Toll Caps in Privatized Road Networks

Tobias Harks^{*}

Marc Schröder[†] Dries Vermeulen[‡]

February 14, 2019

Abstract

We consider a network pricing game on a parallel network with congestion effects in which link owners set tolls for travel so as to maximize profit. A central authority is able to regulate this competition by means of a (uniform) price cap. The first question we want to answer is how such a cap should be designed in order to minimize the total congestion. We provide an algorithm that finds an optimal price cap for networks with affine latency functions and a full support Wardrop equilibrium. Second, we consider the induced network performance at an optimal price cap. We show that for two link networks with affine latency functions, the congestion costs at the optimal price cap are at most 8/7 times the optimal congestion costs. For more general latency functions, this bound goes up to 2 under the assumption that an uncapped Nash equilibrium exists. However, in general such an equilibrium need not exist and this can be used to show that optimal price caps can induce arbitrarily inefficient flows.

Keywords: Game theory, competition regulation, toll caps, Nash equilibrium, Wardrop equilibrium.

1 Introduction

With the ongoing privatization of public road infrastructure, toll charging on highways and roads is common practice in many cities around the world (see e.g. Bergen [45], London [44], Santiago de Chile [22], Singapore [42], and Stockholm [43]). The toll market in the United States (starting already in the 18th century) is built on the key idea that private firms obtain the right to construct the infrastructure (usually via an auction, see Porter and Zona [39]) and as compensation are allowed to charge tolls for road usage. Firms are further obligated to reinvest parts of the revenues to maintain the infrastructure. To date about 35 States use this mechanism and while it has worked well in some cases, several problems arose in the past and still prevail. We report here one incident. In 2012, large toll rate increases have been implemented by the Port Authority of New York and New Jersey (justified in part to finance its World Trade Center project). In response, a bill was introduced in Congress that would allow the Secretary of Transportation to *regulate* tolls on every bridge on the country's interstates and other federally aided highways.

If a regulation authority introduces toll caps, how should this cap be designed in order to induce a socially beneficial outcome and how does this outcome perform? This question is at

^{*}University of Augsburg, Augsburg, Germany, tobias.harks@math.uni-augsburg.de

[†]RWTH Aachen University, Aachen, Germany, marc.schroeder@oms.rwth-aachen.de

[‡]Maastricht University, Maastricht, The Netherlands, d.vermeulen@maastrichtuniversity.nl

the core of our paper. We model the situation using a three-level optimization model, where in the lowest (third) level, there are commuters that want to travel from an origin to a destination and each commuter minimizes a combination of latency costs plus toll costs. We model the route choice of a fixed volume of commuters using the concept of Wardrop equilibrium [47]. We further assume that the underlying network consists of $n \ge 2$ parallel edges connecting a common origin with a common destination. While this assumption is indeed restrictive, it still models relevant situations, for instance, when there are parallel access roads (e.g. tollable bridges) connecting the central district of a city with a suburb, and it is used in recent related work (e.g. Acemoglu and Ozdaglar [1], and Wan [46]).

In the middle (second) layer, there are firms owning the edges and charging tolls on them. Each firm maximizes revenue, but faces competition from other firms. We assume that each firm owns only one edge. A *Nash equilibrium* is a vector of tolls, one toll value per firm, so that no firm can improve by unilaterally deviating to another toll value. A Nash equilibrium takes into account that commuters will (potentially) change their route choices once a firm deviates to a different toll.¹

The authority in the first and highest level can decide on some toll cap. We assume that the authority is not allowed to discriminate between the different firms and puts one (uniform) price cap applicable to all firms. We consider the situation in which the authority is purely interested in minimizing the resulting total latency costs (clearly other objectives may be feasible as well).

1.1 Our Results

We study a three-level optimization problem in which a central authority can impose a (uniform) price cap so as to minimize the total induced equilibrium congestion costs. We motivate this form of competition regulation by showing that without any price regulation, the induced equilibrium congestion costs can be arbitrarily higher than the optimal congestion costs because of too high prices of one of the firms.

For solving the three-level optimization problem, we first derive a set of structural results for instances with affine latency functions. In particular, we show that a Nash equilibrium exists for all price caps c, and is in fact unique whenever the Wardrop equilibrium has full support. Moreover, under the full support assumption, we derive a complete characterization of Nash equilibria using the KKT conditions of the underlying optimization problems of the involved firms. Based on this characterization, we devise an algorithm that finds the optimal price cap in polynomial time. The algorithm computes breakpoints at which the set of firms putting their price equal to the cap changes. We show that these sets only increase (with decreasing cap) so that there are at most n breakpoints, where n denotes the number of firms. The breakpoints essentially divide the real line into at most n intervals and for each of these intervals we show that the cost function is quadratic. Thus, the algorithm needs to solve at most n quadratic one-dimensional minimization problems, which can be done in polynomial time.

We then consider the performance of an optimal price cap. Given that a Nash equilibrium always exists for networks with affine latency functions, we show a tight bound of 8/7 between the congestion costs at an optimal price cap and the optimal congestion costs for duopoly

 $^{^{1}}$ In game theory this equilibrium concept is called subgame perfect Nash equilibrium, but given that the focus of this paper is on the pricing game, we will call it a Nash equilibrium.

instances. For more general latency functions, this bound goes up to 2 if we assume that an uncapped Nash equilibrium exists.

In general, however, the objective function of a firm in the second level is not concave, the set of best replies is not convex and therefore standard techniques for proving existence of a Nash equilibrium fail. In fact, we give an example in which an uncapped subgame perfect Nash equilibrium does not exist. Moreover, we are able to use the nonexistence of uncapped Nash equilibria to construct a sequence of instances in which the ratio between the congestion costs at the optimal price cap and the optimal congestion costs goes to infinity.

1.2 Related Work

The inefficiency of selfish behavior in congestion games has been well recognized since the work of Pigou [38] in economics and Wardrop [47] in transportation networks. A natural way to quantify this inefficiency is the price of anarchy as defined by Koutsoupias and Papdimitrou [33]. For routing games this is first done by Roughgarden and Tardos [41] and later by Correa et al. [15].

One way to restore efficiency is by means of centralized pricing. Beckmann et al. [4] and Dafermos and Sparrow [16] showed that marginal cost pricing induces an equilibrium flow that is optimal. Despite the effectiveness of marginal cost tolls, there are two drawbacks. First, potentially each edge in the network is tolled; an issue that is considered by Hoefer et al. [27] and Harks et al. [24]. Second, the imposed tolls can be arbitrary large; an issue that is considered by Bonifaci et al. [6], Fotakis et al. [21], Jelinek et al. [28] and Kleer and Schäfer [32]. Cole et al. [12] show that optimal tolls also exist when users are heterogeneous with respect to the tradeoff between time and money. Later this result was extended to general networks by Fleischer et al. [20], Karakostas and Kolliopoulos [31] and Yang and Huang [49].

Instead of improving efficiency, tolls can also be used to minimize or maximize the profit of one or multiple leaders. The model that computes tolls that induce the optimal flow at minimal profit was analyzed by Dial [18, 19]. The model with one profit maximizing leader and no congestion effects, was first analyzed by Labbé et al. [34] and later by Briest et al. [7].

Acemoglu and Ozdaglar [1] introduced a model of price competition between link owners in a parallel network that is very similar to ours. The main difference is that in their model users have a reservation value for travel. This implies that if links become too expensive users choose not to travel, whereas in our model users always travel through the network. Their main result is a (tight) bound on the inefficiency of equilibria. Later several generalizations of the model were introduced. The follow-up work of Acemoglu and Ozdaglar [2] allows for a slightly more general topology, namely parallel paths with multiple links, and show that equilibria can be arbitrarily inefficient. Hayrapetyan et al. [26] consider the model with elastic traffic demand. Their bounds on inefficiency, however, are not tight, which is improved upon by Ozdaglar [36].

Recently, following our question on competition regulation of price competition between edge owners, Correa et al. [14] considered the setting in which a central authority is allowed to put different toll caps on different edges of the network. Their main result is that for all network topologies there are toll caps so that firms are willing to put their toll equal to the cap, and the induced equilibrium flow is the optimal flow. An example of such caps are marginal tolls as introduced by Beckmann et al. [4]. Notice that the assumption of individual toll caps is important. In practice, however, toll cap discrimination is often not allowed, which motivates our work on uniform toll caps.

A different extension was proposed by Johari et al. [29]. They extend the analysis of price

competition by including investment and entry decisions. The model of Xiao et al. [48] studies competition in both tolls and capacities and finds that tolls are higher, but capacities are lower than socially desired. Other recent models of Bertrand competition in a network setting that use different ways of modelling congestion effects are Anshelevich and Sekar [3], Chawla and Niu [10], Chawla and Roughgarden [11], Papadimitriou and Valiant [37] and Caragiannis et al. [8].

2 The model

An instance of the three-level optimization problem is given by the tuple $I = (N, (\ell_i)_{i \in N})$, where $N = \{1, \ldots, n\}$, with $n \geq 2$, is a set of parallel links connecting a common source with a common destination. There is one unit of flow to be sent over the links. Let $x_i \in \mathbb{R}_+$ be the total flow on link *i* and $x = (x_1, \ldots, x_n)$ be the vector of flows. Each link has a flow-dependent latency function $\ell_i(x_i)$ that we assume to be strictly increasing, convex and continuously differentiable. We denote this class of functions by \mathcal{L}_c . By \mathcal{L}_d we denote the class of strictly increasing polynomial latency functions with nonnegative coefficients and degree at most *d*. In particular, \mathcal{L}_1 denotes the class of strictly increasing affine latency functions.

Given a flow x, define the total latency costs by

$$C(x) = \sum_{i \in N} \ell_i(x_i) \cdot x_i.$$

A flow x^* is optimal, if $C(x^*) \leq C(x)$ for all flows x.

Let $t_i \in \mathbb{R}_+$ be the toll on link *i* and $t = (t_1, \ldots, t_n)$ be the vector of tolls. The effective cost of a user of link *i* is $\ell_i(x_i) + t_i$. Given a toll vector *t*, a flow *x* is a Wardrop equilibrium for *t*, if for all $i, j \in N$ with $x_i > 0$,

$$\ell_i(x_i) + t_i \le \ell_j(x_j) + t_j.$$

In particular, if x is an equilibrium for t, then all links with positive flow have equal effective costs, i.e., there is some K > 0 with $\ell_i(x_i) + t_i = K$ for all $i \in N$ with $x_i > 0$. It is well-known that given our assumptions on the latency functions, an equilibrium for t exists, is unique and can be described by means of the following inequality (see Beckmann et al. [4] and Dafermos and Sparrow [16]).

Lemma 2.1. A flow x is a Wardrop equilibrium for t if and only if for all feasible flows x',

$$\sum_{i \in N} (\ell_i(x_i) + t_i) \cdot (x_i - x'_i) \le 0.$$

For each toll vector t, we denote by x(t) the unique equilibrium for t. The equilibrium for t = 0 is called the *Wardrop equilibrium*.

2.1 Nash Equilibrium

We assume that each link $i \in N$ is owned by a firm (also denoted by i) and that the objective of each firm is to maximize profit. Given a toll vector $t \in \mathbb{R}^N_+$, define $\prod_i (t_i, t_{-i}) = t_i \cdot x_i(t)$ for all $i \in N$. A toll vector t is a *c*-capped Nash equilibrium, if for all $i \in N$, $t \leq c$ and for all $0 \leq t'_i \leq c$,

$$\Pi_i(t_i, t_{-i}) \ge \Pi_i(t'_i, t_{-i}).$$

Observe that the flow adapts to the toll vector (t'_i, t_{-i}) in the calculation of $\Pi_i(t'_i, t_{-i})$. Given $t_{-i} \in \mathbb{R}^{N \setminus \{i\}}_+$ and $c \in \mathbb{R}_+$, define $B^c_i(t_{-i}) = \underset{0 \le t_i \le c}{\operatorname{arg max}} \prod_i(t_i, t_{-i})$ and $B^\infty_i(t_{-i}) = \underset{t_i \ge 0}{\operatorname{arg max}} \prod_i(t_i, t_{-i})$ for all $i \in N$.

Given $c \in \mathbb{R}_+$, we denote by

 $\mathcal{T}(c) = \{ t \in \mathbb{R}^n_+ \mid t \text{ is a } c\text{-capped Nash equilibrium} \}$

the set of c-capped Nash equilibria. By $\mathcal{T}(\infty)$ we denote the set of uncapped Nash equilibria. Note that there might be instances for which $\mathcal{T}(c)$ is not a singleton for some $c \in \mathbb{R}_+$, and for which $\mathcal{T}(c) = \emptyset$ for some $c \in \mathbb{R}_+$.

2.2 Designing Price Caps

The following example provides the motivation for our research: in terms of latency costs, a Nash equilibrium can be arbitrarily worse than the optimal flow.

Example 2.1. Consider the network of Figure 1.

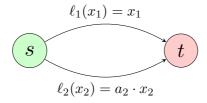


Figure 1: An inefficient Nash equilibrium as $a_2 \to \infty$.

Then

$$B_1^{\infty}(t_2) = \begin{cases} \frac{a_2+t_2}{2}, & \text{if } t_2 \le a_2+2, \\ t_2-1, & \text{if } t_2 > a_2+2, \end{cases}$$

and

$$B_2^{\infty}(t_1) = \begin{cases} \frac{1+t_1}{2}, & \text{if } t_1 \le 2a_2 + 1, \\ t_1 - a_2, & \text{if } t_1 > 2a_2 + 1. \end{cases}$$

Combining $B_1^{\infty}(t_2)$ and $B_2^{\infty}(t_1)$, if $t \in \mathcal{T}(\infty)$, then $t = \left(\frac{2a_2+1}{3}, \frac{a_2+2}{3}\right)$. Since $x(t) = \left(\frac{2a_2+1}{3a_2+3}, \frac{a_2+2}{3a_2+3}\right)$ and $x^* = \left(\frac{a_2}{a_2+1}, \frac{1}{a_2+1}\right)$, we have $C(x(t)) = \frac{a_2^2+7a_2+1}{9a_2+9}$ and $C(x^*) = \frac{a_2}{a_2+1}$. Hence $\frac{C(x(t))}{C(x^*)} = \frac{a_2^2+7a_2+1}{9a_2} \to \infty$ as $a_2 \to \infty$. As $a_2 \to \infty$, the price of firm 1 grows faster than the price of firm 2, which pushes too much flow onto link 2, creating inefficiencies.

The bad performance of the Nash equilibrium calls for competition regulation. The regulation policy we focus on is that of a price cap $c \in \mathbb{R}_+$: no firm is allowed to set its toll above the price cap, i.e., $t_i \leq c$ for all $i \in N$.

Formally, we seek to choose c so as to solve the following three-level optimization problem:

$$\inf_{c \in \mathbb{R}_+} \sup_{t \in \mathcal{T}(c)} C(x(t)).$$
(3L-P)

If $\mathcal{T}(c) = \emptyset$ for some $c \in \mathbb{R}_+$, we assume that $C(x(t)) = \infty$. Observe that $\mathcal{T}(0) \neq \emptyset$.

Note that this is a quite appealing and robust formulation: we minimize total costs for the *worst* possible realization of a Nash equilibrium. It is known, however, that multi-level optimization problems are notoriously hard in terms of proving existence and computability of optimal solution.

In the next section, we derive structural results of Nash equilibria and their corresponding Wardrop equilibria. These structural results contain existence and uniqueness of equilibria for networks with affine latencies and a comparative statics result with respect to different price caps. These insights form the basis of our polynomial time algorithm computing an optimal price cap for the case of affine latency functions.

3 Structural Properties of Nash Equilibria

Assume that $\ell_i(x_i) = a_i \cdot x_i + b_i$ with $a_i > 0$ and $b_i \ge 0$. The first result shows that for affine latency functions, a Nash equilibrium exists for all price caps. A similar result is obtained for the model in which users have a reservation value for travel and there are no price caps by Acemoglu and Ozdaglar [1] and Hayrapetyan et al. [26].

Proposition 3.1. Let $\ell_i \in \mathcal{L}_1$ for all $i \in N$. Then, there exists a c-capped Nash equilibrium for all $c \in \mathbb{R}_+$.

Proof. Assume that $\ell_i(x_i) = a_i \cdot x_i + b_i$ with $a_i > 0$ and $b_i \ge 0$. For all toll vectors $t \in \mathbb{R}^N_+$, define $N(t) = \{i \in N \mid x_i(t) > 0\}$.

Since x(t) is an equilibrium for t, we have $a_i \cdot x_i(t) + b_i + t_i = K$ for all $i \in N(t)$. With $\sum_{i \in N(t)} x_i(t) = 1$, we get $K = \frac{1 + \sum_{j \in N(t)} \frac{b_j + t_j}{a_j}}{\sum_{i \in N(t)} \frac{1}{a_i}}$ and thus

$$x_i(t) = 1, \text{ we get } K = \frac{\sum_{j \in N(t)} \frac{1}{a_j}}{\sum_{j \in N(t)} \frac{b_j + t_j - b_i - t_i}{a_j}} \text{ for all } i \in N(t).$$

$$(1)$$

By the Theorem of the Maximum [5], it follows that $B^c(t) = (B^c_i(t_{-i}))_{i \in N}$ is upper semicontinuous and hence has a closed graph. We show that $B^c_i(t_{-i})$ is a convex set for all t_{-i} .

Assume $t_{-i} \in \mathbb{R}^{N \setminus \{i\}}_+$ and consider the following two cases: (a) $\Pi_i(t_i, t_{-i}) = 0$ for all $t_i \in \mathbb{R}_+$ and (b) $\Pi_i(t_i, t_{-i}) > 0$ for some $t_i \in \mathbb{R}_+$.

Case (a): Suppose $\Pi_i(t_i, t_{-i}) = 0$ for all $t_i \in \mathbb{R}_+$. Obviously $B_i^c(t_{-i})$ is convex.

Case (b): Suppose there is some $t'_i \in \mathbb{R}_+$ with $\Pi_i(t'_i, t_{-i}) > 0$. We will prove that $\Pi_i(t_i, t_{-i})$ is a concave function in t_i until there is some \bar{t}_i such that $\Pi_i(t_i, t_{-i}) = 0$ for all $t_i \ge \bar{t}_i$. Observe the following from (1).

- (i) If $j \in N(t'_i, t_{-i})$, with $j \neq i$, for some $t'_i \in \mathbb{R}_+$, then $j \in N(t_i, t_{-i})$ for all $t_i \geq t'_i$. That is, the set N(t) increases as t_i increases.
- (ii) For a fixed N(t), $\Pi_i(t_i, t_{-i})$ is either linear or quadratic in t_i . $\Pi_i(t)$ is linear if and only if $N(t) = \{i\}$, which can only happen for $t_i \in [0, \hat{t}_i]$ for some $\hat{t}_i \in \mathbb{R}_+$. In all other cases, $\Pi_i(t)$ is quadratic with a decreasing slope.

Now in order to prove that $\Pi_i(t_i, t_{-i})$ is a concave function in t_i , it is sufficient to show that the slope decreases at every toll t_i for which new links start to receive flow, i.e. a new set of players joins $N(t_i, t_{-i})$. To this end, let t_i denote a price at which a new set of firms starts to receive flow and let us denote this set by \overline{N} . Notice that

$$\frac{\partial_{-}\Pi_{i}(t_{i}, t_{-i})}{\partial t_{i}} = \frac{1 + \sum_{j \in N(t): j \neq i} \frac{b_{j} + t_{j} - b_{i} - 2t_{i}}{a_{j}}}{\sum_{j \in N(t)} \frac{a_{i}}{a_{j}}}$$
$$\frac{\partial_{+}\Pi_{i}(t_{i}, t_{-i})}{\partial t_{i}} = \frac{1 + \sum_{j \in N(t) \cup \bar{N}: j \neq i} \frac{b_{j} + t_{j} - b_{i} - 2t_{i}}{a_{j}}}{\sum_{j \in N(t) \cup \bar{N}} \frac{a_{i}}{a_{j}}}$$

where $\frac{\partial_{-}\Pi_{i}(t_{i},t_{-i})}{\partial t_{i}}$ and $\frac{\partial_{+}\Pi_{i}(t_{i},t_{-i})}{\partial t_{i}}$ denote the left and right partial derivative with respect to t_{i} , respectively. So in order to prove that the slope decreases it is sufficient to show that

$$\left(1+\sum_{j\in N(t): j\neq i}\frac{b_j+t_j-b_i-2t_i}{a_j}\right)\cdot\sum_{k\in\bar{N}}\frac{a_i}{a_k}\geq\sum_{k\in\bar{N}}\frac{b_k+t_k-b_i-2t_i}{a_k}\cdot\sum_{j\in N(t)}\frac{a_i}{a_j}.$$
 (2)

We get

$$\begin{split} &\sum_{k\in\bar{N}} \frac{b_k + t_k - b_i - 2t_i}{a_k} \cdot \sum_{j\in N(t)} \frac{a_i}{a_j} \\ &= \sum_{k\in\bar{N}} \frac{a_i \cdot x_i(t) - t_i}{a_k} \cdot \sum_{j\in N(t)} \frac{a_i}{a_j} \\ &= \sum_{k\in\bar{N}} \frac{1 + \sum_{j\in N(t): j\neq i} \frac{b_j + t_j - b_i - t_i}{a_j} - \sum_{j\in N(t)} \frac{t_i}{a_j}}{a_k \cdot \sum_{j\in N(t)} \frac{1}{a_j}} \cdot \sum_{j\in N(t)} \frac{a_i}{a_j} \\ &= \sum_{k\in\bar{N}} \left(1 + \sum_{j\in N(t): j\neq i} \frac{b_j + t_j - b_i - 2t_i}{a_j} \right) \cdot \frac{a_i}{a_k} - \frac{t_i}{a_k} \\ &= \left(1 + \sum_{j\in N(t): j\neq i} \frac{b_j + t_j - b_i - 2t_i}{a_j} - \frac{t_i}{a_i} \right) \cdot \sum_{k\in\bar{N}} \frac{a_i}{a_k}, \end{split}$$

where the first equality follows from $b_k + t_k = a_i \cdot x_i(t) + b_i + t_i$ for all $k \in \overline{N}$, the second equality from the definition of $x_i(t)$, and the last two inequalities from rewriting. Since $\frac{t_i}{a_i} > 0$, (2) holds true. Hence $B_i^c(t_{-i})$ is convex-valued.

Kakutani's fixed point theorem [30] now implies that there exists an equilibrium for all price caps $c \in \mathbb{R}_+$.

The following lemma states a very natural comparative statics result, showing that a unilateral increase in toll by one firm only decreases the flow on the corresponding link while the flow on other links only increases. **Lemma 3.1.** Let $i \in N$ and $t = (t_i, t_{-i}), t' = (t'_i, t_{-i}) \in \mathbb{R}^n_+$ with $t_i \leq t'_i$. Then $x_i(t) \geq x_i(t')$ and $x_j(t) \leq x_j(t')$ for all $j \neq i$.

Proof. It is well known (see Beckmann et al. [4] and Dafermos and Sparrow [16]) that x is an equilibrium for t if x solves the following minimization problem

$$\min_{x} \sum_{i \in N} \int_0^{x_i} (\ell_i(y) + t_i) \, dy$$

Hence

$$\sum_{i \in N} \int_0^{x_i(t)} (\ell_i(y) + t_i) \, dy \leq \sum_{i \in N} \int_0^{x_i(t')} (\ell_i(y) + t_i) \, dy$$
$$\sum_{i \in N} \int_0^{x_i(t')} (\ell_i(y) + t'_i) \, dy \leq \sum_{i \in N} \int_0^{x_i(t)} (\ell_i(y) + t'_i) \, dy$$

Combining these inequalities we get

$$\sum_{j \in N} (t_j - t'_j) \cdot (x_j(t) - x_j(t')) \le 0.$$
(3)

Using $t' = (t'_i, t_{-i})$, we get

$$(t_i - t'_i) \cdot (x_i(t) - x_i(t')) \le 0,$$

and thus $x_i(t) \ge x_i(t')$.

Before we prove that $x_j(t) \leq x_j(t')$ for all $j \neq i$, we first show that $\ell_i(x_i(t)) + t_i \leq \ell_i(x_i(t')) + t'_i$. Since x(t) + (x(t') - x(t)) = x(t') is a feasible flow, x(t') - x(t) is a feasible direction for x(t). By the first-order optimality conditions,

$$\sum_{i \in N} (\ell_i(x_i(t)) + t_i) \cdot (x_i(t) - x_i(t')) \le 0.$$

Analogously, x(t) - x(t') is a feasible direction for x(t'), and thus

$$\sum_{i \in N} (\ell_i(x_i(t')) + t'_i) \cdot (x_i(t') - x_i(t)) \le 0.$$

Adding up these inequalities, we obtain

$$\sum_{j \neq i} (x_j(t) - x_j(t')) \cdot (\ell_j(x_j(t)) - \ell_j(x_j(t'))) + (x_i(t) - x_i(t')) \cdot (\ell_i(x_i(t)) + t_i - \ell_i(x_i(t')) - t'_i) \le 0.$$

Notice that the first summation term is nonnegative, as ℓ_i is increasing for all $i \in N$. Thus,

$$\ell_i(x_i(t)) + t_i \le \ell_i(x_i(t')) + t'_i.$$
(4)

Now we prove that $x_j(t) \leq x_j(t')$ for all $j \neq i$. Notice that we can assume that $x_i(t) > 0$, as otherwise the conclusion follows trivially. We use a proof by contradiction. Suppose that $x_j(t) > x_j(t')$ for some $j \neq i$. Then

$$\ell_j(x_j(t')) + t_j < \ell_j(x_j(t)) + t_j \le \ell_i(x_i(t)) + t_i \le \ell_i(x_i(t')) + t'_i \le \ell_j(x_j(t')) + t_j,$$

where the first inequality follows from $x_j(t) > x_j(t')$, the second inequality from $x_j(t) > 0$, the third inequality from (4), and the fourth inequality from $x_i(t') > 0$. This is a contradiction and finishes the proof.

From now on, we assume that $x_i(0) > 0$ for all $i \in N$. In terms of practical applications, this assumption seems not overly restrictive since a road segment is usually only priced if there is traffic without any tolls. However, the assumption is important for proving uniqueness of Nash equilibria. Example 4.2 shows that multiple Nash equilibria may exist in case the full Wardrop support assumption is not satisfied.

The full support assumption $x_i(0) > 0$ for all $i \in N$ does not imply a priori that any Nash equilibrium $x_i(t), i \in N$ for some $t \in \mathcal{T}(c), c \in \mathbb{R}_+$ also has the full support property. In the following, however, we show that given the full support property for t = 0, it continues to hold for all $t \in \mathcal{T}(c), c \in \mathbb{R}_+$.

Lemma 3.2. Let c > 0 and $t \in \mathcal{T}(c)$. If $x_i(0) > 0$ for all $i \in N$, then $x_i(t) > 0$ and $t_i > 0$ for all $i \in N$.

Proof. By Lemma 3.1, if $x_i(0) > 0$ for all $i \in N$, then $x_i(0, t_{-i}) > 0$ for all $i \in N$ and all t_{-i} . As the profit function is continuous in the toll vector (Hayrapetyan et al. [26]), this implies that for all t_{-i} , there is some t_i so that $\prod_i(t_i, t_{-i}) > 0$. Hence if $t \in \mathcal{T}(c)$, then for all $i \in N$, $\prod_i(t_i, t_{-i}) > 0$ and thus $x_i(t) > 0$ and $t_i > 0$.

In the following lemma we derive an explicit price representation of Nash equilibria.

Lemma 3.3. Let $c \in \mathbb{R}_+ \cup \{\infty\}$ and $t \in \mathcal{T}(c)$. If $x_i(t) > 0$ for all $i \in N$, then for all $i \in N$,

$$t_i = \min\left\{ \left(\ell'_i(x_i(t)) + \frac{1}{\sum_{j \neq i} \frac{1}{\ell'_j(x_j(t))}} \right) \cdot x_i(t), c \right\}.$$

Proof. Assume that $x_i(t) > 0$ for all $i \in N$. Given $t_{-i} \in \mathbb{R}^{N \setminus \{i\}}_+$, each firm $i \in N$ solves the following maximization problem:

$$\max t_i \cdot x_i$$

s.t. $\ell_j(x_j) + t_j = K$ for all $j \in N$,
$$\sum_{j \in N} x_j = 1,$$

 $t_i \le c.$

The corresponding Lagrangian is

$$\mathcal{L}(t_i, x, K, \lambda, \mu, \nu)$$

= $t_i \cdot x_i - \sum_{j \in N} \lambda_j \cdot (\ell_j(x_j) + t_j - K) - \nu \cdot (\sum_{j \in N} x_j - 1) - \mu \cdot (t_i - c)$

So the Kuhn-Tucker conditions reduce to

$$\frac{\partial \mathcal{L}(t_i, x, K, \lambda, \mu, \nu)}{\partial t_i} = x_i - \lambda_i - \mu = 0,$$
(5)

$$\frac{\partial \mathcal{L}(t_i, x, K, \lambda, \mu, \nu)}{\partial x_i} = t_i - \lambda_i \cdot \ell'_i(x_i) - \nu = 0, \tag{6}$$

$$\frac{\partial \mathcal{L}(t_i, x, K, \lambda, \mu, \nu)}{\partial x_j} = -\lambda_j \cdot \ell'_j(x_j) - \nu = 0,$$

$$\frac{\partial \mathcal{L}(t_i, x, K, \lambda, \mu, \nu)}{\partial x_j} = \sum_{i=1}^{N} \lambda_i = 0$$
(8)

$$\frac{\partial \mathcal{L}(t_i, x, K, \lambda, \mu, \nu)}{\partial K} = \sum_{j \in N} \lambda_j = 0.$$
(8)

By (5),

$$\lambda_i = x_i - \mu. \tag{9}$$

By (6) and plugging in (9),

$$t_i = \ell'_i(x_i) \cdot \lambda_i + \nu = \ell'_i(x_i) \cdot (x_i - \mu) + \nu.$$

$$(10)$$

By (7), for all $j \neq i$,

$$\lambda_j = -\nu \cdot \frac{1}{\ell'_j(x_j)}.\tag{11}$$

By (11),

$$\sum_{j \neq i} \lambda_j = -\nu \cdot \sum_{j \neq i} \frac{1}{\ell'_j(x_j)}$$
(12)

By (8), (9) and (12),

$$\nu = \frac{\lambda_i - \sum_{j \in N} \lambda_j}{\sum_{j \neq i} \frac{1}{\ell'_j(x_j)}} = \frac{\lambda_i}{\sum_{j \neq i} \frac{1}{\ell'_j(x_j)}} = \frac{x_i - \mu}{\sum_{j \neq i} \frac{1}{\ell'_j(x_j)}}.$$
(13)

Combining (10) and (13) yields

$$t_i = \left(\ell'_i(x_i) + \frac{1}{\sum_{j \neq i} \frac{1}{\ell'_j(x_j)}}\right) \cdot (x_i - \mu).$$

If $t_i < c$, then we know $\mu = 0$, and thus,

$$t_i = \left(\ell'_i(x_i) + \frac{1}{\sum_{j \neq i} \frac{1}{\ell'_j(x_j)}}\right) \cdot x_i.$$

Now we derive a complete characterization of equilibria.

Theorem 3.1. Let $x_i(0) > 0$ for all $i \in N$ and let $c \in \mathbb{R}_+$. The tuple (t, x) is a c-capped Nash equilibrium if and only if the following conditions hold for some K > 0:

$$a_i x_i + b_i + t_i = K \text{ for all } i \in N, \tag{14}$$

$$\sum_{i\in N} x_i = 1,\tag{15}$$

$$t_i = \min\left\{ \left(a_i + \frac{1}{\sum_{j \neq i} \frac{1}{a_j}} \right) \cdot x_i, c \right\} \text{ for all } i \in N,$$
(16)

$$x_i > 0 \text{ for all } i \in N.$$

$$\tag{17}$$

Proof. We first show \Rightarrow : Conditions (14) and (17) follow from Lemma 3.2 and the Wardrop condition. Condition (15) is trivial. Condition (16) follows from Lemma 3.3.

Now we prove \Leftarrow : let (t', x') be a tuple that satisfies (14)-(17). We want to show that (t', x') = (t, x), where (t, x) is a Nash equilibrium, which by Proposition 3.1 exists. By (16), we get that tolls t' are feasible w.r.t. c. Condition (14) implies that x' is a Wardrop equilibrium with full support with respect to t'. Hence, similarly as in (3), we get

$$\sum_{j \in N} (t_j - t'_j) \cdot (x_j - x'_j) \le 0.$$
(18)

Assume by contradiction that there is $i \in N$ with $t'_i < t_i$ (the case $t'_i > t_i$ follows similarly). By (16), we get

$$t'_i = \left(a_i + \frac{1}{\sum_{j \neq i} \frac{1}{a_j}}\right) \cdot x'_i < t_i \le \left(a_i + \frac{1}{\sum_{j \neq i} \frac{1}{a_j}}\right) \cdot x_i.$$

From this $x'_i < x_i$ follows. Putting things together, we get

$$\sum_{j \in N} (t_j - t'_j)(x_j - x'_j) > 0$$

a contradiction to (18).

Corollary 3.1. There is at most one tuple (t, x) that satisfies (14)-(17), thus the Nash equilibrium is unique.

If $\mathcal{T}(c)$ is singular, we let t(c) denote $t \in \mathcal{T}(c)$, and x(c) denote x(t(c)).

4 Optimal price caps for affine latencies

Assume that $\ell_i(x_i) = a_i \cdot x_i + b_i$ with $a_i > 0$ and $b_i \ge 0$, and $x_i(0) > 0$ for all $i \in N$. Now we have everything together to derive an optimal polynomial time algorithm for networks with affine latencies, which is the main result of this section.

Theorem 4.1. Let $\ell_i \in \mathcal{L}_1$ and $x_i(0) > 0$ for all $i \in N$. Algorithm 1 computes in polynomial time an optimal price cap.

Algorithm 1: An optimal algorithm for affine latencies.

input : $I = (N, (\ell_i)_{i \in N})$ **output:** An optimal price cap c^* 1 initialize 2 $c_0 \leftarrow \infty;$ **3** $A(c_0) \leftarrow \emptyset;$ 4 $j \leftarrow 0;$ $\begin{array}{c|c}
\mathbf{4} & j \leftarrow 0, \\
\mathbf{5} & \text{while } A(c_j) \neq N \text{ do} \\
\mathbf{5} & \text{while } A(c_j) \neq N \text{ do} \\
\mathbf{6} & I + \sum_{k \in A(c_j)} \frac{b_k + c}{a_k} + \sum_{k \in N \setminus A(c_j)} \frac{2a_k + \frac{1}{2a_k + \frac{1}{k \in N \setminus \{k\}} 1/a_l}}{\sum_{k \in A(c_j)} \frac{1}{a_k} + \sum_{k \in N \setminus A(c_j)} \frac{2a_k + \frac{1}{k \in N \setminus \{k\}} 1/a_l}{2a_k + \frac{1}{k \in N \setminus \{k\}}}; \\
\end{array}$ 7 $x_i(c) \leftarrow \frac{K(c) - b_i - c}{a_i} \text{ for all } i \in A(c_j);$ 8 $x_i(c) \leftarrow \frac{K(c) - b_i}{2a_i + \frac{1}{\sum_{j \neq i} 1/a_j}} \text{ for all } i \in N \setminus A(c_j);$ 9 $t_i(c) \leftarrow c \text{ for all } i \in A(c_j);$ 10 $t_i(c) \leftarrow \left(a_i + \frac{1}{\sum_{l \neq i} 1/a_l}\right) \cdot x_i(c) \text{ for all } i \in N \setminus A(c_j);$ 11 $c_{j+1}^i \leftarrow \max\left\{c \left| \left(a_i + \frac{1}{\sum_{l \neq i} 1/a_l}\right) \cdot x_i(c) = c\right\} \text{ for all } i \in N \setminus A(c_j); \right.$ 12 $c_{j+1} \leftarrow \max\left\{c_{j+1}^i \middle| i \in N \setminus A(c_j)\right\};$ $A(c_{j+1}) \leftarrow A(c_j) \cup \arg \max \left\{ c_{j+1}^i \mid i \in N \setminus A(c_j) \right\};$ 13 $j \leftarrow j + 1;$ $\mathbf{14}$ $c_j^* \leftarrow \arg\min_{c \in [c_j, c_{j-1}]} \sum_{i \in N} \ell_i(x_i(c)) \cdot x_i(c)$ $\mathbf{15}$ 16 end 17 **output** $c^* \in \arg\min\{C(x(c_k^*)), k = 1, \dots, j\}$

Proof. We show by induction on the iterations of the while loop of Algorithm 1 (indexed by $k \in \mathbb{N}$) that the algorithm computes breakpoints $c_1 > \cdots > c_k > \cdots > c_j$ so that in the intervals $\mathcal{I}_1 = [c_1, \infty)$ and $\mathcal{I}_k = [c_k, c_{k-1}]$ for $k = 2, \ldots, j+1$ with $c_{j+1} = 0$ the following invariant holds: for all $k = 1, \ldots, j+1$, the flows $x_i(c), i \in N$ as defined in Lines 7 and 8 together with prices $t_i(c), i \in N$ as defined in Lines 9 and 10 constitute the unique Nash equilibrium for all $c \in \mathcal{I}_k$.

Consider the base case k = 1. First, observe that Line 12 is well defined since the maximum in Line 12 obviously exists (it is attained at $c_1 = \max_{i \in N} t_i(\infty)$). The parameterized flow x(c)as defined in lines Line 7 and 8 is a solution of the following system of linear equations:

$$a_i \cdot x_i(c) + b_i + \left(a_i + \frac{1}{\sum_{l \neq i} 1/a_l}\right) \cdot x_i(c) = K(c) \text{ for all } i \in N,$$
$$\sum_{i \in N} x_i(c) = 1.$$

By Theorem 3.1 and Corollary 3.1, x(c), t(c) is the unique Nash equilibrium for all $c \in \mathcal{I}_1$.

For the inductive step $k \to k + 1$, assume x(c), t(c) as defined in Lines 7-10 is the unique Nash equilibrium for all $c \in \mathcal{I}_{\ell}, \ell = 1, \ldots, k$.

First, we show again that the maximum c_{k+1}^i , $i \in N \setminus A(c_k)$ in Line 11 exists so that Lines 11 and 12 are well defined, and, that $c_{k+1} = \max\{c_{k+1}^i | i \in N \setminus A(c_k)\} < c_k$ as claimed. To see this, observe that for $i \in N \setminus A(c_k)$, we have $\left(a_i + \frac{1}{\sum_{l \neq i} 1/a_l}\right) \cdot x_i(c_k) < c_k$. On the other hand, by the assumption that $x_i(0) > 0$ for all $i \in N$ we have $\left(a_i + \frac{1}{\sum_{l \neq i} 1/a_l}\right) \cdot x_i(0) > 0$. As the function $\left(a_i + \frac{1}{\sum_{l \neq i} 1/a_l}\right) \cdot x_i(c)$ is continuous in c, by the intermediate value theorem, there exists \bar{c}_{k+1}^i with

$$\left(a_i + \frac{1}{\sum_{l \neq i} 1/a_l}\right) \cdot x_i(\bar{c}_{k+1}^i) = \bar{c}_{k+1}^i,$$

implying $c_{k+1}^i < c_k$ and thus $c_{k+1} < c_k$.

We next prove that x(c), t(c) as defined in Lines 7-10 is the unique Nash equilibrium for all $c \in \mathcal{I}_{k+1}$. We prove this by showing that x(c), t(c) satisfies the conditions of Theorem 3.1. Observe that x(c) is a solution of the following system of linear equations:

$$\begin{aligned} a_i \cdot x_i(c) + b_i + t_i &= K(c) \text{ for all } i \in N, \\ t_i &= c \text{ for all } i \in A(c_k), \\ t_i &= \left(a_i + \frac{1}{\sum_{l \neq i} 1/a_l}\right) \cdot x_i(c) \text{ for all } i \in N \setminus A(c_k), \\ \sum_{i \in N} x_i(c) &= 1. \end{aligned}$$

By Theorem 3.1, it is suffices to show that $t_i(c)$ satisfies condition (16) for all $i \in N, c \in \mathcal{I}_{k+1}$. From Lines 11 and 12 we get $t_i(c) \leq c$ for all $i \in N \setminus A(c_k)$ and all $c \in \mathcal{I}_{k+1}$. It remains to show that for all $i \in A(c_k)$ and all $c \in \mathcal{I}_{k+1}$,

$$\left(a_i + \frac{1}{\sum_{l \neq i} 1/a_l}\right) \cdot x_i(c) \ge c.$$

We know by the induction hypothesis that $x(c_k)$ is the unique Nash equilibrium, and thus satisfies $\left(a_i + \frac{1}{\sum_{l \neq i} 1/a_l}\right) \cdot x_i(c_k) \ge c_k$ for all $i \in A(c_k)$. For all $c \in (c_{k+1}, c_k)$, we have

$$\frac{\partial K(c)}{\partial c} = \frac{\sum_{i \in A(c_j)} \frac{1}{a_i}}{\sum_{i \in A(c_j)} \frac{1}{a_i} + \sum_{i \in N \setminus A(c_j)} \frac{1}{2a_i + \frac{1}{\sum_{l \in N \setminus \{i\}} \frac{1}{l/a_l}}} \le 1,$$

and thus by Line 7, we get for all $i \in A(c_k)$ and all $c \in (c_{k+1}, c_k)$,

$$\frac{\partial x_i(c)}{\partial c} \le 0$$

So when decreasing c from c_k to c_{k+1} , it follows that $\left(a_i + \frac{1}{\sum_{l \neq i} 1/a_l}\right) \cdot x_i(c)$ increases, and thus $\left(a_i + \frac{1}{\sum_{l \neq i} 1/a_l}\right) \cdot x_i(c) \ge c$ for all $i \in A(c_k)$ and all $c \in \mathcal{I}_{k+1}$, which completes the induction proof.

For each \mathcal{I}_k , with $k = 1, \ldots, j + 1$, the objective function

$$C(x(c)) = \sum_{i \in N} \ell_i(x_i(c)) \cdot x_i(c)$$

is quadratic in c, thus, we can find a local minimum by comparing the two endpoints and a possible interior minimum point. For the interior minimum, we just need to check firstorder optimality conditions, thus, solving a linear equation in c. This way, we have effectively partitioned the search space \mathbb{R}_+ into at most n intervals. As we solved each segment optimally, taking the best solution leads to the optimal c^* .

Remark 4.1. The algorithm is still polynomial for arbitrary demands $d \ge 0$ since solving the linear equation systems appearing in the algorithm can be done in polynomial time in the encoding length of d.

The following example demonstrates the calculation process of the algorithm.

Example 4.1. Consider the network of Figure 2.

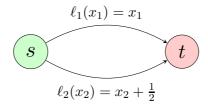


Figure 2: A demonstration of the algorithm.

By Lemma 3.3, the uncapped Nash equilibrium prices can be found by solving the following system of linear equations

$$x_1(c) + 2x_1(c) = K(c),$$

$$x_2(c) + \frac{1}{2} + 2x_2(c) = K(c),$$

$$x_1(c) + x_2(c) = 1,$$

and thus are given by $t(\infty) = (\frac{7}{6}, \frac{5}{6})$. Initialize $c_1 = \frac{7}{6}$ and $A(c_1) = \{1\}$. Solve the following system of linear equations

$$x_1(c) + c = K(c),$$

$$x_2(c) + \frac{1}{2} + 2x_2(c) = K(c),$$

$$x_1(c) + x_2(c) = 1,$$

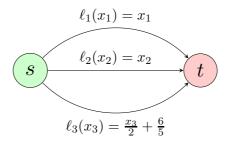
yields $x_2(c) = \frac{2c+1}{8}$. Solving $2 \cdot \frac{2c+1}{8} = c$ yields $c = \frac{1}{2}$. Hence, $c_2 = \frac{1}{2}$ and $A(c_2) = \{1, 2\}$. Therefore,

$$C(x(c)) = \begin{cases} \frac{3}{4}, & \text{if } c \leq \frac{1}{2}, \\ \frac{4c^2 - 9c + 27}{32}, & \text{if } \frac{1}{2} < c \leq \frac{7}{6}, \\ \frac{13}{18}, & \text{if } c > \frac{171}{224}, \end{cases}$$

which is minimized for c = 1. Notice that we induce the optimal flow with a uniform price cap of c = 1.

The last example of this section shows that the full Wardrop support assumption is important. Without this assumption, a Nash equilibrium need not be unique.

Example 4.2. Consider the network of Figure 3. Assume that $t_3 = 0$. Then





$$B_1^{\infty}(t_2, t_3) = \begin{cases} t_1 = \frac{t_2 + 1}{2}, & \text{if } 0 \le t_2 \le \frac{3}{5}, \\ t_1 = \frac{7 - 5t_2}{5}, & \text{if } \frac{3}{5} < t_2 \le \frac{5}{7}, \\ t_1 = \frac{5t_2 + 17}{30}, & \text{if } \frac{5}{7} < t_2 \le \frac{17}{10}, \\ t_1 = \frac{17}{20}, & \text{if } t_2 > \frac{17}{10}. \end{cases}$$

and

$$B_2^{\infty}(t_1, t_3) = \begin{cases} t_2 = \frac{t_1 + 1}{2}, & \text{if } 0 \le t_1 \le \frac{3}{5}, \\ t_2 = \frac{7 - 5t_1}{2}, & \text{if } \frac{3}{5} < t_1 \le \frac{5}{7}, \\ t_2 = \frac{5t_1 + 17}{30}, & \text{if } \frac{5}{7} < t_1 \le \frac{17}{10}, \\ t_2 = \frac{17}{20}, & \text{if } t_1 > \frac{17}{10}. \end{cases}$$

Combining $B_1^{\infty}(t_2, t_3)$ and $B_2^{\infty}(t_1, t_3)$ (see Figure 4) implies that the set $\left\{\left(t_1, \frac{7}{5} - t_1, 0\right) \mid \frac{24}{35} \leq t_1 \leq \frac{5}{7}\right\}$ are Nash equilibria.

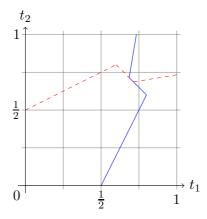


Figure 4: Best reply correspondence: in blue for player 1, and in (dashed) red for player 2.

Remark 4.2. The network pricing game of Example 4.2 exhibits another interesting phenomenon: if $\ell_3(x_3) = a_3 \cdot x_3 + \frac{6}{5}$ with $a_3 \leq 2/3$, then every equilibrium yields a profit of 0 for firm 3, whereas if $\ell_3(x_3) = a_3 \cdot x_3 + \frac{6}{5}$ with $a_3 > 2/3$, then the equilibrium is unique and yields a strictly positive profit for firm 3. This implies that firm 3 is worse off by being congestion free than by being sufficiently congestion dependent. The reason is that in the former case firm 3 is too competitive, and thus its two competitors set prices in such a way that prevents the third firm from entering the market.

5 Optimal price caps for duopolies

In this section, we compare the latency costs of optimal price caps with those of optimal solutions minimizing total congestion (ignoring price competition). The aim of this comparison is to reveal the possible strength of introducing price caps as a mechanism. Our results, however, only hold for duopolies, that is, we assume n = 2 for the following analysis. We conjecture though that our results carry over to general n (in the spirit of Pigou instances).

Let c^* denote an optimal price cap (a solution to problem 3L-P), let x^* denote the optimal flow (ignoring price caps), and let $\mathcal{L} \subseteq \mathcal{L}_c$ be a class of latency functions. We are interested in the ratio $\rho(\mathcal{L})$ between the latency costs at the optimal price cap and the latency costs of the optimal flow, defined by, for all $\ell \in \mathcal{L}$,

$$C(x(c^*)) \le \rho(\mathcal{L}) \cdot C(x^*).$$

Our main results are as follows.

Theorem 5.1. Let n = 2 and $\mathcal{T}(c) \neq \emptyset$ for all $c \in \mathbb{R}_+$. Then

(i)
$$\rho(\mathcal{L}_d) \le (1 - \frac{d}{2(d+1)^{(d+1)/d}})^{-1}$$

(ii)
$$\rho(\mathcal{L}_c) \leq 2$$
.

Suppose that we have a Nash equilibrium in which one firm receives all the flow, say firm 1. The proof of Lemma 3.3 implies that in that case $\ell_1(1) + \ell'_1(1) \leq \ell_2(0)$. This again implies that the duopoly and optimal flow coincide. So in the remainder of this section, we assume that $x_i(t) > 0$ for all i = 1, 2 for $t \in \mathcal{T}(\infty)$.

Before we prove the main result of this section, we introduce some helpful lemmas.

Lemma 5.1. Let $t \in \mathcal{T}(\infty)$. Suppose that $x_i(t) > 0$ for i = 1, 2. Then $x_1(0) \ge x_1(t)$ if and only if $x_1(t) \ge \frac{1}{2}$.

Proof. Suppose that $x_1(0) \ge x_1(t)$. Then

$$\ell_1(x_1(t)) \le \ell_1(x_1(0)) \le \ell_2(x_2(0)) \le \ell_2(x_2(t)),$$

where the first inequality follows from $x_1(t) \leq x_1(0)$ and ℓ_1 increasing, the second inequality from $x_1(0) \geq x_1(t) > 0$, and the third inequality from $x_2(0) \leq x_2(t)$ and ℓ_2 increasing. Since $x_i(t) > 0$ for i = 1, 2, we have that $\ell_1(x_1(t)) + t_1 = \ell_2(x_2(t)) + t_2$ and thus $t_1 \geq t_2$. Hence Lemma 3.3 implies that $x_1(t) \geq x_2(t)$ and thus $x_1(0) \geq x_1(t) \geq \frac{1}{2}$.

Suppose that $x_1(0) < x_1(t)$. Then

$$\ell_2(x_2(t)) < \ell_2(x_2(0)) \le \ell_1(x_1(0)) < \ell_1(x_1(t)),$$

where the first inequality follows from $x_2(t) < x_2(0)$ and ℓ_2 increasing, the second inequality from $x_2(0) > 0$, and the third inequality from $x_1(0) < x_1(t)$ and ℓ_1 increasing. Since $x_i(t) > 0$ for i = 1, 2, we have that $\ell_1(x_1(t)) + t_1 = \ell_2(x_2(t)) + t_2$ and thus $t_1 < t_2$. Hence Lemma 3.3 implies that $x_1(t) < x_2(t)$ and thus $x_1(t) < \frac{1}{2}$.

Define

$$\mu_1(\ell_i) = \sup_{x_i, x_i^* \ge 0} \frac{(\ell_i(x_i) - \ell_i(x_i^*)) \cdot x_i^*}{\ell_i(x_i) \cdot x_i}$$

for each $\ell_i \in \mathcal{L}$ and

$$\mu_1(\mathcal{L}) = \sup_{\ell_i \in \mathcal{L}} \mu_1(\ell_i).$$

The parameter $\mu_1(\mathcal{L})$ is a measure of the steepness of the class of allowable latency functions that is well studied in the context of bounding the price of anarchy in routing games. It is well known that the price of anarchy in routing games can be bounded as a function of the class of allowable latency functions, but usually not in terms of other characteristics of the instance like the topology of the network. For more details, see, for example, Correa et al. [15] and Roughgarden [40]. Observe that $\mu_1(\mathcal{L}) \in [0, 1]$.

Lemma 5.2. If $x_i(0) > x_i^*$ and $x_i(0) \le \frac{1}{2}$ for some i = 1, 2, then $C(x(0)) \le \frac{1}{1-\mu_1(\mathcal{L})/2} \cdot C(x^*)$.

Proof. W.l.o.g. suppose that $x_1(0) > x_1^*$ and $x_1(0) \le \frac{1}{2}$. By Lemma 2.1, we have

$$C(x(0)) \le C(x^*) + \sum_{i=1}^{2} (\ell_i(x_i(0)) - \ell_i(x_i^*)) \cdot x_i^*,$$

$$\le C(x^*) + (\ell_1(x_1(0)) - \ell_1(x_1^*)) \cdot x_1^*,$$

where the second inequality follows from $x_2(0) < x_2^*$ and ℓ_2 increasing. By definition of $\mu_1(\mathcal{L})$,

$$\frac{(\ell_1(x_1(0)) - \ell_1(x_1^*)) \cdot x_1^*}{\ell_1(x_1(0)) \cdot x_1(0)} \le \mu_1(\mathcal{L}).$$

The lemma then follows because $\ell_1(x_1(0)) \leq C(x(0))$ and $x_1(0) \leq \frac{1}{2}$.

Define

$$\mu_2(\ell_i) = \sup_{x_i \ge \frac{1}{2}, 0 \le x_i^* \le x_i} \frac{(\ell_i(x_i) - \ell_i(x_i^*)) \cdot (x_i^* + 1 - 2x_i)}{\ell_i(x_i)}$$

for each $\ell_i \in \mathcal{L}$ and

$$\mu_2(\mathcal{L}) = \sup_{\ell_i \in \mathcal{L}} \mu(\ell_i)$$

The parameter $\mu_2(\mathcal{L})$ is a new smoothness parameter that takes into account the pricing behavior of the firms. Observe that $\mu_2(\mathcal{L}) \in [0, 1]$.

Lemma 5.3. Let $t \in \mathcal{T}(\infty)$ and $x_i(t) > 0$ for all i = 1, 2. If $x_i(t) > x_i^*$ and $x_i(t) \ge \frac{1}{2}$ for some i = 1, 2, then $C(x(t)) \le \frac{1}{1-\mu_2(\mathcal{L})} \cdot C(x^*)$.

Proof. W.l.o.g. suppose that $x_1(t) > x_1^*$ and $x_1(t) \ge \frac{1}{2}$. By Lemma 2.1 and 3.3, we have

$$C(x(t))$$

$$\leq C(x^*) + \sum_{i=1}^{2} (\ell_i(x_i(t)) - \ell_i(x_i^*)) \cdot x_i^* + \sum_{i=1}^{2} t_i(x_i^* - x_i(t)),$$

$$\leq C(x^*) + \sum_{i=1}^{2} (\ell_i(x_i(t)) - \ell_i(x_i^*)) \cdot x_i^*$$

$$- (\ell_1'(x_1(t)) + \ell_2'(x_2(t))) \cdot (x_1(t) - x_1^*) \cdot (2x_1(t) - 1),$$

$$\leq C(x^*) + (\ell_1(x_1(t)) - \ell_1(x_1^*)) \cdot x_1^* - \ell_1'(x_1(t)) \cdot (x_1(t) - x_1^*) \cdot (2x_1(t) - 1),$$

$$\leq C(x^*) + (\ell_1(x_1(t)) - \ell_1(x_1^*)) \cdot (x_1^* + 1 - 2x_1(t)).$$

where the second inequality follows from Lemma 3.3, $x_2(t) \leq x_2^*$ and ℓ_2 increasing, the third from $\ell'_2(x_2(t)) \geq 0$ and the fourth from convexity of ℓ_1 , $\ell'_1(x_1(t)) \cdot (x_1(t) - x_1^*) \geq \ell_1(x_1(t)) - \ell_1(x_1^*)$. Since

$$\frac{(\ell_1(x_1(t)) - \ell_1(x_1^*)) \cdot (x_1^* + 1 - 2x_1(t))}{\ell_1(x_1(t))} \le \mu_2(\mathcal{L}_c),$$

and, by Lemma 3.3, $t_1 \ge t_2$, we have $\ell_1(x_1(t)) \le C(x(t))$ and the lemma follows.

Proof of Theorem 5.1. Let $t \in \mathcal{T}(\infty)$ and $x_i(t) > 0$ for all i = 1, 2. Suppose that either $x_1(t) \leq x_1^* < x_1(0)$, or $x_1(0) < x_1^* < x_1(t)$. By the Theorem of the Maximum (Berge [5]), the profit function of each firm is continuous in c. Given that the profit function of each firm is continuous in the toll vector (Hayrapetyan et al. [26]), and the toll vector continuously changes the induced flow (Beckmann et al. [4]), there exists a price cap c such that $x_1(c) = x_1^*$. So, we can assume that $x_1^* < x_1(0)$ and $x_1^* < x_1(t)$.

Suppose that $x_1(t) \leq \frac{1}{2}$. Then by Lemma 5.1, $x_1(0) \leq x_1(t) \leq \frac{1}{2}$. Since $\mu_1(\mathcal{L}_d) \leq \frac{d}{(d+1)^{(d+1)/d}}$ (Correa et al. [15]) and $\mu_1(\mathcal{L}_c) \leq 1$, the result follows by Lemma 5.2. Suppose that $x_1(t) \geq \frac{1}{2}$. Then

$$\mu_2(\mathcal{L}_d) = \sup_{\substack{x_i(t) \ge \frac{1}{2}, 0 \le x_i^* \le x_i(t) \\ \le \frac{d}{2(d+1)^{(d+1)/d}}}, \frac{(\ell_1(x_1(t)) - \ell_1(x_1^*)) \cdot (x_1^* + 1 - 2x_1(t))}{\ell_1(x_1(t))}$$

and

$$\mu_2(\mathcal{L}_c) = \sup_{\substack{x_i(t) \ge \frac{1}{2}, 0 \le x_i^* \le x_i(t) \\ \le \frac{1}{2} \le x_1(t)}} \frac{(\ell_1(x_1(t)) - \ell_1(x_1^*)) \cdot (x_1^* + 1 - 2x_1(t))}{\ell_1(x_1(t))}$$

$$\leq \sup_{\substack{\frac{1}{2} \le x_1(t) \\ \le \frac{1}{2}, x_1(t)}} 1 - x_1(t) \le \frac{1}{2},$$

and the result follows by Lemma 5.3.

The following example shows that the bound of 8/7 in Theorem 5.1 for affine latencies is tight for duopolies, and is a lower bound for arbitrary parallel graphs.

Example 5.1. Consider the network of Figure 5 with $n \ge 2$.

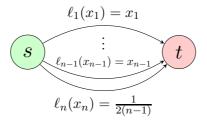


Figure 5: The bound for affine latencies is tight.

Let $t \in \mathcal{T}(\infty)$. By Lemma 3.3, $t = x(t) = x(0) = \left(\frac{1}{2(n-1)}, \dots, \frac{1}{2(n-1)}, \frac{1}{2}\right)$ and $x^* = \frac{1}{2(n-1)}$. $\left(\frac{1}{4(n-1)},\ldots,\frac{1}{4(n-1)},\frac{3}{4}\right).$ If $0 \le c \le \frac{1}{2(n-1)}$, then we have $t_1 = \ldots = t_n = c$ and thus $x(c) = \left(\frac{1}{2(n-1)}, \ldots, \frac{1}{2(n-1)}, \frac{1}{2}\right)$. If $c > \frac{1}{2(n-1)}$, then $t_1 = \ldots = t_2 = \frac{1}{2(n-1)}$ and thus $x(c) = \left(\frac{1}{2(n-1)}, \ldots, \frac{1}{2(n-1)}, \frac{1}{2}\right)$. Hence any price can is optimal. So

price cap is optimal. So

$$\min_{c \in \mathbb{R}_+} \frac{C(x(c))}{C(x^*)} = \frac{8}{7}$$

The main result in Theorem 5.1 assumes that a c-capped Nash equilibrium exists for all $c \in \mathbb{R}_+$. The following example shows that an uncapped Nash equilibrium need not exist for quadratic latency functions.

Example 5.2. Consider the network of Figure 6.

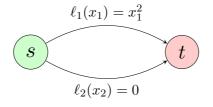


Figure 6: No Nash equilibrium.

Then

$$B_1^{\infty}(t_2) = \begin{cases} t_1 = \frac{2t_2}{3}, & \text{if } 0 \le t_2 \le 3, \\ t_1 = t_2 - 1, & \text{if } t_2 > 3, \end{cases}$$

and

$$B_2^{\infty}(t_1) = \begin{cases} t_2 = \arg\max_{t_2} \left(1 - (t_2 - t_1)^{1/2}\right) \cdot t_2, & \text{if } 0 \le t_1 \le \frac{1}{4}, \\ t_2 = t_1, & \text{if } t_1 \ge \frac{1}{4}. \end{cases}$$

Observe that $B_2^{\infty}(t_1)$ is not convex at $t_1 = \frac{1}{4}$. Combining $B_1^{\infty}(t_2)$ and $B_2^{\infty}(t_1)$ (see Figure 7) implies that there is no uncapped Nash equilibrium. In fact, the result is even stronger: there is no c-capped Nash equilibrium whenever $c \geq \frac{1}{2}$.

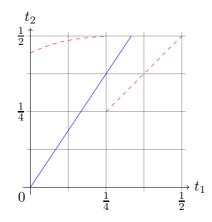


Figure 7: Best reply correspondence: in blue for player 1, and in (dashed) red for player 2.

Remark 5.1. The nonexistence result in Example 5.2 violates the assumption that latency functions are strictly increasing. Nevertheless, we can change the latency function of link 2 into $\ell_2(x_2) = a_2 \cdot x_2$ with $a_2 < \sqrt{17} - 4$ and obtain the same result: an uncapped Nash equilibrium need not exist.

The next and final example shows that the latency costs at the optimal price cap can be arbitrarily worse than the optimal latency costs due to nonexistence of an uncapped Nash equilibrium.

Example 5.3. Consider the network of Figure 8, where $d \ge 3$.

Then

$$B_1^{\infty}(t_2) = \begin{cases} t_1 = \frac{d \cdot (t_2 + \left(\frac{d-1}{d}\right)^d)}{d+1}, & \text{if } 0 \le t_2 \le d+1 - \left(\frac{d-1}{d}\right)^d, \\ t_1 = t_2 + \left(\frac{d-1}{d}\right)^d - 1, & \text{if } t_2 > d+1 - \left(\frac{d-1}{d}\right)^d, \end{cases}$$

and

$$B_2^{\infty}(t_1) = \begin{cases} t_2 = \underset{t_2}{\arg\max} \left(1 - (t_2 - t_1 + b_2)^{1/d}\right) \cdot t_2, & \text{if } 0 \le t_1 \le \frac{d \cdot (d-1)^{d-1}}{d^d}, \\ t_2 = t_1 - \left(\frac{d-1}{d}\right)^d, & \text{if } t_1 \ge \frac{d \cdot (d-1)^{d-1}}{d^d}. \end{cases}$$

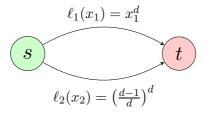


Figure 8: No uncapped Nash equilibrium if $d \geq 3$.

Combining $B_1^{\infty}(t_2)$ and $B_2^{\infty}(t_1)$ implies that there is only an equilibrium for all price caps $c \leq \left(\frac{d-1}{d}\right)^{d-1}$. If $c \leq \left(\frac{d-1}{d}\right)^{d-1}$, then $t_1(c) = t_2(c) = c$. Similarly as in Example 5.2, an important reason for the nonexistence of equilibria seems to be the nonconvexity of the set of best replies for player 2 at $t_1 = \left(\frac{d-1}{d}\right)^{d-1}$. Since $x(c) = x(0) = \left(\frac{d-1}{d}, \frac{1}{d}\right)$ for all $c \leq \left(\frac{d-1}{d}\right)^{d-1}$, and $x^* = \left(\frac{d-1}{d \cdot (d+1)^{1/d}}, 1 - \frac{d-1}{d \cdot (d+1)^{1/d}}\right)$, we have

$$\rho(\mathcal{L}_d) \ge \min_{0 \le c \le \left(\frac{d-1}{d}\right)^{d-1}} \frac{C(x(c))}{C(x^*)} = \frac{(d+1)^{(d+1)/d}}{(d+1)^{(d+1)/d} - (d-1)} \to \infty \ as \ d \to \infty.$$

6 Discussion

We consider a network pricing game in which, in the first stage, edge owners set prices so as to maximize profit, and, in the second stage, users choose paths that minimize their total costs. The problem with these games is that Nash equilibria might not exist, and if they exist, they can induce arbitrarily inefficient flows. We therefore allow for competition regulation and consider a (uniform) price cap regulation policy. Our main goal is, firstly, to find a price cap that minimizes the inefficiency of the induced flow, and, secondly, to quantify the loss in efficiency due to competition even in the presence of competition regulation. Our main results are the following. For parallel link networks with affine latency functions and a full support Wardrop equilibrium, we provide an algorithm that finds the optimal price cap in polynomial time. Due to multiplicity of Nash equilibria, the algorithm is not valid for instances without a full support Wardrop equilibrium. Then we show that the ratio between the congestion costs at an optimal price cap and the optimal congestion costs are at most 2 for duopoly instances with an uncapped Nash equilibrium. This bound lowers down to 8/7 for affine latency functions. However, due to the nonexistence of Nash equilibria, we are able to construct a sequence of instances such that the performance of the induced flow at an optimal price cap is arbitrarily bad.

The following questions remain open. First, is there a (polynomial) algorithm that finds the optimal price cap for more general instances? In particular, for parallel link networks with affine latency functions that have no full support Wardrop equilibrium. Secondly, we have provided an instance that shows that the flow at an optimal price cap can be 8/7 times as costly as the optimal flow. Can the upper bound in Theorem 5.1 be extended from two link networks to arbitrary parallel link networks with affine latencies? Thirdly, for two link networks with affine latency functions, a worst-case guarantee of 8/7 also holds for the simple algorithm that selects the best flow from the Wardrop flow and the uncapped Nash equilibrium flow. How does this simple algorithm perform for arbitrary parallel link networks? Fourthly, how do results change if we assume a different user behavior, like elastic users (Chau and Sim [9], Hayrapetyan et al. [26] or Ozdaglar [36]), atomic splittable users (Haurie and Marcotte [25], Orda et al. [35] or Cominetti et al. [13]), or stochastic users (Daganzo and Sheffi [17] or Guo et al. [23])?

References

References

- D. Acemoglu and A. Ozdaglar. Competition and efficiency in congested markets. Mathematics of Operations Research, 32(1):1–31, 2007.
- [2] D. Acemoglu and A. Ozdaglar. Competition in parallel-serial networks. *IEEE Journal on Selected Areas in Communications*, 25(6):1180–1192, 2007.
- [3] E. Anshelevich and S. Sekar. Price competition in networked markets: How do monopolies impact social welfare? International Conference on Web and Internet Economics, pages 16–30, 2015.
- [4] M. Beckmann, C. McGuire, and C. Winsten. Studies in the economics of transportation. Yale University Press, 1956.
- [5] C. Berge. Topological Spaces: Including a treatment of multi-valued functions, vector spaces, and convexity. Courier Corporation, 1963.
- [6] V. Bonifaci, M. Salek, and G. Schäfer. Efficiency of restricted tolls in non-atomic network routing games. *International Symposium on Algorithmic Game Theory*, pages 302–313, 2011.
- [7] M. Briest, P. Hoefer and P. Krysta. Stackelberg network pricing games. Algorithmica, 62(3-4):733-753, 2012.
- [8] I. Caragiannis, X. Chatzigeorgiou, P. Kanellopoulos, G. Krimpas, N. Protopapas, and A. Voudouris. Efficiency and complexity of price competition among single-product vendors. *Artificial Intelligence*, 248:9–25, 2017.
- [9] C. Chau and K. Sim. The price of anarchy for non-atomic congestion games with symmetric cost maps and elastic demands. *Operations Research Letters*, 31(5):327–334, 2003.
- [10] S. Chawla and F. Niu. The price of anarchy in Bertrand games. ACM Conference on Electronic Commerce, pages 305–314, 2009.
- [11] S. Chawla and T. Roughgarden. Bertrand competition in networks. International Symposium on Algorithmic Game Theory, pages 70–82, 2008.
- [12] R. Cole, Y. Dodis, and T. Roughgarden. Pricing network edges for heterogeneous selfish users. Symposium on Theory of Computing, 3(72):444–467, 2003.
- [13] R. Cominetti, J. Correa, and N. Stier-Moses. The impact of oligopolistic competition in networks. Operations Research, 57(6):1421–1437, 2009.

- [14] J. Correa, C. Guzmán, T. Lianeas, E. Nikolova, and M. Schröder. Network pricing: How to induce optimal flows under strategic link operators. ACM Conference on Economics and Computation, pages 375–392, 2018.
- [15] J. Correa, A. Schulz, and N. Stier-Moses. Selfish routing in capacitated networks. Mathematics of Operations Research, 29(4):961–976, 2004.
- [16] S. Dafermos and F. Sparrow. The traffic assignment problem for a general network. Journal of Research of the National Bureau of Standards, 73(2):91–118, 1969.
- [17] C. Daganzo and Y. Sheffi. On stochastic models of traffic assignment. Transportation Science, 11(3):253–274, 1977.
- [18] R. Dial. Minimal-revenue congestion pricing part i: A fast algorithm for the single-origin case. *Transportation Research*, 33(B):189–202, 1999.
- [19] R. Dial. Minimal-revenue congestion pricing part ii: An efficient algorithm for the general case. Transportation Research, 34(B):645–665, 2000.
- [20] L. Fleischer, K. Jain, and M. Mahdian. Tolls for heterogeneous selfish users in multicommodity networks and generalized congestion games. Symposium on Foundations of Computer Science, pages 277–285, 2004.
- [21] D. Fotakis, D. Kalimeris, and T. Lianeas. Improving selfish routing for risk-averse players. International Conference on Web and Internet Economics, pages 328–342, 2015.
- [22] A. Gonzalez. Controversias por reducciones en la demanda en las concesiones de carreteras en chile. *Revista de Derecho Económico*, 76:61–97, 2016.
- [23] X. Guo, H. Yang, and T. Liu. Bounding the inefficiency of logit-based stochastic user equilibrium. European Journal of Operational Research, 201(2):463–469, 2010.
- [24] T. Harks, I. Kleinert, M. Klimm, and R. Möhring. Computing network tolls with support constraints. *Networks*, 65(3):262–285, 2015.
- [25] A. Haurie and P. Marcotte. On the relationship between Nash-Cournot and Wardrop equilibria. Networks, 15(3):295–308, 1985.
- [26] A. Hayrapetyan, E. Tardos, and T. Wexler. A network pricing game for selfish traffic. Distributed Computing, 19(4):255–266, 2007.
- [27] M. Hoefer, L. Olbrich, and A. Skopalik. Taxing subnetworks. International Workshop on Internet and Network Economics, pages 286–294, 2008.
- [28] T. Jelinek, M. Klaas, and G. Schäfer. Computing optimal tolls with arc restrictions and heterogeneous players. *International Symposium on Theoretical Aspects of Computer Science*, pages 433–444, 2014.
- [29] R. Johari, G. Weintraub, and B. Van Roy. Investment and market structure in industries with congestion. Operations Research, 58(5):1303–1317, 2010.
- [30] S. Kakutani. A generalization of Brouwer's fixed point theorem. Duke Mathematical Journal, 8(3):457–459, 1941.

- [31] G. Karakostas and S. Kolliopoulos. Edge pricing of multicommodity networks for heterogeneous selfish users. *Foundations of Computer Science*, 4:268–276, 2004.
- [32] P. Kleer and G. Schäfer. The impact of worst-case deviations in non-atomic network routing games. *International Symposium on Algorithmic Game Theory*, pages 129–140, 2016.
- [33] E. Koutsoupias and C. Papadimitrou. Worst-case equilibria. Annual Symposium on Theoretical Aspects of Computer Science, pages 404–413, 1999.
- [34] P. Labbé, M. Marcotte and G. Savard. A bilevel model of taxation and its application to optimal highway pricing. *Management Science*, 44(12):1608–1622, 1998.
- [35] A. Orda, R. Rom, and N. Shimkin. Competitive routing in multiuser communication networks. *IEEE/ACM Transactions on Networking*, 1(5):510–521, 1993.
- [36] A. Ozdaglar. Price competition with elastic traffic. *Networks*, 52(3):141–155, 2008.
- [37] C. Papadimitrou and E. Valiant. A new look at selfish routing. Innovations in Computer Science, pages 178–187, 2010.
- [38] A. Pigou. The economics of welfare. Macmillan, 1920.
- [39] R. Porter and J. Zona. Detection of bid rigging in procurement auctions. Journal of Political Economy, 101(3):518–538, 1993.
- [40] T. Roughgarden. The price of anarchy is independent of the network topology. *Journal* of Computer and System Sciences, 67(2):341–364, 2003.
- [41] T. Roughgarden and E. Tardos. How bad is selfish routing? Journal of the ACM, 49(2):236–259, 2002.
- [42] Singapur Government. Electronic road pricing. Technical report, Singapur Government, 2010. http://www.lta.gov.sg/motoring_matters/index_motoring_erp.htm.
- [43] Swedish Road Administration. Congestion tax in Stockholm. Technical report, Swedish Road Administration, 2007.
- [44] Transport for London. Central London congestion charging. Technical report, Transport for London, 2008. http://www.tfl.gov.uk/assets/downloads/sixth-annual-impacts-monitoring-report-2008-07.pdf
- [45] A. Vingan, L. Fridstrom, and K. Johansen. Congestion charging in Bergen and Trondheim an alternative 20 years ahead? Technical report, Norwegian Transportation Institute, 2007.
- [46] C. Wan. Strategic decentralization in binary choice composite congestion games. European Journal of Operational Research, 250(2):531–542, 2016.
- [47] J. Wardrop. Some theoretical aspects of road traffic research. Proceedings of the Institute of Civil Engineers, 1(3):325–362, 1952.

- [48] F. Xiao, H. Yang, and D. Han. Competition and efficiency of private toll roads. Transportation Research Part B: Methodological, 41(3):292–308, 2007.
- [49] H. Yang and H.-J. Huang. The multi-class, multi-criteria traffic network equilibrium and systems optimum problem. *Transportation Research Part B: Methodological*, 38(1):1–15, 2004.