

# Truthful Mechanisms for Selfish Routing and Two-Parameter Agents

Clemens Thielen

Department of Mathematics, University of Kaiserslautern, Paul-Ehrlich-Str. 14,  
D-67663 Kaiserslautern, Germany  
`thielen@mathematik.uni-kl.de`

**Abstract.** We prove a general monotonicity result about Nash flows in directed networks and use it for the design of *truthful mechanisms* in the setting where each edge of the network is controlled by a different selfish agent, who incurs costs when her edge is used. The costs for each edge are assumed to be linear in the load on the edge. To compensate for these costs, the agents impose tolls for the usage of edges. When nonatomic selfish network users choose their paths through the network independently and each user tries to minimize a weighted sum of her latency and the toll she has to pay to the edges, a Nash flow is obtained. Our monotonicity result implies that the load on an edge in this setting can not increase when the toll on the edge is increased, so the assignment of load to the edges by a Nash flow yields a *monotone algorithm*. By a well-known result, the monotonicity of the algorithm then allows us to design truthful mechanisms based on the load assignment by Nash flows.

Moreover, we consider a mechanism design setting with *two-parameter agents*, which is a generalization of the case of one-parameter agents considered by Archer and Tardos [1]. While the private data of an agent in the one-parameter case consists of a single nonnegative real number specifying the agent's cost per unit of load assigned to her, the private data of a *two-parameter agent* consists of a *pair* of nonnegative real numbers, where the first one specifies the cost of the agent per unit load as in the one-parameter case, and the second one specifies a fixed cost, which the agent incurs independently of the load assignment.

We give a complete characterization of the set of output functions that can be turned into truthful mechanisms for two-parameter agents. Namely, we prove that an output function for the two-parameter setting can be turned into a truthful mechanism if and only if the load assigned to every agent is nonincreasing in the agent's bid for her per unit cost and, for almost all fixed bids for the agent's per unit cost, the load assigned to her is independent of the agent's bid for her fixed cost. When the load assigned to an agent is continuous in the agent's bid for her per unit cost, it must be completely independent of the agent's bid for her fixed cost. These results motivate our choice of linear cost functions without fixed costs for the edges in the selfish routing setting, but the results also seem to be interesting in the context of algorithmic mechanism design themselves.

*Keywords:* algorithmic game theory, mechanism design, selfish routing

## 1 Introduction

The emergence of the Internet as a new platform for distributed computing at the end of the last century has greatly influenced the point of view of algorithm designers. The machines participating on the Internet are controlled by different users that are more likely to selfishly maximize their own profit rather than any global objective. Hence, the old implicit assumption that an algorithm can make definitive decisions that are always carried out by the machines can no longer be taken for granted since lying and manipulating the system could increase the profit of individual users. This possible introduction of false information can lead to a severe loss of performance in classical optimization algorithms. A common approach for dealing with this problem is to study algorithmic problems using methods from the microeconomic field of *mechanism design* (also known as *implementation theory*), which studies how privately known preferences of several people can be aggregated towards a social choice. The resulting research area is commonly termed *algorithmic mechanism design*.

The main idea in algorithmic mechanism design is to overcome the selfishness of the individual users (called *agents*) by making it a (weakly) dominant strategy for each agent to truthfully reveal her private information (called the agent's *type*). This is achieved by paying the agents according to a suitable *payment scheme*. A *mechanism* is then defined as a pair  $M = (\mathcal{A}, \mathcal{P})$  consisting of an (optimization) algorithm  $\mathcal{A}$  and a payment scheme  $\mathcal{P}$ , and the profit of each agent in the mechanism is given as the sum of the profit she makes from the solution of the global optimization problem chosen by the algorithm and the payment she receives. A mechanism is called *truthful with dominant strategies* (in the sequel simply *truthful*) or *strategyproof* if truthtelling is a dominant strategy for every agent, i.e., it maximizes the profit of the agent for every possible behavior of the other agents. In a truthful mechanism, the rational agents can be assumed to give correct information about their types, so the underlying optimization problem can be solved on the correct data. Thus, the design of truthful mechanisms is the central goal in algorithmic mechanism design. However, not every algorithm admits a payment scheme such that the resulting mechanism is truthful. Hence, the question “Which algorithms can be turned into truthful mechanisms?” is one of the central questions in algorithmic game theory these days.

*Selfish routing* is another important research area in algorithmic game theory. It is motivated by the study of routing in road networks or the Internet, where the traffic is not controlled by a global authority but by many different selfish users. Hence, the classical problem of network routing is considered from the perspective of game theory, which gives rise to a wide range of interesting issues. Different players are assumed to selfishly route traffic through a network in an attempt to optimize their own objective functions (e.g., their own latency or cost). However, as in the case of mechanism design, selfishness may lead to outcomes that are suboptimal from a global perspective.

## 1.1 Previous Results

Mechanism design is a classical area of research with many results. It is a subfield of noncooperative game theory [2] and microeconomics [3]. An introduction to the subject can be found in Chapter 23 of [3]. An important result known as the *revelation principle* (cf. [3], page 871) states that it imposes no loss of generality to consider only mechanisms in which the agents' strategies are to simply report their type (though not necessarily in a truthful manner) as we implicitly assumed in the introduction. Such mechanisms are called *direct revelation mechanisms*. The systematic study of algorithmic problems in the context of mechanism design was initiated by a seminal paper of Nisan and Ronen [4], who also introduced the formal framework for algorithmic mechanism design presented above.

Archer and Tardos [1] considered the important case of algorithmic mechanism design for *one-parameter agents*. In this setting, the type of each agent  $i$  is a single nonnegative real number  $t_i$ . Each feasible solution  $x$  of the global optimization problem results in an amount of *work*  $w_i(x)$  being assigned to agent  $i$ , who incurs a *cost* of  $t_i \cdot w_i(x)$  for completing this amount of work. The profit of agent  $i$  is defined as the payment she receives from the mechanism minus her cost. Archer and Tardos [1] showed that an algorithm  $\mathcal{A}$  for an optimization problem with one-parameter agents can be used in a truthful mechanism  $M = (\mathcal{A}, \mathcal{P})$  if and only if  $\mathcal{A}$  is *monotone*.

**Definition 1.** *An algorithm for an optimization problem with one-parameter agents is called monotone if, for every agent, the amount of work assigned to it does not increase if its bid increases. More formally, an algorithm is monotone if, given two vectors  $b, b'$  of length  $m = \#$ agents that represent sets of  $m$  bids and differ only in one component  $i$ , i.e.,  $b_i > b'_i$  and  $b_j = b'_j$  for  $j \neq i$ , the amount of work (the load) that agent  $i$  gets from the algorithm if the bid vector is  $b$  is never higher than if the bid vector is  $b'$ .*

A similar result for a weaker notion of truthfulness called *truthfulness in Bayesian Nash equilibrium* was previously obtained by Myerson [5], who considered auctions of a single object as Bayesian games, where the players are the bidders, and the private information of each bidder is a nonnegative real number specifying the maximal price she would be willing to pay for the object.

The result of Archer and Tardos is a strong motivation for considering monotone algorithms since it implies that every monotone algorithm for a problem with one-parameter agents immediately yields a truthful mechanism. Using this result, monotone algorithms (and, hence, truthful mechanisms) were designed for several classical problems like scheduling related machines to minimize the makespan, where the bid of a machine is the inverse of its speed [1, 6–9].

Archer and Tardos also considered the concept of *voluntary participation* (also known as *participation constraints* or *individual rationality constraints* in the mechanism design literature). A mechanism satisfies voluntary participation if agents who bid truthfully never incur a net loss, which means that it makes sense for the agents to participate in the mechanism without being forced to do

so. As proved in [1], a monotone algorithm admits a truthful payment scheme satisfying voluntary participation if and only if, for every  $i$  and every fixed vector of bids of all agents except  $i$ , the integral of the work curve of agent  $i$  is finite.

Selfish routing and, more general, network games are active research areas these days. A survey of these topics can be found in [10]. The book [11] gives a good introduction to selfish routing. Much work on selfish routing and network games in recent years has focused on quantifying the loss of efficiency due to selfishness. The most common game-theoretic approach in this area is to consider *Nash equilibria* [12], i.e., solutions of noncooperative games in which no player has an incentive to unilaterally change her strategy. In *nonatomic* models of selfish routing, where the traffic routed by the selfish network users is modeled as a network flow, Nash equilibria are commonly referred to as *Nash flows*. To measure the loss of efficiency caused by selfishness, Papadimitriou [13] introduced the term *price of anarchy* for the ratio between the worst-case Nash equilibrium and the globally optimal solution in game-like situations. Results concerning the price of anarchy for network games can be found in [11, 14–16]. Other work related to this paper investigates how Nash equilibria in selfish routing can be influenced by tolls on the network edges [14, 17–20]. Monotonicity results about Nash flows have previously been obtained in [21].

Note that most of the given references concentrate on *nonatomic* models of selfish routing as studied in this paper. A survey of results for a common *atomic* model introduced by Koutsoupias and Papadimitriou [22] and a comparison to the nonatomic case can be found in [23].

## 1.2 Our Results

We prove a general monotonicity result about Nash flows in directed networks, which states that the Nash flow on an edge can not increase when the cost of the edge to the network users is increased. This result generalizes a result of Dafermos and Nagurney [21], who studied a model of selfish routing equivalent to the one considered in this paper. However, the analysis in [21] crucially relies on the so called *strong monotonicity condition* for the cost functions of the network, which is a rather strong assumption. In particular, Dafermos and Nagurney [21] only proved the monotonicity of Nash flows for the case of *strictly increasing* cost functions on the network edges. We do *not* use the strong monotonicity condition and our monotonicity result holds true in the more general setting of nondecreasing cost functions.

We use our monotonicity result for the design of truthful mechanisms in a setting with two classes of selfish agents. The first class consists of the owners of the network edges. The owner of each edge incurs a cost when her edge is used, and this cost is assumed to be linear in the load on the edge. To compensate for these costs, the owner of an edge imposes a toll on her edge, which every single user of the edge has to pay independently of the load on the edge. The second class of selfish agents is given by the users of the network. Each selfish user tries choose a path through the network that minimizes the (weighted) sum

of the user's latency and the overall toll she has to pay to the edges. Our result about Nash flows implies that, when considering the toll defined by each owner of an edge as a bid for her privately known cost per unit load, the assignment of load to the edges by a Nash flow yields a *monotone* algorithm. Hence, by the monotonicity result of Archer and Tardos [1], this algorithm can be used in a truthful mechanism. Thus, our result connects the research areas of mechanism design and selfish routing, which are two of the main research topics in algorithmic game theory these days. Moreover, we show that our monotonicity result about Nash flows can be extended to the more general setting of Nash equilibria in *nonatomic congestion games* without modification in the proof. Hence, the result may be used for the construction of further truthful mechanisms in this more general setting in the future.

We present two truthful mechanisms for our selfish routing setting. The first mechanism assumes strictly increasing latency functions, where the Nash flow in a network is essentially unique. The second one is a randomized mechanism, which is truthful in expectation and works in the more general setting where the latency functions of the edges are only assumed to be nondecreasing. Both mechanisms guarantee complete cooperation of both kinds of selfish agents (the edges and the network users) with the mechanism in the sense that no agent has an incentive to manipulate the mechanism.

Moreover, we show that, under some additional assumptions, it is possible to enforce *optimal tolls* by making it a dominant strategy for each owner of an edge to set her toll in such a way that the Nash flow induced by the tolls on the edges has minimal total latency. We show how our selfish routing mechanism for the case of strictly increasing cost functions can be modified to enforce optimal tolls and give bounds on the payments to the edges in this mechanism.

We motivate the choice of linear cost functions for the edges by proving results about mechanisms for the case of *two-parameter agents*. This setting generalizes the one-parameter setting considered by Archer and Tardos by allowing a fixed cost component. Namely, we consider the situation where the costs of an agent are given as a constant per unit cost times the load assigned to the agent plus an additional fixed cost, which the agent incurs independently of the load assignment. These two real numbers defining an agent's cost function are the agent's private data. This kind of cost functions would be desirable, e.g., when the edges of the network in the selfish routing setting above are thought of as roads or links of a telecommunication network, where a large part of the total cost of an edge is given as the fixed cost for building and maintaining the road or link.

We show that, for almost all fixed bids for an agent's per unit cost, the load assigned to the agent in a truthful mechanism has to be independent of the agent's bid for her fixed cost. Moreover, when the load assigned to an agent is continuous in the agent's bid for her per unit cost, it must be completely independent of the agent's bid for her fixed cost, so the situation essentially reduces to the one-parameter setting. Together with the monotonicity of the load assigned to an agent in the agent's bid for her per unit cost as in the one-parameter setting our necessary condition for truthfulness turns out to be

sufficient as well, so we obtain a complete characterization of the set of output functions that can be turned into truthful mechanisms for two-parameter agents. Furthermore, this characterization implies that no truthful mechanism in the two-parameter setting can satisfy voluntary participation. Hence, our results imply that considering additional fixed costs for the edges in the selfish routing setting does not allow the design of more general truthful mechanisms, but rather prevents any truthful mechanism from satisfying voluntary participation. These results motivate our choice of linear cost functions without fixed costs in the selfish routing setting. Moreover, the results themselves are of theoretic interest in algorithmic mechanism design.

This paper is organized as follows: Section 2 contains our monotonicity result about Nash flows in directed networks. In Section 3, we show how this result can be used in the context of algorithmic mechanism design and Section 4 contains our results on enforcing optimal tolls. In Section 5, the monotonicity result is extended to the more general setting of nonatomic congestion games. Section 6 contains our results about two-parameter agents, which motivate our choice of linear cost functions without fixed costs for the edges in the selfish routing part.

## 2 The Monotonicity of Nash Flows

In the selfish routing part of this paper, we consider a directed network with a continuous, nondecreasing cost function for every edge, which maps the amount of flow on the edge (which we call the *load* on this edge) to a nonnegative real number specifying the cost for transversing the edge. We consider a model of selfish routing with infinitely splittable flow, i.e., the users of our network are infinitesimally small, so the effect of a single users actions on the other users is negligible. Each infinitely small user wants to travel from one of the sources in the network to one of the sinks and is assumed to selfishly route her flow on the cheapest path available to her given the load induced by the other users. This setting has already been extensively studied [11, 14–21]. As usual in the literature, we assume the flow arising from the selfish behavior of the network users to be a *Nash flow*, i.e., a flow in which no infinitesimally small user has an incentive to unilaterally change her route.

We start the discussion by introducing our notation. We are given a directed network  $G = (V, E)$  with vertex set  $V$ , edge set  $E$ , and  $K$  source-destination pairs  $(s_1, t_1), \dots, (s_K, t_K) \in V^2$ .  $\mathcal{P}_i$  denotes the set of (simple)  $s_i$ - $t_i$  paths and is assumed to be nonempty for every  $i \in \{1, \dots, K\}$ . We write  $\mathcal{P} := \cup_i \mathcal{P}_i$ .

A *flow* is a function  $F : \mathcal{P} \rightarrow \mathbb{R}_{\geq 0}$ . For a fixed flow  $F$ , we denote the value of  $F$  at  $p \in \mathcal{P}$  by  $F_p$ . Every flow  $F$  induces a nonnegative *load*  $f_e := \sum_{p \in \mathcal{P}: e \in p} F_p \geq 0$  on every edge  $e \in E$ , and we call the vector  $f = (f_e)_{e \in E}$  the *load vector* of  $F$ . We write  $F^i$  for the restriction of  $F$  to  $\mathcal{P}_i$  and denote the load produced by commodity  $i$  on an edge  $e \in E$  by  $f_e^i := \sum_{p \in \mathcal{P}_i: e \in p} F_p$ . Each commodity  $i$  has a finite and positive *demand*  $d_i > 0$ , i.e.,  $d_i$  units of flow have to be sent from

source  $s_i$  to destination  $t_i$ . Since all demands are finite, we may assume without loss of generality that  $\sum_{i=1}^K d_i = 1$ . A flow  $F$  is *feasible* if  $\sum_{p \in \mathcal{P}_i} F_p = d_i$  for all  $i \in \{1, \dots, K\}$ . A load vector  $f$  is *feasible* if it is the load vector of some feasible flow  $F$ . The set of all feasible flows will be denoted by  $D$  and can be considered as a (compact) subset of  $\mathbb{R}^{|\mathcal{P}|}$ .

Every edge  $e \in E$  is given a nonnegative *cost function*  $c_e : [0, 1] \ni f_e \mapsto c_e(f_e) \in \mathbb{R}_{\geq 0}$ , which specifies the cost for using edge  $e$  when the load on  $e$  is  $f_e$ . We assume the cost functions  $c_e$  to be *continuous* and *nondecreasing* and denote the vector of all cost functions by  $c = (c_e)_{e \in E}$ . The cost of a path  $p \in \mathcal{P}_i$  to commodity  $i$  is the sum of the costs of the edges in the path, denoted by  $c_p(f) = \sum_{e \in p} c_e(f_e)$ . We write  $c(F) = (c_p(f))_{p \in \mathcal{P}}$  to denote the vector of all costs of paths  $p \in \mathcal{P}$  under the flow  $F$ . Similarly, we denote the vector of costs of edges  $e \in E$  when the load vector is  $f$  by  $c(f) = (c_e(f_e))_{e \in E}$ . We call the triple  $(G, d, c)$  an *instance*. In what follows, we will always assume the network  $G$  and the demand vector  $d$  to be fixed, so an instance is defined by only the vector  $c$  of cost functions.

**Definition 2.** A feasible flow  $F \in D$  with load vector  $f$  is at Nash equilibrium (or is a Nash flow) for costs  $c$  if for all  $i \in \{1, \dots, K\}$ ,  $p_1, p_2 \in \mathcal{P}_i$  with  $F_{p_1} > 0$ , and  $\delta \in (0, F_{p_1}]$ , we have  $c_{p_1}(f) \leq c_{p_2}(\tilde{f})$ , where  $\tilde{f}$  is the load vector of the feasible flow  $\tilde{F}$  defined by

$$\tilde{F}_p := \begin{cases} F_p - \delta & \text{if } p = p_1 \\ F_p + \delta & \text{if } p = p_2 \\ F_p & \text{else} \end{cases}$$

Before we state results about the existence and uniqueness of Nash flows, we note that an argument of Roughgarden [14] shows that Nash flows coincide with flows at *Wardrop equilibrium* in our model:

**Definition 3.** A feasible flow  $F \in D$  is a Wardrop equilibrium for costs  $c$  if, for every  $i \in \{1, \dots, K\}$ , the following holds:

$$c_{p_1}(f) > c_{p_2}(f) \text{ for } p_1, p_2 \in \mathcal{P}_i \text{ implies } F_{p_1} = 0.$$

We use the terms ‘‘Nash equilibrium’’ and ‘‘Nash flow’’ in the rest of this paper, but we included the characterization given in Definition 3 since we use the following result due to Smith [24], whose proof is based on this characterization:

**Proposition 1.** A feasible flow  $F$  is at Nash equilibrium for costs  $c$  if and only if its load vector  $f$  satisfies the variational inequality:

$$c(f)^T \cdot (f' - f) \geq 0 \quad \text{for all feasible load vectors } f'$$

Here,  $c(f)^T$  denotes the transpose of the vector  $c(f)$  of edge costs under the load vector  $f$ , and ‘‘ $\cdot$ ’’ denotes the usual product of a row vector and a column vector.

The following theorem guarantees the existence and essential uniqueness of Nash flows in our model:

**Theorem 1.** ([14]) *There exists a flow at Nash equilibrium for every vector  $c$  of continuous, nondecreasing cost functions. Moreover, if  $F, \tilde{F}$  are flows at Nash equilibrium for costs  $c$  and  $\tilde{c}$ ,  $f, \tilde{f}$  are the respective load vectors, then  $c_e(f_e) = c_e(\tilde{f}_e)$  for each edge  $e$ . If, in addition, all cost functions  $c_e$  are strictly increasing, then  $f_e = \tilde{f}_e$  for all  $e \in E$ .*

We now prove the main theorem of this section:

**Theorem 2.** *Let  $c, \tilde{c}$  be two vectors of continuous, nondecreasing cost functions, and let  $f, \tilde{f}$  be load vectors of Nash flows for costs  $c$  and  $\tilde{c}$ , respectively. If  $\tilde{c}_{e_0} \leq c_{e_0}$  for a fixed edge  $e_0 \in E$ ,  $\tilde{c}_{e_0}(f_{e_0}) < c_{e_0}(f_{e_0})$ , and  $\tilde{c}_e = c_e$  for all  $e \neq e_0$ , then  $\tilde{f}_{e_0} \geq f_{e_0}$ .*

*Proof.* By Proposition 1,  $f$  and  $\tilde{f}$  satisfy the variational inequality, so

$$\begin{aligned} c(f)^T \cdot (f' - f) &\geq 0 \\ \tilde{c}(\tilde{f})^T \cdot (f' - \tilde{f}) &\geq 0 \end{aligned}$$

for all feasible load vectors  $f'$ . Choosing  $f' = \tilde{f}$  in the first inequality and  $f' = f$  in the second one and adding yields

$$\left( c(f) - \tilde{c}(\tilde{f}) \right)^T \cdot (f - \tilde{f}) \leq 0.$$

Now we set

$$\epsilon := c_{e_0}(f_{e_0}) - \tilde{c}_{e_0}(f_{e_0}).$$

We then have  $\epsilon > 0$  since  $\tilde{c}_{e_0}(f_{e_0}) < c_{e_0}(f_{e_0})$ , so we obtain:

$$\begin{aligned} 0 &\geq \left( c(f) - \tilde{c}(\tilde{f}) \right)^T \cdot (f - \tilde{f}) \\ &= \sum_{e \in E} \left( c_e(f_e) - \tilde{c}_e(\tilde{f}_e) \right) \cdot (f_e - \tilde{f}_e) \\ &= \sum_{e \neq e_0} \underbrace{\left( c_e(f_e) - c_e(\tilde{f}_e) \right)}_{\geq 0} \cdot (f_e - \tilde{f}_e) + \left( \underbrace{c_{e_0}(f_{e_0})}_{=\epsilon + \tilde{c}_{e_0}(f_{e_0})} - \tilde{c}_{e_0}(\tilde{f}_{e_0}) \right) \cdot (f_{e_0} - \tilde{f}_{e_0}) \\ &\geq \epsilon \cdot (f_{e_0} - \tilde{f}_{e_0}) + \underbrace{\left( \tilde{c}_{e_0}(f_{e_0}) - \tilde{c}_{e_0}(\tilde{f}_{e_0}) \right)}_{\geq 0} \cdot (f_{e_0} - \tilde{f}_{e_0}) \\ &\geq \underbrace{\epsilon}_{>0} \cdot (f_{e_0} - \tilde{f}_{e_0}) \end{aligned}$$

Thus, it follows that  $\tilde{f}_{e_0} \geq f_{e_0}$  as claimed.  $\square$



Note that we do not assume the cost functions to be *strictly* increasing in Theorem 2. The result holds for all continuous, nondecreasing cost functions even though the load vector of a Nash flow is not unique in this setting.

Also note that, in the case where all cost functions are *strictly* increasing, the assumption  $\tilde{c}_{e_0}(f_{e_0}) < c_{e_0}(f_{e_0})$  is not needed: If  $\tilde{c}_{e_0}(f_{e_0}) = c_{e_0}(f_{e_0})$ , then  $\tilde{c}(f) = c(f)$  and it follows from the characterization given in Definition 3 that  $f$  is also the load vector of a Nash flow for costs  $\tilde{c}$ . Hence, the uniqueness implies that  $\tilde{f} = f$  and, in particular,  $\tilde{f}_{e_0} = f_{e_0}$ , so we obtain the following corollary:

**Corollary 1.** *Let  $c, \tilde{c}$  be two vectors of continuous, strictly increasing cost functions, and let  $f, \tilde{f}$  be the unique load vectors of Nash flows for costs  $c$  and  $\tilde{c}$ , respectively. If  $\tilde{c}_{e_0} \leq c_{e_0}$  for a fixed edge  $e_0 \in E$  and  $\tilde{c}_e = c_e$  for all  $e \neq e_0$ , then  $\tilde{f}_{e_0} \geq f_{e_0}$ .*

We now mention a special case of Theorem 2, where the costs of an edge are given as a weighted sum of latencies and tolls. This special case will be used for the design of truthful mechanisms in the next section. Namely, we are given a nondecreasing, continuous latency function  $l_e : [0, 1] \rightarrow \mathbb{R}_{\geq 0}$  and a nonnegative toll  $\tau_e$  for every edge  $e \in E$ . The vector of all tolls is denoted by  $\tau = (\tau_e)_{e \in E}$ . The cost function  $c_e$  of edge  $e \in E$  is given by  $c_e(x) := l_e(x) + \alpha \cdot \tau_e$ , where  $\alpha > 0$  is a constant factor describing the sensitivity of the agents to tolls. The total latency on a path  $p \in \mathcal{P}_i$  is denoted by  $l_p(f) = \sum_{e \in p} l_e(f_e)$ , and the total toll on  $p$  is denoted by  $\tau_p = \sum_{e \in p} \tau_e$ . Using this model, Theorem 2 immediately yields the following corollary:

**Corollary 2.** *Let the costs be given as a (weighted) sum of latencies and tolls as above, where the latency functions are continuous and nondecreasing. Let  $\tau, \tilde{\tau}$  be toll vectors such that  $\tilde{\tau}_{e_0} < \tau_{e_0}$  for a fixed edge  $e_0 \in E$  and  $\tilde{\tau}_e = \tau_e$  for all  $e \neq e_0$ . If  $f, \tilde{f}$  are load vectors of Nash flows for tolls  $\tau$  and  $\tilde{\tau}$ , respectively, then  $\tilde{f}_{e_0} \geq f_{e_0}$ .*

Note that the assumption that  $\alpha > 0$ , i.e., that the agents have a nonzero sensitivity to tolls, is crucial here.

### 3 Application in Mechanism Design

We now show how Corollary 2 can be used in the context of algorithmic mechanism design. We are given a multicommodity network with costs given by latencies and tolls as at the end of the last section, where the latencies are assumed to be continuous and nondecreasing. There are two kinds of selfish agents interacting in this network: The owners of the edges and the users of the network. The users want to travel through the network, and the latency function of each edge specifies how long it takes them to traverse the edge. Every agent owning an edge incurs costs when her edge is used. The toll on an edge of the network specifies the amount of money that the owner of the edge asks from a network user for traversing the edge, i.e., the tolls are paid by the network users to

the owners of the edges to compensate them for the costs they incur due to the usage of their edges.

Our goal is to design a mechanism that ensures a certain amount of cooperation of both classes of selfish agents, e.g., to make sure that the owners of the edges do not exploit the network users by setting the tolls too high. A *mechanism* in this setting is a pair  $(\mathcal{A}, \mathcal{P})$  consisting of an algorithm  $\mathcal{A}$ , which determines an assignment of load to the edges, and a payment scheme  $\mathcal{P}$ , which specifies the payments to the edges by the mechanism.

The costs that the owner of an edge  $e$  incurs are assumed to be linear in the load on the edge, so they are of the form  $t_e \cdot g_e$ , where  $t_e \geq 0$  is a nonnegative constant, and  $g_e \geq 0$  is the load on edge  $e \in E$ . The constant  $t_e$  specifying the costs for the agent controlling edge  $e$  is only known to the agent herself. Since we assume that every agent only controls a single edge in the network, we will slightly abuse notation by identifying the agent controlling edge  $e \in E$  with the edge  $e$  itself.

The users traveling through the network are assumed to be infinitely small, so the effect of a single users actions on the other users is negligible. Every selfish user of commodity  $i$  will choose a path  $p$  between her source vertex  $s_i$  and destination vertex  $t_i$  minimizing her cost given by  $l_p(g) + \alpha \cdot \tau_p$ . Thus, the traffic pattern arising will be a Nash flow with respect to the given latencies and the tolls defined by the edges. Note that the costs of an edge  $e$  for the network users are *different* from the costs that the edge incurs due to its usage: The costs that the edge incurs per unit load are given by the value  $t_e$  known only to the edge itself, whereas the costs that each user of edge  $e$  incurs on  $e$  are given by  $l_e(g_e) + \alpha \cdot \tau_e$ , i.e., by the weighted sum of the latency of  $e$  under the current load and the toll on  $e$ .

The mechanism considers the toll  $\tau_e$  defined by edge  $e$  as a claimed value (a *bid*) of edge  $e$  for  $t_e$ . Based on these bids, the mechanism hands out a payment  $P_e$  to every edge  $e$ . These payments are used to motivate the edges to set the tolls according to their true values  $t_e$ , which define the costs the edges have to compensate for.

To see how we can use Corollary 2 for the design of a *truthful* mechanism, we first assume the latencies to be *strictly* increasing, so that the load vector of a Nash flow is uniquely determined by the latencies and tolls (cf. Theorem 1). Under this assumption, the situation fits into the framework of mechanism design with one-parameter agents: The agents are the edges, and the private value of edge  $e \in E$  is the constant  $t_e$  defining its cost per unit flow/load on the edge. The selfish behavior of the network users does not have to be considered anymore since it is taken into account by assuming the multicommodity flow arising from the given latency functions and the tolls defined by the edges to be a Nash flow. The result obtained in Corollary 2 states that the load on an edge of the network can not increase when the toll on the edge is increased, so the algorithm described above, which just takes the Nash flow with the given latencies and the tolls defined by the edges as the assignment of load to the edges, is a monotone algorithm. Hence, as shown in [1], our mechanism will be truthful if and only

if the total amount of money that an edge  $e \in E$  receives when the toll vector (bid vector) is  $\tau$  is given as

$$h_e(\tau_{-e}) + \tau_e \cdot f_e(\tau_{-e}, \tau_e) - \int_0^{\tau_e} f_e(\tau_{-e}, u) du,$$

where the  $h_e$  are arbitrary functions, and  $f_e(\tau_{-e}, \tau_e) = f_e(\tau)$  denotes the load on edge  $e$  in the Nash flow with the given latencies and the tolls  $\tau$ .  $\tau_{-e}$  denotes the vector of all tolls except for  $\tau_e$ . More generally, we define:

**Definition 4.** For any vector  $v = (v_1, \dots, v_n)$  with  $n$  entries and  $k \in \{1, \dots, n\}$ , we denote the  $(n-1)$ -dimensional vector that consists of all entries of  $v$  except  $v_k$  by  $v_{-k} := (v_1, \dots, v_{k-1}, v_{k+1}, \dots, v_n)$ . Moreover, we write  $v = (v_1, \dots, v_n) = (v_{-k}, v_k)$ .

In our situation, every edge  $e$  already gets  $\tau_e \cdot f_e(\tau_{-e}, \tau_e)$  units of money from the users traveling on  $e$  via the tolls. Thus, the (additional) payment it has to receive from the mechanism in order to obtain a truthful mechanism has to be of the form  $P_e(\tau_{-e}, \tau_e) = h_e(\tau_{-e}) - \int_0^{\tau_e} f_e(\tau_{-e}, u) du$ . Hence, we obtain the following result:

**Theorem 3.** If all latency functions are strictly increasing, the following is a truthful mechanism:

1. Let every edge  $e \in E$  set the toll  $\tau_e$  itself.
2. Let the selfish users of the network choose their paths themselves, so that a Nash flow is obtained. The users pay the tolls on the edges directly to the edges.
3. Consider the toll  $\tau_e$  set by edge  $e \in E$  as a bid for the private value  $t_e$  of  $e$ .
4. Hand out the payment  $P_e(\tau_{-e}, \tau_e) = h_e(\tau_{-e}) - \int_0^{\tau_e} f_e(\tau_{-e}, u) du$  to edge  $e \in E$ , where the  $h_e$  are arbitrary functions.

Note that Nash flows in the setting of Theorem 3 can be computed in polynomial time via convex programming [25, 14].

Also note that truthfulness of a mechanism in our setting implies that the tolls paid by the users of the network are as low as possible in the sense that the payments of the users to every edge  $e \in E$  are exactly equal to the costs of  $e$  (but not any higher). Hence, a truthful mechanism ensures that the network users are not exploited by the edges via too high tolls as mentioned at the beginning of this section. In fact, the mechanism from Theorem 3 guarantees complete cooperation of both classes of selfish agents with the mechanism in the sense that no agent of either class has an incentive to change her strategy in order to manipulate the mechanism: No user of the network has an incentive to change her route since the flow generated in the mechanism is a Nash flow, and no edge has an incentive to misreport her cost per unit load (via changing her toll) since the mechanism is truthful.

When all latency functions are *linear*, the results of [15] imply that the *total cost*  $\sum_{e \in E} (l_e(f_e) f_e + \alpha \cdot \tau_e \cdot f_e)$  (i.e., the average cost experienced by the users

of the network) in a Nash flow  $F$  with load vector  $f$  is at most  $\frac{4}{3}$  times that of an optimal flow, i.e., of a feasible flow with minimal total cost (such a flow exists since the set  $D$  of all feasible flows is compact and the function mapping a feasible flow to its total cost is continuous). Hence, the total cost of the flow produced by the mechanism from Theorem 3 is at most  $\frac{4}{3}$  times optimal in this case. However, a simple example in [15] also shows that without the assumption of linearity of the latencies the total cost of a Nash flow can not be bounded by any constant factor times the minimal total cost.

We now consider voluntary participation. As proved in [1], a monotone algorithm admits a truthful payment scheme satisfying voluntary participation if and only if, for every  $e$  and every fixed vector of bids of all agents except  $e$ , the integral of the work curve of agent  $e$  is finite, i.e., if  $\int_0^\infty f_e(\tau_{-e}, u) du < \infty$  in our setting. It is easy to see that the functions  $h_e$  can be chosen such that the mechanism from Theorem 3 has this property under one additional assumption. Namely, we have to assume that, for each commodity  $i$ , there exist (at least) two edge disjoint  $s_i$ - $t_i$  paths in the network  $G$ . Otherwise, all users of commodity  $i$  are forced to use the only existing  $s_i$ - $t_i$  path, so the load on each edge  $e$  in this path is at least  $d_i$  no matter what the bid of the edge is, so  $\int_0^\infty f_e(\tau_{-e}, u) du \geq \int_0^\infty d_i du = \infty$ . On the other hand, if there are two node disjoint  $s_i$ - $t_i$  paths (say  $p_1^i$  and  $p_2^i$ ) for every commodity  $i$ , all users of commodity  $i$  will use  $p_2^i$  if the toll on an edge  $e \in p_1^i$  gets too high, assuming that all other bids (and, thus, all other tolls) are unchanged. Hence, we have  $\int_0^\infty f_e(\tau_{-e}, u) du < \infty$  in this case, and the choice  $h_e(\tau_{-e}) := \int_0^\infty f_e(\tau_{-e}, u) du$ , i.e.,  $P_e(\tau_{-e}, \tau_e) := \int_{\tau_e}^\infty f_e(\tau_{-e}, u) du$ , ensures voluntary participation.

**Theorem 4.** *The functions  $h_e$  can be chosen such that the mechanism described in Theorem 3 satisfies voluntary participation if and only if there exist two edge disjoint  $s_i$ - $t_i$  paths in  $G$  for every  $i = 1, \dots, K$ . In this case, we can choose  $h_e(\tau_{-e}) := \int_0^\infty f_e(\tau_{-e}, u) du$ .*

In the case where the latencies are only assumed to be nondecreasing rather than strictly increasing, there can exist Nash flows for a given toll vector that induce different load vectors. This makes the problem of designing truthful mechanism that use Nash flows for the assignment of load to the edges more difficult.

However, when we assume that the mechanism can make the network users choose their paths according to a certain Nash flow chosen in advance, the problem becomes essentially the same as in the case of a unique Nash flow. For example, the mechanism can then choose the lexicographically smallest Nash flow (with respect to some fixed ordering of the edges) among all Nash flows for the given toll vector  $\tau$  with minimal total cost and make the network users use the paths given by this Nash flow. Hence, under this assumption, we obtain a truthful mechanism similar to the one from Theorem 3. However, this mechanism does no longer have the nice and natural property that the users choose their paths completely by themselves as in the mechanism from Theorem 3.

When we do not assume that the mechanism can make the network users choose their paths according to a certain Nash flow chosen in advance, the mechanism has to deal with the uncertainty about the load assignment resulting from the selfish behavior of the network users. In order to motivate truthful bidding by the edges, the mechanism needs at least *some* information about which of the possible Nash flows will be obtained for a given toll vector since the Nash flow determines the loads assigned to the edges and, thus, the edges' costs.

In the rest of this section, we show how a *randomized mechanism* truthful in expectation can be obtained in this setting under the assumption that there is a commonly known *probability distribution* of the possible load vectors of Nash flows for every toll vector  $\tau$ . In our setting, we define randomized mechanisms as follows:

**Definition 5.** A randomized mechanism is a pair  $M = (\mathcal{A}, \mathcal{P})$ , where  $\mathcal{A}$  is a randomized algorithm, which determines a (random) assignment of load to the edges, and  $\mathcal{P}$  is a randomized payment scheme, i.e., the payment  $\mathcal{P}_e$  to each edge  $e \in E$  is a random variable. A randomized mechanism is called truthful in expectation if truthtelling maximizes the expected profit of every edge regardless of what the other edges bid and it satisfies voluntary participation if the expected profit of an edge bidding truthfully is always nonnegative.

Note that there is also a more restrictive definition of truthfulness for randomized mechanisms, which requires truthtelling to maximize each agent's profit, regardless of the outcome of the algorithm's random coin flips. Randomized mechanisms that are truthful with respect to this more restrictive definition are often referred to as *universally truthful* in the mechanism design literature [4, 26].

The characterization of truthful mechanisms for one-parameter agents by Archer and Tardos implies the following for randomized mechanisms defined as above: A randomized algorithm  $\mathcal{A}$  can be used in a randomized truthful in expectation mechanism if and only if the expected load on each edge is a decreasing function of the edges bid/toll, for every fixed vector of bids/tolls of the other edges. The payments must then be random variables whose *expectation* is given by the same formula as in the deterministic case [1].

Given the probability distribution  $\text{Pr}_\tau$  of the possible load vectors of Nash flows for every nonnegative toll vector  $\tau$ , we can design a randomized mechanism as follows: We let every edge  $e \in E$  set the toll  $\tau_e$  itself and let the network users choose their paths completely by themselves as in the mechanism from Theorem 3. Thus, when the toll vector defined by the edges is  $\tau$ , every load vector of a Nash flow for tolls  $\tau$  is obtained with the probability given by  $\text{Pr}_\tau$ . Hence, we obtain a randomized algorithm for assigning the load to the edges. We now denote the random variable that specifies the load on edge  $e \in E$  when the toll vector is  $\tau$  by  $f_e(\tau)$ . Corollary 2 then implies that the expected value of  $f_e(\tau)$  is decreasing in the toll/bid of  $e$ , so the randomized algorithm can be used

in a randomized mechanism that is truthful in expectation. The payments to the edges can be defined by the same formula as in the mechanism from Theorem 3, but are now random variables since the values  $f_e(\tau)$  are random variables. Thus, we obtain the following result:

**Theorem 5.** *Let all latency functions be continuous and nondecreasing and assume that, for every nonnegative toll vector  $\tau$ , there is a commonly known probability distribution  $Pr_\tau$  of the possible load vectors of Nash flows with respect to the tolls  $\tau$ . Then the following randomized mechanism is truthful in expectation:*

1. *Let every edge  $e \in E$  set the toll  $\tau_e$  itself.*
2. *Let the selfish users of the network choose their paths themselves, so that every load vector of a Nash flow for tolls  $\tau$  is obtained with the probability given by  $Pr_\tau$ , and let  $f_e(\tau)$  denote the random variable that specifies the load on edge  $e \in E$ . The users pay the tolls on the edges directly to the edges.*
3. *Consider the toll  $\tau_e$  set by edge  $e \in E$  as a bid for the private value  $t_e$  of edge  $e$ .*
4. *Hand out the payment  $P_e(\tau_{-e}, \tau_e) = h_e(\tau_{-e}) - \int_0^{\tau_e} f_e(\tau_{-e}, u) du$  to edge  $e \in E$ , where the  $h_e$  are arbitrary random variables.*

This randomized mechanism ensures cooperation of both classes of selfish agents with the mechanism in the same sense as the mechanism from Theorem 3 (assuming that each edge tries to maximize her expected profit).

Note that the mechanism does not use randomization to obtain truthfulness or a lower total cost. Randomization is only used to deal with the problem of uncertainty about the Nash flow and load assignment resulting from the selfish behavior of the network users in the situation where the mechanism can not simply enforce a certain Nash flow, but is only given probability distributions over possible load vectors of Nash flows.

Also note that Theorem 5 holds for every possible choice of probability distributions over load vectors of Nash flows, but, to the best of our knowledge, there is so far no theory that answers the question which probability distributions are “realistic” for a given network and its cost structure. In practice, the probability distributions could, e.g., be obtained from statistical data about the behavior of the selfish users in the network considered.

When considering voluntary participation, the definition of voluntary participation for randomized mechanisms and the arguments preceding Theorem 4 immediately yield the following result:

**Theorem 6.** *The random variables  $h_e$  defining the payments in the randomized mechanism from Theorem 5 can be chosen such that the mechanism satisfies voluntary participation if and only if there exist two edge disjoint  $s_i$ - $t_i$  paths in  $G$  for every  $i = 1, \dots, K$ . In this case, we can choose  $h_e(\tau_{-e}) := \int_0^\infty f_e(\tau_{-e}, u) du$ .*

## 4 Enforcing Optimal Tolls

In the last section, we specified the role of the mechanism as to avoid too high tolls being set by the edges. In this section, we want to change the objective of the mechanism as follows: The mechanism should now make sure that the Nash flow arising from the tolls defined by the edges has a low *total latency*. Here, we define the total latency of a feasible flow  $F$  with load vector  $f$  as  $\sum_{p \in \mathcal{P}} l_p(f) F_p = \sum_{e \in E} l_e(f_e) f_e$  and call  $F$  a *latency optimal flow* if it minimizes the total latency among the set  $D$  of all feasible flows. Notice that there does always exist a latency optimal flow since  $D$  is compact and the function mapping a feasible flow to its total latency is continuous. Moreover, the total latency of a feasible flow  $F$  is exactly equal to the average latency that the network users experience under  $F$ .

We again assume the latency functions of the edges to be strictly increasing and investigate the question to which extent it is possible to modify the mechanism described in Theorem 3 in order to motivate the edges to set *optimal* tolls. Here, we call a toll vector  $\tau$  *optimal* if the Nash flow induced by the given latencies and the tolls  $\tau$  is a latency optimal flow. The existence of such tolls was shown for example by Fleischer et. al. [19]. Moreover, the results in [19] imply that such a toll vector can be computed in polynomial time.

Note that the goal here is *not* the design of a truthful mechanism, but we want every agent  $e \in E$  to bid exactly the value  $\tau_e^{\text{opt}}$ , where  $\tau^{\text{opt}}$  is an optimal toll vector. This requires a small change in the definition of voluntary participation. Namely, we now require that an agent  $e \in E$  who bids  $\tau_e^{\text{opt}}$  does never incur a net loss. This captures the basic idea behind the concept of voluntary participation: An agent who behaves like the mechanism wants does not incur a loss. Unfortunately, the following proposition shows that this is too much to hope for:

**Proposition 2.** *The payments in the mechanism described in Theorem 3 can not be chosen such that an agent  $e \in E$  bidding  $\tau_e^{\text{opt}}$  never incurs a net loss.*

*Proof.* An agent  $e \in E$  bidding  $\tau_e^{\text{opt}}$  receives  $\tau_e^{\text{opt}} \cdot f_e(\tau_{-e}, \tau_e^{\text{opt}})$  units of money from the network users through the tolls and her costs are  $t_e \cdot f_e(\tau_{-e}, \tau_e^{\text{opt}})$ . Hence, the utility of such an agent is

$$\begin{aligned} & \tau_e^{\text{opt}} \cdot f_e(\tau_{-e}, \tau_e^{\text{opt}}) - t_e \cdot f_e(\tau_{-e}, \tau_e^{\text{opt}}) + P_e(\tau_{-e}, \tau_e^{\text{opt}}) \\ &= (\tau_e^{\text{opt}} - t_e) \cdot f_e(\tau_{-e}, \tau_e^{\text{opt}}) + P_e(\tau_{-e}, \tau_e^{\text{opt}}) \end{aligned}$$

But this can never be nonnegative for all nonnegative values of  $t_e$  and  $\tau_{-e}$  since  $P_e(\tau_{-e}, \tau_e^{\text{opt}})$  does not depend on  $t_e$  and  $(\tau_e^{\text{opt}} - t_e) \cdot f_e(\tau_{-e}, \tau_e^{\text{opt}}) \xrightarrow{t_e \rightarrow \infty} -\infty$  as long as  $f_e(\tau_{-e}, \tau_e^{\text{opt}}) > 0$ , which is always the case for large enough tolls on all paths not containing edge  $e$  since, by our assumptions on  $G$ , there does exist at least one path  $p \in \mathcal{P}$  containing  $e$ .  $\square$

As a consequence, the question “How much do we need to pay to the edges to enforce optimal tolls?” does not make sense in this setting: A mechanism where an agent  $e$  bidding  $\tau_e^{\text{opt}}$  is not required to have nonnegative utility can always keep the overall payment arbitrarily low by making the agents pay money to the mechanism even if they behave as desired.

To escape the above impossibility result, we need an additional assumption in our model. The fact which makes nonnegative utilities for “well-behaved” agents impossible is that the true cost factor  $t_e$  of an agent  $e \in E$  can be arbitrarily large. Hence, we now consider the model as before with the additional assumption that the true costs are bounded. Namely, we assume that, for each  $e \in E$ , there exists a commonly known upper bound  $T_e \in \mathbb{R}_{\geq 0}$  such that  $t_e \leq T_e$ .

To make bidding  $\tau_e^{\text{opt}}$  an optimal strategy for agent  $e$  we need to compare the utility when  $e$  bids  $\tau_e^{\text{opt}}$  with the utility when she bids an arbitrary value  $\tau_e \geq 0$ .

**Definition 6.** Given an optimal toll vector  $\tau^{\text{opt}}$  and a vector  $\tau_{-e}$  of bids of all agents except  $e$ , the function  $\text{diff}_e : [0, T_e] \times \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}$  defined by

$$\text{diff}_e(t_e, \tau_e) := (\tau_e - t_e) \cdot f_e(\tau_{-e}, \tau_e) - (\tau_e^{\text{opt}} - t_e) \cdot f_e(\tau_{-e}, \tau_e^{\text{opt}})$$

is called the utility difference function of agent  $e$ .

To make bidding  $\tau_e^{\text{opt}}$  a dominant strategy for agent  $e$  the payment  $P_e(\tau_{-e}, \tau_e^{\text{opt}})$  to agent  $e$  when she bids  $\tau_e^{\text{opt}}$  and the other agents bid  $\tau_{-e}$  must satisfy

$$P_e(\tau_{-e}, \tau_e^{\text{opt}}) \geq \sup_{(t_e, \tau_e) \in [0, T_e] \times \mathbb{R}_{\geq 0}} \text{diff}_e(t_e, \tau_e)$$

Otherwise, there would exist a cost factor  $t_e \in [0, T_e]$  and a bid  $\tau_e \geq 0$  for agent  $e$  such that the utility when  $e$  has type  $t_e$  and bids  $\tau_e$  is strictly larger than when she bids  $\tau_e^{\text{opt}}$  (assuming that  $e$  gets zero payment when she does not bid  $\tau_e^{\text{opt}}$ ).

In Section 3 we saw that, under the assumption that there exist (at least) two edge disjoint  $s_i$ - $t_i$  paths in the graph  $G$  for each commodity  $i$ , there exists a *threshold bid*  $\tau_e^{\text{thres}}(\tau_{-e})$  for every edge  $e \in E$  and every vector  $\tau_{-e}$  of bids of the other agents, i.e., we have  $f_e(\tau_{-e}, \tau_e) = 0$  for all  $\tau_e \geq \tau_e^{\text{thres}}(\tau_{-e})$ . Using this and the monotonicity of  $f_e(\tau_{-e}, \tau_e)$  in  $\tau_e$  proved in Corollary 2, we obtain

$$\begin{aligned} \text{diff}_e(t_e, \tau_e) &= (\tau_e - t_e) \cdot f_e(\tau_{-e}, \tau_e) - (\tau_e^{\text{opt}} - t_e) \cdot f_e(\tau_{-e}, \tau_e^{\text{opt}}) \\ &\leq \tau_e^{\text{thres}}(\tau_{-e}) \cdot f_e(\tau_{-e}, 0) - (\tau_e^{\text{opt}} - t_e) \cdot f_e(\tau_{-e}, \tau_e^{\text{opt}}) \end{aligned}$$

for all  $\tau_e \geq 0$ . Thus

$$\begin{aligned} &\sup_{(t_e, \tau_e) \in [0, T_e] \times \mathbb{R}_{\geq 0}} \text{diff}_e(t_e, \tau_e) \\ &\leq \sup_{t_e \in [0, T_e]} \tau_e^{\text{thres}}(\tau_{-e}) \cdot f_e(\tau_{-e}, 0) - (\tau_e^{\text{opt}} - t_e) \cdot f_e(\tau_{-e}, \tau_e^{\text{opt}}) \\ &= \tau_e^{\text{thres}}(\tau_{-e}) \cdot f_e(\tau_{-e}, 0) - (\tau_e^{\text{opt}} - T_e) \cdot f_e(\tau_{-e}, \tau_e^{\text{opt}}) \\ &= \tau_e^{\text{thres}}(\tau_{-e}) \cdot f_e(\tau_{-e}, 0) + (T_e - \tau_e^{\text{opt}}) \cdot f_e(\tau_{-e}, \tau_e^{\text{opt}}) \end{aligned}$$



To guarantee that an agent  $e \in E$  bidding  $\tau_e^{\text{opt}}$  does not incur a net loss we need to pay such an agent an amount at least as high as her maximal loss before payments, which is given by

$$\max_{t_e \in [0, T_e]} -(\tau_e^{\text{opt}} - t_e) \cdot f_e(\tau_{-e}, \tau_e^{\text{opt}}) = (T_e - \tau_e^{\text{opt}}) \cdot f_e(\tau_{-e}, \tau_e^{\text{opt}})$$

Thus, choosing

$$\begin{aligned} P_e(\tau_{-e}, \tau_e^{\text{opt}}) &:= \tau_e^{\text{thres}}(\tau_{-e}) \cdot f_e(\tau_{-e}, 0) + (T_e - \tau_e^{\text{opt}}) \cdot f_e(\tau_{-e}, \tau_e^{\text{opt}}) \\ &\geq (T_e - \tau_e^{\text{opt}}) \cdot f_e(\tau_{-e}, \tau_e^{\text{opt}}) \end{aligned}$$

makes bidding  $\tau_e^{\text{opt}}$  a dominant strategy for each agent  $e \in E$  and guarantees a nonnegative utility for agents  $e \in E$  who bid  $\tau_e^{\text{opt}}$ . We obtain:

**Theorem 7.** *Let the true costs of the agents be bounded, i.e., for each  $e \in E$ , there exists a commonly known upper bound  $T_e \in \mathbb{R}_{\geq 0}$  such that  $t_e \leq T_e$ , and assume that there exist (at least) two edge disjoint  $s_i$ - $t_i$  paths in the graph  $G$  for each commodity  $i$ . Let  $\tau^{\text{opt}}$  denote an optimal toll vector. Then bidding  $\tau_e^{\text{opt}}$  is a dominant strategy for every agent  $e \in E$  in the following mechanism and an agent  $e \in E$  bidding  $\tau_e^{\text{opt}}$  does never incur a net loss:*

1. Let every edge  $e \in E$  set the toll  $\tau_e$  itself.
2. Let the selfish users of the network choose their paths themselves, so that a Nash flow is obtained. The users pay the tolls on the edges directly to the edges.
3. Consider the toll  $\tau_e$  set by edge  $e \in E$  as a bid for the private value  $t_e$  of  $e$ .
4. Hand out the payment

$$P_e(\tau_{-e}, \tau_e^{\text{opt}}) = \tau_e^{\text{thres}}(\tau_{-e}) \cdot f_e(\tau_{-e}, 0) + (T_e - \tau_e^{\text{opt}}) \cdot f_e(\tau_{-e}, \tau_e^{\text{opt}})$$

to every edge  $e \in E$  with  $\tau_e = \tau_e^{\text{opt}}$ . All other edges receive zero payment.

To conclude this section, we now want to show how the lower bound on the payments in Theorem 7 can be weakened by slightly changing the model. In the model used so far, only the edge owners can impose tolls on the usage of their edges. However, it is a realistic assumption that the mechanism can also impose (additional) tolls on edges, e.g., in the case where the mechanism is applied by the government, which wants to optimize the traffic pattern on roads owned by some private companies.

Given this additional possibility, a straightforward change in the mechanism presented in Theorem 7 yields lower payments: If an edge  $e \in E$  sets a toll  $\tau_e < \tau_e^{\text{opt}}$  the mechanism defines an additional toll  $\tau'_e := \tau_e^{\text{opt}} - \tau_e > 0$  on edge  $e$ , so the total toll to be paid by the network users for traveling on edge  $e$  is  $\tau_e + \tau'_e = \tau_e^{\text{opt}}$  in this case. With this modification to the mechanism, it is clear that agent  $e$  is worse off when bidding any value  $\tau_e < \tau_e^{\text{opt}}$  than when she bids  $\tau_e^{\text{opt}}$ , since the load she receives (and, hence, her cost) is the same in both cases

but she gets less money from the network users in the former case (under the assumption that the network users do not care to whom they pay the tolls). As a consequence, our lower bound on the payment an agent  $e \in E$  has to receive when bidding  $\tau_e^{\text{opt}}$  relaxes to

$$\begin{aligned} P_e(\tau_{-e}, \tau_e^{\text{opt}}) &\geq \tau_e^{\text{thres}}(\tau_{-e}) \cdot f_e(\tau_{-e}, \tau_e^{\text{opt}}) + (T_e - \tau_e^{\text{opt}}) \cdot f_e(\tau_{-e}, \tau_e^{\text{opt}}) \\ &= (\tau_e^{\text{thres}}(\tau_{-e}) + T_e - \tau_e^{\text{opt}}) \cdot f_e(\tau_{-e}, \tau_e^{\text{opt}}) \end{aligned}$$

and paying exactly this amount still guarantees a nonnegative utility for every agent  $e$  who bids  $\tau_e^{\text{opt}}$ . We obtain:

**Theorem 8.** *Suppose that the assumptions of Theorem 7 hold but the mechanism is allowed to impose (additional) tolls on edges. Then bidding  $\tau_e^{\text{opt}}$  is a dominant strategy for every agent  $e \in E$  in the following mechanism and an agent  $e \in E$  bidding  $\tau_e^{\text{opt}}$  does never incur a net loss:*

1. Let every edge  $e \in E$  set the toll  $\tau_e$  itself.
2. Consider the toll  $\tau_e$  set by edge  $e \in E$  as a bid for the private value  $t_e$  of  $e$ .
3. If  $\tau_e < \tau_e^{\text{opt}}$ , define an additional toll  $\tau'_e := \tau_e^{\text{opt}} - \tau_e$  on edge  $e$ .
4. Let the selfish users of the network choose their paths themselves, so that a Nash flow is obtained. The users pay the tolls  $\tau_e$  directly to the edges and the tolls  $\tau'_e$  to the mechanism.
5. Hand out the payment

$$P_e(\tau_{-e}, \tau_e^{\text{opt}}) = (\tau_e^{\text{thres}}(\tau_{-e}) + T_e - \tau_e^{\text{opt}}) \cdot f_e(\tau_{-e}, \tau_e^{\text{opt}})$$

to every edge  $e$  with  $\tau_e = \tau_e^{\text{opt}}$ . All other edges receive zero payment.

## 5 Extension to Nonatomic Congestion Games

In this section, we show that our monotonicity result about Nash flows can be extended to the more general setting of Nash equilibria in *nonatomic congestion games*.

Congestion games model situations in which several selfish users (or *players*) share a finite number of resources. Each user can choose among different subsets of the resources and her cost (or profit) depends on the choices of all users. The selfish routing setting considered in the previous sections is a special case of a congestion game. Here, the resources are the network edges and the sets of resources that a certain class of users may choose correspond to paths between the source-destination pair of the commodity representing this user class.

Congestion games represent an active research area these days. *Atomic* congestion games with a finite number of discrete players each of which controls one unit of unsplittable demand were introduced by Rosenthal [27], who showed that every such game admits a pure-strategy Nash equilibrium. We study *nonatomic* congestion games, in which infinitely many players interact and each player has a negligible ability to affect the others. Nonatomic congestion games have recently

been studied in [28–31]. They are a subclass of the class of nonatomic games introduced by Schmeidler [32].

We consider the model of nonatomic congestion games studied in [30, 31], which is a nonatomic version of the congestion games defined by Rosenthal [27]. Our notations are chosen analogously to the selfish routing setting from the previous sections.

We are given a finite set  $E$  of *resources*. Similar to the network edges in the selfish routing setting, each resource  $e \in E$  is given a nonnegative *cost function*  $c_e : [0, 1] \rightarrow \mathbb{R}_{\geq 0}$ . The vector of all cost functions is again denoted by  $c = (c_e)_{e \in E}$  and every cost function is assumed to be *continuous* and *nondecreasing*. There are  $K$  classes of *users* or *players* sharing the resources. For each class  $i$  of players, there is a finite set  $\mathcal{S}_i \subseteq 2^E$  of subsets of  $E$ , which represents the *strategy set* of the players of class  $i$ . An element of  $\mathcal{S}_i$  is called a *strategy* for player class  $i$  and we write  $S := \cup_i \mathcal{S}_i$ . The continuum of players of class  $i$  is represented by the interval  $[0, d_i]$  endowed with the Lebesgue measure. As in the selfish routing part, we may assume without loss of generality that  $\sum_{i=1}^K d_i = 1$ . For each player class  $i$ , strategy  $S \in \mathcal{S}_i$ , and resource  $e \in E$ , there is an associated *rate of consumption*  $r_{S,e}$  that specifies the amount of congestion caused on resource  $e$  by players of class  $i$  selecting strategy  $S$ . The 5-tuple  $(E, c, \mathcal{S}, d, r)$  is called an *instance* of the nonatomic congestion game.

A *strategy distribution* is a function  $F : \mathcal{S} \rightarrow \mathbb{R}_{\geq 0}$  and we denote the value of  $F$  at  $S \in \mathcal{S}$  by  $F_S$ .  $F_S$  can be interpreted as the measure of the set of players selecting strategy  $S$  in the strategy distribution  $F$ . A strategy distribution  $F$  is *feasible* if  $\sum_{S \in \mathcal{S}_i} F_S = d_i$  for all  $i \in \{1, \dots, K\}$ . The *congestion* or *load* induced by a strategy distribution  $F$  on a resource  $e \in E$  is denoted by  $f_e = \sum_{i=1}^K \sum_{S \in \mathcal{S}_i} r_{S,e} F_S$  and the vector  $f = (f_e)_{e \in E}$  is called the *congestion vector* or *load vector* of  $F$ . The cost incurred by the players of class  $i$  selecting strategy  $S \in \mathcal{S}_i$  is defined as the sum of the costs they incur on the resources in  $S$ , denoted by  $c_S(f) = \sum_{e \in S} r_{S,e} c_e(f_e)$ .

Nash equilibria in our model of nonatomic congestion games are defined similarly to Nash/Wardrop equilibria in the selfish routing setting (cf. Definition 3):

**Definition 7.** *A feasible strategy distribution  $F$  is a Nash equilibrium for costs  $c$  if, for every  $i \in \{1, \dots, K\}$ , the following holds:*

$$c_{S_1}(f) > c_{S_2}(f) \text{ for } S_1, S_2 \in \mathcal{S}_i \text{ implies } F_{S_1} = 0.$$

The existence of Nash equilibria follows by straightforward generalizations of techniques used to prove the existence of Nash flows in the selfish routing setting.

The variational inequality characterization of Nash equilibria presented in Proposition 1 is well-known to hold in our setting of nonatomic congestion games (cf. for example [31]):

**Proposition 3.** *A feasible strategy distribution  $F$  is a Nash equilibrium for costs  $c$  if and only if its load vector  $f$  satisfies the variational inequality:*

$$c(f)^T \cdot (f' - f) \geq 0 \quad \text{for all feasible load vectors } f'$$

Using Proposition 3, the proof of Theorem 2 immediately yields the analogous result for our model of nonatomic congestion games:

**Theorem 9.** *Let  $c, \tilde{c}$  be two vectors of continuous, nondecreasing cost functions, and let  $f, \tilde{f}$  be load vectors of Nash equilibria for costs  $c$  and  $\tilde{c}$ , respectively. If  $\tilde{c}_{e_0} \leq c_{e_0}$  for a fixed resource  $e_0 \in E$ ,  $\tilde{c}_{e_0}(f_{e_0}) < c_{e_0}(f_{e_0})$ , and  $\tilde{c}_e = c_e$  for all  $e \neq e_0$ , then  $\tilde{f}_{e_0} \geq f_{e_0}$ .*

## 6 Truthful Mechanisms for Two-Parameter Agents

In this section, we prove our results on truthful mechanisms for two-parameter agents, which also motivate our choice of cost functions in the selfish routing part of this paper.

When thinking of the edges in the network as roads or lines of a telecommunication network, one might certainly want to consider fixed costs for building and maintaining the road or link. The owner of a link would incur these fixed costs independently of the amount of traffic on the edge.

These considerations motivate to study the situation where the costs for an edge are given as the load assigned to it times its private cost per unit load plus some fixed cost, which occurs independently of the load assignment. This fixed cost would then also be part of the edge's private data. Thus, the situation would now fit into the more general setting of *two-parameter* agents described in the introduction. However, we show in this section that the load assigned to any agent in a truthful mechanism for two-parameter agents has to be almost independent of the agent's bid for her fixed cost in the sense that, for almost every fixed bid for the agent's per unit cost, the load assigned to her has to be independent of the agent's bid for her fixed cost. Furthermore, when the load assigned to an agent is continuous in the agent's bid for her per unit cost, it must be completely independent of the agent's bid for her fixed cost. Thus, considering additional fixed costs for the edges does not allow the design of more general truthful mechanisms. This motivates our choice of cost functions without fixed costs for the edges in the selfish routing setting.

Together with the monotonicity of the load assigned to an agent in the agent's bid for her per unit cost as in the one-parameter setting our necessary condition for truthfulness turns out to be sufficient as well, so we obtain a complete characterization of the set of output functions that can be turned into truthful mechanisms for two-parameter agents. This characterization also implies that no truthful mechanism in the two-parameter setting can satisfy voluntary participation, which provides another motivation for not considering fixed costs for the edges in the selfish routing setting.

We now formally introduce the setting of two-parameter agents. We consider  $m$  agents indexed by the numbers  $1, \dots, m$ . Every agent  $i$  has some private data, which is known neither to the mechanism nor to the other agents. Everything except the agents' private data is public knowledge. Agent  $i$ 's private data is a pair  $(\alpha_i, \beta_i)$  of nonnegative real numbers, also called the agent's *true values*. Each agent reports a pair  $(a_i, b_i) \in \mathbb{R}_{\geq 0}^2$  of claimed values for her true values to the mechanism. This pair  $(a_i, b_i)$  is also called the *bid* of agent  $i$ . Based on the vectors  $a, b$  given by the bids of the agents, the mechanism's *output algorithm* computes an *output*  $o = o(a, b)$ , where the *output function*  $o$  takes values in a given allowable set  $O$ . Note that we do *not* assume the set  $O$  of possible outcomes to be finite as is needed for the characterization of truthful mechanisms/social choice functions by the *weak monotonicity condition* [33, 34] in more general settings. Each agent  $i$  incurs a cost  $\text{cost}_i(a, b) = \text{cost}_i(o(a, b))$ , which depends on her private data and the outcome chosen by the mechanism. To compensate the agents for these costs, the mechanism makes a *payment*  $\mathcal{P}_i(a, b)$  to each agent  $i$ , which depends on the bids. The objective of every agent  $i$  is to maximize her profit given by  $\text{profit}_i(a, b) = \mathcal{P}_i(a, b) - \text{cost}_i(a, b)$ .

As mentioned earlier, we assume the cost functions of the agents to have a special form. Namely, the outcome function  $o$  assigns an amount  $w_i(a, b) = w_i(o(a, b))$  of *load* or *work* to each agent  $i$ . The cost functions are given as  $\text{cost}_i(a, b) = \alpha_i \cdot w_i(a, b) + \beta_i$ . That is, the private value  $\alpha_i$  measures agent  $i$ 's cost per unit load and  $\beta_i$  is the fixed cost she incurs independently of the load assigned to her.

Our aim is to design *truthful* mechanisms, i.e., mechanisms for which truth-telling is a (weakly) dominant strategy for every agent, which means that  $\text{profit}_i((a_{-i}, \alpha_i), (b_{-i}, \beta_i)) \geq \text{profit}_i((a_{-i}, \alpha_i), (b_{-i}, b_i))$  for all values of  $(a_{-i}, b_{-i})$  and  $(a_i, b_i)$ . In this setting, a mechanism is a pair  $\mathcal{M} = (o, \mathcal{P})$  consisting of an output function  $o$  and a vector  $\mathcal{P}$  of payment functions. An output function  $o$  is said to *admit* a truthful payment scheme if there exists a vector  $\mathcal{P}$  of payments such that the mechanism  $\mathcal{M} = (o, \mathcal{P})$  is truthful.

We now prove our main result on two-parameter agents. Parts of the proof are similar to the proof for the one-parameter monotonicity result of Archer and Tardos given in [6].

**Theorem 10.** *An output function  $o = o(a, b)$  admits a truthful payment scheme if and only if, for every  $i$  and every pair  $(a_{-i}, b_{-i})$  of vectors of bids of all agents except  $i$ , the following holds:*

1. *For every fixed value of  $b_i$ , the load  $w_i(a, b) = w_i(o(a, b))$  assigned to agent  $i$  is nonincreasing in  $a_i$ .*
2. *For almost all values of  $a_i$ ,  $w_i(a, b)$  is a constant function of  $b_i$  (where "for almost all" means "for all but a set of Lebesgue measure zero").*

If these conditions hold, the truthful payments must be given as

$$P_i(a_i, b_i) = P_i(0, 0) + a_i \cdot w_i(a_i, b_i) - \int_0^{a_i} w_i(x, b_i) dx,$$

which is independent of  $b_i$  for almost all values of  $a_i$ .

*Proof.* First assume that the output function  $o$  admits a truthful payment scheme, i.e., there exist payments  $\mathcal{P}$  such that the mechanism  $\mathcal{M} = (o, \mathcal{P})$  is truthful. We fix an agent  $i$  and the other agents' bids  $(a_{-i}, b_{-i})$ . Then, we can consider  $\mathcal{P}_i$  and  $w_i$  as functions of just agent  $i$ 's bid  $(a_i, b_i)$ . We define a function  $p_i : \mathbb{R}_{\geq 0}^2 \rightarrow \mathbb{R}$  by

$$p_i(x, y) := \mathcal{P}_i(x, y) - x \cdot w_i(x, y) - y,$$

so  $p_i(\alpha_i, \beta_i) = \mathcal{P}_i(\alpha_i, \beta_i) - \alpha_i \cdot w_i(\alpha_i, \beta_i) - \beta_i$  is the profit of agent  $i$  when bidding her true values  $(\alpha_i, \beta_i)$ . Truthfulness of the mechanism is then equivalent to

$$\begin{aligned} p_i(\alpha_i, \beta_i) &\geq \mathcal{P}_i(a_i, b_i) - \alpha_i \cdot w_i(a_i, b_i) - \beta_i \\ \Leftrightarrow p_i(\alpha_i, \beta_i) &\geq \mathcal{P}_i(a_i, b_i) - a_i \cdot w_i(a_i, b_i) - b_i \\ &\quad - \alpha_i \cdot w_i(a_i, b_i) + a_i \cdot w_i(a_i, b_i) - \beta_i + b_i \\ \Leftrightarrow p_i(\alpha_i, \beta_i) &\geq p_i(a_i, b_i) + (\alpha_i - a_i) \cdot (-w_i(a_i, b_i)) + (\beta_i - b_i) \cdot (-1) \quad (*) \end{aligned}$$

for all  $\alpha_i, \beta_i, a_i, b_i \geq 0$ . In particular, choosing  $b_i$  and  $\beta_i$  to be equal in  $(*)$  yields

$$p_i(\alpha_i, b_i) \geq p_i(a_i, b_i) + (\alpha_i - a_i) \cdot (-w_i(a_i, b_i))$$

for all  $\alpha_i, a_i, b_i \geq 0$ . Hence, for every fixed  $b_i$ ,  $p_i(a_i, b_i)$  is a convex function of  $a_i$  and  $-w_i(a_i, b_i)$  is a subgradient at  $a_i$ . As standard results from analysis show, this implies that, for every  $b_i$ ,  $p_i(a_i, b_i)$  is a continuous function of  $a_i$ , differentiable almost everywhere, and equal to the integral of its derivative. Moreover, we have  $\frac{\partial p_i}{\partial a_i}(a_i, b_i) = -w_i(a_i, b_i)$  whenever  $p_i$  is differentiable with respect to  $a_i$ . Thus,  $p_i(a_i, b_i) = p_i(0, b_i) - \int_0^{a_i} w_i(x, b_i) dx$  and using the definition of  $p_i$  this is equivalent to

$$P_i(a_i, b_i) = P_i(0, b_i) + a_i \cdot w_i(a_i, b_i) - \int_0^{a_i} w_i(x, b_i) dx \quad (**)$$

Moreover, since  $p_i(a_i, b_i)$  is a convex function of  $a_i$  for fixed  $b_i$ , the partial derivative  $\frac{\partial p_i}{\partial a_i}(a_i, b_i) = -w_i(a_i, b_i)$  is nondecreasing, so  $w_i(a_i, b_i)$  is a nonincreasing function of  $a_i$  for all values  $b_i \geq 0$ , which proves condition 1.

If we choose  $a_i$  and  $\alpha_i$  to be equal in  $(*)$ , we obtain

$$p_i(a_i, \beta_i) \geq p_i(a_i, b_i) + (\beta_i - b_i) \cdot (-1)$$

for all  $\beta_i, a_i, b_i \geq 0$ . Thus, for every fixed  $a_i$ ,  $p_i(a_i, b_i)$  is a convex function of  $b_i$  and  $-1$  is a subgradient at  $b_i$ . Again by results from analysis, this implies that, for every  $a_i$ ,  $p_i(a_i, b_i)$  is a continuous function of  $b_i$ , differentiable almost everywhere,

and equal to the integral of its derivative. Furthermore, we have  $\frac{\partial p_i}{\partial b_i}(a_i, b_i) = -1$  whenever  $p_i$  is differentiable with respect to  $b_i$ . Hence,  $p_i(a_i, b_i) = p_i(a_i, 0) - b_i$  and we calculate

$$\begin{aligned} p_i(a_i, b_i) &= p_i(a_i, 0) - b_i \\ &= P_i(a_i, 0) - a_i \cdot w_i(a_i, 0) - b_i \\ &= P_i(0, 0) + a_i \cdot w_i(a_i, 0) - \int_0^{a_i} w_i(x, 0) dx - a_i \cdot w_i(a_i, 0) - b_i \\ &= P_i(0, 0) - \int_0^{a_i} w_i(x, 0) dx - b_i, \end{aligned}$$

so using that  $p_i(a_i, b_i) = P_i(a_i, b_i) - a_i \cdot w_i(a_i, b_i) - b_i$  we obtain

$$P_i(a_i, b_i) - a_i \cdot w_i(a_i, b_i) = P_i(0, 0) - \int_0^{a_i} w_i(x, 0) dx.$$

In particular, for  $a_i = 0$ , we get  $P_i(0, b_i) = P_i(0, 0)$  for all  $b_i \geq 0$ , so by (\*\*) the payments are given by the formula in the claim. Plugging in we obtain

$$\begin{aligned} p_i(a_i, b_i) &= p_i(0, b_i) - \int_0^{a_i} w_i(x, b_i) dx \\ &= P_i(0, b_i) - b_i - \int_0^{a_i} w_i(x, b_i) dx \\ &= P_i(0, 0) - b_i - \int_0^{a_i} w_i(x, b_i) dx. \end{aligned}$$

On the other hand, we have already seen that

$$p_i(a_i, b_i) = P_i(0, 0) - \int_0^{a_i} w_i(x, 0) dx - b_i,$$

so we obtain

$$\int_0^{a_i} w_i(x, 0) dx = \int_0^{a_i} w_i(x, b_i) dx \quad (***)$$

for all  $a_i, b_i$ . Moreover, since  $w_i(a_i, b_i)$  is nonincreasing in  $a_i$  for every fixed value of  $b_i$ , we have

$$\int_0^{a_i} w_i(x, b_i) dx - \int_0^{\tilde{a}_i} w_i(x, b_i) dx = \int_{\tilde{a}_i}^{a_i} w_i(x, b_i) dx \geq (a_i - \tilde{a}_i) \cdot w_i(a_i, b_i)$$

for all  $a_i, \tilde{a}_i \geq 0$ , i.e., the function  $\varphi$  defined by  $\varphi(a_i) := \int_0^{a_i} w_i(x, b_i) dx$  (which is well-defined by equality (\*\*\*)) is concave and  $w_i(a_i, b_i)$  is a supergradient of  $\varphi$  at  $a_i$  for every  $b_i \geq 0$ . Similar to the convex case, this implies that  $\varphi$  is continuous, differentiable almost everywhere, and equal to the integral of its derivative. Moreover, we have  $\varphi'(a_i) = w_i(a_i, b_i)$  for every  $b_i \geq 0$  whenever  $\varphi$  is differentiable with respect to  $a_i$ . Thus, for almost all  $a_i$ , we can differentiate equation (\*\*\*) and obtain

$$w_i(a_i, 0) = w_i(a_i, b_i)$$

for all  $b_i \geq 0$ , which proves condition 2 and completes the proof of the first direction.

Now suppose that conditions 1. and 2. are satisfied for a given output function  $o$ . As before, we fix an agent  $i$  and the other agents' bids  $(a_{-i}, b_{-i})$  and consider  $\mathcal{P}_i$  and  $w_i$  as functions of just agent  $i$ 's bid  $(a_i, b_i)$ . We claim that the formula in the claim defines a truthful payment scheme for  $o$ . To prove this, we have to show that inequality (\*) is satisfied for all  $\alpha_i, \beta_i, a_i, b_i \geq 0$ , which is equivalent to

$$p_i(\alpha_i, \beta_i) - \mathcal{P}_i(a_i, b_i) + \alpha_i \cdot w_i(a_i, b_i) + \beta_i \geq 0$$

for all  $\alpha_i, \beta_i, a_i, b_i \geq 0$ .

Using that  $w_i(\cdot, b_i)$  is nonincreasing for every  $b_i$  by 1. and that  $\int_0^{\alpha_i} w_i(x, b_i) dx$  is independent of  $b_i$  by 2., we calculate

$$\begin{aligned} & p_i(\alpha_i, \beta_i) - \mathcal{P}_i(a_i, b_i) + \alpha_i \cdot w_i(a_i, b_i) + \beta_i \\ &= \mathcal{P}_i(\alpha_i, \beta_i) - \alpha_i \cdot w_i(\alpha_i, \beta_i) - \beta_i - \mathcal{P}_i(a_i, b_i) + \alpha_i \cdot w_i(a_i, b_i) + \beta_i \\ &= \mathcal{P}_i(0, 0) + \alpha_i \cdot w_i(\alpha_i, \beta_i) - \int_0^{\alpha_i} w_i(x, \beta_i) dx - \alpha_i \cdot w_i(\alpha_i, \beta_i) - \beta_i \\ &\quad - \mathcal{P}_i(0, 0) - a_i \cdot w_i(a_i, b_i) + \int_0^{a_i} w_i(x, b_i) dx + \alpha_i \cdot w_i(a_i, b_i) + \beta_i \\ &= (\alpha_i - a_i) \cdot w_i(a_i, b_i) + \int_{\alpha_i}^{a_i} w_i(x, b_i) dx \\ &\geq (\alpha_i - a_i) \cdot w_i(a_i, b_i) + (a_i - \alpha_i) \cdot w_i(a_i, b_i) \\ &= 0 \end{aligned}$$

which completes the proof.  $\square$

When  $w_i(a_i, b_i)$  is continuous in  $a_i$  for fixed  $b_i$ , we see that the function  $\varphi$  in the proof is differentiable everywhere, so differentiating equation (\*\*\*) yields the independence of  $w_i(a_i, b_i)$  of  $b_i$  for every fixed value of  $a_i$ . Hence, we obtain the following corollary:

**Corollary 3.** *An output function  $o(a, b)$  for which the load  $w_i(a, b)$  assigned to each agent  $i$  is continuous in  $a_i$  for every fixed  $a_{-i}, b$  admits a truthful payment scheme if and only if  $w_i(a, b)$  is independent of  $b_i$  and nonincreasing in  $a_i$ . In this case, the payments must be given as*

$$P_i(a_i, b_i) = P_i(0, 0) + a_i \cdot w_i(a_i, b_i) - \int_0^{a_i} w_i(x, b_i) dx,$$

which is also independent of  $b_i$ .

As a particular consequence of Theorem 10 there does not exist a mechanism for two parameter agents which is *strongly truthful*, i.e., a mechanism in which



truth-telling is the *only* dominant strategy for every agent: Whenever the true per unit cost  $\alpha_i$  of an agent  $i$  is such that  $w_i(\alpha_i, b_i)$  (and, hence, also  $\mathcal{P}_i(\alpha_i, b_i)$ ) is independent of  $b_i$ , any bid of the form  $(\alpha_i, b_i)$  represents a dominant strategy for agent  $i$ .

Furthermore, Theorem 10 implies that the voluntary participation condition can never be satisfied in a truthful mechanism for two-parameter agents: To guarantee a nonnegative profit for every agent  $i$  bidding truthfully, we must have

$$\text{profit}_i(\alpha_i, \beta_i) = \mathcal{P}_i(\alpha_i, \beta_i) - \alpha_i \cdot w_i(\alpha_i, \beta_i) - \beta_i \geq 0$$

for all  $\alpha_i, \beta_i \geq 0$ . But by Theorem 10, the value  $\mathcal{P}_i(\alpha_i, \beta_i) - \alpha_i \cdot w_i(\alpha_i, \beta_i)$  is independent of  $\beta_i$  for almost every  $\alpha_i$ , so the profit of agent  $i$  is unbounded from below as  $\beta_i \rightarrow \infty$  for every such  $\alpha_i$ . Thus, without any a priori upper bound for  $\beta_i$ , it is impossible to guarantee a nonnegative profit for agent  $i$  when she bids truthfully. Hence, we obtain:

**Theorem 11.** *The voluntary participation condition can never be satisfied in a truthful mechanism for two-parameter agents.*

## 7 Directions for Further Work

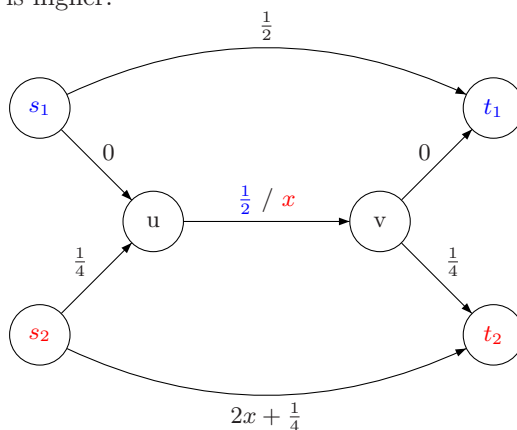
Although our monotonicity result holds in the very general setting of nonatomic congestion games presented in Section 5, a natural question is whether it can be extended to even more general settings of selfish routing and congestion games. For example, one could consider settings with nonseparable cost functions, i.e., models in which the cost of an edge or resource  $e$  is a function not only of the load on  $e$  itself, but also of the load on other edges or resources. Alternatively, one could allow different cost functions for different user classes. However, the following example shows that the result from Theorem 2 becomes invalid in this setting:

*Example 1.* Consider the multicommodity network in Fig. 1. There are two commodities with demands  $d_1 = d_2 = \frac{1}{2}$ . The costs of the middle edge  $(u, v)$  differ for the two commodities. The costs of  $(u, v)$  to the first commodity are indicated in blue and the costs to the second commodity in red.

The flow in which commodity 1 sends all its flow on the path  $(s_1, u, v, t_1)$  and commodity 2 sends  $\frac{5}{12}$  units of flow on the path  $(s_2, t_2)$  and  $\frac{1}{12}$  on the path  $(s_2, u, v, t_2)$  is a Nash flow.

Now suppose that the cost function of the edge  $(s_2, t_2)$  is changed to  $2x$ , i.e., edge  $(s_2, t_2)$  becomes *cheaper*. Then the flow in which commodity 1 sends all its flow on the path  $(s_1, t_1)$  and commodity 2 sends  $\frac{1}{3}$  units of flow on the path  $(s_2, t_2)$  and  $\frac{1}{6}$  on the path  $(s_2, u, v, t_2)$  is a Nash flow. Hence, the load on edge  $(s_2, t_2)$  is *lower* in the new Nash flow even though the edge became cheaper.

Observe that the given Nash flows are *not* unique. In fact, the flow in which commodity 1 sends all its flow on the path  $(s_1, u, v, t_1)$  and commodity 2 sends all its flow on the path  $(s_2, t_2)$  is also a Nash flow for costs  $2x$  on edge  $(s_2, t_2)$ , and the load on edge  $(s_2, t_2)$  in this flow is *higher* than in the Nash flow given for costs  $2x + \frac{1}{4}$  on edge  $(s_2, t_2)$ . Thus, a natural question for this setting is whether there does always exist at least one Nash flow for the cheaper costs in which the load on the edge is higher.



**Fig. 1.** A multicommodity network with different cost functions for different commodities

A different direction of research could be the design of further truthful mechanisms based on our monotonicity result for other congestion games than selfish routing.

## References

1. Archer, A., Tardos, É.: Truthful mechanisms for one-parameter agents. In: Proceedings of the 42nd Annual IEEE Symposium on the Foundations of Computer Science. (2001) 482–491
2. Osborne, M.J., Rubinstein, A.: A Course in Game Theory. MIT Press (1994)
3. Mas-Colell, A., Whinston, M.D., Green, J.R.: Microeconomic Theory. Oxford University Press (1995)
4. Nisan, N., Ronen, A.: Algorithmic mechanism design. Games and Economic Behavior **35** (2001) 166–196
5. Myerson, R.B.: Optimal auction design. Mathematics of Operations Research **6**(1) (1981) 58–73
6. Archer, A.: Mechanisms for Discrete Optimization with Rational Agents. PhD thesis, Cornell University (2004)
7. Auletta, V., Prisco, R.D., Penna, P., Persiano, G.: Deterministic truthful approximation mechanisms for scheduling related machines. In: Proceedings of the 21st International Symposium on Theoretical Aspects of Computer Science. Volume 2996 of Lecture Notes in Computer Science. (2004) 608–619

8. Andelman, N., Azar, Y., Sorani, M.: Truthful approximation mechanisms for scheduling selfish related machines. In: Proceedings of the 22nd International Symposium on Theoretical Aspects of Computer Science. Volume 3404 of Lecture Notes in Computer Science. (2005) 69–82
9. Kovács, A.: Fast monotone 3-approximation algorithm for scheduling related machines. In: Proceedings of the 13th Annual European Symposium on Algorithms. Volume 3669 of Lecture Notes in Computer Science. (2005) 616–627
10. Tardos, É.: Network games. In: Proceedings of the 36th Annual ACM Symposium on the Theory of Computing. (2004) 341–342
11. Roughgarden, T., ed.: Selfish Routing and the Price of Anarchy. MIT Press (2005)
12. Nash, J.: Non-cooperative games. *Annals of Mathematics* **54**(2) (1951) 286–295
13. Papadimitriou, C.: Algorithms, games, and the internet. In: Proceedings of the 33rd Annual ACM Symposium on the Theory of Computing. (2001) 749–753
14. Roughgarden, T.: Selfish Routing. PhD thesis, Cornell University (2002)
15. Roughgarden, T., Tardos, É.: How bad is selfish routing? *Journal of the ACM* **49**(2) (2002) 236–259
16. Roughgarden, T.: The price of anarchy is independent of the network topology. *Journal of Computer and System Sciences* **67**(2) (2003) 341–364
17. Cole, R., Dodis, Y., Roughgarden, T.: Pricing network edges for heterogeneous selfish users. In: Proceedings of the 35th Annual ACM Symposium on the Theory of Computing. (2003) 521–530
18. Cole, R., Dodis, Y., Roughgarden, T.: How much can taxes help selfish routing? In: Proceedings of the 4th ACM Conference on Electronic Commerce. (2003) 98–107
19. Fleischer, L., Jain, K., Mahdian, M.: Tolls for heterogeneous selfish users in multicommodity networks and generalized congestion games. In: Proceedings of the 45th Annual IEEE Symposium on the Foundations of Computer Science. (2004) 277–285
20. Fleischer, L.: Linear tolls suffice: New bounds and algorithms for tolls in single source networks. *Theoretical Computer Science* **348** (2005) 217–225
21. Dafermos, S.C., Nagurney, A.: Sensitivity analysis for the asymmetric network equilibrium problem. *Mathematical Programming* **28** (1984) 174–184
22. Koutsoupias, E., Papadimitriou, C.: Worst-case equilibria. In: Proceedings of the 16th International Symposium on Theoretical Aspects of Computer Science. Volume 1563 of Lecture Notes in Computer Science. (1999) 404–413
23. Feldmann, R., Gairing, M., Lücking, T., Monien, B., Rode, M.: Selfish routing in non-cooperative networks: A survey. In: Proceedings of the 28th International Symposium on Mathematical Foundations of Computer Science. Volume 2747 of Lecture Notes in Computer Science. (2003) 21–45
24. Smith, M.J.: The existence, uniqueness, and stability of traffic equilibria. *Transportation Research* **13B** (1979) 295–304
25. Beckmann, M.J., McGuire, C.B., Winsten, C.B.: *Studies in the Economics of Transportation*. Yale University Press (1956)
26. Mu’alem, A., Schapira, M.: Setting lower bounds on truthfulness. In: Proceedings of the 18th Annual ACM-SIAM Symposium on Discrete Algorithms. (2007) 1143–1152
27. Rosenthal, R.W.: A class of games possessing pure-strategy nash equilibria. *International Journal of Game Theory* **2**(1) (1973) 65–67
28. Milchtaich, I.: Generic uniqueness of equilibrium in large crowding games. *Mathematics of Operations Research* **25**(3) (2000) 349–364
29. Milchtaich, I.: Social optimality and cooperation in nonatomic congestion games. *Journal of Economic Theory* **114**(1) (2004) 56–87

30. Roughgarden, T., Tardos, É.: Bounding the inefficiency of equilibria in nonatomic congestion games. *Games and Economic Behavior* **47** (2004) 389–403
31. Correa, J.R., Schulz, A.S., Stier-Moses, N.E.: On the inefficiency of equilibria in congestion games. In: *Proceedings of the 11th Mathematical Programming Society Conference on Integer Programming and Combinatorial Optimization*. Volume 3509 of *Lecture Notes in Computer Science*. (2005) 167–181
32. Schmeidler: Equilibrium points of nonatomic games. *Journal of Statistical Physics* **7**(4) (1973) 295–300
33. Bikhchandani, S., Chatterji, S., Lavi, R., Mu'alem, A., Nisan, N., Sen, A.: Weak monotonicity characterizes deterministic dominant strategy implementation. *Econometrica* **74**(4) (2006) 1109–1132
34. Saks, M., Yu, L.: Weak monotonicity suffices for truthfulness on convex domains. In: *Proceedings of the 6th ACM Conference on Electronic Commerce*. (2005) 286–293