On the Existence of Optimal Taxes for Network Congestion Games with Heterogeneous Users [*]

Dimitris Fotakis[1], George Karakostas[2], and Stavros G. Kolliopoulos[3]

[1] School of Electrical and Computer Engineering,
National Technical University of Athens, 157 80 Athens, Greece.
Email: `fotakis@cs.ntua.gr`
[2] Department of Computing and Software, and
School of Computational Engineering and Science,
McMaster University, Hamilton, ON, Canada.
E-mail: `karakos@mcmaster.ca`
[3] Department of Informatics and Telecommunications,
National and Kapodistrian University of Athens, Athens 157 84, Greece.
URL: `wwww.di.uoa.gr/~sgk`

Abstract. We consider network congestion games in which a finite number of non-cooperative users select paths. The aim is to mitigate the inefficiency caused by the selfish users by introducing taxes on the network edges. A tax vector is *strongly (weakly)-optimal* if all (at least one of) the equilibria in the resulting game minimize(s) the total latency. The issue of designing optimal tax vectors for selfish routing games has been studied extensively in the literature. We study for the first time taxation for networks with atomic users which have unsplittable traffic demands and are *heterogeneous,* i.e., have different sensitivities to taxes. On the positive side, we show the existence of weakly-optimal taxes for single-source network games. On the negative side, we show that the cases of homogeneous and heterogeneous users differ sharply as far as the existence of strongly-optimal taxes is concerned: there are parallel-link games with linear latencies and heterogeneous users that do not admit strongly-optimal taxes.

1 Introduction

We consider atomic network congestion games with unsplittable traffic demands, where a finite number of non-cooperative users select each a path from a specified source to a sink in an underlying network. The users experience a load-dependent latency on their chosen paths. Being selfish, they want to choose a minimum-latency path. The solution concept we study is that of a *pure Nash equilibrium*, where no user has an incentive to unilaterally switch to a different path. It is well-known that this type of game always has at least one pure Nash equilibrium [13].

The users induce a *social cost* to the system, which in this work we define as the total latency. Selfish behavior leads typically to suboptimal social cost at equilibrium. A long series of papers has studied the inefficiency of Nash equilibria for congestion games as quantified by the price of anarchy. See the surveys [10, 8] for an introduction to the very rich literature on the topic.

In order to offset the inefficiency of uncoordinated users, a common approach is to introduce fixed *taxes* (or *tolls*) on the edges of the network. The users will experience the taxes as part of their individual disutility, in addition to their latency. The aim is to design an *optimal* tax vector steering the selfish users to an equilibrium with desirable characteristics; in our case the desired target is minimum total latency.

Related Work. In the non-atomic setting, where there is an infinite number of users and each user controls an infinitesimal amount of traffic demand, the problem of designing optimal tax vectors has been studied extensively. A classic result going all the way back to Pigou [12] states that marginal cost taxes induce the optimal traffic pattern for homogeneous users [2]. A significant volume of recent work on optimal taxes for non-atomic congestion games considers the more intriguing and realistic case of *heterogeneous* users, which may have different valuations of time (latency) in terms of money (taxes). Yang and Huang [17] established the existence of optimal taxes for non-atomic asymmetric network congestion games[4] with heterogeneous users. Subsequently, their result was rediscovered by Fleischer, Jain, and Mahdian [5], and Karakostas and Kolliopoulos [9]. Previously the the single-source special case had been investigated by Cole, Dodis, and Roughgarden [4]. The existence of optimal taxes for non-atomic congestion games with heterogeneous users follows from Linear Programming duality, and thus an optimal tax vector can be computed efficiently by solving a linear program.

For non-atomic games, under mild assumptions on the latency functions the edge flow at equilibrium is unique. Hence the taxes of [2, 4, 5, 9, 17] induce the optimal solution as the unique edge flow of the equilibria of the game with taxes. On the other hand, atomic congestion games, even with splittable traffic, may admit many different Nash equilibria, possibly with different edge flows. Therefore, when considering atomic games, one has to distinguish between *weakly-optimal* tax vectors, for which at least one Nash equilibrium of the game with taxes minimizes the total latency, and *strongly-optimal* tax vectors, for which all Nash equilibria of the game with taxes minimize the total latency.

For atomic congestion games with splittable traffic and heterogeneous players, Swamy [14] proved that weakly-optimal tax vectors exist and can be computed efficiently by solving a convex program. As for atomic congestion games with unsplittable traffic, the existence and efficient computation of optimal taxes has been studied only in the restricted setting of homogeneous users. Caragiannis, Kaklamanis, and Kanellopoulos [3] considered atomic games with linear latency functions and homogeneous users, and investigated how much taxes can

---

[4] A network congestion game is symmetric if all users share the same source and sink and, in the case of atomic games, have the same traffic demand.

improve the price of anarchy. On the negative side, they established that if the users either do not share the same source and sink or have different traffic demands, then strongly-optimal taxes may not exist. In particular, Caragiannis et al. presented a non-symmetric game for which any tax vector induces a Nash equilibrium of total latency at least 6/5 times the optimum, and a parallel-link game with user-specific traffic demands for which any tax vector induces an equilibrium of total latency at least 9/8 times the optimum. On the positive side, they presented an efficient construction of strongly-optimal taxes for parallel-link games with linear latencies and unit-demand users. Subsequently, Fotakis and Spirakis [7] proved that weakly-optimal taxes exist and can be computed efficiently for atomic symmetric network congestion games, and that such taxes are strongly-optimal if the network is series-parallel.

Contribution. Despite the considerable interest in optimal taxes for non-atomic games with heterogeneous users and for atomic games with homogeneous users, it is unknown whether weakly- or strongly-optimal taxes exist for *atomic* network games with *heterogeneous* users. The case of heterogeneous users is substantially different, and more complicated, than that of homogeneous users, since the game with taxes is a congestion game with player-specific additive constants [11].

In this work, we study for the first time the existence of optimal taxes for atomic network games with heterogeneous users, and present two complementary and essentially best-possible results. On the positive side, we prove the existence of weakly-optimal taxes in single-source network congestion games with heterogeneous users (cf. Section 3). To establish this result, we follow the proof technique of [9], and show that any acyclic traffic pattern is induced as a Nash equilibrium of the game with the taxes calculated as in [9, Theorem 1]. Our result is significantly stronger that any previously known positive result on weakly-optimal taxes for atomic congestion games. In particular, our result generalizes previous results of [3,7] not only in the direction of considering heterogeneous users, but also in the direction of considering non-symmetric games on single-source multiple-sink networks.

On the negative side, we show that users' heterogeneity precludes the existence of strongly-optimal taxes even on the simplest topology of parallel-link networks. More specifically, we present a parallel-link game with linear latency functions and heterogeneous users for which any tax vector induces an equilibrium with total latency at least 28/27 times the optimum. Hence, we establish a dichotomy between the general case of heterogeneous users and the special case of homogeneous ones, as far as the existence of strongly-optimal taxes is concerned.

To the best of our knowledge, this is the first time in congestion games that a dichotomy is established (i) between the cases of homogeneous and heterogeneous users with respect to the existence of optimal taxes, and (ii) between the cases of non-atomic and atomic users on parallel links with respect to the efficiency of a price-of-anarchy-reducing mechanism. For the latter, we note that the worst-case price of anarchy for atomic games on parallel links is the same as the worst-case price of anarchy for non-atomic congestion games (see e.g. [15,6]), and

that the two classes of congestion games have similar behaviour with respect to their worst-case price of anarchy under some common price-of-anarchy-reducing mechanisms, such as Stackelberg strategies (see e.g. the bounds in [14, 6] on the efficiency of Stackelberg strategy LLF for non-atomic and atomic parallel-link games) and taxes for homogeneous users.

2   Preliminaries

We consider a network congestion game $\mathcal{G}(l)$ defined on a directed graph $G = (V, E)$ with a nondecreasing latency function $l_e : \mathbb{R}_+ \to \mathbb{R}_+$ on each edge $e \in E$. A set N of *users* is given, each with an amount of traffic (flow) to be routed from an origin node to a destination node of G. The users are *non-atomic* if each has infinitesimal demand and *atomic* otherwise. The game is single-source (resp. single-sink) if all users share the same origin (resp. destination) node, and *symmetric* if all users share the same origin-destination pair and have the same traffic demand.

Each user α has a positive *tax-sensitivity* factor $a(\alpha) > 0$. We will assume that the tax-sensitivity factors for all users come from a finite set of possible positive values. We call the users *heterogeneous* if there are at least two distinct sensitivity values and *homogeneous* otherwise. Unless we declare them explicitly to be heterogeneous, the users are assumed to be homogeneous. We can bunch together into a single *user class* all the users with the same origin-destination pair and with the same tax-sensitivity factor; let k be the number of different such classes. We denote by $d_i, \mathcal{P}_i, a(i)$ the total traffic demand of class i, the paths that can be used by class i, and the tax-sensitivity of class i, for all $i = 1, \ldots, k$ respectively. Thus each user in class i selects a path in $\mathcal{P}_i$ and routes her traffic though it. We set $\mathcal{P} \doteq \cup_{i=1,\ldots,k} \mathcal{P}_i$ the union of paths used by all classes. In the following, we assume that the game is single-source and the users are atomic and have unit demands, unless it is stated otherwise.

A configuration f is a tuple $f = (f^j)_{j \in N}$ consisting of a path f^j from the corresponding origin node to the corresponding destination node for each user j. Given a configuration f, we let f_P denote the total traffic routed through any path $P \in \mathcal{P}$, and let $f_e = \sum_{e \ni P} f_P$ denote the total traffic routed through any edge $e \in E$. Given a configuration f, we refer to the traffic vector $(f_e)_{e \in E}$ as the (edge-)*flow* induced by f. We note that different configurations may induce the same edge-flow. We say that a flow f is *feasible* (with respect to an atomic network congestion game $\mathcal{G}(l)$) if there is a configuration f of $\mathcal{G}(l)$ which routes traffic f_e through any edge e. We slightly abuse the notation by letting the same symbol denote both a configuration and the feasible flow induced by it. A configuration (or the corresponding flow) f is acyclic if for any cycle C in the underlying network G, there is an edge $e \in C$ with $f_e = 0$.

The *latency function* $l_e : \mathbb{R}_+ \to \mathbb{R}_+$ assigned to each edge e gives the latency experienced by any user on e due to the congestion caused by the traffic routed through e. We assume that the functions l_e are nondecreasing, and that $l_e(f_e) > 0$ when $f_e > 0$, i.e., the function l_e is *positive*.

For any configuration f and path $P \in \mathcal{P}$, the latency of P is $l_P(f) = \sum_{e \in P} l_e(f_e)$. The *individual cost* of a user j in a configuration f is $c^j(f) = \sum_{e \in f^j} l_e(f_e)$, i.e., the latency on her path in f. A configuration f is a *pure Nash equilibrium* of $\mathcal{G}(l)$ if no user can improve her individual cost by unilaterally deviating from f. Formally, for a tuple $x = (x_1, \ldots, x_n)$, let $x_{-i} = (x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)$ and $(x_{-i}, x_i') = (x_1, \ldots, x_{i-1}, x_i', x_{i+1}, \ldots, x_n)$. Configuration f is a pure Nash equilibrium if $c^j(f) \leq c^j(f_{-j}, P)$ for any user j in any class i and any path $P \in \mathcal{P}_i$.

A flow f satisfies the *Wardrop principle* [16] if for each class i, the latency on all paths in $\mathcal{P}_i$ used by f is no greater than the latency on any other path in $\mathcal{P}_i$. A non-atomic (atomic) *Wardrop equilibrium* is a (feasible) flow f that satisfies the Wardrop principle. We distinguish between atomic and non-atomic Wardrop equilibria, depending on whether the users are atomic or not. An atomic Wardrop equilibrium is also a pure Nash equilibrium, while the converse may not be true.

If every edge is assigned a *tax* (also called *toll*) $\beta_e \geq 0$, the resulting game is denoted as $\mathcal{G}(l + \beta)$. Given a configuration f in $\mathcal{G}(l + \beta)$, the individual cost of a user j included in a class i is: $c^j_\beta(f) = \sum_{e \in f^j} l_e(f_e) + a(i) \sum_{e \in f^j} \beta_e$.

Let $\hat{f}$ be a configuration that minimizes the total latency $\sum_e f_e l_e(f_e)$ over all configurations of $\mathcal{G}(l)$. Although in certain cases (e.g., when the functions $f_e l_e(f_e)$ are convex) the flow $\hat{f}$ can be computed efficiently, for more general latency functions it may be intractable to compute $\hat{f}$. We will assume that $\hat{f}$ is given to us off-line and that it induces a finite latency on every edge. A tax vector β *weakly induces* a feasible (non-atomic) flow f if f is a pure Nash (non-atomic Wardrop) equilibrium of $\mathcal{G}(l + \beta)$. A tax vector β is called *weakly-optimal* if it weakly induces a pure Nash equilibrium f whose total latency $\sum_{e \in E} f_e l_e(f_e)$ is equal to the optimal total latency $\sum_e \hat{f}_e l_e(\hat{f}_e)$. A tax vector β is called *strongly-optimal,* if every pure Nash equilibrium it induces in $\mathcal{G}(l + \beta)$ has total latency equal to the optimal total latency $\sum_e \hat{f}_e l_e(\hat{f}_e)$.

Let $F(x) = (F_1(x), F_2(x), \ldots, F_n(x))$ be a vector-valued function from the n-dimensional space $\mathbb{R}^n$ into itself. Then the *nonlinear complementarity problem* of mathematical programming is to find a vector x that satisfies the following system:

$$x^T F(x) = 0, \quad x \geq 0, \quad F(x) \geq 0.$$

3   Existence of Weakly-Optimal Taxes

In this section we consider networks with a single-source s and heterogeneous users. Each user class i consists of a single user who wishes to route d_i units of traffic through a single $s - t_i$ path. We show that if $d_i = 1$ (or more generally, if d_i are arbitrary and the optimal configuration is acyclic), there exists a vector of weakly-optimal taxes. In particular, we establish the existence of a tax vector that weakly induces any acyclic flow $\hat{f}$ as an atomic Wardrop equilibrium. Since single-source network congestion games with unit-demand users admit an acyclic optimal flow $\hat{f}$, this implies the existence of weakly optimal taxes for such games.

The proof follows closely [9], where the existence of weakly-optimal taxes is shown for the non-atomic case, and here we give a sketch with the new elements added for our case. In [9], it is shown that, if we add to the network artificial *capacity constraints,* $f_e \le \hat{f}_e$, $\forall e \in E$, there is a tax-vector β^* that induces as a *non-atomic* Wardrop equilibrium a flow f^* that satisfies demands d_i and respects the capacities. In particular, [9] shows that the following nonlinear complementarity problem always has a solution (details omitted). Moreover, if $\hat{f}$ is given offline, this solution can be computed in polynomial time.

$$f_P(T_P(f) - u_i) = 0 \qquad \forall i, \, \forall P \in \mathcal{P}_i \qquad\qquad \text{(BIG CP)}$$
$$T_P(f) \ge u_i \qquad \forall i, \, \forall P \in \mathcal{P}_i$$
$$u_i\Big(\sum_{P \in \mathcal{P}_i} f_P - d_i\Big) = 0 \qquad \forall i$$
$$\sum_{P \in \mathcal{P}_i} f_P \ge d_i \qquad \forall i$$
$$\beta_e(f_e - \hat{f}_e) = 0 \qquad \forall e \in E$$
$$f_e \le \hat{f}_e \qquad \forall e \in E$$
$$f_P, \beta_e, u_i \ge 0 \qquad \forall P, e, i$$

Here the function $T_P(f)$ is set to $l_P(f)/a(i) + \sum_{e \in P} \beta_e^*$, $\forall P \in \mathcal{P}_i$, $\forall i$.

Lemma 1. *Let $\hat{f}$ be an acyclic feasible flow for demands d_i, and let (f^*, β^*, u^*) be any solution of (BIG CP). Then $\sum_{P \in \mathcal{P}_i} f_P^* = d_i$, $\forall i$ and $f_e^* = \hat{f}_e$, $\forall e \in E$.*

Proof. The proof of the first part is essentially the same as the proof by contradiction of Proposition 4.1 in [1] and is omitted.

Vector f^* is a non-atomic flow, that satisfies the following set of constraints:

$$\sum_{P \in \mathcal{P}_i} f_P = d_i \qquad\qquad \forall i \in \{1, \dots, k\} \qquad\qquad (1)$$

$$f_e = \sum_{P \in \mathcal{P}: e \in P} f_P \quad \forall e \in E \qquad\qquad (2)$$

$$f_e \le \hat{f}_e \qquad\qquad \forall e \in E \qquad\qquad (3)$$
$$f_P \ge 0 \qquad\qquad \forall P \in \mathcal{P} \qquad\qquad (4)$$

Consider the network which consists only of the edges e of G with $\hat{f}_e > 0$. Augment this network by adding a super-sink t and an edge (t_i, t) from each of the old sinks to t. Call $G_{\hat{f}}$ the resulting network. Extend f^* to an s-t flow in $G_{\hat{f}}$ by setting $f^*_{(t_i,t)} = d_i$. Let (S, T) be any cut that separates s from t in $G_{\hat{f}}$. Since $\hat{f}$ is acyclic, it must be that $\sum_{e \in \delta(S)} \hat{f}_e = \sum_{i=1}^{k} d_i$. Because of (1), it must be that $\sum_{e \in \delta(S)} f_e^* \ge \sum_{i=1}^{k} d_i = \sum_{e \in \delta(S)} \hat{f}_e$. By the capacity constraints (3), we conclude that $\sum_{e \in \delta(S)} f_e^* = \sum_{e \in \delta(S)} \hat{f}_e$, and in particular, that $f_e^* = \hat{f}_e$

for all edges e that cross the cut. The only edges of G on which f^* might send positive flow are the edges of $G_{\hat{f}}$. Any such edge e belongs to at least one s-t cut in $G_{\hat{f}}$. By applying the previous argument to such a cut, it follows that $f_e^* = \hat{f}_e, \ \forall e \in E$. □

Then [9] establishes that we can use β^* as a tax vector to weakly induce f^* as a non-atomic Wardrop equilibrium in the original network *without* the capacity constraints. This follows from the fact that we can use β^* as a tax vector to ensure that (f^*, u^*) is also a solution to the following complementarity problem:

$$
\begin{aligned}
(T_P(f) - u_i) f_P &= 0 & \forall i, \ \forall P \in \mathcal{P}_i & \qquad \text{(CP)}\\
T_P(f) - u_i &\geq 0 & \forall i, \ \forall P \in \mathcal{P}_i &\\
u_i\Big(\sum_{P \in \mathcal{P}_i} f_P - d_i\Big) &= 0 & \forall i &\\
\sum_{P \in \mathcal{P}_i} f_P &\geq d_i & \forall i &\\
f, u &\geq 0 & &
\end{aligned}
$$

If the path latency functions are continuous and positive, Aashtiani and Magnanti [1] show that the Wardrop equilibria of the game $\mathcal{G}(l + \beta)$ can be described as the solutions to (CP). T_P above denotes the cost of a user that uses path P, f_P is the flow through path P, and $u = (u_1, \ldots, u_k)$ is the vector of shortest travel times for the commodities. The first two equations model Wardrop's principle by requiring that for any origin-destination pair i, the travel cost for all paths in $\mathcal{P}_i$ with nonzero flow is the same and equal to u_i. The remaining equations ensure that the demands are met and that the variables are nonnegative.

The fact that for all $e \in E$, $f_e^* = \hat{f}_e$ proves that the tax vector β^* we compute weakly induces as an equilibrium the *atomic* solution $\hat{f}$ as well. We have thus shown the following theorem, which is the main result of this section.

Theorem 1. *Let all atomic heterogeneous users share the same source, and let $\hat{f}$ be any acyclic feasible flow. If for every edge $e \in E$, $l_e()$ is a nondecreasing positive function, then there is a tax vector $\beta \in \mathbb{R}_+^{|E|}$ such that, there is an atomic Wardrop traffic equilibrium $\bar{f}$ for the game $\mathcal{G}(l + \beta)$, where $\bar{f}_e = \hat{f}_e, \ \forall e \in E$. Given $\hat{f}$, β can be computed in polynomial time.*

If the latency functions are strictly increasing, the uniqueness results from [1] yield that $\hat{f}$ is the only Wardrop atomic equilibrium induced by the tax vector of the theorem.

Single-source network congestion games with unit-demand users and nondecreasing latency functions admit an acyclic optimal flow $\hat{f}$. Moreover, if for all $e \in E$, $xl_e(x)$ are convex, such an optimal flow can be computed in polynomial time by a min-cost flow computation. Therefore, we obtain the following corollary of Theorem 1:

Corollary 1. *Let $\mathcal{G}(l)$ be an atomic network congestion with nondecreasing latency functions and heterogeneous users, where all users share the same source and have the same traffic demand. Then $\mathcal{G}(l)$ admits a weakly-optimal tax vector β. Furthermore, if for all edges e, $xl_e(x)$ is convex, β can be computed in polynomial time.*

Theorem 1 states that computing the weakly-optimal tax vector β for an acyclic optimal flow $\hat{f}$ is not substantially harder than computing $\hat{f}$: if $\hat{f}$ can be computed in polynomial time, β can also be computed in polynomial time. An interesting question is whether computing the tax vector β of Theorem 1 is substantially *easier* than computing the corresponding acyclic optimal flow $\hat{f}$. The following theorem practically excludes this possibility. In particular, we show that given the weakly-optimal tax vector β of Theorem 1, we can decide in polynomial time whether the optimal total latency is bounded from above by a given number. So the problem of computing the weakly-optimal tax vector β is at least as difficult as the problem of determining the optimal total latency.

Theorem 2. *For atomic games with user-specific demands, if the optimal flow $\hat{f}$ is not given, it is NP-hard to compute the taxes whose existence is established by Theorem 1. This holds even for parallel-link games with homogeneous users.*

Proof. We employ a Turing reduction from PARTITION. We consider an instance of the decision version of PARTITION, i.e., a set of integers $\{a_1, \ldots, a_n\}$ whose total sum is $2B$ for some $B > 0$. For every integer a_i, we create a user with demand a_i and tax-sensitivity 1. Every partition of the users into two sets, one with total sum $B - T$ and the other with total sum $B + T$ for some $T \geq 0$, induces in the network with two parallel links and latency function $l(x) = x$, a corresponding acyclic routing of the users whose total cost is

$$(B - T)^2 + (B + T)^2 = 2B^2 + 2T^2.$$

This quantity is minimized for $T = 0$, i.e., when the PARTITION instance is a YES-instance. Assume now that you can compute in polynomial time the taxes of Theorem 1. Because the latency functions are strictly increasing, the Wardrop equilibrium is unique in terms of edge flows [1]. Moreover, it is well-known that the equilibrium solution f can be computed in polynomial time by solving a convex quadratic mathematical program [2]. By Theorem 1, on each of the two parallel edges, the value of f will be equal to value of the optimal unsplittable solution. Checking these values, we can determine whether the PARTITION instance is a YES-instance.   □

Unfortunately, it is known that the taxes of Theorem 1 are not in general strongly-optimal. Note that for homogeneous users, our taxes are cost-balancing in the sense of Fotakis and Spirakis [7]. They give an example of a symmetric network congestion game, with homogeneous users, where the cost-balancing taxes induce an a pure Nash equilibrium of total latency 1.13 times the optimum. In the full version of the paper we give another such example where the cost-balancing taxes induce a pure Nash equilibrium of total latency $(1.2 - \epsilon)$ times the optimum.

4   Inexistence of Strongly-Optimal Taxes

We proceed to show that that atomic congestion games with heterogeneous users may not admit strongly-optimal taxes, even for parallel-link games with linear latencies and unit-demand users.

Theorem 3. *There exists a parallel-link game with linear latencies and heterogeneous unit-demand users, for which any tax vector induces an equilibrium with total latency at least 28/27 times the optimal total latency.*

Proof. We consider a game $\mathcal{G}(l)$ on 3 parallel links with latency functions $l_1(x) = 7$, $l_2(x) = 2x$, and $l_3(x) = x + 1$. There are 6 unit-demand users, 2 users with tax-sensitivity 1 and 4 users with tax-sensitivity $1/2$. The unique optimal flow assigns a single user to link 1, 2 users to link 2, and 3 users to link 3, and achieves a total latency of 27. Any other feasible flow has total latency at least 28. In the following, we show that any weakly-optimal tax vector β induces an equilibrium of total latency at least 28, and thus this game does not admit strongly-optimal taxes. The proof proceeds by considering different cases depending on the 5 optimal allocations of heterogeneous users.

Case I: We consider an optimal flow that assigns a user with tax-sensitivity 1 to link 1, the other user with tax-sensitivity 1 and a user with tax-sensitivity $1/2$ to link 2, and 3 users with tax-sensitivity $1/2$ to link 3 (we denote such a configuration as $\langle (1), (1, 1/2), (1/2, 1/2, 1/2) \rangle$). Let $\beta = (\beta_1, \beta_2, \beta_3)$ be any (weakly-optimal) tax vector that induces the particular configuration as an equilibrium of $\mathcal{G}(l + \beta)$. No user has an incentive to deviate from its assigned link; writing down the corresponding inequalities, we obtain that β must satisfy the following:

$$1 + \beta_1 \leq \beta_2 \leq 3 + \beta_1 \tag{5}$$
$$\beta_2 - 1 \leq \beta_3 \leq 4 + \beta_2 \tag{6}$$
$$2 + \beta_1 \leq \beta_3 \leq 6 + \beta_1 \tag{7}$$

If β is strongly-optimal, configuration $\langle (1, 1), (1/2, 1/2), (1/2, 1/2) \rangle$ is not an equilibrium of $\mathcal{G}(l + \beta)$. Therefore at least one user in that configuration has an incentive to deviate, and β must satisfy *at least one* of the following:

$$\beta_2 < 1 + \beta_1 \quad (8) \qquad\qquad\qquad \beta_3 < \beta_2 \quad (11)$$
$$\beta_3 < 3 + \beta_1 \quad (9) \qquad\qquad 8 + \beta_1 < \beta_3 \quad (12)$$
$$6 + \beta_1 < \beta_2 \quad (10) \qquad\qquad 6 + \beta_2 < \beta_3 \quad (13)$$

We observe that (8) contradicts (5), (10) contradicts (5), (12) contradicts (7), and (13) contradicts (6). Hence, if β is strongly optimal, either $\beta_3 < 3 + \beta_1$ or $\beta_3 < \beta_2$ (ie. β_3 must be "small").

Moreover, if β is strongly-optimal, $\langle (\ ), (1, 1), (1/2, 1/2, 1/2, 1/2) \rangle$ is not an equilibrium of $\mathcal{G}(l + \beta)$, and β must satisfy at least one of the following:

$$3 + \beta_1 < \beta_2 \quad\qquad (14) \qquad\qquad 4 + \beta_1 < \beta_3 \quad\qquad (16)$$
$$\beta_3 < \beta_2 - 2 \quad\qquad (15) \qquad\qquad 2 + \beta_2 < \beta_3 \quad\qquad (17)$$

We observe that (14) contradicts (5) and (15) contradicts (6). Hence, if β is strongly optimal, either $\beta_3 > 4 + \beta_1$ or $\beta_3 > 2 + \beta_2$ (ie. β_3 must be "large").

If $\beta_3 < 3 + \beta_1$, neither $\beta_3 > 4 + \beta_1$ nor $\beta_3 > 2 + \beta_2$ is possible (note that $3 + \beta_1 > \beta_3 > 2 + \beta_2$, which contradicts (5)). If $\beta_3 < \beta_2$, neither $\beta_3 > 2 + \beta_2$ nor $\beta_3 > 4 + \beta_1$ is possible (note that $\beta_2 > \beta_3 > 4 + \beta_1$, which contradicts (5)). Therefore, any tax vector that induces optimal configuration $\langle (1), (1, 1/2), (1/2, 1/2, 1/2) \rangle$ as an equilibrium of $\mathcal{G}(l + \beta)$ also induces either $\langle (1, 1), (1/2, 1/2), (1/2, 1/2) \rangle$ or $\langle (\,), (1, 1), (1/2, 1/2, 1/2, 1/2) \rangle$ (both of total latency 28) as an equilibrium.

Case II: We consider optimal configuration $\langle (1), (1/2, 1/2), (1, 1/2, 1/2) \rangle$. Working as in Case I, we obtain that any tax vector $\beta = (\beta_1, \beta_2, \beta_3)$ that induces this configuration as an equilibrium of $\mathcal{G}(l + \beta)$ must satisfy the following:

$$1 + \beta_1 \leq \beta_2 \leq 5 + \beta_1 \tag{18}$$

$$\beta_2 - 2 \leq \beta_3 \leq 2 + \beta_2 \tag{19}$$

$$2 + \beta_1 \leq \beta_3 \leq 3 + \beta_1 \tag{20}$$

In fact, the right-hand side of (18) follows from $\beta_2 - 2 \leq \beta_3 \leq 3 + \beta_1$.

Considering configuration $\langle (1, 1), (1/2, 1/2), (1/2, 1/2) \rangle$ and working as in Case I, we obtain that if β is strongly-optimal, either $\beta_3 < 3 + \beta_1$ or $\beta_3 < \beta_2$. Considering configuration $\langle (\,), (1, 1), (1/2, 1/2, 1/2, 1/2) \rangle$, we obtain that if β is strongly-optimal, either $\beta_2 > 3 + \beta_1$ (note that (14) does not contradict (18)), or $\beta_3 > 4 + \beta_1$, or $\beta_3 > 2 + \beta_2$.

Working as in Case I, we show that if β is strongly-optimal, it must satisfy both $\beta_2 > 3 + \beta_1$ and $\beta_3 < \beta_2$ (since $\beta_2 > 3 + \beta_1$, $\beta_3 < 3 + \beta_1$ implies $\beta_3 < \beta_2$, so β_3 must be smaller than β_2 in any case), in addition to (18), (19), (20).

Moreover, if β is strongly-optimal, configuration $\langle (1, 1), (1/2), (1/2, 1/2, 1/2) \rangle$ is not an equilibrium of $\mathcal{G}(l + \beta)$, and β must satisfy at least one of the following:

$$\beta_2 < 3 + \beta_1 \quad (21) \qquad\qquad 6 + \beta_3 < \beta_2 \quad (24)$$

$$\beta_3 < 2 + \beta_1 \quad (22) \qquad\qquad 6 + \beta_1 < \beta_3 \quad (25)$$

$$10 + \beta_1 < \beta_2 \quad (23) \qquad\qquad \beta_2 < \beta_3 \quad (26)$$

We observe that (21) contradicts $\beta_2 > 3 + \beta_1$, (22) contradicts (20), (23) contradicts (18), (25) contradicts (20), and (26) contradicts $\beta_3 < \beta_2$. Furthermore, (18) and (19) imply that $\beta_2 \leq 5 + \beta_1 \leq 3 + \beta_3$, which contradicts (24). Hence, any tax vector that induces optimal configuration $\langle (1), (1/2, 1/2), (1, 1/2, 1/2) \rangle$ as an equilibrium also induces either configuration $\langle (1, 1), (1/2, 1/2), (1/2, 1/2) \rangle$, or $\langle (\,), (1, 1), (1/2, 1/2, 1/2, 1/2) \rangle$, or $\langle (1, 1), (1/2), (1/2, 1/2, 1/2) \rangle$ (all of total latency 28) as an equilibrium.

Case III: We consider optimal configuration $\langle (1/2), (1, 1), (1/2, 1/2, 1/2) \rangle$. Any tax vector $\beta = (\beta_1, \beta_2, \beta_3)$ that induces this configuration as an equilibrium of $\mathcal{G}(l + \beta)$ must satisfy the following:

$$2 + \beta_1 \leq \beta_2 \leq 3 + \beta_1 \tag{27}$$

$$\beta_2 - 1 \leq \beta_3 \leq 4 + \beta_2 \tag{28}$$

$$4 + \beta_1 \leq \beta_3 \leq 6 + \beta_1 \tag{29}$$

Therefore, any tax vector that induces $\langle(1/2),(1,1),(1/2,1/2,1/2)\rangle$ as an equilibrium of $\mathcal{G}(l+\beta)$ must satisfy (5), (6), and (7), and by Case I, is not strongly-optimal.

Cases IV and V: Optimal configurations $\langle(1/2),(1,1/2),(1,1/2,1/2)\rangle$ and $\langle(1/2),(1/2,1/2),(1,1,1/2)\rangle$ are not induced as an equilibrium of $\mathcal{G}(l+\beta)$ by any tax vectors. In particular, applying the inequalities for possible deviations between link 1 and link 3, we obtain that any tax vector β that induces any of the configurations above as an equilibrium must satisfy $4+\beta_1 \le \beta_3 \le 3+\beta_1$.

Thus we have considered all optimal allocations of heterogeneous users and all weakly-optimal tax vectors β, and have shown that any of them induces a configuration of total latency at least 28 as an equilibrium of $\mathcal{G}(l+\beta)$. $\square$

Remark 1. For the atomic game with homogeneous users corresponding to the parallel-link game in the proof of Theorem 3, the tax vector $(0,3-\delta,3-\delta)$, for a sufficiently small $\delta > 0$, is a strongly-optimal tax vector (a slightly different strongly-optimal tax vector is given by [3, Theorem 1]). For the corresponding non-atomic game with heterogeneous users, the tax vector $(0,3,3)$ is a strongly-optimal one.

5 Open problems

It is known that for homogeneous users with unit-demands on multicommodity networks there exist no strongly-optimal taxes [3]. Series-parallel networks is the largest class for which such taxes have been shown so far to exist [7]. In this work, we established that when the users are heterogeneous, there are no strongly-optimal taxes even on the very specialized topology of parallel links. The challenging open problem stated in [3] remains for future work: determine the largest class of network congestion games for which strongly-optimal taxes exist. The candidate class is that of symmetric network games [3], i.e., when users are homogeneous, have identical demands, and share the same source and destination on a general-topology network.

References

1. H. Z. Aashtiani and T. L. Magnanti. Equilibria on a congested transportation network. *SIAM Journal of Algebraic and Discrete Methods*, 2:213–226, 1981.
2. M. Beckmann, C. B. McGuire, and C. B. Winsten. *Studies in the Economics of Transportation*. Yale University Press, 1956.
3. I. Caragiannis, C. Kaklamanis, and P. Kanellopoulos. Taxes for linear atomic congestion games. In *Proceedings of the 14th Annual European Symposium on Algorithms*, pages 184–195, 2006.

4. R. Cole, Y. Dodis, and T. Roughgarden. Pricing network edges for heterogeneous selfish users. In *Proceedings of the 35th Annual ACM Symposium on Theory of Computing*, pages 521–530, 2003.
5. L. Fleischer, K. Jain, and M. Mahdian. Tolls for heterogeneous selfish users in multicommodity networks and generalized congestion games. In *Proceedings of the 45th Annual IEEE Symposium on Foundations of Computer Science*, pages 277–285, 2004.
6. D. Fotakis. Stackelberg strategies for atomic congestion games. *Theory of Computing Systems*, 47(1): 218–249, 2010.
7. D. Fotakis and P. G. Spirakis. Cost-balancing tolls for atomic network congestion games. *Internet Mathematics*, 5(4): 343–364, 2008.
8. M. Gairing, T. Lücking, B. Monien, and K. Tiemann. Nash Equilibria, the Price of Anarchy and the Fully Mixed Nash Equilibrium Conjecture. In *Proceedings of the 32th International Colloquium on Automata, Languages and Programming (ICALP)*, pages 51–65, 2005.
9. G. Karakostas and S. G. Kolliopoulos. Edge pricing of multicommodity networks for heteregoneous selfish users. In *Proceedings of the 45th Annual IEEE Symposium on Foundations of Computer Science*, pages 268–276, 2004.
10. S. C. Kontogiannis and P. G. Spirakis. Atomic selfish routing in networks: a survey. In *Proceedings of the 1st International Workshop, Internet and Network Economics, (WINE)*, pages 989–1002, 2005.
11. M. Mavronicolas, I. Milchtaich, B. Monien, and K. Tiemann. Congestion games with player-specific constants. In *Proceedings of the 32nd Symposium on Mathematical Foundations of Computer Science, (MFCS)*, pages 633–644, 2007.
12. A. C. Pigou. *The Economics of Welfare*. Macmillan and Co., London, 1920.
13. R. W. Rosenthal. A class of games posessing pure Nash strategy equilibria. *International Journal of Game Theory*, 2:65–67, 1973.
14. C. Swamy. The effectiveness of Stackelberg strategies and tolls for network congestion games. In *Proceedings of the 18th ACM-SIAM Symposium on Discrete Algorithms, (SODA)*, pages 1133–1142, 2007.
15. T. Roughgarden. The price of anarchy is independent of the network topology. *Journal of Computer and System Sciences*, 67(2):341–364, 2003.
16. J. G. Wardrop. Some theoretical aspects of road traffic research. *Proc. Inst. Civil Engineers, Part II*, 1:325–378, 1952.
17. H. Yang and H.-J. Huang. The multi-class, multi-criteria traffic network equilibrium and systems optimum problem. *Transportation Research B*, 38:1–15, 2004.