and 3 in Figure 1, or it can increase flow on link 1 (a bad link) beyond f_1 resulting in its marginal latency being even farther from L^* than $l_1^*(f_1)$ originally was (increasing its deviation from L^*). It can also control flow on a good link as long as it does not control more than the Nash flow on that good link; for example, the strategy can possibly control less than f_2 flow on link 2 (in Figure 1). Any of these strategies cannot improve the quality of the Nash equilibrium.

To prove the second key lemma, we start from a Stackelberg strategy that induces the Nash equilibrium as its Stackelberg equilibrium, and gradually modify it to the given strategy. We show that as long as we do not have enough flow to influence flow on a good link, the rate of change in the social cost throughout the above said modification is nonnegative. Then Lemma 6 guarantees that the cost of the equilibrium induced by the given strategy must be no better than that of the Nash.

The proofs of Lemmas 9 and 10 are mainly built on the intuition above. We proceed to state some lemmas which will be helpful in the proof of two key lemmas.

4. Some useful lemmas

In this section, we state some lemmas which we need for the proofs of Lemma 9 and 10. Most of the proofs are omitted due to lack of space.

LEMMA 12. *If $c(G, r, l, 0) > c(G, r, l, 1)$, then (i) the latency of the constant link is positive, (ii) $L > 0$.*

OBSERVATION 13. *Suppose a Nash equilibrium f has common latency L and an optimum solution g has common marginal latency L^*. Then $L \le L^*$.*

LEMMA 14. (Uniqueness of Nash equilibrium and optimum solution). *If $c(G, r, l, 0) > c(G, r, l, 1)$, then there is a unique Nash equilibrium, and a unique optimum solution.*

The following lemma relates the amount of Nash flow on the constant link to amount of optimum flow on it.

LEMMA 15. *Suppose $c(G, r, l, 0) > c(G, r, l, 1)$. Then for the constant link z, if $f_z > 0$, then it must be the case that $f_z < g_z$. Also, if $f_z = 0$ then $f_z \le g_z$.*

LEMMA 16. *For the Nash flow f with common latency L and the optimum flow g with common marginal latency L^*, using the notation $l_j^*(f_j) = L^* + \varepsilon_j^*$, ($z$ is the constant link)*

$$(f_z - g_z) + \sum_{j:j\neq z, f_j>0} \frac{\varepsilon_j^*}{2a_j} = \sum_{j:j\neq z, L\le b_j \le L^*} \frac{L^* - b_j}{2a_j}.$$

PROOF. Let $P = \{j : j \neq z, f_j > 0, g_j > 0\}$ and $N = \{j : j \neq z, f_j = 0, g_j > 0\}$. We have $g_z + \sum_{j:j\neq z, g_j>0} g_j = g_z + \sum_{j\in P} g_j + \sum_{j\in N} g_j = r$.

For each $j \in P$, $2a_j f_j + b_j = L^* + \varepsilon_j^*$ (from the definition of ε_j^*) and $2a_j g_j + b_j = L^*$ (because $g_j > 0$). Subtracting the second equality from the first one and rearranging the terms, we get for all $j \in P$, $g_j = f_j - \frac{\varepsilon_j^*}{2a_j}$.

A necessary and sufficient condition for $j \in N$ is $j \neq z$ and $L \le b_j \le L^*$. (Actually, it should be $L \le b_j < L^*$, but the case $b_j = L^*$ does not change the equation.) For these links, $g_j = \frac{L^* - b_j}{2a_j}$. Using these expressions for all j, we get

$$g_z + \sum_{\substack{j:j\neq z,\\ f_j>0}} \left(f_j - \frac{\varepsilon_j^*}{2a_j}\right) + \sum_{\substack{j:j\neq z,\\ L\le b_j\le L^*}} \frac{L^* - b_j}{2a_j} = f_z + \sum_{\substack{j:j\neq z,\\ f_j>0}} f_j,$$

since both sides are equal to r. Rearranging the terms and subtracting $\sum_{j:j\neq z, f_j>0} f_j$ from both sides gives the result.

We have changed the index set in the first summation on the left hand side from $\{j : j \neq z, f_j > 0, g_j > 0\}$ to $\{j : j \neq z, f_j > 0\}$ above because $\{j : j \neq z, f_j > 0, g_j > 0\} = \{j : j \neq z, f_j > 0\}$. The $\subseteq$ direction is easy. For the $\supseteq$ direction, notice that for $j \neq z$, $f_j > 0 \implies b_j < L \implies b_j < L^* \implies g_j > 0$. ☐

The following lemma relates the amounts of flow on links of a network in two different Nash equilibria routing different amounts of total flow.

LEMMA 17. *Let f^r be a Nash flow with common latency L^r for the network (G, r, l) and let $q < r$. Then there exists a Nash flow f^q with common latency L^q for the network (G, q, l) with the property that $f_j^q \le f_j^r$ for all $j \in [k]$ and $L^q \le L^r$.*

5. Relating Stackelberg equilibrium to Nash equilibrium

In this section, we describe a continuous time process, called the *Stackelberg process P*, which does the following. Given an instance (G, r, l) and a Stackelberg strategy $\bar{h}$ controlling $\beta(\bar{h}) = \sum_{j\in[k]} \bar{h}_j / r = (\bar{h}_{[k]})/r$ fraction of flow, it starts at time $t = 0$ with a Nash equilibrium f of the instance (G, r, l) and ends at time $t = 1$ with an induced Stackelberg equilibrium for the instance (G, r, l, β) with Stackelberg strategy $\bar{h}$. We use the notation $f_A = \sum_{i\in A} f_i$.

5.1 The Stackelberg process

The process P is a continuous time process that transforms a Nash equilibrium f into a particular induced Stackelberg equilibrium for strategy $\bar{h}$. The time varies from $t = 0$

to $t = 1$ and the state of the process at time t is denoted by P^t. For any time $t \in [0, 1]$, P^t carries the information about the Stackelberg strategy $\bar{h}^t$ at time t, an induced Stackelberg equilibrium h^t for the strategy $\bar{h}^t$, and the fraction of flow centrally controlled by the strategy $\bar{h}^t$, which we call $\beta(\bar{h}^t)$. The common latency of the selfish flow in h^t is denoted by L_{h^t}; that is, if $h_j^t > 0$ then $l_j(\bar{h}_j^t + h_j^t) = L_{h^t}$. We now give the details of the process.

We start with a Nash equilibrium f for the network (G, r, l) (choose an arbitrary one if there are many choices). Recall the definitions of $S(\bar{h})$ and $U(\bar{h})$ from Definition 8; $S(\bar{h}) = \{j : \bar{h}_j > f_j\}$ and $U(\bar{h}) = \{j : 0 < \bar{h}_j \le f_j\}$. When $\bar{h}$ is clear from the context, we will call them S and U respectively for brevity. We first give the description of P^0, the state of the process in the beginning. $\bar{h}^0$ is defined as: $\bar{h}_j^0 = \min\{\bar{h}_j, f_j\}$ for $j \in [k]$, which gives rise to $\beta(\bar{h}^0) = \sum_{j=1}^k \bar{h}_j^0 / r = \sum_{j=1}^k \min\{\bar{h}_j, f_j\}/r$. We define $h^0 = f - \bar{h}^0$, which is easily seen to be a Stackelberg equilibrium for $\bar{h}^0$.

At time $t \in (0, 1]$, the state P^t of the process P has the following specification. $\bar{h}^t$ is described as

$$\bar{h}_j^t = \min\{\bar{h}_j, f_j\} + t\left(\bar{h}_j - \min\{\bar{h}_j, f_j\}\right) = \bar{h}_j^0 + t\left(\bar{h}_j - \bar{h}_j^0\right),$$

which gives rise to $\beta(\bar{h}^t) = \sum_{j=1}^k \bar{h}_j^t / r$. It is clear that for an edge j in $S(\bar{h})$, P monotonically increases flow on it at a constant rate of $(\bar{h}_j - f_j)$, while for an edge j in $U(\bar{h})$, it keeps the flow constant on j, maintaining at the constant amount $\bar{h}_j > 0$. We have that

$$\bar{h}_j^t = f_j + t(\bar{h}_j - f_j) \text{ for } j \in S(\bar{h}); \quad \bar{h}_j^t = \bar{h}_j \text{ for } j \in U(\bar{h}).$$

Given h^s for all $s < t$, we will now see how to find h^t (if h^t is not unique, we will find a particular one which will suffice for our purpose). It is clear that h^s is a Nash equilibrium in the network $(G, r - \sum_{j \in [k]} \bar{h}_j^s, \{l_j(\bar{h}_j^s + x)\}_{j=1}^k)$ and h^t is a Nash equilibrium in the network $(G, r - \sum_{j \in [k]} \bar{h}_j^t, \{l_j(\bar{h}_j^t + x)\}_{j=1}^k)$. Since $\bar{h}_j^s \le \bar{h}_j^t$ for $s \le t$, we have that latencies (due to $\bar{h}^t$) at time t are at least as large as the latencies (due to $\bar{h}^s$) at time s and the total selfish flow routed in h^t is less than or equal to the total selfish flow in h^s. Also, h^s restricted to links in $[k] - S(\bar{h})$ with latency functions $l_j(\bar{h}_j^s + x_j)$ is a Nash equilibrium, and so is h^t with latency functions $l_j(\bar{h}_j^t + x_j)$ restricted to links in $[k] - S(\bar{h})$. Since $\bar{h}_j^s = \bar{h}_j^t$ for $j \in [k] - S(\bar{h})$, it follows from Lemma 17 that there exists a Stackelberg equilibrium h^t for $\bar{h}^t$ such that $h_j^s \ge h_j^t$ for all $j \in [k] - S(\bar{h})$. This Stackelberg equilibrium (resulting from Lemma 17) is what we call h^t. Therefore, we have the following for all $s \le t$

$$\bar{h}_j^s \le \bar{h}_j^t \text{ for } j \in S(\bar{h}) \text{ and } h_j^s \ge h_j^t \text{ for } j \in [k] - S(\bar{h}). \quad (3)$$

This can alternatively be viewed as the following: in moving from time s to $t > s$, the amount of altruistic flow increases on links in $S(\bar{h})$ by an infinitesimal amount $\varepsilon = (t - s)(\bar{h}_{S(\bar{h})} - f_{S(\bar{h})}) > 0$, resulting in the decrease of the selfish flow (by exactly the same amount ε). The amount of selfish flow on links j with $h_j^s > 0$ decreases to $h_j^t \le h_j^s$ so as to keep the latency on all links with $h_j^t > 0$ equal. This finishes the description of the Stackelberg process.

We wish to remark that the exact rule for how we increase flow on links in S and U from time 0 to time 1 is immaterial. Another rule for increasing the altruistic flow will be equally valid as long as it increases flow on links in S monotonically (strictly increasing fashion), it gives rise to a Nash equilibrium at time $t = 0$ and a Stackelberg equilibrium corresponding to $\bar{h}$ at time $t = 1$. The rule that we give is easy to describe, and we work with it in the rest of the paper.

To recap the process P, it starts at time $t = 0$ controlling $f_S + \bar{h}_U$ amount of flow (this is strategy $\bar{h}^0$), giving rise to a particular Stackelberg equilibrium $h^0 = f - \bar{h}_0$. The process continuously increases the altruistic flow on links in S linearly with time, and the (selfish) flow on all other links responds. For links outside the set S, if there is a positive selfish flow on a link, it decreases by a small amount to respond to the deficit of selfish flow on links outside of S, else the only altruistic flow on the link remains at the same level. This process continues until it makes altruistic flow amount equal to the amount in the Stackelberg strategy, that is $\bar{h}_S + \bar{h}_U$ at time $t = 1$.

We claim that this process gives rise to a valid Stackelberg equilibrium for strategy $\bar{h}$ at time 1. It is easy to see that all selfish flow is on shortest latency paths (we have chosen the Stackelberg equilibrium guaranteed by Lemma 17) and the amount of altruistic flow is also respected, so $h \overset{\text{def}}{=} h^1$ is a valid Stackelberg equilibrium. There might be other equilibria induced by strategy $\bar{h}$ than the one the process finds, but from Lemma 6 the cost of all equilibria is the same.

We denote by Z^t the set of links which have zero h^t-flow at time t. Z^0 is of course $\{j : j \in S(\bar{h})\} \cup \{j : f_j = \bar{h}_j\} = \{j : f_j < \bar{h}_j\} \cup \{j : f_j = \bar{h}_j\}$. In particular, $S(\bar{h}) \subseteq Z^0$. As the time increases from $t = 0$ to $t = 1$, the selfish flow on links in $[k] - Z^0$ decreases, and the link j is added to Z^t (and to all $Z^{t'}$ for all $t' \ge t$) at the smallest time instance t such that $h_j^t = 0$ (but $h_j^s > 0$ for all $s < t$). Note that $t \mapsto Z^t$ is a monotone set function.

We give the formal definition of Z^t now. For each link j, define $t_j = \inf\{t : h_j^t = 0\}$ if the set is nonempty, and equal to 2 otherwise. We claim that $h_j^{t_j} = 0$ if $t_j \le 1$. This is because the segment $\{t : h_j^t > 0\}$ is open on the upper boundary, hence its complement is closed on the lower boundary (hence the limit point on the lower boundary is contained in the set).

Definition 18. *The set Z^t is defined to be the set of links with zero amount of selfish flow at time $t \in [0, 1]$. Formally, $Z^t = \{j \in [k] : t_j \le t\} = \{j \in [k] : h_j^t = 0\}$.*

6. Proof of Lemma 9

Proof of Lemma 9: We first consider the case when $\beta = \frac{f_* + \varepsilon}{r}$ for infinitesimally small $\varepsilon > 0$. We will extend the same idea to larger values of β at the end of the proof.

We will prove a slightly more general following claim: for any good link i, the Stackelberg strategy of routing $f_i + \varepsilon$ flow on link i induces a Stackelberg strategy of strictly smaller cost than the cost of the Nash equilibrium. Taking this link to be the link i_* will prove the lemma.

Let i be a good link (a link with $f_i < g_i$). We consider the Stackelberg strategy $\bar{h}$ with $\bar{h}_i = f_i + \varepsilon$ for small $\varepsilon > 0$ and $\bar{h}_j = 0$ for $j \ne i$. Clearly, $S(\bar{h}) = \{i\}$ and $U(\bar{h}) = \emptyset$. The idea is to first determine the rate of increase of the social cost as the Stackelberg strategy increases flow on link i (we call it *the rate of increase* for brevity), then to determine the rate of decrease of the social cost because of the decrease of flow on links other than i (we call it *the rate of decrease*), and then show that the rate of decrease is strictly more than the rate of increase.

We consider the Stackelberg process described in Section 5. In the process, the selfish flow does not increase on any link in $[k] - \{i\}$ when we increase the altruistic flow on link i. Indeed, the amount of selfish flow decreases on links $\{j : j \neq i, f_j > 0\}$ as we increase the amount of altruistic flow on link i. Note that the set $[k] - Z^0$ as defined in the description of the process P in Section 5 is equal to $\{j : j \neq i, f_j > 0\}$.

The rate of increase of the social cost with increase of flow on link i is $l_i^*(f_i) = 2a_i f_i + b_i = L^* + \varepsilon_i^*$.

On the other hand, the rate of decrease of social cost with decrease of flow on links in $[k] - Z^0$ depends on whether $z \in [k] - Z^0$. We consider these cases separately.

Case 1: $z \in [k] - Z^0$**.** If $z \in [k] - Z^0$, then the constant link carries a non-zero amount of flow and the decrease of flow occurs only on the constant link. This gives rise to the rate of decrease of social cost equal to $l_z^*(f_z) = b_z = L = L^* + \varepsilon_z^*$. We have that $L^* \leq b_z = L \leq L^*$, or $L^* = b_z$ or $\varepsilon_z^* = 0$. Therefore, the rate of decrease becomes L^*.

Case 2: $z \in Z^0$**.** Alternatively, if $z \in Z^0$, then the flow decreases on non-constant links. The decrease of flow on links is such that the decreases in latency on all of them is equal, say l. Since the decrease in latency (on link $j \neq i$ with $f_j > 0$) is l, the decrease in amount of flow is l/a_j. Therefore the total decrease in the social cost is $\sum_{j \neq i, f_j > 0} \frac{l}{a_j}(2a_j f_j + b_j)$ since the rate of decrease of social cost on link j is $l_j^*(f_j) = 2a_j f_j + b_j$. Moreover, the total decrease in flow on links $\{j : f_j > 0, j \neq i\}$ is equal to the total increase of flow on link i, that is ε. Therefore, $\sum_{j:j\neq i, f_j > 0} \frac{l}{a_j} = \varepsilon$ or equivalently, $l = \varepsilon / (\sum_{j:j \neq i, f_j > 0} 1/a_j)$. The rate of decrease of the social cost is total decrease in the social cost divided by ε, which is (by using the fact $2a_j f_j + b_j = L^* + \varepsilon_j^*$)

$$\sum_{j:j\neq i, f_j > 0} \frac{1}{\sum_{j:j\neq i, f_j>0} 1/a_j} \frac{L^* + \varepsilon_j^*}{a_j}.$$

We now prove that the rate of increase is strictly less than the rate of decrease of the social cost.

Case 1: $z \in [k] - Z^0 = \{j : j \neq i, f_j > 0\}$**.** In this case, we need to prove that $L^* + \varepsilon_i^* < L^*$ or $\varepsilon_i^* < 0$. This is true since $2a_i f_i + b_i = L^* + \varepsilon_i^*$ and $2a_i g_i + b_i = L^*$ with $f_i < g_i$ (i is a good link). This gives $\varepsilon_i^* = 2a_i(f_i - g_i) < 0$. The inequality follows.

Case 2: $z \notin [k] - Z^0 = \{j : j \neq i, f_j > 0\}$**.** In this case, we need to prove that

$$L^* + \varepsilon_i^* < \sum_{j:j\neq i, f_j > 0} \frac{1}{\sum_{j:j\neq i, f_j>0} 1/a_j} \frac{L^* + \varepsilon_j^*}{a_j}.$$

After some manipulation, we are left to prove $\sum_{j:f_j>0, j\neq i} \frac{\varepsilon_i^*}{2a_j} < \sum_{j:f_j>0, j\neq i} \frac{\varepsilon_j^*}{2a_j}$. If $i = z$ (other case will be considered later), then it reduces to proving (using Lemma 16)

$$\sum_{j:f_j>0, j\neq i} \frac{\varepsilon_i^*}{2a_j} < (g_z - f_z) + \sum_{j:L \leq b_j \leq L^*, j \neq z} \frac{L^* - b_j}{2a_j}.$$

In this case the left hand side is 0. This follows from the fact that $i = z$ is a good link, $f_z < g_z$, and $L \leq b_z = L^*$ which implies $\varepsilon_i^* = \varepsilon_z^* = 0$. The right hand side is strictly positive since $f_z < g_z$. The inequality hence follows in this case. On the other hand, if $i \neq z$, then we are left to prove (again using Lemma 16)

$$\sum_{j:f_j>0, j\neq i} \frac{\varepsilon_i^*}{2a_j} < (g_z - f_z) + \frac{-\varepsilon_i^*}{2a_i} + \sum_{j:L \leq b_j \leq L^*, j \neq z} \frac{L^* - b_j}{2a_j}.$$

In this case, we have $f_i < g_i$. We also have $2a_i f_i + b_i = L^* + \varepsilon_i^*$ and $2a_i g_i + b_i = L^*$. This gives $\varepsilon_i^* = 2a_i(f_i - g_i) < 0$. The left hand side is therefore strictly negative. On the right hand side, the first term in nonnegative from Lemma 15, the second term is positive, and the last term is again nonnegative. Therefore, the right hand side is strictly positive, proving that the inequality holds in this case too. This finishes the proof the the lemma for the case when $\beta = \frac{f_* + \varepsilon}{r}$ for sufficiently small $\varepsilon > 0$.

Let $\beta \in (f_*/r, 1]$ now. We find a small enough $\varepsilon_0 > 0$ (with $f_* + \varepsilon_0 \leq \beta r$) from the previous part such that we can find a Stackelberg strategy $\bar{h}$ that induces a Stackelberg equilibrium h with cost strictly less than the cost of the Nash equilibrium. To get a strategy for βr altruistic flow, the intuition is the following. We route $f_* + \varepsilon_0$ amount of flow on edge i_* and pretend that the rest of the flow is selfish and let it route on minimum latency paths. At the end, we declare some of the selfish flow (exactly $\beta r - (f_* + \varepsilon_0)$ amount) altruistic and output the corresponding strategy for βr flow. The old induced Stackelberg equilibrium h is still a Stackelberg equilibrium if we remove from h the flow that we declared altruistic in the end. This intuition can be turned into a proof in a straightforward manner, which is omitted due to lack of space. The lemma then follows. $\square$

7. Proof of Lemma 10

We want to prove that if the amount of flow controlled by a Stackelberg strategy is at most f_*, then the resulting Stackelberg equilibrium cannot be cheaper than the Nash solution. Let us fix a Stackelberg strategy $\bar{h}$ such that the total flow controlled centrally by $\bar{h}$ is at most f_*. We then derive from this assumption (and the assumption that $\{j : f_j = 0, g_j > 0\} = \emptyset$ which is part of Lemma 10) that the cost of the resulting Stackelberg equilibrium h has cost at least as much as the cost of the Nash equilibrium, that is $c(G, r, l, \beta, \bar{h}) \geq c(G, r, l, 0)$. Recall the definitions of $S(\bar{h})$ and $U(\bar{h})$ from Definition 8; $S(\bar{h}) = \{j : \bar{h}_j > f_j\}$ and $U(\bar{h}) = \{j : 0 < \bar{h}_j \leq f_j\}$. When $\bar{h}$ is clear from the context, we call them S and U respectively. Let $S = \{s_1, s_2, \ldots, s_{|S|}\}$ and $U = \{u_1, u_2, \ldots, u_{|U|}\}$ for this particular $\bar{h}$.

7.1 The plan for proving Lemma 10

We wish to show that the cost of $\bar{h} + h$ is at least as much as the cost of the Nash equilibrium. We will use the Stackelberg process defined in Section 5 in the following way. At time $t = 0$, the cost of the Nash equilibrium is equal to the cost of the Stackelberg equilibrium $(\bar{h}^0 + h^0)$. For any arbitrary time instance $t \in [0, 1)$, when the process increases an infinitesimal amount of flow on links in S, the flow on other links decreases. There is an increment in the social cost because of increase of flow on links in S and there is some decrease in social cost because of decrease of flow on links not in S (on links in $[k] - Z^t$ to be precise). We determine the rate of increase of social cost with increase of flow on links in S and also determine the rate of decrease of social cost with decrease of flow on links in $[k] - Z^t$. We compare this rate of increase with the rate of decrease and prove that the rate of increase is no less than the rate of decrease for all $t \in [0, 1)$.

This proves that the cost of $(h^t + \bar{h}^t)$ is nondecreasing with t and that $c(h + \bar{h}) = c(h^t + \bar{h}^t)|_{t=1} \geq c(h^t + \bar{h}^t)|_{t=0} = c(f)$ where $c(\cdot)$ denotes the cost of the flow. The uniqueness of the costs of Nash and Stackelberg equilibrium in Lemmas 5 and 6 will then establish Lemma 10.

7.2 The rate of increase of the social cost

At time $t \in [0, 1)$, if the increase in the (centrally controlled) flow amount on links in S is $\varepsilon > 0$, let it be divided among links in S in the ratio $\sigma_1 : \sigma_2 : \cdots : \sigma_{|S|}$ with $\sigma_1 + \cdots + \sigma_{|S|} = 1$. (According to the rules described in the description of the process, $\sigma_j = (\bar{h}_j - f_j)/(\sum_{j \in S}(\bar{h}_j - f_j))$ for $j \in S$.) On link $s_i \in S$, the rate of increase of social cost with increase in flow on this link is $l^*_{s_i}((h^t + \bar{h}^t)_{s_i}) = 2a_{s_i}\bar{h}^t_{s_i} + b_{s_i}$ since $h_{s_i} = 0$ for all $s_i \in S$. Therefore the rate of increase of social cost is $\varepsilon\big(\sigma_1(2a_{s_1}\bar{h}^t_{s_1} + b_{s_1}) + \cdots + \sigma_{|S|}(2a_{s_{|S|}}\bar{h}^t_{s_{|S|}} + b_{s_{|S|}})\big)/\varepsilon = \sum_{s \in S}\sigma_s(2a_s\bar{h}^t_s + b_s)$. This can be lower bounded as

$$\begin{aligned}\sum_{s \in S}\sigma_s(2a_s\bar{h}^t_s + b_s) &\geq \sum_{s \in S}\sigma_s(2a_s\bar{h}^0_s + b_s) \\ &= \sum_{s \in S}\sigma_s(L^* + \varepsilon^*_s) = L^* + \sum_{s \in S}\sigma_s\varepsilon^*_s. \qquad (4)\end{aligned}$$

Here we have used the fact that $\bar{h}^0_s = f_s$ for $s \in S$. The inequality above follows since $\bar{h}^t_j \geq \bar{h}^0_j$ for $j \in S$ and $t \geq 0$.

7.3 The rate of decrease of the social cost

With increase of altruistic flow on links in S, the selfish flow on other links responds by decreasing on links in $[k] - Z^t$ to keep the latencies on all links with positive selfish flow the same (see Definition 18 for the definition of Z^t). The flow decreases precisely on links on which the selfish flow is already positive (as a result those links are not in Z^t). Formally, all links in $\{j : h^t_j > 0\}$ contribute to the decrease of flow at time t. Note that $\{j : h^t_j > 0\} = [k] - Z^t$. We will use the set $\{j : h^t_j > 0, j \in [k] - Z^t\}$ as the set of links on which the flow decreases at time t; we do not need both the conditions ($h^t_j > 0$ and $j \in [k] - Z^t$) since they are equivalent, but we use them to facilitate the proofs later. Recall that $S \subseteq Z^t$ for $t \in [0, 1]$.

We now consider in the following case analysis several cases depending on which links participate in the decrement of selfish flow.

Case 1: $L_{h^t} > 0$ and $z \in \{j : h^t_j > 0, j \in [k] - Z^t\}$. In this case, there is some positive flow on the constant link and the flow decrease occurs only on the constant link. In going from time t to a time infinitesimally greater than t, the total amount of selfish flow decreases, say be amount δ and that decreases occurs on the links in $\{j : h^t_j > 0, j \in [k] - Z^t\}$. If the constant link z is contained in this set, then δ amount of flow decreases on z to keep the latencies on all links with positive selfish flow at the equal value. Therefore, in this case, the rate of decrease becomes equal to $\varepsilon \cdot l^*_z(\bar{h}^t_z + h^t_z)/\varepsilon = \varepsilon \cdot b_z/\varepsilon = L^* + \varepsilon^*_z$.

Case 2: $L_{h^t} > 0$ and $z \notin \{j : h^t_j > 0, j \in [k] - Z^t\}$. In this case, the flow does not decrease on the constant link: it decreases collectively on linear links. The rate of decrease is (similar to the proof of Lemma 9)

$$\begin{aligned}&\frac{1}{\varepsilon} \cdot \frac{\varepsilon}{\displaystyle\sum_{j:h^t_j > 0, j \in [k] - Z^t} 1/a_j} \sum_{j:h^t_j > 0, j \in [k] - Z^t} \frac{2a_j(h^t + \bar{h}^t)_j + b_j}{a_j} \\ &\leq \frac{1}{\sum_{j:h^t_j > 0, j \in [k] - Z^t} 1/a_j} \sum_{j:f_j > 0, j \in [k] - Z^t} \frac{2a_j f_j + b_j}{a_j} \\ &= \frac{1}{\sum_{j:f_j > 0, j \in [k] - Z^t} 1/a_j} \sum_{j:f_j > 0, j \in [k] - Z^t} \frac{L^* + \varepsilon^*_j}{a_j}.\end{aligned}$$

The first inequality follows from the next lemma which states that the two index sets in the summation are the same and the fact that $(h^t + \bar{h}^t)_j \leq f_j$ for $j \in [k] - Z^t \subseteq [k] - S$. The second equality also follows because the two index sets in the summation are the same.

LEMMA 19. *Let $\beta(\bar{h})r = \bar{h}_S + \bar{h}_U \leq f_*$. Then $\{j : h^t_j > 0, j \in [k] - Z^t\} = \{j : f_j > 0, j \in [k] - Z^t\}$ for all $t \in [0, 1)$.*

Our next step in the proof is to show that the rate of increase of social cost is at least as large as the rate of its decrease (in all cases considered above). The next section will prove some facts which will be useful.

7.4 Some sufficient conditions

We state some facts about the network below.

LEMMA 20. *Let f be the the Nash equilibrium in the network and g be the optimum solution. Let $\bar{h}$ be a Stackelberg strategy with $\beta(\bar{h})r = \bar{h}_{[k]} \leq f_*$. Then*

1. *All links in S are bad links.*
2. *$\varepsilon^*_j \geq 0$ for all $j \in S$.*
3. *$f_z > 0$ implies $\varepsilon^*_z = 0$, and $f_z = 0$ implies $\varepsilon^*_z \geq 0$.*
4. *If j is such that $f_j > 0$ and $j \in Z^t$ for $t \in [0, 1)$, then j is a bad link.*
5. *For all $t \in [0, 1)$, and $j \in Z^t$, $\varepsilon^*_j \geq 0$.*

PROOF. We will prove the first claim by contradiction. Assume the contrary: j is a good link and $j \in S$. We have $\bar{h}_{[k]} \geq \bar{h}_S \geq \bar{h}_j > f_j \geq f_*$ where the second to the last inequality follows from the definition of S and the last one follows from the definition of f_* (f_* is the minimum Nash flow on any good link). This is a contradiction to the assumption $\bar{h}_{[k]} \leq f_*$. Therefore, links in S are bad links.

For the second claim, let us consider the case $j \in S$ and $j \neq z$. If $f_j = 0$, then $g_j = 0$ from the conditions in Lemma 10. We have $g_j = 0 \iff b_j \geq L^* \iff L^* + \varepsilon^*_j \geq L^* \iff \varepsilon^*_j \geq 0$. If $f_j > 0$ on the other hand, then $f_j > 0 \iff b_j < L \implies b_j < L^* \iff g_j > 0$. We have $2a_j f_j + b_j = L^* + \varepsilon^*_j$ from the definition of ε^*_j and $2a_j g_j + b_j = L^*$ from $g_j > 0$. Subtracting the second one from the first one we get $\varepsilon^*_j = 2a_j(f_j - g_j) \geq 0$ since $a_j > 0$ and $f_j \geq g_j$ (j is a bad link). For the case $j \in S$ and $j = z$, $\varepsilon^*_z \geq 0$ from Observation 4.

For the third claim, $f_z > 0$ implies $b_z = L \leq L^* \leq b_z$ (the last inequality follows from Observation 4). Therefore, $b_z = L^*$ which is equivalent to $\varepsilon^*_z = 0$. The second part follows from Observation 4.

We will prove the fourth claim by contradiction; assume j is a good link. We will now show that the amount of flow controlled by strategy $\bar{h}^t$ is at least f_*, a contradiction since the amount of flow controlled by $\bar{h}^1$ is at most f_* and the centrally controlled flow is monotonically increasing with t, so the amount of flow controlled by $\bar{h}^t$ for $t < 1$ must be strictly less than f_*. (We are assuming that the amount of flow controlled by $\bar{h}^t$ is strictly increasing in t, since otherwise $S = \emptyset$ and the Stackelberg equilibrium is the same as the Nash equilibrium.)

If $j \in S$, then we are done from previous part which proves that all links in S are bad links. So we assume that $j \notin S$.

We have $h_j^0 = f_j - \bar{h}_j^0 \geq 0$ since $j \notin S$. Also $h_j^t = 0$ since $j \in Z^t$. Since the amount of selfish flow is nonincreasing on all links, its amount has decreased by at least $h_j^0 - h_j^t = f_j - \bar{h}_j^0 = f_j - \bar{h}_j$ from time 0 to t. By the conservation of total flow, the amount of altruistic flow has increased by at least this amount from time 0 to t. Therefore $\beta(\bar{h}^t)r \geq \beta(\bar{h}^0)r + (f_j - \bar{h}_j) \geq \bar{h}_j + (f_j - \bar{h}_j) = f_j \geq f_*$ where the last inequality follows by the definition of f_* and our assumption that j is a good link. This is a contradiction since $\beta(\bar{h}^t)r$ must be less than f_* for $t < 1$. Hence j is a bad link.

For the fifth claim, recall the definition of Z^t; $Z^t = \{j : h_j^t = 0\}$. For $t \in [0, 1)$ let us fix an arbitrary j in Z^t. We will prove that for this j, $\varepsilon_j^* \geq 0$. If $j = z$, then Observation 4 shows that $\varepsilon_z^* \geq 0$. If j is a nonconstant link on the other hand, then we consider two cases. If $f_j = 0$, then $g_j = 0$ from the condition of Lemma 10. $g_j = 0 \implies b_j \geq L^* \iff L^* + \varepsilon_j^* \geq L^* \iff \varepsilon_j^* \geq 0$. If $f_j > 0$, then we claim some properties of the link j:

1. $g_j > 0$. This is because $f_j > 0 \implies b_j < L \implies b_j < L^* \implies g_j > 0$.
2. $j \notin S$. If $j \in S$, then Part 2 proves that $\varepsilon_j^* \geq 0$ and we are done.
3. j is a bad link. This follows from previous part (Part 4) and the fact that $f_j > 0$ and $j \in Z^t$.

With these properties of the link j, we have $2a_j f_j + b_j = L^* + \varepsilon_j^*$ and $2a_j g_j + b_j = L^*$ since $g_j > 0$. Subtracting the second inequality from the first one, we get $\varepsilon_j^* = 2a_j(f_j - g_j) \geq 0$ since $a_j > 0$ and j is a bad link. □

We now want to prove that the rate of increase is at least the rate of decrease for all cases. We will derive some sufficient conditions for this and then show that the sufficient conditions hold. The rate of increase is always given by the same expression in (4), but rate of decrease varies depending on which links are involved in the decrement. We will consider these cases separately.

Case 1: $L_{h^t} > 0$ and $z \in \{j : h_j^t > 0, j \in [k] - Z^t\}$. In this case, the inequality to be proven becomes $L^* + \sigma_1 \varepsilon_{s_1}^* + \cdots + \sigma_s \varepsilon_{s_{|S|}}^* \geq L^* + \varepsilon_z^*$. A sufficient condition for this inequality to hold is $\varepsilon_j^* \geq 0, \forall j \in S$ and $\varepsilon_z^* = 0$. These conditions also hold directly from Lemma 20 (Part 2 and 3), and the fact that $f_z > 0$ since $z \in \{j : h_j^t > 0, j \in [k] - Z^t\}$ and $\{j : h_j^t > 0, j \in [k] - Z^t\} = \{j : f_j > 0, j \in [k] - Z^t\}$ from Lemma 19.

Case 2: $L_{h^t} > 0$ and $z \notin \{j : h_j^t > 0, j \in [k] - Z^t\}$. In this case, either $h_z^t = 0$ or $z \in Z^t$. The inequality to be proven becomes

$$\sum_{j:f_j>0, j\in[k]-Z^t} \frac{L^* + \sigma_1 \varepsilon_{s_1}^* + \cdots + \sigma_s \varepsilon_{s_{|S|}}^*}{a_j} \geq \sum_{j:f_j>0, j\in[k]-Z^t} \frac{L^* + \varepsilon_j^*}{a_j}.$$

Canceling the L^* terms, dividing by 2 and using Lemma 16, this is equivalent to proving

$$\sum_{\substack{j:f_j>0,\\ j\in[k]-Z^t}} \frac{\sigma_1 \varepsilon_{s_1}^* + \cdots + \sigma_s \varepsilon_{s_{|S|}}^*}{2a_j} \geq \sum_{\substack{j:f_j>0,\\ j\neq z, j\in Z^t}} \frac{-\varepsilon_j^*}{2a_j} + (g_z - f_z).$$

The $\sum_{j:j\neq z, L\leq b_j \leq L^*} \frac{L^* - b_j}{2a_j}$ term equals zero since $\{j : f_j = 0, g_j > 0\} = \emptyset$. A sufficient condition for the above inequality to hold is the following:

$$\varepsilon_j^* \geq 0, \forall j \in S \text{ and } \varepsilon_j^* \geq 0, \forall j \in Z^t - \{z\} \text{ and } f_z = g_z. \quad (5)$$

We show that these conditions also hold. The first two conditions hold from Lemma 20 (Part 2 and 5). For the last condition, note that $z \notin \{j : h_j^t > 0, j \in [k] - Z^t\} = \{j : f_j > 0, j \in [k] - Z^t\}$, so either $f_z = 0$ or $z \in Z^t$.

If $f_z = 0$, we have $f_z = g_z = 0$ from the assumption that there does not exist any link with zero Nash flow and positive optimum flow. All the conditions of (5) are satisfied.

If $f_z > 0$ and $z \in Z^t$, then z must be a bad link (see Lemma 20 Part 4). It follows that $f_z \geq g_z$. But from Lemma 15, $f_z > 0$ implies that $f_z < g_z$. Both these conditions ($f_z \geq g_z$ and $f_z < g_z$) cannot hold simultaneously, so, this case cannot happen. This proves the entire claim.

To recap the proof, we started with the Nash equilibrium and transformed it, using the continuous time process P, into a Stackelberg equilibrium. We proved that all through this modification, the net rate of increase of the social cost was nonnegative (all the cases above dealt with various possibilities for this). It hence follows that the social cost of the resulting end product (which is the Stackelberg equilibrium) is no less than the cost of the Nash equilibrium. The result of Lemma 10 then follows.

8. The price of optimum

In this section, we give an alternate (and simpler) proof of a result from Kaporis and Spirakis [6]. Kaporis and Spirakis consider a slightly different problem than we have considered here: they characterize the minimum fraction of flow that must be centrally controlled to induce the solution with cost of the optimum solution. This fraction is called β_M in [6]. We assume that the graph is a parallel links graph with *strictly increasing* and *differentiable* latency functions on the edges (as assumed in [6]). Note that the latency functions are not assumed to be linear any more. They could be arbitrary as long as they are increasing and differentiable. Because of the strictly increasing nature of the latency functions, there is a unique Nash equilibrium and a unique optimum solution in the network. This follows from the fact that costs of all Nash equilibria are same (see Lemma 5), and for strictly increasing functions each latency value has a unique flow associated with it. Functions $x_i l_i(x_i)$ are also assumed to be convex since that ensures that we can compute the optimum flow in polynomial time.

Here is the algorithm from [6] to compute β_M.

Algorithm: OpTop.

1. Let $\Gamma = 0$. Let $\bar{h}_j = 0$ for all j (this is the current Stackelberg strategy). Compute the optimum solution for the network, sending flow $g_i \geq 0$ on link i.
2. Compute the Stackelberg equilibrium h corresponding to strategy $\bar{h}$. Let $G = \{j : (\bar{h} + h)_j < g_j\}$ (set of good links).
3. If $G = \emptyset$, go to Step 4. Otherwise, set $\bar{h}_j \leftarrow g_j$ for all $j \in G$. Set $\Gamma \leftarrow \Gamma + \sum_{j\in G} g_j$. Go to Step 2.
4. Output Γ/r as the value of β_M.

Looking at the Stackelberg strategy from the perspective presented in our paper gives a short proof for the correctness of this algorithm. We proceed to present the proof now.

THEOREM 21. *Algorithm OpTop correctly computes β_M.*

PROOF. Let $\gamma = \Gamma/r$. We prove $\gamma = \beta_M$ by proving $\gamma \geq \beta_M$ and then $\gamma \leq \beta_M$. The first inequality is simple: if we control $\gamma r = \Gamma$ units of flow, the corresponding Stackelberg strategy $\bar{h}$ produced by the algorithm induces an equilibrium in which total flow on each link is equal to its flow in the optimum solution. This follows since if $(\bar{h}+h)_j > g_j$ on any link, it must be the case that $(\bar{h}+h)_j < g_j$ on some other link, and hence the algorithm would not have terminated. Thus controlling γr flow indeed can give rise to an optimum solution, proving $\gamma \geq \beta_M$.

For the other direction, order the links in the order that the flow $\bar{h}$ on them becomes positive (i.e., the algorithm OpTop controls positive flow on that link). The links controlled in the first iteration are ordered first (in arbitrary order), then the links controlled in the second iteration and so on. Uncontrolled links are placed at the end of the arrangement. Note that for all j, $\bar{h}_j \in \{g_j, 0\}$. We prove $\gamma \leq \beta_M$ by a contradiction; for the sake of contradiction, assume that there is a Stackelberg strategy $\bar{h}^*$ controlling strictly less than γ fraction of the flow that induces a Stackelberg equilibrium of optimal cost. Then there must be a link, say j_0, such that $\bar{h}_{j_0} > \bar{h}^*_{j_0}$. Choose such a j_0 which occurs earliest in the above order. Suppose the link j_0 was controlled in i_0-th iteration of the algorithm. That means that for all links controlled in up to $i_0 - 1$ iterations (call these links $S^{(i_0-1)}$), $\bar{h}^*$ controls at least as much flow as $\bar{h}$ ($\bar{h}^*_j \geq \bar{h}_j = g_j$ for all $j \in S^{(i_0-1)}$). We claim that $\bar{h}^*_j \leq g_j$ for all j, since otherwise, the resulting Stackelberg equilibrium will have flow more than g_j on link j, and by the uniqueness of the optimum solution, the induced equilibrium will not be the optimum solution. For links in $S^{(i_0-1)}$, we get $\bar{h}_j \leq \bar{h}^*_j$ and $\bar{h}_j \geq \bar{h}^*_j$. Therefore, for all links controlled by the algorithm up to the (i_0-1)-th iteration (that is, for links in $S^{(i_0-1)}$), $\bar{h}_j = \bar{h}^*_j$.

Let $\bar{h}^{(i_0-1)}$ be the Stackelberg strategy just after the (i_0-1)-th iteration of the algorithm and let $h^{(i_0-1)}$ be the corresponding equilibrium. Since link j_0 was controlled in the i_0-th iteration, we have $\bar{h}^{(i_0-1)}_j + h^{(i_0-1)}_j < g_{j_0}$. Let $\bar{h}^{*(i_0-1)}$ be same as $\bar{h}^*$ for links in $S^{(i_0-1)}$, and 0 on other links. We have $\bar{h}^{*(i_0-1)} = \bar{h}^{(i_0-1)}$; therefore, the corresponding Stackelberg equilibria are also same, that is $h^{*(i_0-1)} = h^{(i_0-1)}$. Therefore, $\bar{h}^{*(i_0-1)}_{j_0} + h^{*(i_0-1)}_{j_0} < g_{j_0}$. ($h^{*(i_0-1)}$ is the equilibrium corresponding to $\bar{h}^{*(i_0-1)}$.) Note that $\bar{h}^*_{j_0} < g_{j_0}$ too, since $\bar{h}^*_j < \bar{h}_{j_0} = g_{j_0}$. Consider the Stackelberg equilibrium corresponding to the Stackelberg strategy which controls $\bar{h}^*_j$ flow for $j \in S^{(i_0-1)} \cup \{j_0\}$ and 0 flow on other links. In this Stackelberg equilibrium, the total flow on link j_0 is at most $\max\{\bar{h}^{*(i_0-1)}_{j_0} + h^{*(i_0-1)}_{j_0}, \bar{h}^*_{j_0}\} < g_{j_0}$. Since the amount of centrally controlled flow on link j_0 is not going to increase as the Stackelberg strategy $\bar{h}^*$ controls more flow on links in $S^{(i_0-1)} \cup \{j_0\}$, the total amount of flow on link j_0 can only decrease (as stated in Section 5, see discussion around Equation (3)). Therefore, $(\bar{h}^* + h^*)_{j_0} < g_{j_0}$. From the uniqueness of the optimum solution, we see that $(\bar{h}^* + h^*)$ cannot be the optimum solution (since it differs from the optimum solution on link j_0). This is a contradiction to the fact that the Stackelberg strategy $\bar{h}^*$ gives rise to the optimum solution. $\square$

9. Conclusions

We believe it would be interesting to try to find characterizations of the Stackelberg thresholds for other classes of graphs: for parallel link graphs with more general latency functions, or for linear latency functions in more general graphs, such as series-parallel graphs. We would also like to carry forward the agenda of considering tradeoffs in the size of altruistic coalitions versus the overall social cost in a variety of games.

10. REFERENCES

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

[2] R. Cole, Y. Dodis, and T. Roughgarden. How much can taxes help selfish routing? In *ACM Conference on Electronic Commerce*, pages 98–107, 2003.

[3] J. Correa and N. Stier Moses. A note on Stackelberg routing. Unpublished Manuscript, Sept. 2006.

[4] J. Feigenbaum, C. H. Papadimitriou, and S. Shenker. Sharing the cost of muliticast transmissions (preliminary version). In *STOC*, pages 218–227, 2000.

[5] L. Fleischer. Linear tolls suffice: New bounds and algorithms for tolls in single source networks. *Theor. Comput. Sci.*, 348(2-3):217–225, 2005.

[6] A. C. Kaporis and P. G. Spirakis. The price of optimum in stackelberg games on arbitrary single commodity networks and latency functions. In *SPAA*, pages 19–28, 2006.

[7] G. Karakostas and S. G. Kolliopoulos. Stackelberg strategies for selfish routing in general multicommodity networks. Technical report, McMaster University, June 2006.

[8] Y. A. Korilis, A. A. Lazar, and A. Orda. Achieving network optima using Stackelberg routing strategies. *IEEE/ACM Trans. Netw.*, 5(1):161–173, 1997.

[9] E. Koutsoupias and C. H. Papadimitriou. Worst-case equilibria. In *STACS*, pages 404–413, 1999.

[10] V. S. A. Kumar and M. V. Marathe. Improved results for Stackelberg scheduling strategies. In *ICALP*, pages 776–787, 2002.

[11] N. Nisan. Algorithms for selfish agents. In *STACS*, pages 1–15, 1999.

[12] N. Nisan and A. Ronen. Algorithmic mechanism design. In *STOC*, pages 129–140, 1999.

[13] A. C. Pigou. *The Economics of Welfare*. MacMillan, 1920.

[14] T. Roughgarden. Designing networks for selfish users is hard. In *FOCS*, pages 472–481, 2001.

[15] T. Roughgarden. *Selfish Routing*. PhD thesis, Cornell University, 2002.

[16] T. Roughgarden. Stackelberg scheduling strategies. *SIAM J. Comput.*, 33(2):332–350, 2004.

[17] T. Roughgarden. *Selfish Routing and the Price of Anarchy*. The MIT Press, 2005.

[18] T. Roughgarden. Selfish routing and the price of anarchy. *OPTIMA*, 2006.

[19] T. Roughgarden and É. Tardos. How bad is selfish routing? *J. ACM*, 49(2):236–259, 2002.

[20] C. Swamy. The effectiveness of Stackelberg strategies and tolls for network congestion games. In *SODA*, Jan. 2007.