

Stackelberg thresholds in network routing games

or

The value of altruism

Yogeshwer Sharma*

David P. Williamson[†]

2006-08-22 14:38

Abstract

We study the problem of determining the minimum amount of flow required to be centrally controlled in a Stackelberg routing game in order to improve the social cost of a Nash equilibrium. We consider the special case of routing on a parallel link graph with linear delays and give a closed form expression for the above quantity.

*Department of Computer Science, Cornell University, Ithaca, NY 14853. Email: yogi@cs.cornell.edu. Supported by NSF grant CCF-0514628.

[†]School of Operations Research and Industrial Engineering, Cornell University, Ithaca, NY 14853. Email: dpw@cs.cornell.edu. Supported by NSF grant CCF-0514628.

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 [Pig20] shows that this can indeed be the case. The ratio of the cost of the Nash equilibrium and the optimum solution is called *Price of Anarchy* [KP99]. The idea of bounding the price of anarchy in network routing games has become well-studied after the groundbreaking work of Roughgarden and Tardos [RT02]. 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 has been 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. Three 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, [NR99, Nis99]), (ii) *Taxes and tolls* on the network links to discourage users from using some links which lead to inefficient equilibria, (see, for example, [BMW56, CDR03, Fle05, FPS00]) and (iii) *Designing the network* in such a way that the network has good Nash to optimum ratio to start with (see, for example, [Rou01, KLO97]). All these methods necessitate either a change in the way game is played in the existing network or the design of new networks from scratch.

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 [Rou02, Chapter 6] is that of a network where there may be two different prices, clients paying the *higher* price get to choose their own route through the network and those paying the *bargain* price do not get a choice of routes and they are controlled centrally by the network administrator. Roughgarden [Rou04] considers the problem of routing a β fraction of flow centrally in such a way that if the rest $1 - \beta$ fraction choose 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 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 $0 \leq \beta \leq 1$ 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 $G = (\{s, t\}, \{e_1, e_2, \dots, e_k\})$ with two nodes, a source s and a sink t , and k directed parallel links from s to t . Each edge e_i is equipped with a latency function $l_i(x) : \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}_{\geq 0}$ 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 $f = (f_1, f_2, \dots, f_k)$ 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 β is a flow $\bar{h} = (\bar{h}_1, \dots, \bar{h}_k)$ such that $\sum_{i=1}^k \bar{h}_i = \beta r$. Define $\tilde{l}_i(x) = l_i(x + \bar{h}_i)$ for $i = 1, \dots, k$. Then the Stackelberg equilibrium induced by \bar{h} is a Nash equilibrium in the graph G with latencies \tilde{l} routing flow $(1 - \beta)r$.

For $0 \leq \beta \leq 1$, 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 [Rou02, Chapter 6]. Roughgarden [Rou04] has shown that $c(G, r, l, \beta) \leq \frac{1}{\beta} c(G, r, l, 1)$, and when the latency functions l_i are linear, then $c(G, r, l, \beta) \leq \frac{4}{3+\beta} c(G, r, l, 1)$.

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 [KM02] give a polynomial-time algorithm for finding a Stackelberg strategy \bar{h} such that $c(G, r, l, \beta, \bar{h}) \leq (1 + \varepsilon) c(G, r, l, \beta)$ for any given $\varepsilon > 0$. Swamy [Swa06] extends the results for Roughgarden's Largest Latency First strategy [Rou04] 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. For general topology networks equipped with linear delay functions, and multicommodity users, Karakostas and Kolliopoulos [KK06] show that the cost of a particular Stackelberg equilibrium (SCALE strategy as suggested in [Rou04]) 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 $\sigma(G, r, l)$. 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 $\varepsilon > 0$. 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 [RT02] show that $c(G, r, l, 0)/c(G, r, l, 1) \leq 4/3$ (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

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

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, and $\rho(G, r, l) = 4/3$, 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 of the form $l(x) = \max\{0, ax + b\}$ with $a \geq 0$.

¹Correa and Stier Moses [CS06] have in personal communication told us of similar, independently obtained results.

If f is a Nash flow, and 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. It is worth pointing out that even though our latency functions are linear, since we allow $b < 0$ in $ax + b$ and take maximum of $ax + b$ and zero for determining the latency, the price of anarchy bound of $4/3$ of Roughgarden and Tardos [RT02] does not hold for our case. There are examples for which the price of anarchy can be $2 - \varepsilon$ for any small $\varepsilon > 0$.

Kaporis et. al. [KPS05] 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 parallel links networks equipped with continuous, differentiable, and strictly increasing latency functions.

Before turning to the proof of our result, 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 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. We describe a process to convert a Nash equilibrium to a Stackelberg equilibrium in Section 5.1. The remaining sections prove the lemmas.

2 Some introductory notation, 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 $l_i(x_i) = \max\{0, a_i x_i + b_i\}$ with $a_i \geq 0$; we will sometimes refer to this as the *delay* of the link. Note that we allow the latencies which remain zero up to some flow and then increase linearly. We assume without loss of generality that there is at most one link whose latency function is a constant (by removing all but the one with minimum latency if there are many). We let z denote the index of this link (if it exists); we will refer to it as the *constant link*.

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 . Clearly, $\sum_{i=1}^k f_i = \sum_{i=1}^k g_i = r$. 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 the 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

$\sum_j h_j = (1 - \beta)r$ where β is the fraction of flow centrally controlled by \bar{h} , while it is the case that $\sum_j (\bar{h}_j + h_j) = r$. We call the flow controlled by a Stackelberg strategy *centrally controlled flow* or *altruistic flow*. The equilibrium flow h we sometimes call *selfish flow*.

We first recall the following well-known lemmas and definitions specialized to the case of parallel links.

Lemma 1. *A flow f is at Nash equilibrium if and only if for every i with $f_i > 0$ and every $j \neq i$, $l_i(f_i) \leq l_j(f_j)$.*

Proof. See Proposition 2.2.2 in [Rou05]. □

Corollary 2 (to Lemma 1). For a Stackelberg strategy \bar{h} , h is an induced Stackelberg equilibrium if and only if for all i with $h_i > 0$ and $j \neq i$, $l_i(\bar{h}_i + h_i) \leq l_j(\bar{h}_j + h_j)$.

Proof. Since h is a Nash equilibrium in the network G with *shifted* delay functions $\tilde{l}_j(x_j) = l_j(\bar{h}_j + x_j)$, the claim follows directly from Lemma 1. □

Definition 3. 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 4. *A flow g is optimal if and only if for every i with $g_i > 0$ and every $j \neq i$, $l_i^*(g_i) \leq l_j^*(g_j)$.*

Proof. See Corollary 2.4.6 in [Rou05]. □

By Lemma 1, each link j with $f_j > 0$ must have the same latency; 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 , the 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 5. 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, $\varepsilon_z^* \geq 0$.

Proof. By Lemmas 1 and 4. Since the marginal latency L^* in the optimum solution is at most the marginal latency of the constant link, we have $L^* \leq b_z \iff L^* \leq L^* + \varepsilon_z^* \iff 0 \leq \varepsilon_z^*$. □

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

Lemma 6. *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 $L_{f'}$ respectively denote the common latency experienced by selfish flow in f and f' respectively.*

Proof. For (i) and (ii), see Proposition 2.2.6 and Corollary 2.6.2 in [Rou05]. For (iii), note that the cost of f is $L_f \cdot r$ since whole flow experiences the same latency. Similarly the cost of f' is $L_{f'} \cdot r$. The equality of costs implies $L_f = L_{f'}$. □

Lemma 7. *Let h and h' be two Stackelberg equilibria induced by Stackelberg strategy \bar{h} . Then the costs of two equilibria are equal.*

Proof. See Proposition 6.2.3 in [Rou05]. □

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

Definition 8. 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 pushing more than the Nash flow on bad links gives rise to Stackelberg equilibria with increased social cost (subject to some caveats). We let f_* denote the minimum of Nash flows on good links. We denote the corresponding link by i_* . Formally,

$$f_* \stackrel{\text{def}}{=} \min_{i: f_i < g_i} f_i, \quad i_* \stackrel{\text{def}}{=} \operatorname{argmin}_{i: f_i < g_i} f_i. \quad (1)$$

If there are many indices for which the flow is equal to f_* , define i_* to be an arbitrary, say the lowest indexed, such link. We have $f_* = f_{i_*}$.

We need some preliminary notions to introduce our main theorem.

Definition 9. 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 central flow controlled by Stackelberg strategy \bar{h} , that is $\beta(\bar{h}) = (\sum_j \bar{h}_j) / r$.

The set of all other links is denoted by $[k] - S(\bar{h}) - U(\bar{h})$. We now state the lemmas which are central to the proof of our main theorem.

Lemma 10. *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 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 11. *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 the cost 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_*); see Section 7 for details. The above lemma in essence states that if the strategy does not have enough flow to control on a good link, 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 12. *If $c(G, r, l, 0) > c(G, r, l, 1)$, then for a parallel link network G with k links and linear delay functions, Nash flow f , and optimum flow g ,*

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

Proof. If there is a link with zero Nash flow and positive optimum flow, then $f_* = 0$. Lemma 10 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 is also 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, $\sigma(G, r, l) = 0 = f_*$ in this case.

We can now assume that there is no link with zero Nash flow and positive optimum flow. Lemma 10 states that $\sigma(G, r, l) \leq f_*$ while Lemma 11 states that $\sigma(G, r, l) \geq f_*$. It follows that $\sigma(G, r, l) = f_*$. \square

We comment here on the definition of f_* . The quantity f_* is defined with respect to a Nash equilibrium and an optimum solution. But the condition $c(G, r, l, 0) > c(G, r, l, 1)$ together with Lemmas 17 and 18 below guarantees that the Nash equilibrium is unique, as is the optimum solution, so the quantity is well defined even without any reference to a particular Nash equilibrium and optimum solution.

4 Some useful lemmas

In this section, we prove some lemmas which we need for the proof of Lemma 10 and 11.

Definition 13. For Nash flow $f_j > 0$ on link j , we say that the flow is *in the increasing segment* of the latency function if $l_j(f_j) = \max\{0, a_j f_j + b_j\} > 0$ and $a_j > 0$. We say that the flow is *in the zero delay segment* or *in the initial zero delay segment* of the latency function if $l_j(f_j) = \max\{0, a_j f_j + b_j\} = 0$ and $a_j > 0$.

Lemma 14. *If the latency of the constant link is positive, then the following conditions are equivalent. (i) $L > 0$, (ii) $r > \sum_{j \in [k], j \neq z} \max\{0, -b_j/a_j\}$, (iii) the flow on all links $j \neq z$ with $f_j > 0$ is in the increasing segment of their latency functions.*

Proof. We prove (i) \iff (ii) and (i) \iff (iii). For (i) \iff (ii), if $L > 0$, then the total flow must be more than $\sum_{j \in [k], j \neq z} \max\{0, -b_j/a_j\}$ since otherwise all flow can be routed on the zero delay segment of the latency functions, and this routing is a valid Nash. This gives rise to $L = 0$, a contradiction. If $L = 0$, then all flow must be routed on the initial zero delay segment and at most only $\sum_{j \in [k], j \neq z} \max\{0, -b_j/a_j\}$ amount of flow can be routed on the zero delay segments. So, $r \leq \sum_{j \in [k], j \neq z} \max\{0, -b_j/a_j\}$.

For (i) \iff (iii), if $L > 0$, then for all nonconstant links $j \in [k] - \{z\}$ with $f_j > 0$, $0 < L = \max\{0, a_j f_j + b_j\}$. Hence the flow on link j is in the increasing segment. If $L = 0$, all flow is routed on the initial zero segment on all link, since that is the part that can give rise to zero latency. The claim is now proved. \square

Lemma 15. *If $c(G, r, l, 0) > c(G, r, l, 1)$, then (i) the latency of the constant link is positive, (ii) $L > 0$, (iii) $r > \sum_{j \in [k], j \neq z} \max\{0, -b_j/a_j\}$, (iv) the flows on all links $j \neq z$ with $f_j > 0$ are in the increasing segment of their latency functions.*

Proof. For (i), if the latency of the constant link were zero, then the cost of the Nash equilibrium will be zero, giving rise to $c(G, r, l, 0) = c(G, r, l, 1) = 0$, a contradiction.

For (ii), if $L = 0$, then the cost of the Nash equilibrium is equal to zero which again contradicts $c(G, r, l, 0) > c(G, r, l, 1)$. Using (i) and Lemma 14, (ii), (iii), and (iv) are equivalent. This finishes the proof of the lemma. \square

Observation 16. For a Nash equilibrium f with common latency L and an optimum solution g with common marginal latency L^* , $L \leq L^*$.

Proof. If $L = 0$, then $L \leq L^*$ since L^* is nonnegative.

If $L > 0$ on the other hand, then all flow is on the increasing segment of their respective latency functions. We consider two cases. If $f_j = g_j$ for all links, consider a link with $f_j = g_j > 0$. It follows that $L = a_j f_j + b_j \leq 2a_j f_j + b_j = 2a_j g_j + b_j = L^*$ where the inequality follows from the fact that $a_j \geq 0$. If Nash flow and optimum flow differ on some link, there must exist a link j with $f_j < g_j$. For this link, $L \leq a_j f_j + b_j \leq 2a_j f_j + b_j < 2a_j g_j + b_j = L^*$, where the last equality follows from $g_j > 0$. The claim $L \leq L^*$ follows in both cases. \square

Lemma 17 (Uniqueness of Nash equilibrium). *If $c(G, r, l, 0) > c(G, r, l, 1)$, then there is a unique Nash equilibrium.*

Proof. If condition in the lemma holds, then $L > 0$ and all flows $f_j > 0$ are in the increasing segment of their respective latency functions from Lemma 15. Let us assume for the sake of contradiction that there are two different Nash equilibria namely f and f' with common delays L_f and $L_{f'}$. We have $L_f = L_{f'}$ from Lemma 6. There exists a link, say i , such that $f_i > f'_i$ and another link, say j , with $f_j < f'_j$. If j is a nonconstant link, $f_j < f'_j$ implies $L_f \leq a_j f_j + b_j < a_j f'_j + b_j = L_{f'}$ where the strict inequality follows since $a_j > 0$. This is a contradiction to $L_f = L_{f'}$. Similarly, if i is a nonconstant link, we have $f_i > f'_i$ which implies $L_f = a_i f_i + b_i > a_i f'_i + b_i \geq L_{f'}$, which again contradicts $L_f = L_{f'}$. Hence we have that $f = f'$. \square

Lemma 18 (Uniqueness of optimum solution). *If $c(G, r, l, 0) > c(G, r, l, 1)$, then there is a unique optimum solution.*

Proof. The condition in the lemma implies that $L^* \geq L > 0$. Using the fact that an optimum solution is a Nash equilibrium with marginal latency functions, an argument similar to the one in the proof of Lemma 17 proves the lemma. \square

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

Lemma 19. *Suppose $c(G, r, l, 0) > c(G, r, l, 1)$. Then for the constant link z (if it exists), if $f_z > 0$, then it must be the case that $f_z < g_z$. Also, if $f_z = 0$ then $f_z \leq g_z$.*

Proof. The second part is trivial since flow on the constant link is nonnegative. For the first part, the condition $f_z > 0$ implies that for all links $j \neq z$ with $f_j > 0$, $L = b_z = a_j f_j + b_j < 2a_j f_j + b_j$ where the last inequality follows since $a_j > 0$ (there is only one constant link). For links j with $f_j = 0$, we have $L = b_z \leq b_j = 2a_j f_j + b_j$. Therefore, for all links $j \neq z$, $b_z \leq 2a_j f_j + b_j$. From Observation 5, all marginal latencies in the optimum flow must be equal and at most b_z . It therefore follows that $2a_j g_j + b_j \leq b_z$ for all $j \neq z$. Combining the two inequalities, it follows for all $j \neq z$ that $g_j \leq f_j$. Since the sum of the flows on all links is r in f as well as in g , we have $g_z \geq f_z$.

The equality $g_z = f_z$ holds if and only if equality holds in $g_j \leq f_j$ for all $j \neq z$. But if that were the case, the Nash equilibrium is same as the optimum solution, contradicting our assumption that $c(G, r, l, 0) > c(G, r, l, 1)$. This proves that $g_z > f_z$. \square

Lemma 20. For the Nash flow f with common latency L and the optimum flow g with common marginal latency L^* , using the notation $l_j^*(f_j) = L^* + \varepsilon_j^*$,

$$(f_z - g_z) + \sum_{j:j \neq z, f_j > 0} \frac{\varepsilon_j^*}{2a_j} = \sum_{j:j \neq z, L \leq b_j \leq L^*} \frac{L^* - b_j}{2a_j}. \quad (3)$$

Here, z is the index of the constant link.

Proof. We have $g_z + \sum_{j:j \neq z, g_j > 0} g_j = g_z + \sum_{j:j \neq z, f_j > 0, g_j > 0} g_j + \sum_{j:j \neq z, f_j = 0, g_j > 0} g_j = r$. Let us call $P = \{j : j \neq z, f_j > 0, g_j > 0\}$ and $N = \{j : j \neq z, f_j = 0, g_j > 0\}$. For each $j \in P$, $2a_j f_j + b_j = L^* + \varepsilon_j^*$ and $2a_j g_j + b_j = L^*$. The first equality holds since $f_j > 0$ and the second one holds because $g_j > 0$. Subtracting the second equality from the first one and rearranging the terms, we get for all $j \in P$, $g_j = f_j - \frac{\varepsilon_j^*}{2a_j}$.

The necessary and sufficient condition for $j \in N$ is $j \neq z$ and $L \leq b_j \leq L^*$. For these links, $g_j = \frac{L^* - b_j}{2a_j}$. Using these expressions for all j , we get

$$g_z + \sum_{j:j \neq z, f_j > 0} \left(f_j - \frac{\varepsilon_j^*}{2a_j} \right) + \sum_{j:j \neq z, L \leq b_j \leq L^*} \frac{L^* - b_j}{2a_j} = r = f_z + \sum_{j:j \neq z, f_j > 0} f_j. \quad (4)$$

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

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

The following lemma relates the flow of links in two different Nash equilibria in the same network with different amounts of total flow.

Lemma 21. 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 $f_j^q \leq f_j^r$ for all $j \in [k]$ and $L^q \leq L^r$.

Proof. Let us consider the case $L^r > 0$. We claim that any Nash equilibrium f^q satisfies the conditions of the lemma. We prove the claim by contradiction. Fix a Nash equilibrium for (G, q, l) . Note that L^q could be equal to zero. Let there be a link with more flow in f^q than in f^r . If this link is a constant link, then we have $f_z^q > f_z^r$. Since the total flow amount is reduced, there exists a link $j \in [k] - \{z\}$ with $f_j^q < f_j^r$ and $a_j > 0$. The first inequality implies $L^r \leq b_z = L^q$ or $L^r \leq L^q$. The second inequality implies $L^q \leq a_j f_j^q + b_j < a_j f_j^r + b_j = L^r$ or $L^q < L^r$. This is a contradiction.

If the link carrying more flow in f^q were a nonconstant link, then have $f_i^q > f_i^r$ with $a_i > 0$. There also exists a link j with $a_j \geq 0$ with $f_j^q < f_j^r$. The first inequality implies that $L^q = \max\{0, a_i f_i^q + b_i\} > \max\{0, a_i f_i^r + b_i\} \geq L^r > 0$ or $L^q > L^r > 0$. The first strict inequality follows from the fact that $L^r > 0: \max\{0, a_i f_i^r + b_i\} = a_i f_i^r + b_i$. The second inequality implies $L^q \leq a_j f_j^q + b_j \leq a_j f_j^r + b_j = L^r$ or $L^q \leq L^r$, a contradiction. Hence the claim follows.

Now consider the case $L^r = 0$. In this case, all flow is routed on the initial zero segment of the latency functions. Consider the flow f^q with $f_j^q = f_j^r \cdot \frac{q}{r}$. It is straightforward to check that f^q is a Nash equilibrium (all flow is routed with zero latency).

The claim $L^q \leq L^r$ follows directly from the fact that $f_j^q \leq f_j^r$. \square

5 Relating Stackelberg equilibrium to Nash equilibrium

In this section, we describe a continuous time process, called the Stackelberg process or simply P , which does the following. Given an instance (G, r, l) and a Stackelberg strategy \bar{h} controlling $\beta(\bar{h}) = \sum_{j \in [k]} h_j / r = (\bar{h}_{[k]}) / r$ 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} . As mentioned above, 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 centrally controlled flow 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 will specify this information below.

We start with a Nash equilibrium f for the network (G, r, l) (choose one if there are many choices). Recall the definition of $S(\bar{h})$ and $U(\bar{h})$ from Definition 9; $S(\bar{h}) = \{j : \bar{h}_j > f_j\}$ and $U(\bar{h}) = \{j : 0 < \bar{h}_j \leq f_j\}$. We will sometimes 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 $\beta(\bar{h}^0) = \sum_{j=1}^k \min\{\bar{h}_j, f_j\} / r = \sum_{j=1}^k \bar{h}_j^0 / r$. We define $h^0 = f - \bar{h}^0$, which is easily seen to be a Stackelberg equilibrium.

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

$$\bar{h}_j^t = \min\{\bar{h}_j, f_j\} + t(\bar{h}_j - \min\{\bar{h}_j, f_j\}) = \bar{h}_j^0 + t(\bar{h}_j - \bar{h}_j^0), \quad (5)$$

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 $(f_j - \bar{h}_j)$, 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

$$\bar{h}_j^t = f_j + t(\bar{h}_j - f_j) \quad \text{for } j \in S(\bar{h}), \quad \bar{h}_j^t = \bar{h}_j \quad \text{for } j \in U(\bar{h}). \quad (6)$$

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 $\bar{h}_j^s \leq \bar{h}_j^t$ for $s \leq t$, we have the total selfish flow routed in h^t is less than or equal to the total selfish flow in h^s and latencies (due to \bar{h}^t) at time t are at least as large as the latencies (due to \bar{h}^s) at s . It is clear that 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 21 that there exists a Stackelberg equilibrium h^t for \bar{h}^t such that $h_j^s \geq h_j^t$ for all $j \in [k] - S(\bar{h})$. The Stackelberg equilibrium alluded to above is what we call h^t . Therefore, for all $s \leq t$, $h_j^s \leq h_j^t$ for $j \in S(\bar{h})$ and $h_j^s \geq h_j^t$ for $j \in [k] - S(\bar{h})$.

This can alternatively be viewed as the following. In moving from time s to $t > s$, the amount of altruistic flow increases on $S(\bar{h})$ by an infinitesimal amount $\varepsilon = (t - s)(\bar{h}_{S(\bar{h})} - f_{S(\bar{h})}) > 0$, resulting in the decrease of amount 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 $h_j^t \leq h_j^s$ so as to keep the latency on all links with $h_j^t > 0$ equal.

We now give a notation for the links which have zero h^t flow at time t . We call this subset of links Z^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 $t' \geq t$) at the smallest time instance t such $h_j^t = 0$ (but $h_j^s > 0$ for all $s < t$). Note that $t \mapsto Z^t$ is a monotone set function. We define U^t to be $U \cap Z^t$.

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 $h_j^t = 0$ if $t_j \leq 1$. 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). With this notation, $Z^t = \{j \in [k] : t_j \leq t\}$.

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 $h^0 = f - \bar{h}$. The process continuously increases the altruistic flow on links in S linearly with time, and the (selfish) flow on all other links responds. 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 21) and the amount of altruistic flow is also respected, so $h \stackrel{\text{def}}{=} h^1$ is a valid Stackelberg equilibrium. Now, there might be other possible equilibria induced by strategy \bar{h} than the one the process finds, but from Lemma 7 the cost of all equilibria is the same.

6 Proof of Lemma 10

Proof of Lemma 10: 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.

Let i be a link with $f_i < g_i$ as in the statement of the lemma. We consider the Stackelberg strategy \bar{h} with $\bar{h}_i = f_i + \varepsilon$ for small $\varepsilon > 0$ and $\bar{h}_j = 0$ for $j \neq i$. 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 sometimes 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 sometimes call it *the rate of decrease*, again for brevity), and then show that the rate of decrease is strictly more than the rate of increase.

As we saw in Section 5.1, the flow does not increase on any link in $[k] - \{i\}$ when we increase the 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 5.1 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 $z \in [k] - Z^0$. 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^* .

Alternatively, if $z \in Z^0$, then the flow decreases on non-constant links. The condition $c(G, r, l, 0) > c(G, r, l, 1)$ implies that the flow on all non-constant links is on the increasing segments of their re-

spective latency functions (see Lemma 15). 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 $\sum_{j \neq i, f_j > 0} \frac{1}{a_j} (2a_j f_j + b_j)$ 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

$$\sum_{j: j \neq i, f_j > 0} \frac{1}{\sum_{j: j \neq i, f_j > 0} 1/a_j} \frac{2a_j f_j + b_j}{a_j} = \sum_{j: j \neq i, f_j > 0} \frac{1}{\sum_{j: j \neq i, f_j > 0} 1/a_j} \frac{L^* + \varepsilon_j^*}{a_j}. \quad (7)$$

Note that $z \notin [k] - Z^0$ in this case. We now prove that the rate of increase is strictly less than the rate of decrease of the social cost. There are several cases to consider here.

$z \notin Z^0$ or $z \in [k] - Z^0 = \{j : j \neq i, f_j > 0\}$ 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.

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

$$L^* + \varepsilon_i^* < \sum_{j: j \neq i, f_j > 0} \frac{1}{\sum_{j: j \neq i, f_j > 0} 1/a_j} \frac{L^* + \varepsilon_j^*}{a_j}. \quad (8)$$

Multiplying both sides by $\sum_{j: j \neq i, f_j > 0} 1/a_j$ and canceling L^* term from both sides, we are left with proving

$$\sum_{j: f_j > 0, j \neq i} \frac{\varepsilon_i^*}{2a_j} < \sum_{j: f_j > 0, j \neq i} \frac{\varepsilon_j^*}{2a_j}. \quad (9)$$

If $i = z$, then it reduces to proving (using Lemma 20)

$$\sum_{j: f_j > 0, j \neq i} \frac{\varepsilon_i^*}{2a_j} < (g_z - f_z) + \sum_{j: L \leq b_j \leq L^*, j \neq z} \frac{L^* - b_j}{2a_j}. \quad (10)$$

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 too. On the other hand, if $i \neq z$, then we are left with proving

$$\sum_{j: f_j > 0, j \neq i} \frac{\varepsilon_i^*}{2a_j} < (g_z - f_z) + \frac{-\varepsilon_i^*}{2a_i} + \sum_{j: L \leq b_j \leq L^*, j \neq z} \frac{L^* - b_j}{2a_j}. \quad (11)$$

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 is nonnegative from Lemma 19, 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 of 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 $f_* + \varepsilon$ 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 $\beta r - (f_* + \varepsilon)$ 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. We turn this intuition into a straightforward proof.

We define the new Stackelberg strategy \bar{h}' as follows,

$$\bar{h}'_j = \begin{cases} \bar{h}_{i_*} & \text{if } j = i_* \\ (\beta r - \bar{h}_{i_*}) \frac{h_j}{\sum_{j:h_j > 0} h_j} & \text{otherwise.} \end{cases} \quad (12)$$

It is easy to see that $\bar{h}'_j \leq h_j$ for $j \neq i_*$ since $(\beta r - \bar{h}_{i_*}) \leq r - \bar{h}_{i_*} = \sum_{j:h_j > 0} h_j$. Therefore, the flow h' defined by

$$h'_j = \begin{cases} 0 & \text{if } j = i_* \\ h_j - (\beta r - \bar{h}_{i_*}) \frac{h_j}{\sum_{j:h_j > 0} h_j} & \text{otherwise.} \end{cases} \quad (13)$$

is a Stackelberg equilibrium for \bar{h}' . This is because $\bar{h} + h = \bar{h}' + h'$ and hence $h'_j > 0$ implies $h_j > 0$ which implies that j is the minimum latency link in $\bar{h} + h$ and therefore in $\bar{h}' + h$ too. Therefore, the selfish flow goes on the minimum latency paths, hence h' is a Stackelberg equilibrium.

Also, the costs of $\bar{h} + h$ and $\bar{h}' + h'$ are equal. Therefore, we have established the existence of a Stackelberg strategy \bar{h}' with $\beta(\bar{h}') = \beta$ and an induced equilibria h' such that the resulting cost is strictly less than the cost of the Nash equilibrium. The lemma now follows. \square

7 Proof of Lemma 11

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 11) 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 9; $S(\bar{h}) = \{j : \bar{h}_j > f_j\}$ and $U(\bar{h}) = \{j : 0 < \bar{h}_j \leq f_j\}$. We sometimes call them S and U respectively. Let $S = \{s_1, s_2, \dots, s_{|S|}\}$ and $U = \{u_1, u_2, \dots, u_{|U|}\}$ for this particular \bar{h} .

7.1 The plan

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 5.1 in the following way. Note that 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 increases 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 to 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 6 and 7 now establishes Lemma 11.

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 $\sigma_1 : \sigma_2 : \dots : \sigma_{|S|}$ with $\sigma_1 + \dots + \sigma_{|S|} = 1$. (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 $l_{s_i}^*((h^t + \bar{h}^t)_{s_i}) = 2a_{s_i} \bar{h}_{s_i}^t + b_{s_i}$ since $h_{s_i} = 0$ for all $s_i \in S$. Therefore the rate of increase of social cost is $\varepsilon(\sigma_1(2a_{s_1} \bar{h}_{s_1}^t + b_{s_1}) + \dots + \sigma_{|S|}(2a_{s_{|S|}} \bar{h}_{s_{|S|}}^t + b_{s_{|S|}})) / \varepsilon = \sum_{s \in S} \sigma_s(2a_s \bar{h}_s^t + b_s)$. This can be lower bounded as

$$\begin{aligned} \sum_{s \in S} \sigma_s(2a_s \bar{h}_s^t + b_s) &\geq \sum_{s \in S} \sigma_s(2a_s \bar{h}_s^0 + b_s) = \sum_{s \in S} \sigma_s(2a_s f_s + b_s) \\ &= \sum_{s \in S} \sigma_s(L^* + \varepsilon_s^*) = L^* + \sum_{s \in S} \sigma_s \varepsilon_s^*. \end{aligned} \quad (14)$$

The inequality here follows from the fact that $\bar{h}_j^t \geq \bar{h}_j^s$ for all $j \in S$ and $s \leq t$.

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 Section 5.1 for a definition of Z^t). The flow decreases precisely on links on which the selfish flow is already positive (and as a result that link is not in Z^t). Formally, all links in $\{j : h_j^t > 0\}$ contribute to the decrease of flow at time t . Note that $\{j : h_j^t > 0\} = [k] - Z^t$. We will use the set $\{j : h_j^t > 0, j \in [k] - Z^t\}$ as set of links on which the flow decreases at time t , we do not need both the conditions, but we will use them since it facilitates the proofs later. Recall that $S \subseteq Z^t$ for all $t \in [0, 1]$.

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

$L_{h^t} = 0$ In this case, all selfish flow is on the initial zero segment of the latency function, so the latency as well as the rate of decrease of social cost is zero. Note that the function $l_j(x_j) = \max\{0, a_j x_j + b_j\}$ is not differentiable at $x_j = -b_j/a_j$ if $b_j < 0$, but this fact does not affect the rate of decrease mentioned here. If the flow is decreasing on a link, the rate of change of social cost is equal to the *left* derivative of $x_j \cdot l_j(x_j)$, which always exists for the function $l_j(x_j)$.

$L_{h^t} > 0$ **and** $z \in \{j : h_j^t > 0, j \in [k] - Z^t\}$ (Equivalently, the flow decrease occurs only on the constant link.) In this case, there is some selfish positive flow 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 $\{j : h_j^t > 0, j \in [k] - Z^t\}$. 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 $\varepsilon \cdot l_z^*(\bar{h}_z^t + h_z^t)/\varepsilon = \varepsilon \cdot b_z/\varepsilon = L^* + \varepsilon_z^*$.

$L_{h^t} > 0$ and $z \notin \{j : h_j^t > 0, j \in [k] - Z^t\}$ (Equivalently, the flow decrease does not occur on the constant link.) In this case, the flow does not decrease on the constant link (it decreases collectively on linear links). The condition $c(G, r, l, 0) > c(G, r, l, 1)$ implies that the flow on all nonconstant links is on the increasing segments of their respective latency functions from Lemma 15. The rate of decrease is (in a way similar to the proof of Lemma 10)

$$\begin{aligned} & \frac{\varepsilon}{\sum_{j:h_j^t>0, j \in [k]-Z^t} 1/a_j} \sum_{j:h_j^t>0, j \in [k]-Z^t} \frac{2a_j(h^t + \bar{h}^t)_j + b_j}{a_j} \bigg/ \varepsilon \\ & \leq \frac{1}{\sum_{j:h_j^t>0, j \in [k]-Z^t} 1/a_j} \sum_{j:f_j>0, j \in [k]-Z^t} \frac{2a_j f_j + b_j}{a_j} \\ & = \frac{1}{\sum_{j:f_j>0, j \in [k]-Z^t} 1/a_j} \sum_{j:f_j>0, j \in [k]-Z^t} \frac{L^* + \varepsilon_j^*}{a_j}. \end{aligned} \quad (15)$$

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

Proof. If j is contained in the left hand side, then $h_j^t > 0$ and $j \notin S$ since $S \subseteq Z^t$. Since the selfish flow on all links is nonincreasing with time, we have $h_j^t \leq h_j^0$. We have $f_j = \bar{h}_j^0 + h_j^0 \geq h_j^t > 0$. The first equality holds since $j \notin S$.

If j is contained in the right hand side on the other hand, then $j \notin Z^t = \{j : h_j^t = 0\}$. Hence $h_j^t > 0$ and j is contained in the right hand side too. \square

7.4 Some sufficient conditions

We state some easy facts about the network in the following lemma.

Lemma 23. *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 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 11. $g_j = 0 \iff b_j \geq L^* \iff L^* + \varepsilon_j^* \geq L^* \iff \varepsilon_j^* \geq 0$. 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). Let us now consider the case $j \in S$ and $j = z$. In this case $\varepsilon_z^* \geq 0$ from Observation 5.

For the third claim, $f_z > 0$ implies $b_z = L \leq L^* \leq b_z$ (the last inequality follows from Observation 5). Therefore, $b_z = L^*$ which is equivalent to $\varepsilon_z^* = 0$. The second part follow from Observation 5.

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 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 $h_j^0 = f_j - \bar{h}_j^0 \geq 0$ from the fact that $j \notin S$. 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 some $t \in [0, 1)$ 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 5 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 11. $g_j = 0 \implies b_j \geq L^* \iff L^* + \varepsilon_j^* \geq L^* \iff \varepsilon_j^* \geq 0$. 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 the previous part proves that $\varepsilon_j^* \geq 0$ and we are done.
3. j is a bad link. This follows from the previous part 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. \square

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 (14), but rate of decrease varies depending on which links are involved in the decrement. We will consider these cases separately.

Case 1: $L_{h^t} = 0$ In this case, we need to prove that

$$L^* + \sigma_1 \varepsilon_{s_1}^* + \dots + \sigma_s \varepsilon_{s_{|S|}}^* \geq 0. \quad (16)$$

A sufficient condition for (16) to hold is

$$\varepsilon_j^* \geq 0 \quad \forall j \in S. \quad (17)$$

This conditions always holds from Lemma 23.

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

$$L^* + \sigma_1 \varepsilon_{s_1}^* + \cdots + \sigma_s \varepsilon_{s_1|s_1}^* \geq L^* + \varepsilon_z^*. \quad (18)$$

A sufficient condition for inequality in (18) to hold is

$$\varepsilon_j^* \geq 0 \quad \forall j \in S \quad \text{and} \quad \varepsilon_z^* = 0. \quad (19)$$

These conditions also hold directly from Lemma 23, 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 22.

Case 3: $L_{h^t} > 0$ 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

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

Canceling the L^* terms, dividing by 2 and using Lemma 20, this is equivalent to proving

$$\sum_{j: f_j > 0, j \in [k] - Z^t} \frac{\sigma_1 \varepsilon_{s_1}^* + \cdots + \sigma_s \varepsilon_{s_1|s_1}^*}{2a_j} \geq \sum_{j: f_j > 0, j \neq z, j \in Z^t} \frac{-\varepsilon_j^*}{2a_j} + (g_z - f_z). \quad (20)$$

The $\sum_{j: j \neq z, L \leq b_j \leq L^*} \frac{L^* - b_j}{2a_j}$ equals zero since $\{j : f_j = 0, g_j > 0\} = \emptyset$. A sufficient condition for (20) to hold is the following:

$$\varepsilon_j^* \geq 0, \quad \forall j \in S \quad \text{and} \quad \varepsilon_j^* \geq 0 \quad \forall j \in Z^t - \{z\} \quad \text{and} \quad f_z = g_z. \quad (21)$$

We show that these conditions also hold. The first two conditions hold from Lemma 23. 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 (21) are satisfied.

If $f_z > 0$ and $z \in Z^t$, then z must be a bad link (see Lemma 23). It follows that $f_z \geq g_z$. But from Lemma 19, $f_z \leq g_z$. It follows that $f_z = g_z$ and all the conditions of (21) are again satisfied.

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 11 hence follows.

8 Conclusion

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. 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.

References

- [BMW56] M. J. Beckmann, C. B. McGuire, and C. B. Winsten. *Studies in the Economics of Transportation*. Yale University Press, 1956.
- [CDR03] Richard Cole, Yevgeniy Dodis, and Tim Roughgarden. How much can taxes help selfish routing? In *ACM Conference on Electronic Commerce*, pages 98–107. ACM, 2003.
- [CS06] Jose Correa and Nicholas Stier Moses. Personal Communication, August 2006.
- [Fle05] Lisa Fleischer. Linear tolls suffice: New bounds and algorithms for tolls in single source networks. *Theor. Comput. Sci.*, 348(2-3):217–225, 2005.
- [FPS00] Joan Feigenbaum, Christos H. Papadimitriou, and Scott Shenker. Sharing the cost of multicast transmissions (preliminary version). In *STOC*, pages 218–227, 2000.
- [KK06] George Karakostas and Stavros G. Kolliopoulos. Stackelberg strategies for selfish routing in general multicommodity networks. Technical report, McMaster University, June 2006.
- [KLO97] Yannis A. Korilis, Aurel A. Lazar, and Ariel Orda. Achieving network optima using Stackelberg routing strategies. *IEEE/ACM Trans. Netw.*, 5(1):161–173, 1997.
- [KM02] V. S. Anil Kumar and Madhav V. Marathe. Improved results for Stackelberg scheduling strategies. In Peter Widmayer, Francisco Triguero Ruiz, Rafael Morales Bueno, Matthew Hennessy, Stephan Eidenbenz, and Ricardo Conejo, editors, *ICALP*, volume 2380 of *Lecture Notes in Computer Science*, pages 776–787. Springer, 2002.
- [KP99] Elias Koutsoupias and Christos H. Papadimitriou. Worst-case equilibria. In *STACS*, pages 404–413, 1999.
- [KPS05] Alexis C. Kaporis, Efpraxia Politopoulou, and Paul G. Spirakis. The price of optimum in Stackelberg games. *Electronic Colloquium on Computational Complexity (ECCC)*, TR05(056), 2005.
- [Nis99] Noam Nisan. Algorithms for selfish agents. In Christoph Meinel and Sophie Tison, editors, *STACS*, volume 1563 of *Lecture Notes in Computer Science*, pages 1–15. Springer, 1999.
- [NR99] Noam Nisan and Amir Ronen. Algorithmic mechanism design (extended abstract). In *STOC*, pages 129–140, 1999.
- [Pig20] A. C. Pigou. *The Economics of Welfare*. MacMillan, 1920.
- [Rou01] Tim Roughgarden. Designing networks for selfish users is hard. In *FOCS*, pages 472–481, 2001.

- [Rou02] Tim Roughgarden. *Selfish Routing*. PhD thesis, Cornell University, 2002.
- [Rou04] Tim Roughgarden. Stackelberg scheduling strategies. *SIAM J. Comput.*, 33(2):332–350, 2004.
- [Rou05] Tim Roughgarden. *Selfish Routing and the Price of Anarchy*. The MIT Press, 2005.
- [RT02] Tim Roughgarden and Éva Tardos. How bad is selfish routing? *J. ACM*, 49(2):236–259, 2002.
- [Swa06] Chaitanya Swamy. The effectiveness of Stackelberg strategies and tolls for network congestion games. Unpublished manuscript, July 2006.