

Fundamental Limits of Locally-Computed Incentives in Network Routing

Philip N. Brown and Jason R. Marden

Abstract—We ask if it is possible to positively influence social behavior with no risk of unintentionally incentivizing pathological behavior. In network routing problems, if network traffic is composed of many individual agents (such as drivers in a city’s road network), it is known that self-interested behavior among the agents can lead to suboptimal network congestion. To mitigate this, a system planner may charge monetary tolls for the use of network links in an effort to incentivize efficient routing choices by the users. We study situations in which these tolls are computed locally on each edge, as in the classical case of marginal-cost taxation, but that the users’ sensitivity to tolls is not known. We seek locally-computed tolls that are guaranteed not to incentivize worse network routing than in the un-influenced case. Our results are twofold: first, we give a full characterization of all non-perverse locally-computed tolls for parallel networks with arbitrary convex delay functions, and show that they are all a generalized version of traditional marginal-cost tolls. Second, we exhibit a type of pathological network in which all locally-computed tolling functions can cause perverse incentives for heterogeneous price-sensitive user populations. That is, in general networks, the only locally-computed tolling functions that *do not* incentivize pathological behavior on some network are effectively zero tolls. Finally, we show that our results have interesting implications for the theory of altruistic behavior.

I. INTRODUCTION

Modern computational and infrastructure systems are becoming increasingly linked with the social systems that they serve. Accordingly, engineers must be aware of the ways in which social behavior affects the performance of engineered systems; this has spurred recent research on influencing social behavior to achieve engineering objectives [1]–[4], [21]. Examples of interesting problems in this context include ridesharing systems [5], transportation networks [6], and power grids [7].

In this paper, we focus on a well-studied model of network traffic congestion known as a “non-atomic congestion game,” in which traffic needs to be routed across a network from a source node to a destination node in a way that minimizes the average delay experienced by the traffic. If a central authority can control the traffic explicitly, it is typically straightforward to compute the optimal assignment of traffic; unfortunately, if the mass of traffic is composed of individual decision-makers, the aggregate network flows that emerge

from individual localized decision-making may be far from optimal [8].

Accordingly, much research has focused on methods of influencing the routing choices made by individual users; one promising set of methodologies involves charging specially-designed tolls to network links in an effort to incentivize more-efficient network flows [9], [10]. In [11], [12] it is shown that a special type of tolling function called a *marginal-cost* toll levied on each network link incentivizes optimal network routing – provided that all network users trade off time and money equally. An attractive feature of marginal-cost tolls is that they can be computed locally on each network link; that is, a link’s toll depends *only* on that link’s congestion characteristics and traffic flow. Thus, the optimality guaranteed by these tolls is intrinsically robust to variations of network structure. This local-computation property is known as *network agnosticity*; in essence, marginal-cost tolls only “know” their own edge – they are agnostic to global network specifications [21].

Weak robustness is defined in [19] as a guarantee that a given behavior-influencing mechanism never creates perverse incentives. Unfortunately, the authors of [19] also show that marginal-cost tolls are *not* weakly robust to variations of user toll-sensitivity. That is, if some users value their time more than others, networks exist on which the routing incentivized by marginal-cost tolls is worse than un-influenced routing.

Despite this negative result for traditional marginal-cost tolls, it has remained an open question whether some other network-agnostic taxation mechanism exists which can be weakly robust to variations of user toll-sensitivity. This question was partially addressed in [13] for the simplified case of parallel networks with linear-affine latency functions, where it was shown that any weakly-robust network agnostic taxation mechanism is essentially a generalization of traditional marginal-cost tolls. However, this restriction to parallel networks is not without loss of generality, as it is also proved that there is no weakly robust network-agnostic taxation mechanism for general asymmetric networks (that is, networks with more than one source and/or destination).

In this paper, we present an extension of the positive results of [13] to the case of all convex latency functions. First, in Theorem 4.1 we define the *generalized marginal-cost* taxation mechanism, and show that it is the only non-trivial network-agnostic taxation mechanism that is weakly robust on the class of parallel-path networks. Thus, a system planner can apply generalized marginal-cost tolls on any parallel network without fear of causing perverse incentives.

However, we also strengthen the negative results of [13], and exhibit a family of symmetric networks (that is, having

This work is supported ONR Grant #N00014-17-1-2060 and NSF Grant #ECCS-1638214.

P. N. Brown is a graduate student researcher with the Department of Electrical and Computer Engineering, University of California, Santa Barbara, CA pnbrown@ece.ucsb.edu. Corresponding author.

J. R. Marden is with the Department of Electrical and Computer Engineering, University of California, Santa Barbara, jmarden@ece.ucsb.edu.

a single source/destination pair and all agents having access to the same set of paths) on which generalized marginal-cost taxes can never be weakly-robust. Thus, in Theorem 4.2 we rule out the possibility of a network-agnostic taxation mechanism being weakly-robust on general symmetric networks. Note that this does not mean that generalized marginal-cost tolls are not weakly-robust on *every* general network; merely that without *a priori* knowledge of network structure, the possibility of perverse incentives cannot be ruled out.

Finally, we show that our results imply corresponding facts about the behavior of altruistic network users. In particular, in the altruism model of [14], our characterization result in Theorem 4.1 implies that in every parallel network, users acting altruistically is socially beneficial. On the other hand, our negative result in Theorem 4.2 implies the somewhat paradoxical statement that there exist symmetric networks in which it can actually be socially harmful for some (but not all) users to act altruistically; we include these two results in Corollary 4.3.

II. MODEL AND ROBUSTNESS

A. Routing Game

Consider a network routing problem for a network (V, E) comprised of vertex set V and edge set E . A mass of r units of traffic needs to be routed from a common source $s \in V$ to a common destination $t \in V$. We write \mathcal{P} to denote the set of *paths* available to the traffic, where each path $p \in \mathcal{P}$ consists of a set of edges connecting s to t . Note that this paper considers only the case of *symmetric* (or single-commodity) routing problems, in which all traffic can access the same set of paths. A network is called a *parallel-path* network if all paths are disjoint; i.e., for all paths $p, p' \in \mathcal{P}$, $p \cap p' = \emptyset$.

A *feasible flow* $f \in \mathbb{R}^{|\mathcal{P}|}$ is an assignment of traffic to various paths such that $\sum_{p \in \mathcal{P}} f_p = r$, where $f_p \geq 0$ denotes the mass of traffic on path p . Given a flow f , the flow on edge e is given by $f_e = \sum_{p: e \in p} f_p$. To characterize transit delay as a function of traffic flow, each edge $e \in E$ is associated with a specific latency function $\ell_e : [0, r] \rightarrow [0, \infty)$; $\ell_e(f_e)$ denotes the delay experienced by users of edge e when the edge flow is f_e . We adopt the standard assumptions that each latency function is nondecreasing, convex, continuous, and continuously differentiable. We measure the cost of a flow f by the *total latency*, given by

$$\mathcal{L}(f) = \sum_{e \in E} f_e \cdot \ell_e(f_e) = \sum_{p \in \mathcal{P}} f_p \cdot \ell_p(f_p), \quad (1)$$

where $\ell_p(f) = \sum_{e \in p} \ell_e(f_e)$ denotes the latency on path p . We denote the flow that minimizes the total latency by

$$f^* \in \underset{f \text{ is feasible}}{\operatorname{argmin}} \mathcal{L}(f). \quad (2)$$

Due to the convexity of ℓ_e , $\mathcal{L}(f^*)$ is unique.

A *routing problem* is given by $G = (V, E, r, \{\ell_e\})$. The set of all routing problems is written \mathcal{G} , and we denote by \mathcal{G}_p the set of all parallel-path routing problems.

To study the effect of taxes on self-interested behavior, we model the above routing problem as a heterogeneous non-atomic congestion game. We assign each edge $e \in E$ a flow-dependent taxation function $\tau_e : \mathbb{R}^+ \rightarrow \mathbb{R}^+$. To characterize

users' taxation sensitivities, let each user $x \in [0, r]$ have a taxation sensitivity $s_x \in [S_L, S_U] \subseteq \mathbb{R}^+$, where $S_L \geq 0$ and $S_U \leq +\infty$ are lower and upper sensitivity bounds, respectively. Note that we allow S_U to take the value $+\infty$. If all users in s have the same sensitivity (i.e., $s_x = s_y$ for all $x, y \in [0, r]$), the population is said to be *homogeneous*; otherwise it is *heterogeneous*. Given a flow f , the cost that user x experiences for using path $\tilde{p} \in \mathcal{P}$ is of the form

$$J_{\tilde{p}}^x(f) = \sum_{e \in \tilde{p}} [\ell_e(f_e) + s_x \tau_e(f_e)], \quad (3)$$

and we assume that each user selects the lowest-cost path from the available source-destination paths. We call a flow f a *Nash flow* if all users are individually using minimum-cost paths given the choices of other users, or if for all users $x \in [0, r]$ we have

$$J^x(f) = \min_{p \in \mathcal{P}} \sum_{e \in p} [\ell_e(f_e) + s_x \tau_e(f_e)]. \quad (4)$$

It is well-known that a Nash flow exists for any non-atomic congestion game of the above form [15]. If the population is homogeneous or the taxes are constant functions, these flows are unique in cost (that is, all Nash flows have the same total latency) [16], [17].

Let the agents be ordered according to sensitivity; that is, let s_x be a nondecreasing function of x . The set of possible sensitivity distributions is the set of monotone-increasing functions $S = \{s : [0, r] \rightarrow [S_L, S_U]\}$.

B. Taxation Mechanisms and Robustness

To model locally-computed tolls, we consider so-called *network-agnostic* taxation mechanisms. Here, each edge's taxation function is computed using only locally-available information. That is, $\tau_e(f_e)$ depends only on ℓ_e , not on edge e 's location in the network, the network topology, the overall traffic rate, or the properties of any other edge. A network-agnostic taxation mechanism T is thus a mapping from latency functions to taxation functions, and the taxation function associated with latency function ℓ_e is given by

$$\tau_e(\cdot) = T(\ell_e). \quad (5)$$

To evaluate the performance of taxation mechanisms, we write $\mathcal{L}^{\text{nf}}(G, s, T)$ to denote the total latency of a Nash flow for routing problem G and population s induced by taxation mechanism T . If more than one Nash flow exists, let $\mathcal{L}^{\text{nf}}(G, s, T)$ denote the total latency of the worst Nash flow. We write $\mathcal{L}^{\text{nf}}(G, \emptyset)$ to denote the total latency of an un-influenced Nash flow; note that when there are no tolls, the sensitivity distribution plays no role.

In the robustness framework of [19], taxation mechanism T is said to be *weakly robust* if for every network and sensitivity distribution, the total latency induced by T never exceeds the total latency of an un-influenced Nash flow; i.e., for all $G \in \mathcal{G}$,

$$\sup_{s \in S} \mathcal{L}^{\text{nf}}(G, s, T) \leq \mathcal{L}^{\text{nf}}(G, \emptyset). \quad (6)$$

Loosely speaking, if a taxation mechanism is weakly robust, this means that it will never create perverse incentives on any routing problem. Note that at a minimum, the zero toll is weakly robust.

III. RELATED WORK AND EXAMPLE

The following is a brief survey of relevant work on the robustness of taxation mechanisms in congestion games.

The classical example of a network-agnostic taxation mechanism is that of the *marginal-cost* or *Pigovian* taxation mechanism T^{mc} . For any edge e with latency function ℓ_e , the accompanying marginal-cost toll is

$$\tau_e^{\text{mc}}(f_e) = f_e \cdot \ell'_e(f_e), \quad \forall f_e \geq 0, \quad (7)$$

where ℓ' represents the flow derivative of ℓ . In [11] the authors show that for any $G \in \mathcal{G}$, it is true that $\mathcal{L}^*(G) = \mathcal{L}^{\text{nf}}(G, s, T^{\text{mc}})$, provided that all users have a sensitivity equal to 1.

Recent research has identified several new network-agnostic taxation mechanisms, which are all in some sense generalizations of T^{mc} . For example, [18] exhibits a universal taxation mechanism which achieves weak robustness with large tolls. In parallel networks, [21] studies scaled marginal-cost tolls for parallel networks under a utilization constraint. In [19], it is shown that constant tolling functions can never be weakly robust, even on parallel-path networks. In [13], the authors show that for linear latency functions,

$$\tau_e(f_e) = \kappa_a a_e f_e + \kappa_b b_e, \quad (8)$$

can be weakly-robust for parallel networks if $\kappa_a \geq 0$ and $\kappa_b \geq 0$ are chosen carefully. However, [13] contains an impossibility result showing that *no* non-trivial network-agnostic taxation mechanism can be weakly robust on general asymmetric networks.

Here, we exhibit a new type of pathology for marginal-cost tolls. In contrast to the pathology reported in [13] for asymmetric networks, our example occurs for symmetric networks as well.

Example 3.1: Consider the network in Figure 1, with the well-known Braess's Paradox network [20] in parallel with a constant-latency edge. Charge marginal-cost tolls (7) on the network. If the user population has 2 units of traffic and a homogeneous toll sensitivity of $s \in [0, 1]$, it is easy to verify that that unique Nash flow on this network is the one labeled "Efficient Nash Flow" in Figure 1. In this flow, all have a cost of $2 + s$ since deviating to the zig-zag path would yield a larger cost of $2 + 2s$, and deviating to the constant-latency link would yield a cost of 3. This gives a total latency of $2 \times 2 = 4$.

Now consider a heterogeneous population with 1 unit of sensitivity $s_1 = 0$ traffic (orange in Figure 1), and 1 unit of sensitivity $s_2 = 1$ traffic. A new Nash flow emerges: one with all insensitive traffic using the zig-zag path, and all sensitive traffic using the constant-latency link; this is labeled "Inefficient Nash Flow" in Figure 1. In this flow, any agent on the zig-zag path has a delay of 2, but any agent on the constant-latency path has a delay of 3, for a total latency of $2 + 3 = 5$, greater than the un-tolled total latency of 4.

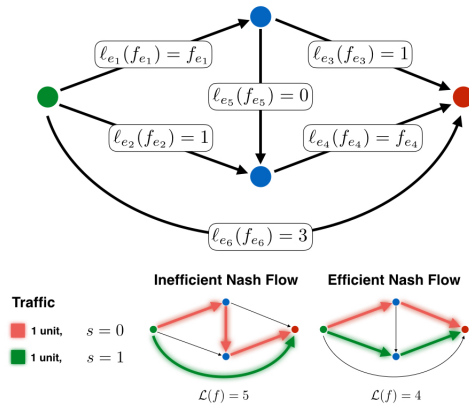


Fig. 1. Example 3.1: Marginal-cost tolls are not weakly robust to user heterogeneity. The user population has total mass 2, half of which has sensitivity $s_1 = 0$, half of which has sensitivity $s_2 = 1$. Here, marginal-cost tolls induce multiple Nash flows; two of these are depicted in the figure. On the left, all of the insensitive traffic (orange) is using the zig-zag path, with a latency of 2; all the sensitive traffic (green) is using the constant-latency edge, with a latency of 3 – for a total latency of 5. On the right flow, all traffic is experiencing a latency of 2, for a total latency of 4. Here, *any* homogeneous population using this network has the right-hand flow as a unique Nash flow. Thus, marginal-cost tolls incentivize a flow that is 25% worse than the un-influenced Nash flow; i.e., they are not weakly robust.

IV. OUR CONTRIBUTIONS

A. Weakly Robust Network Agnostic Taxation Mechanisms for Parallel Networks

Theorem 4.1 characterizes the space of weakly-robust network-agnostic taxation mechanisms for parallel-path networks and arbitrary admissible latency functions. Specifically, we show that all weakly-robust network-agnostic taxation mechanisms can be expressed as a simple generalization of classical marginal-cost tolls. Thus, perverse incentives can be systematically avoided on parallel-path networks by applying T^{gmcc} .

Theorem 4.1: A network agnostic taxation mechanism is weakly robust on \mathcal{G}_p if and only if it assigns *generalized marginal-cost* tolls

$$\tau_e^{\text{gmcc}}(f_e) = \kappa_1 \ell_e(f_e) + \kappa_2 f_e \ell'_e(f_e), \quad (9)$$

where $\kappa_1 > -1/S_U$, $\kappa_2 \geq 0$, and $\kappa_2 \leq \kappa_1 + 1/S_U$.

The proof of Theorem 4.1 appears in the appendix.

B. Impossibility for General Symmetric Networks

Given the characterization of weakly-robust tolls for parallel networks given in Theorem 4.1, it is an attractive goal to extend the analysis beyond parallel networks. Unfortunately, one need not go far before even T^{gmcc} fails to be weakly robust. In particular, we show in Theorem 4.2 that even the relatively restrictive condition of symmetry (all agents have access to the same set of paths) is not sufficient to guarantee the existence of nontrivial weakly-robust taxation mechanisms. This means that if network structure is unknown, the *only* way to avoid perverse incentives is effectively to do nothing.

Theorem 4.2: Let \mathcal{G} denote the class of all symmetric networks. If $S_L = 0$ and $S_U > 0$, a network-agnostic taxation

mechanism T is weakly robust on \mathcal{G} if and only if it is trivial; that is, for every network $G \in \mathcal{G}$ and every population s it satisfies

$$\mathcal{L}^{\text{nf}}(G, s, T) = \mathcal{L}^{\text{nf}}(G, \emptyset). \quad (10)$$

Note that the ‘‘trivial tolls’’ of Theorem 4.2 are any tolls satisfying $\kappa \ell_e(f_e)$ for $\kappa > 0$; these tolls have no effect on any Nash flow, and are thus strategically equivalent to tolls of $\tau_e(f_e) = 0$.

Proof: Theorem 4.1 gives necessary conditions for weak robustness (since $\mathcal{G}_p \subset \mathcal{G}$), so suppose we are given a taxation mechanism assigning taxes of $\tau_e(f_e) = \kappa_1 \ell_e(f_e) + \kappa_2 f_e \ell'_e(f_e)$, where $\kappa_1 > -1/S_U$, and $\kappa_2 \leq \kappa_1 + 1/S_U$. If $\kappa_2 = 0$, this taxation mechanism satisfies (10), so let $\kappa_2 > 0$. It suffices to exhibit a user population s (that is, a distribution of tax-sensitivities) and a network G such that $\mathcal{L}^{\text{nf}}(G, s, T) > \mathcal{L}^{\text{nf}}(G, \emptyset)$. We will do this with a population having two sensitivity values $s_1 < s_2$ and a network resembling that in Figure 1. Let s_2 satisfy $0 < s_2 \leq S_U$, and let $s_1 = 0$ to model the extreme case.¹ Construct the population as follows: let a unit mass of users have sensitivity s_1 and a unit mass have s_2 , for a total of 2 units of traffic. Define $\gamma_2 \triangleq \frac{s_2 \kappa_2}{1 + s_2 \kappa_1} \in (0, 1]$, so any agent with sensitivity s_2 sees an effective cost function² on edge e of

$$J_e(f_e) = \ell_e(f_e) + \gamma_2 f_e \ell'_e(f_e). \quad (11)$$

Now, let G be the network depicted in Figure 1, but let the latency function on edge e_6 be $\ell_{e_6}(f_{e_6}) = 2 + \gamma_2$. Enumerate the paths as follows: denote the ‘‘zig-zag’’ path $p_1 = \{e_1, e_5, e_4\}$, the remaining two paths in the upper sub-network $p_2 = \{e_1, e_3\}$ and $p_3 = \{e_2, e_4\}$, and the isolated constant-latency path $p_4 = \{e_6\}$; and denote the path flow of p_i by f_i . We will refer to paths p_1, p_2 , and p_3 as the ‘‘Braess subnetwork.’’

This population has at least two distinct Nash flows, corresponding to the two Nash flows depicted in Figure 1. We will write the inefficient flow (in which only half the agents use the Braess subnetwork) as $f^{\text{perverse}} \triangleq (1, 0, 0, 1)$ and the efficient flow (in which all agents use the Braess subnetwork) as $f^{\text{efficient}} \triangleq (0, 1, 1, 0)$. Note that in either flow, the delay experienced by agents on the Braess subnetwork is 2. However, in f^{perverse} , half of the agents (those on p_4) experience a delay of $2 + \gamma_2 > 2$. Thus, $\mathcal{L}(f^{\text{perverse}}) = 4 + \gamma_2$, while $\mathcal{L}(f^{\text{efficient}}) = 4$.

It can easily be verified that if tolls are removed, only $f^{\text{efficient}}$ remains as a Nash flow, which means that $\mathcal{L}(f^{\text{efficient}}) = \mathcal{L}^{\text{nf}}(G, \emptyset)$, or

$$\mathcal{L}^{\text{nf}}(G, s, T) > \mathcal{L}^{\text{nf}}(G, \emptyset) \quad (12)$$

and the considered tolls are not weakly-robust. ■

C. Implications for Altruistic Behavior

All of the foregoing has assumed that users are selfish and act with the sole objective of minimizing personal cost.

¹Here, it is also possible to show perversities by letting s_1 be any sufficiently-small positive number; we let $s_1 = 0$ for the sake of parsimony.

²See the argument in the proof of Lemma 5.2 in the Appendix.

However, real users may act altruistically, with the public good in mind. Recent research has investigated this in the α -altruism model, which assigns each user x an *altruism level* $\alpha_x \in [0, 1]$; a user with $\alpha = 0$ is totally selfish, whereas a user with $\alpha = 1$ is totally altruistic [14]. This is modeled by assuming that user x (with corresponding altruism level α_x) on edge e experiences a cost of

$$\begin{aligned} J_e^x(f_e) &= (1 - \alpha_x) \ell_e(f_e) + \alpha_x \frac{d}{df_e} (f_e \ell_e(f_e)) \\ &= \ell_e(f_e) + \alpha_x f_e \ell'_e(f_e). \end{aligned} \quad (13)$$

In other words, a totally-altruistic user fully accounts for the marginal effects that his actions have on those around him.

By comparing the cost functions induced by marginal-cost tolls (7) with the cost functions experienced by altruistic players (13), it is clear that there is a deep connection between this model of α -altruism and the theory of marginal-cost taxation. In essence, marginal-cost taxes are designed to induce artificial altruism in the user population.

The authors of [14] exhibit two contexts in non-atomic congestion games in which worst-case performance improves with increasing levels of altruism: the first is in general networks with homogeneous altruism, and the second is in parallel networks with heterogeneous altruism. In both cases, if the average level of altruism in the population increases, worst-case performance improves.

Given the equivalence of marginal-cost taxation and altruism, our Corollary 4.3 strengthens the parallel-network result of [14], showing that *on any network*, the worst-case flows are realized by a low-altruism homogeneous population. Note that [14] proves that altruism helps in worst-case over all parallel networks; our result is a network-by-network analysis and shows that altruism also helps on each individual network.

On the other hand, given our impossibility result in Theorem 4.2, Corollary 4.3 shows that increased altruism does not, in general, improve performance. That is, on the network in Figure 1, a totally-selfish population is associated with the efficient Nash flow, but a partially-altruistic population is associated with the inefficient Nash flow.

In the following, $\mathcal{L}_{\text{alt}}^{\text{nf}}(G, \alpha)$ denotes the worst-case Nash flow total latency on G for a given altruism distribution α , where users in α take altruism levels in the interval $[A_L, A_U] \subseteq [0, 1]$. A homogeneous altruism distribution in which all users have value A_L is denoted α^L .

Corollary 4.3: For any $G \in \mathcal{G}_p$,

$$\mathcal{L}_{\text{alt}}^{\text{nf}}(G, \alpha) \leq \mathcal{L}_{\text{alt}}^{\text{nf}}(G, \alpha^L). \quad (14)$$

However, for some $G \notin \mathcal{G}_p$, there exists an altruism distribution α satisfying

$$\mathcal{L}_{\text{alt}}^{\text{nf}}(G, \alpha) > \mathcal{L}_{\text{alt}}^{\text{nf}}(G, \alpha^L). \quad (15)$$

Proof: Any Nash flow induced by T^{gmc} is a Nash flow for some altruism distribution (see, e.g., the argument in the proof of Lemma 5.2 in the Appendix). Thus, Corollary 4.3 is implied by Theorems 4.1 and 4.2. ■

V. CONCLUSION

This paper has fully characterized the weakly-robust network-agnostic taxation mechanisms for parallel networks, and ruled them out entirely for general networks. We have shown that except in very limited settings (e.g., parallel networks), local computation of incentives carries a risk of causing harm. Among other things, this seems to indicate that information about the structure of the network is crucial for avoiding perverse incentives; characterizing these types of informational dependencies is the subject of ongoing work.

REFERENCES

- [1] M. Alizadeh, Y. Xiao, A. Scaglione, and M. van der Schaar, "Dynamic Incentive Design for Participation in Direct Load Scheduling Programs," *IEEE Journal of Selected Topics in Signal Processing*, vol. 8, no. 6, pp. 1111–1126, 2014.
- [2] R. Gopalakrishnan, J. R. Marden, and A. Wierman, "Potential games are necessary to ensure pure nash equilibria in cost sharing games," *Proc. 14th ACM Conf. on Electronic Commerce*, pp. 563–564, 2013.
- [3] V. Bonifaci, M. Salek, and G. Schäfer, "Efficiency of restricted tolls in non-atomic network routing games," in *Algorithmic Game Theory*, vol. 6982 LNCS, pp. 302–313, 2011.
- [4] M. Gairing, "Covering games: Approximation through non-cooperation," in *Internet and Economics*, pp. 184–195, 2009.
- [5] A. Kleiner, B. Nebel, and V. A. Ziparo, "A mechanism for dynamic ride sharing based on parallel auctions," in *International Joint Conference on Artificial Intelligence*, pp. 266–272, 2011.
- [6] I. Arieli, "Transfer Implementation in Congestion Games," *Dynamic Games and Applications*, vol. 5, no. 2, pp. 228–238, 2015.
- [7] C. Silva, "Application of Mechanism Design to Electric Power Markets (Republished)," *IEEE Trans. on Power Systems*, vol. 16, no. 4, pp. 862–869, 2001.
- [8] T. Roughgarden, "How Bad Is Selfish Routing?," *J. ACM*, vol. 49, no. 2, pp. 236–259, 2002.
- [9] L. Fleischer, K. Jain, and M. Mahdian, "Tolls for Heterogeneous Selfish Users in Multicommodity Networks and Generalized Congestion Games," in *Proc. 45th IEEE Symp. on Foundations of Computer Science*, (Rome, Italy), pp. 277–285, 2004.
- [10] G. Karakostas and S. Kolliopoulos, "Edge pricing of multicommodity networks for heterogeneous selfish users," in *Proc. 45th IEEE Symp. on Foundations of Computer Science*, (Rome, Italy), pp. 268–276, 2004.
- [11] M. J. Beckmann, C. B. McGuire, and C. B. Winsten, "Studies in the Economics of Transportation," 1955.
- [12] W. Sandholm, "Evolutionary Implementation and Congestion Pricing," *The Review of Economic Studies*, vol. 69, no. 3, pp. 667–689, 2002.
- [13] P. N. Brown and J. R. Marden, "Avoiding Perverse Incentives in Affine Congestion Games," in *55th IEEE Conf. on Decision and Control*, 2016.
- [14] P.-a. Chen, B. De Keijzer, D. Kempe, and G. Shaefer, "Altruism and Its Impact on the Price of Anarchy," *ACM Trans. Economics and Computation*, vol. 2, no. 4, 2014.
- [15] J. Wardrop, "Some Theoretical Aspects of Road Traffic Research," in *Proc. of the Institution of Civil Engineers, Part II, Vol. 1, No. 36*, pp. 352–362, 1952.
- [16] R. Cole, Y. Dodis, and T. Roughgarden, "Pricing network edges for heterogeneous selfish users," in *Proc. of the 35th ACM symp. on Theory of computing*, (New York, New York, USA), pp. 521–530, 2003.
- [17] I. Milchtaich, "Topological Conditions for Uniqueness of Equilibrium in Networks," *Mathematics of Operations Research*, vol. 30, no. 1, pp. 225–244, 2005.
- [18] P. N. Brown and J. R. Marden, "Optimal Mechanisms for Robust Coordination in Congestion Games," *to be published*, 2017.
- [19] P. N. Brown and J. R. Marden, "Studies on Robust Social Influence Mechanisms: Incentives for Efficient Network Routing in Uncertain Settings," *IEEE Control Systems Magazine*, vol. 37, no. 1, pp. 98–115, 2017.
- [20] D. Braess, A. Nagurney, and T. Wakolbinger, "On a Paradox of Traffic Planning," *Transportation Science*, vol. 39, pp. 446–450, nov 2005.
- [21] P. N. Brown and J. R. Marden, "The Robustness of Marginal-Cost Taxes in Affine Congestion Games," *Transactions on Automatic Control*, vol. PP, no. 99, 2016.

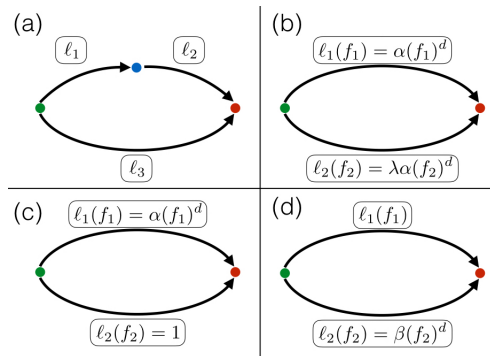


Fig. 2. Example networks used to prove Lemma 5.2.

APPENDIX: PROOF THEOREM 4.1

As a first step toward proving Theorem 4.1, we prove in Lemma 5.1 that all weakly-robust network-agnostic taxation mechanisms are essentially a generalized form of marginal-cost tolls. This lemma shows that tolls of the T^{gmc} form are necessary for weak robustness.

Lemma 5.1: If network-agnostic taxation mechanism T is weakly robust on \mathcal{G}_p , then for every edge e , it assigns taxation functions satisfying the conditions of Theorem 4.1.

Proof: Consider the four networks in Figure 2. The general approach is to exhibit networks whose un-tolled Nash flow is optimal; on these networks, weakly-robust tolls must not cause any deviations from the un-tolled flows. Thus, these networks can be used to deduce the necessary form of any weakly-robust tolls. In the interest of space, we give only an outline of the argument.

In the following, T represents a network-agnostic weakly-robust taxation mechanism. If the latency functions on the network in Figure 2(a) satisfy $\ell_1 + \ell_2 = \ell_3$, the corresponding tolling functions must as well; or T is additive: $T(\ell_1) + T(\ell_2) = T(\ell_1 + \ell_2)$. Similarly, by setting ℓ_1 , ℓ_2 , and ℓ_3 to constant functions, it can be shown that $T(\ell)$ is constant whenever ℓ is constant.

The network in Figure 2(b) shows that degree- d monomial latency functions of $\ell(f) = \alpha(f)^d$ must be assigned degree- d monomial tolling functions of $T(f) = \eta\alpha(f)^d$. Figure 2(c) shows that the tolls must not be too aggressive; that is, $\eta = \kappa_1 + \kappa_2 d$, where $\kappa_2 \leq \kappa_1 + 1/S_U$ is required to prevent users with sensitivity S_U from using edge 2 when r is low. The tolls must effectively be marginal-cost tolls "tuned" for users with S_U sensitivity.

Finally, applying the previously-proved facts, the network in Figure 2(d) is constructed with an arbitrary convex latency function and a monomial latency function specially-designed so that the uninfluenced flow is equal to the optimal flow. Since the monomial latency function's toll has already been specified, this allows direct computation of the toll assigned to the arbitrary convex latency function. ■

Next, Lemma 5.2 shows that Nash flows on parallel-path networks behave nicely under the influence of T^{gmc} . Specifically, Lemma 5.2 proves that the worst-case total latency on a parallel network with T^{gmc} is realized by a

low-sensitivity homogeneous population.

Lemma 5.2: Let s^L denote a homogeneous population in which every user has sensitivity $S_L \geq 0$, and denote by T^{gmc} a taxation mechanism satisfying the conditions of Lemma 5.1. For any $G \in \mathcal{G}_p$ and any population s in which every user has a sensitivity no less than S_L ,

$$\mathcal{L}^{\text{nf}}(G, s^L, T^{\text{gmc}}) \geq \mathcal{L}^{\text{nf}}(G, s, T^{\text{gmc}}). \quad (16)$$

Proof: When $\kappa_2 \geq 0$ and $\kappa_1 \geq \kappa_2 - 1/S_U$, the expression $\frac{s_x \kappa_2}{1 + s_x \kappa_1} \in [0, 1]$ and is monotone increasing in s_x . Thus, it can be assumed without loss of generality that $\kappa_1 = 0$, $\kappa_2 = 1$ and cost functions are simply given by

$$J_e^x(f_e) = \ell_e(f_e) + s_x f_e \ell'_e(f_e), \quad (17)$$

where $s_x \in [0, 1]$ for all x .

Let $\ell_e^*(f_e) \triangleq f_e \ell'_e(f_e)$, let $\ell_e^s(f_e) \triangleq \ell_e(f_e) + s \ell_e^*(f_e)$, and $\ell_e^{\text{mc}}(f_e) \triangleq \ell_e^1(f_e)$ to denote the nominal marginal-cost toll, the cost experienced by user with sensitivity s , and the marginal-cost function of edge e , respectively.

Following the arguments for Claim 1.1.2 in [21], it can be shown that on \mathcal{G}^p , for any population s under the influence of T^{gmc} , the following facts hold for any two paths satisfying $\ell_i(0) \leq \ell_j(0)$, $f_j^{\text{nf}} > 0$, and where a user x is on p_i and user y on p_j :

$$\ell_i^{\text{mc}}(f_i^{\text{nf}}) \geq \ell_j^{\text{mc}}(f_j^{\text{nf}}), \quad (18)$$

$$s_x \leq s_y. \quad (19)$$

When $\ell_i(0) < \ell_j(0)$, inequality (18) is strict. That is, the marginal costs in a Nash flow are ordered in such a way that lower-sensitivity agents always use higher marginal-cost paths than higher-sensitivity agents.

The remainder of the proof consists of showing that reducing agents' sensitivities (thereby making the population "more homogeneous") shifts agents from low marginal-cost paths to high marginal-cost paths, increasing the total latency. Formally, we define a mapping $\Sigma : [0, 1] \times \mathcal{S} \rightarrow \mathcal{S}$. For any starting population s^0 and any α , we will define $\Sigma(\alpha; s^0)$ as a right-shift of s^0 by α units. The sensitivity of user x in population $\Sigma(\alpha, s^0)$ is given by

$$\Sigma(\alpha, s^0)_x = \begin{cases} s_0(0) & \text{if } x \leq \alpha \\ s_0(x - \alpha) & \text{if } x > \alpha. \end{cases} \quad (20)$$

Because s is defined to be a nondecreasing function, this is equivalent to converting a mass of α of the most-sensitive users to a mass α of the least-sensitive users.

Again following the arguments in [21], assume without loss of generality that Nash flows have a finite number of sensitivity types. To be precise, given a Nash flow f^{nf} , we will assume for each path $p_i \in \mathcal{P} \setminus p_1$, each user has a sensitivity s_i that makes her indifferent between paths p_i and path p_{i-1} .

For brevity, let $f^{\text{nf}}(\alpha) \triangleq f^{\text{nf}}(\Sigma(\alpha; s_0))$. Our central goal will be to characterize the effect of marginal increases in α , which is expressed $\frac{\partial}{\partial \alpha} f^{\text{nf}}(\alpha)$. The following definition will be helpful in the proof:

Definition 1: In a Nash flow f^{nf} , paths p_i and p_j with $i < j$ are *strategically coupled* if s_i satisfies $\ell_i^{s_i}(f_i^{\text{nf}}) = \ell_j^{s_i}(f_j^{\text{nf}})$. That is, agents on the lower-index path are indifferent

between the two paths. $\mathcal{P}_i(f^{\text{nf}})$ denotes the set of paths that are strategically coupled to path p_i in f^{nf} .³

First, we show that the primary effect of an increase in α is to shift traffic from \mathcal{P}_n to \mathcal{P}_1 .

Proposition 5.3: For every path $p_i \in \mathcal{P}_1$, $\frac{\partial}{\partial \alpha} f_i^{\text{nf}}(\alpha) \geq 0$. For every path $p_j \in \mathcal{P}_n$, $\frac{\partial}{\partial \alpha} f_j^{\text{nf}}(\alpha) \leq 0$.

Proof: Let s_1 denote the sensitivity of agents using p_1 . Increasing α changes the sensitivity of a small fraction of high-sensitivity users to s_1 . By Definition 1 and (18)-(19), these users strictly prefer the paths in \mathcal{P}_1 to any other paths and thus switch to \mathcal{P}_1 . It is straightforward to show that this increases the flow on all paths in \mathcal{P}_1 , proving the first statement; a parallel argument proves the second, showing that these switching users must *leave* paths in \mathcal{P}_n . ■

Proposition 5.4: For any α , if $p_j \notin \mathcal{P}_1(\alpha)$ and $p_j \notin \mathcal{P}_n(\alpha)$, it holds that $\frac{\partial}{\partial \alpha} f_j^{\text{nf}}(\alpha) = 0$.

Proof: Definition 1 has been constructed to ensure that all users on \mathcal{P}_i strictly prefer it to \mathcal{P}_{i+1} ; thus, a marginal flow increase on \mathcal{P}_i cannot cause users to deviate to \mathcal{P}_{i+1} . Similarly, a flow decrease on \mathcal{P}_{i+1} cannot cause users on \mathcal{P}_i to deviate to \mathcal{P}_{i+1} . Considering these facts, Proposition 5.3 implies that shifting traffic from \mathcal{P}_n to \mathcal{P}_1 affects flow only on those two sets of paths, implying Proposition 5.4. ■

Proof of Lemma 5.2: In the following, $\nabla_f L(f)$ represents the gradient vector of L with respect to flow f given by $\{\ell_p^{\text{mc}}\}_{p \in \mathcal{P}}$, which by (18)-(19) is ordered descending. Let p_j be the highest-index path in \mathcal{P}_1 , and p_k be the lowest-index path in \mathcal{P}_n :

$$\begin{aligned} \frac{\partial}{\partial \alpha} L(f^{\text{nf}}(\alpha)) &= \nabla_f L(f^{\text{nf}}(\alpha)) \cdot \frac{\partial}{\partial \alpha} f^{\text{nf}}(\alpha) \\ &= \sum_{i \in \mathcal{P}_1 \cup \mathcal{P}_n} \ell_i^{\text{mc}}(f_i^{\text{nf}}(\alpha)) \frac{\partial}{\partial \alpha} f_i^{\text{nf}}(\alpha) \\ &\geq [\ell_j^{\text{mc}}(f_j^{\text{nf}}(\alpha)) - \ell_k^{\text{mc}}(f_k^{\text{nf}}(\alpha))] \geq 0. \end{aligned}$$

Since at every Nash flow $f^{\text{nf}}(\alpha)$ it is true that $\frac{\partial}{\partial \alpha} L(f^{\text{nf}}(\alpha)) \geq 0$, the definition of $\Sigma(\alpha, s_0)$ implies that for any initial sensitivity distribution s_0 ,

$$\mathcal{L}(f^{\text{nf}}(\Sigma(1, s_0))) \geq \mathcal{L}(f^{\text{nf}}(\Sigma(0, s_0))), \quad (21)$$

or that $\mathcal{L}^{\text{nf}}(G, s^L, T^{\text{gmc}}) \geq \mathcal{L}^{\text{nf}}(G, s, T^{\text{gmc}})$. ■

Proof of Theorem 4.1: Let s be any arbitrary sensitivity distribution, s^L be a homogeneous population with sensitivity S_L , and apply T^{gmc} . Lemma 5.2 ensures that

$$\mathcal{L}^{\text{nf}}(G, s^L, T^{\text{gmc}}) \geq \mathcal{L}^{\text{nf}}(G, s, T^{\text{gmc}}). \quad (22)$$

Let s^0 be a homogeneous population with sensitivity 0. Since s^0 is itself a low-sensitivity homogeneous population and s is a population in which all users have sensitivity no less than 0; apply Lemma 5.2 a second time to obtain

$$\mathcal{L}^{\text{nf}}(G, s^0, T^{\text{gmc}}) \geq \mathcal{L}^{\text{nf}}(G, s^L, T^{\text{gmc}}). \quad (23)$$

The left-hand side of (23) is simply the un-tolled total latency on G , so combining inequalities (22) and (23), we obtain

$$\mathcal{L}^{\text{nf}}(G, \emptyset) \geq \mathcal{L}^{\text{nf}}(G, s, T^{\text{gmc}}). \quad (24)$$

Since G and s were arbitrary, this implies that T^{gmc} is weakly-robust on \mathcal{G}_p . ■

³When clear from context, we write $\mathcal{P}_i(f^{\text{nf}})$ simply as \mathcal{P}_i .