The Tradeoff Between Altruism and Anarchy in Transportation Networks

Colton Hill and Philip N. Brown

Abstract— **In this paper, we ask how a system designer should endow autonomous vehicles with general routing policies that are guaranteed to prove beneficial in a wide variety of networks and adoption rates. Previous work has found that programming autonomous vehicles to be *altruistic* (i.e., choosing routes in consideration of their impact on aggregate congestion) can guarantee improvements in traffic congestion, provided that enough of the vehicles are autonomous. On the other hand, it is known that if not all vehicles are autonomous, altruistic autonomous vehicles can actually cause significant increases in traffic congestion. Moreover, the benefits of altruistic autonomous vehicles depend in complex ways on the fraction of vehicles that are autonomous, complicating the designer's decision. In this paper, we derive the optimal altruism levels for autonomous vehicles which obtain significant benefits while limiting the perverse effects of partial adoption, all without requiring the designer to know the fraction of vehicles that are autonomous. We demonstrate that our proposed altruism levels ensure significant improvements in routing efficiency with respect to previously-known worst-case guarantees.**

I. INTRODUCTION

As society and technology continue to become more intertwined, there is a growing need to comprehend and enhance the coordination between human social conduct and the technological functionality within these systems [1]. Optimally routing traffic in transportation networks is a canonical problem to study this interplay; it has been well-established that if individuals select routes to minimize their own travel time, suboptimal congestion can occur [2].

A popular way to measure the proximity between a given traffic flow and that of optimal traffic is the ratio between their total latencies. This concept is known as the *price of anarchy* and is defined as the ratio between the total latency in a system when agents act selfishly (modeled by a *Nash equilibrium*) with the total latency that can be achieved if a system designer coordinates all agents' routing for the overall benefit of the system [3]. The *price of anarchy* is used to better understand the degree to which selfish routing undermines network efficiency. In a given network, the price of anarchy is an upper bound on how much a system designer can improve traffic if she is able to implement a centralized routing policy that considers aggregate total latency, compared to the total latency that arises from a selfishly routed scheme. Because centralized routing control is typically not feasible, many studies have aimed to improve the price of anarchy in a less-centralized way, employing various techniques such as direct fleet routing strategies [4], information dissemination [5], and monetary rewards or penalties [6].

*This material is based upon work supported by the Air Force Office of Scientific Research under award number FA9550-23-1-0171 and the National Science Foundation under award number ECCS-2013779.

The authors are with the University of Colorado at Colorado Springs, CO 80918, USA. {chill13,pbrown2}@uccs.edu

One promising possibility is to design autonomous vehicles to be *altruistic*, choosing routes in consideration of their impact on aggregate road congestion [7], [8]. It is known that if the system designer could make all vehicles altruistic, all inefficiencies would be eliminated. However, it has recently been shown that if the system designer can make only a fraction of traffic altruistic (as would be the case with partial autonomous vehicle adoption) and the autonomous fraction can access only a subset of routes available, this heterogeneous altruism can paradoxically have a negative effect on congestion [9]. The potential harm stemming from altruism in a heterogeneous population can be modelled similarly to the price of anarchy. Here we take the ratio between the total latency of congestion resulting from a heterogeneous population, where a fraction of agents route altruistically, with that of the total latency stemming from a homogeneous selfish population; this ratio is known as the *perversity index* [9]. If the perversity index is 1, heterogeneous partial altruism never causes any harm relative to homogeneous selfishness. However, as the perversity index exceeds 1 for a class of networks, the total latency of heterogeneous populations can exceed their homogeneous analogs.

Thus, the key motivation for our work is that when a designer is able to centrally design the routing objectives of *all* agents in a network, purely altruistic latency functions always optimize aggregate congestion [7]. However when the designer can design only a fraction of agents' latency functions in the network, and if those agents have only partial network routing access, altruism can lead to significantly worsened overall congestion [10]. It remains an open question as to whether perversity arises when all traffic has access to all paths, but only a portion of agents' latency functions are designed.

In this paper, we ask the following question: suppose a fleet of autonomous traffic has access to a subset of a transportation network, and suppose a system designer were able to assign partially-altruistic routing policies to the autonomous agents, then how altruistic should the routing policies be? Can she select an altruism level to improve overall congestion without any risk of harm? Our model considers the impact of imbuing autonomous traffic with a level of altruism that improves congestion relative to all-selfish traffic while minimizing the risk of increasing congestion. This is done by supplying autonomous agents with altruistic latency functions that consider the impact their

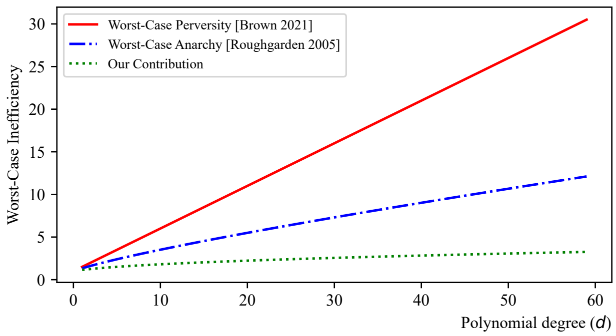


Fig. 1: The solid plot shows the perversity from seeking only to limit anarchy. The dashed plot represents the anarchy that arises when the only concern is to cause no harm. The dotted plot presents the main result of this paper, showing a marked decrease in network inefficiency when a tradeoff is sought between anarchy and perversity.

routing choices have on other agents in the network, but remain cognizant of their potential harm. A summary of our contributions is as follows:

1) Theorem 3.1 provides the unique altruism level that optimally balances between reducing anarchy and limiting perversity (given modest restrictions on network structure). The bound on worst-case congestion that results from this tradeoff between perversity and anarchy is presented relative to related work in Figure 1; note that our improved bound is far lower than previous guarantees.
2) Lemmas 3.2 and 3.3 provide a tight upper bound on the perversity index when the autonomous population is less than half the total population.

II. Model and Related Work

A. Routing Problem

We consider a routing problem for network (V, E), consisting of vertex set V and edge set E. We write $\mathcal{P} \subseteq 2^E$ to denote the set of *paths* accessible to agents, where each path $p \in \mathcal{P}$ comprises a set of edges connecting common origin o to common destination t. We restrict network topology to *series-parallel*; a network is series-parallel if it is

1) a single edge,
2) two series-parallel networks connected in series, or
3) two series-parallel networks connected in parallel [11].

A unit mass of traffic is routed from o to t and is composed of two types, those belonging to an *autonomous* fleet and the remaining uninfluenced traffic is referred to as *selfish*. Autonomous agents comprise mass r^{a}, and selfish users make up mass r^{s}, such that $r^{\mathrm{a}} + r^{\mathrm{s}} = 1$. Autonomous traffic can access an arbitrary subset of paths $\mathcal{P}^{\mathrm{a}} \subseteq \mathcal{P}$, and selfish traffic can access the entire path set $\mathcal{P}$; for convenience, we denote $\mathcal{P}^{\mathrm{s}} = \mathcal{P}$ to refer to the selfish path set.

For each type $\theta \in \{\mathrm{a}, \mathrm{s}\}$, x_p^θ denotes the flow of agents of type θ using path $p \in \mathcal{P}^\theta$. A *feasible flow* for type θ is an assignment of r^θ mass of traffic to paths in $\mathcal{P}^\theta$, denoted by $x^\theta \in \mathbb{R}_{\geq 0}^{|\mathcal{P}^\theta|}$, such that $\sum_{p \in \mathcal{P}} x_p^\theta = r^\theta$. A *network flow* is a combined allocation of autonomous agents and selfish users to paths, denoted $x \in \mathbb{R}_{\geq 0}^{|\mathcal{P}|}$, such that $x_p := \sum_{\theta : p \in \mathcal{P}^\theta} x_p^\theta$, where the flows x^θ are feasible for their respective types.

Provided a network flow x, the flow on edge $e \in E$ is given by $x_e = \sum_{p : e \in p} x_p$, and we denote the flow of type $\theta = \{\mathrm{a}, \mathrm{s}\}$ on edge e by x_e^θ. For each edge $e \in E$, commute time is expressed as a function of traffic flow and is associated with a latency function $\ell_e : [0, 1] \to [0, \infty)$, where $\ell_e(x_e)$ is the cost experienced by agents on edge e with edge flow x_e. We assume the latency function for each edge is a non-decreasing, convex posynomial. For clarity, a posynomial is defined to be a polynomial with non-negative coefficients. So, for every $e \in E$,

$$\ell_e(x_e) = \sum_{i=0}^{d} a_{e,i} x_e^i, \tag{1}$$

where $d \in \mathbb{N}$, and $a_{e,i} \in \mathbb{R}_{\geq 0}$.

We measure the cost of a flow by the *total latency*, given by

$$\mathcal{L}(x) = \sum_{e \in E} x_e \ell_e(x_e) = \sum_{p \in \mathcal{P}} x_p \ell_p(x), \tag{2}$$

where we define the latency of path p, given flow x, as $\ell_p(x) := \sum_{e \in p} \ell_e(x_e)$.

An instance of a *routing problem* is fully specified by the tuple $G = (V, E, \{\ell_e\}, \mathcal{P}^{\mathrm{s}}, \mathcal{P}^{\mathrm{a}}, r^{\mathrm{a}})$, and we write $\mathcal{G}(d, r^{\mathrm{a}})$ to denote the set of all routing problems on series-parallel networks with posynomial latency functions of degree at most d and autonomous population r^{a}.

B. Heterogeneous Routing Game

In order to understand how autonomous vehicles can affect congestion within the context of a heterogeneous population, we model the routing problem as a nonatomic congestion game. That is, each type of traffic is composed of infinitely many infinitesimal agents, where the cost selfish users experience and autonomous agents are programmed to interpret are determined by their type. Given a flow x, the cost a selfish user experiences for using path $p \in \mathcal{P}$ is the latency of the path:

$$\ell_p(x) := \sum_{e \in p} \ell_e(x_e), \tag{3}$$

Intuitively, (3) assumes selfish users are uniform with regard to their routing policy.

To quantify how the system designer would influence an autonomous fleet to improve overall congestion, we introduce an altruism design parameter, $\alpha \in [0, 1]$ that the system designer uses to modulate the autonomous agents' interest in the impact their routing choices have on others. The altruism parameter is used in concert with the well known marginal-cost function [12], so that autonomous traffic is programmed to interpret the *α-marginal cost* function. For a given edge, the α-marginal cost function, denoted $\ell_e^{\mathrm{mc}_\alpha}$, is given by

$$\ell_e^{\mathrm{mc}_\alpha}(x_e) := \ell_e(x_e) + \alpha x_e \ell_e'(x_e), \tag{4}$$

where ℓ' denotes the derivative of ℓ. Thus, autonomous traffic is programmed to interpret the cost of using path $p \in \mathcal{P}^{\mathrm{a}}$ as the α-marginal cost of the path:

$$\ell_p^{\mathrm{mc}_\alpha}(x) := \sum_{e\in p} \ell_e^{\mathrm{mc}_\alpha}(x_e). \tag{5}$$

This path cost can be interpreted as the sum of the latency of the path, and the agent's sensitivity to their marginal effect on other agents on the path. An altruistic agent accounts for the negative congestion they impose on others relative to their altruism level. Note that altruistic latency functions of this type are essential in the use of taxation to influence behavior in congestion games [9]. When $\alpha = 0$, autonomous latency functions resolve to the cost selfish users experience. When $\alpha = 1$, autonomous agents fully account for their negative congestion [10], and (5) resolves to the well-known marginal-cost function for a path:

$$\ell_p^{\mathrm{mc}}(x) := \sum_{e\in p} \left[\ell_e(x_e) + x_e \ell_e'(x_e)\right]. \tag{6}$$

We assume each agent travels from origin o to destination t using the minimum-cost path from those available in their path set. We call a flow x a *Nash flow* if all agents are individually using minimum-cost paths relative to the choices of others. That is, there exists a feasible x^{a} such that the following condition is satisfied for autonomous traffic:

$$\forall p, p' \in \mathcal{P}^{\mathrm{a}}, x_p^{\mathrm{a}} > 0 \Longrightarrow \ell_p^{\mathrm{mc}_\alpha}(x) \leq \ell_{p'}^{\mathrm{mc}_\alpha}(x), \tag{7}$$

and there exists a feasible x^{s} such that the following condition is satisfied for selfish traffic:

$$\forall p, p' \in \mathcal{P}^{\mathrm{s}}, x_p^{\mathrm{s}} > 0 \Longrightarrow \ell_p(x) \leq \ell_{p'}(x). \tag{8}$$

Further, the existence of a Nash flow for any nonatomic congestion game of the aforementioned structure is well known [13]. An instance of a *heterogeneous routing game* with routing problem G and altruism level α is fully specified by the tuple (G, α).

C. Performance Metrics: Price of Anarchy and Perversity Index

A system designer may wish to select altruism levels that suppress the total latency of routing relative to optimal traffic routing. We write $\mathcal{L}^{\mathrm{nf}}(G, \alpha)$ to denote the total latency of a worst-case Nash flow for heterogeneous routing problem G and altruism parameter α, and let $\mathcal{L}^*(G)$ denote the optimal flow on G. Then we define the *price of anarchy* of a class of games $\mathcal{G}(d, r^{\mathrm{a}})$ as the worst-case ratio of the total latency of a heterogeneous Nash flow with the total latency of an optimal flow:

$$\mathrm{PoA}(d, r^{\mathrm{a}}, \alpha) := \sup_{G\in\mathcal{G}(d,r^{\mathrm{a}})} \frac{\mathcal{L}^{\mathrm{nf}}(G, \alpha)}{\mathcal{L}^*(G)}. \tag{9}$$

Given a heterogeneous routing game (G, α), we write $(G, 0)$ to denote the *homogeneous* version of the routing game, where all agents in traffic behave selfishly. $(G, 0)$ is identical to (G, α) with the exception that autonomous traffic is now programmed with the selfish latency function (3); path sets and network topology remain the same.

The system designer will seek altruism levels that limit the worst-case total latency relative to uninfluenced traffic routing. We write $\mathcal{L}^{\mathrm{nf}}(G, 0)$ to denote the total latency of a Nash flow for $(G, 0)$. To capture this worst-case harm, we study the *perversity index* to characterize the worst-case effects of partial altruism in heterogeneous networks [10]. The perversity index captures the potential harm caused when autonomous traffic is partially altruistic relative to when the entire population is completely selfish. Thus, it is similar to the *deviation ratio* of [14] and the *price of risk aversion* of [15]. The perversity index is defined as the worst-case ratio of the total latency of a heterogeneous Nash flow with the total latency of a homogeneous selfish Nash flow:

$$\mathrm{PI}(d, r^{\mathrm{a}}, \alpha) := \sup_{G\in\mathcal{G}(d,r^{\mathrm{a}})} \frac{\mathcal{L}^{\mathrm{nf}}(G, \alpha)}{\mathcal{L}^{\mathrm{nf}}(G, 0)}. \tag{10}$$

Intuitively, if $\mathcal{G}$ has a large perversity index, there exist networks in that class of games for which heterogeneous Nash flows are much worse than the corresponding homogeneous Nash flows. In the case that $r^{\mathrm{a}} = 0$, it trivially holds that since no traffic is autonomous, no harm can be done, so that $\mathrm{PI}(d, 0, \alpha) = 1$. Furthermore, the perversity index is bounded above by the price of anarchy: $\mathrm{PI}(d, r^{\mathrm{a}}, \alpha) \leq \mathrm{PoA}(d, r^{\mathrm{a}}, \alpha)$ since $\mathcal{L}^*(G) \leq \mathcal{L}^{\mathrm{nf}}(G, 0)$ for any G.

For $\alpha = 1$, it is known that when all agents are autonomous, the price of anarchy is minimized; likewise, it is known that when exactly half of traffic is autonomous, the perversity index is maximized. The aim of this paper is to use the known results on anarchy and perversity to find an altruism level that minimizes the price of anarchy while not allowing the perversity index to become too high. That is, we use the parameters that characterize best-case anarchy ($r^{\mathrm{a}} = 1$) and worst-case perversity ($r^{\mathrm{a}} = 1/2$), to find the level of altruism that best balances the tradeoff between anarchy and perversity, given that r^{a} may not be known to the system designer a priori and may change over time. To be precise, for any posynomial of degree at most d, we define the *robust altruism level* $\alpha^*(d)$ as one which satisfies the following:

$$\alpha^*(d) \in \operatorname*{arg\,inf}_{\alpha\in[0,1]} \max\left\{\mathrm{PoA}\left(d, 1, \alpha\right), \mathrm{PI}\left(d, \frac{1}{2}, \alpha\right)\right\}. \tag{11}$$

That is, $\alpha^*(d)$ minimizes the maximum between the price of anarchy where all traffic is autonomous, and the perversity index when exactly half of traffic is autonomous.

D. Related Work

1) Altruism and Alternate Experienced Costs in Congestion Games: It is known that in series-parallel networks with affine latency functions and $r^{\mathrm{a}} = 1$, the total latency in the presence of altruism is always less than the total latency when the population is all-selfish [16]. Further, it is known that in parallel networks, altruism can produce unbounded improvements over selfishness [17]. In parallel networks when $\mathcal{P}^{\mathrm{a}} = \mathcal{P}^{\mathrm{s}} = \mathcal{P}$, the price of anarchy is improved under heterogeneous altruism compared to homogeneous selfishness; additionally, the perversity index in these networks is unity [7], [18]. However, in series-parallel networks with general latency functions and altruistic fraction $r^{\mathrm{a}} \in (0,1)$ where $\mathcal{P}^{\mathrm{a}} \neq \mathcal{P}$, perversity is unbounded as the degree of latency functions $d \to \infty$ [10]. The worst-case perversity outside these networks is an open question.

The concept of perversity has been studied as a measure of how alterations in experienced latency affects total congestion compared to uninfluenced Nash flows; these works study "deviation ratio" [14] and the "price of risk aversion" [19] as analogs to the perversity index here. Other works study the effects altered payoff biases such as pessimism [20] and uncertainty [16] have on the price of anarchy.

2) Marginal-Cost Pricing: If for each edge e, all users are homogeneously charged a price of

$$\tau_e^{\mathrm{mc}}(x_e) = x_e \ell_e'(x_e), \tag{12}$$

Nash flows preferable to those when users are charged no toll are produced regardless of network topology [16], [21]. However, these guarantees vanish if traffic is heterogeneous in its toll-sensitivity [9]. The results of this paper regarding the negative impacts resulting from the design of heterogeneous partially altruistic systems imply analogous outcomes in terms of the potential harm caused by sub-optimal marginal-cost pricing.

III. Contribution: An Optimal Altruism Parameter for System Design

A system designer that creates the latency functions informing the routing policies of an autonomous fraction of traffic will likely be concerned with how to reduce overall congestion. However, it is known that heterogeneous altruistic traffic can be perverse. Thus, a tradeoff must be sought that balances the goal of reducing congestion with the risk of harm, especially when a system designer may have no prior knowledge of the magnitude of r^{a}. Our main result optimizes this tradeoff (presented graphically in Figure 2), and is proven in the following theorem.

Theorem 3.1: Consider the class of series-parallel nonatomic congestion games with posynomial path latencies of degree at most d. Then, the robust altruism level $\alpha^*(d)$ as defined by (11) is the unique solution $\alpha \in [0,1]$ of the following equation:

$$1 + \frac{\alpha d}{2} = \frac{1}{1 + \alpha d - d\left(\frac{1+\alpha d}{1+d}\right)^{\left(1+\frac{1}{d}\right)}}. \tag{13}$$

The proof of Theorem 3.1 is presented in detail at the end of this section; all Lemma proofs appear in the Appendix. We first offer a brief outline of the proof, and discuss the required lemmas. The proof is completed in three steps:

1) Lemmas 3.2 and 3.3 provide a tight bound for the perversity index when $r^{\mathrm{a}} = 1/2$.
2) Lemma 3.4 is derived from [7, Theorem 6.7], and provides a tight bound on the price of anarchy for $r^{\mathrm{a}} = 1$.

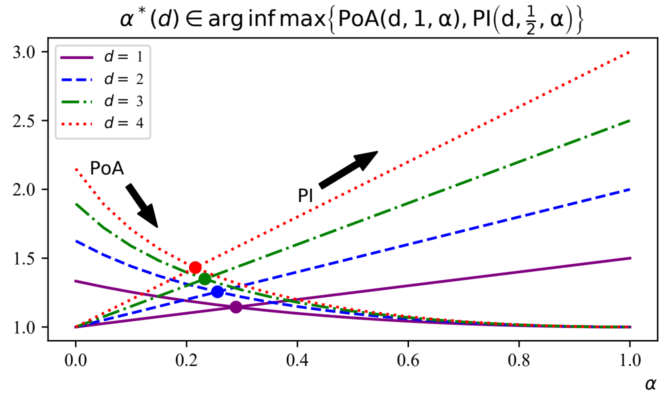


Fig. 2: Price of anarchy (decreasing) plotted with respect to $r^{\mathrm{a}} = 1$, and perversity index (increasing) plotted with respect to $r^{\mathrm{a}} = 1/2$, each for $\alpha \in [0,1]$ with posynomial latency functions of degree $d \in \{1,\ldots,4\}$.

3) The proof is completed by showing $\mathrm{PoA}\,(d,1,\alpha)$ and $\mathrm{PI}\left(d,\frac{1}{2},\alpha\right)$ intersect at a unique α.

We begin with Lemma 3.2, which provides an upper bound on the perversity index for any r^{a}. Intuitively, it demonstrates that as the fraction of autonomous traffic grows smaller, the heterogeneous Nash flow is increasingly similar to a homogeneous selfish Nash flow.

Lemma 3.2: Let $\mathcal{G}(d, r^{\mathrm{a}})$ be the class of series-parallel nonatomic congestion games with posynomial path latencies of degree at most d and autonomous population r^{a}. Assume $G \in \mathcal{G}(d, r^{\mathrm{a}})$, and let x be a heterogeneous Nash flow for (G, α) and $\bar{x}$ be a Nash flow for homogenized network $(G, 0)$. Then the following holds:

$$\mathcal{L}(x) \leq (1 + \alpha d r^{\mathrm{a}})\mathcal{L}(\bar{x}). \tag{14}$$

Next, Lemma 3.3 shows that the bound in (14) is in fact tight for $r^{\mathrm{a}} \in \left[0, \frac{1}{2}\right]$. Thus, for any posynomial degree d, we can construct a network that has a worst-case perversity index.

Lemma 3.3: For any d, and any $r^{\mathrm{a}} \in \left[0, \frac{1}{2}\right]$, there exists a nonatomic congestion game $G \in \mathcal{G}(d, r^{\mathrm{a}})$ such that the following holds:

$$\mathcal{L}(x) \geq (1 + \alpha d r^{\mathrm{a}})\mathcal{L}(\bar{x}), \tag{15}$$

where x is a heterogeneous Nash flow for (G, α) and $\bar{x}$ is a Nash flow for homogenized network $(G, 0)$.

Our final lemma is borrowed from [7, Theorem 6.7] and modified specifically for our context. It furnishes a tight bound on the price of anarchy for any α and any d, provided all traffic is autonomous.

Lemma 3.4 (Chen et al., 2014): Let $\mathcal{G}(d, 1)$ be the class of series-parallel nonatomic congestion games with posynomial path latencies of degree at most d, where all traffic is autonomous. Assume $G \in \mathcal{G}(d, 1)$, then the price of anarchy is given by

$$\mathrm{PoA}\,(d,1,\alpha) = \frac{1}{1 + \alpha d - d\left(\frac{1+\alpha d}{1+d}\right)^{\left(1+\frac{1}{d}\right)}}. \tag{16}$$

With the necessary lemmas in hand, we proceed with the proof of our main result.

Proof of Theorem 3.1: Lemma 3.2 shows that $\mathrm{PI}(d, r^{\mathrm{a}}, \alpha) \leq 1 + \alpha d r^{\mathrm{a}}$ for all $r^{\mathrm{a}} \in [0,1]$, and Lemma 3.3 provides a set of problem instances where $\mathrm{PI}(d, r^{\mathrm{a}}, \alpha) \geq 1 + \alpha d r^{\mathrm{a}}$ for any $r^{\mathrm{a}} \in \left[0, \frac{1}{2}\right]$. Thus we can see that for $r^{\mathrm{a}} = \frac{1}{2}$, the perversity index has a tight bound:

$$\mathrm{PI}\left(d, \frac{1}{2}, \alpha\right) = 1 + \frac{\alpha d}{2}. \tag{17}$$

Similarly, given $r^{\mathrm{a}} = 1$, Lemma 3.4 provides a tight bound for the price of anarchy (16), for any α. It is trivial to show that (17) is increasing as a function of α. To show that (16) is decreasing as a function of α, it can be demonstrated that the partial derivative of $\mathrm{PoA}(d, 1, \alpha)$ with respect to α is negative. To complete the proof, we must show that for any d, there exists an α such that $\mathrm{PI}\left(d, \frac{1}{2}, \alpha\right) = \mathrm{PoA}\left(d, 1, \alpha\right)$ is guaranteed. We make use of the following facts: it is clear that $\mathrm{PI}\left(d, \frac{1}{2}, 0\right) = 1$, and

$$\mathrm{PoA}\left(d, 1, 0\right) = \frac{1}{1 - d\left(\frac{1}{1+d}\right)^{\left(1+\frac{1}{d}\right)}} \geq 1, \tag{18}$$

where (18) follows from the fact that the denominator is not greater than 1. Also, $\mathrm{PI}\left(d, \frac{1}{2}, 1\right) = 1 + \frac{d}{2}$, and $\mathrm{PoA}\left(d, 1, 1\right) = 1$ Hence, we have that $\mathrm{PoA}\left(d, 1, 0\right) \geq \mathrm{PI}\left(d, \frac{1}{2}, 0\right)$, and $\mathrm{PoA}\left(d, 1, 1\right) \leq \mathrm{PI}\left(d, \frac{1}{2}, 1\right)$; it was also shown that $\mathrm{PoA}\left(d, 1, \alpha\right)$ is decreasing as a function of α, and $\mathrm{PI}\left(d, \frac{1}{2}, \alpha\right)$ is increasing as a function of α. Thus, it is guaranteed that a unique α exists such that $\mathrm{PI}\left(d, \frac{1}{2}, \alpha\right) = \mathrm{PoA}\left(d, 1, \alpha\right)$; that is, the robust altruism level $\alpha^*(d)$ defined by (11) is the unique solution to (13). ■

IV. CONCLUSIONS

We have shown that a system designer is guaranteed a unique level of altruism exists and can be found that enables an autonomous fraction of traffic to maximize the benefit to society, while minimizing the potential negative impacts that altruism can produce on series-parallel networks with posynomial latency functions. Future work will focus on extending these results to populations without topological restrictions, wide-ranging altruism levels, and arbitrary latency functions; future work will also investigate fully bounding the perversity that arises from partial altruism with respect to any fraction of altruistic traffic.

APPENDIX

Here we include the proofs of all supporting lemmas. We often write $\Lambda_s(\bar{x})$ to denote the common latency selfish users experience given homogeneous Nash flow $\bar{x}$.

Our final lemma presents an upper bound on the latency autonomous traffic experiences in a heterogeneous Nash flow.

Lemma 4.1: Let G be a series-parallel nonatomic congestion game with posynomial latencies of degree at most d. Assume x is a heterogeneous Nash flow for (G, α), $\bar{x}$ is a Nash flow for homogenized network $(G, 0)$, and that $\mathrm{PI}(d, r^{\mathrm{a}}, \alpha) > 1$. For any $p \in \mathcal{P}^{\mathrm{a}}$, if $x_p^{\mathrm{a}} > 0$, then we have the following:

$$\ell_p^{\mathrm{mc}_\alpha}(x) \leq (1 + \alpha d)\Lambda_{\mathrm{s}}(\bar{x}). \tag{19}$$

Proof of Lemma 4.1: To prove the bound in (19), we will show the following:

$$\ell_{\overline{p}}^{\mathrm{mc}_\alpha}(x) \leq (1 + \alpha d)\Lambda_{\mathrm{s}}(\bar{x}), \tag{20}$$

where $\overline{p} := \arg\max\{\ell^{\mathrm{mc}_\alpha}(x); p \in \mathcal{P}^{\mathrm{a}}, x_p^{\mathrm{s}} > 0\}$.

We first show it is without loss of generality that $\overline{p}$ exists. Assume first that $\overline{p}$ does not exist. Then denote $\delta_p := x_p - \bar{x}_p$; for type $\theta \in \{\mathrm{a}, \mathrm{s}\}$, denote $\delta_p^\theta := x_p^\theta - \bar{x}_p^\theta$, and define $\mathcal{P}_{\mathrm{c}}^{\mathrm{a}} := \mathcal{P}^{\mathrm{s}} \setminus \mathcal{P}^{\mathrm{a}}$. By convexity of $\mathcal{L}$, we have

$$\begin{aligned}
\mathcal{L}(x) - \mathcal{L}(\bar{x}) &\leq \sum_{p \in \mathcal{P}} \delta_p \ell_p^{\mathrm{mc}}(x) \\
&= \sum_{p \in \mathcal{P}^{\mathrm{a}}} \delta_p \ell_p^{\mathrm{mc}}(x) + \sum_{p \in \mathcal{P}_{\mathrm{c}}^{\mathrm{a}}} \delta_p^{\mathrm{s}} \ell_p^{\mathrm{mc}}(x) \\
&= \sum_{p \in \mathcal{P}^{\mathrm{a}}} \left[\delta_p^{\mathrm{a}} \ell_p^{\mathrm{mc}}(x) + \delta_p^{\mathrm{s}} \ell_p^{\mathrm{mc}}(x)\right] + \sum_{p \in \mathcal{P}_{\mathrm{c}}^{\mathrm{a}}} \delta_p^{\mathrm{s}} \ell_p^{\mathrm{mc}}(x) \\
&\leq \sum_{p \in \mathcal{P}^{\mathrm{a}}} \delta_p^{\mathrm{s}} \ell_p^{\mathrm{mc}}(x) + \sum_{p \in \mathcal{P}_{\mathrm{c}}^{\mathrm{a}}} \delta_p^{\mathrm{s}} \ell_p^{\mathrm{mc}}(x),
\end{aligned} \tag{21}$$

where (21) follows from the fact that, for any paths p, q, if $\delta_p^{\mathrm{a}} > 0$ and $\delta_q^{\mathrm{a}} > 0$ ($\delta_q^{\mathrm{a}} \leq 0$, respectively), it holds that $\ell_p^{\mathrm{mc}} = \ell_q^{\mathrm{mc}}$ ($\ell_p^{\mathrm{mc}} \leq \ell_q^{\mathrm{mc}}$, respectively) so that $\sum_{p \in \mathcal{P}^{\mathrm{a}}} \delta_p^{\mathrm{a}} \ell_p^{\mathrm{mc}}(x) \leq 0$. Since $\overline{p}$ is assumed to not exist for the moment, it should be clear that for each $p \in \mathcal{P}^{\mathrm{a}}$, $\delta_p^{\mathrm{s}} \leq 0$; so that $\sum_{p \in \mathcal{P}^{\mathrm{a}}} \delta_p^{\mathrm{s}} \ell_p^{\mathrm{mc}}(x) \leq 0$.

For any path p, define $\delta_p^{\mathrm{s}+} = \delta_p^{\mathrm{s}}$ if $\delta_p^{\mathrm{s}} > 0$, and $\delta_p^{\mathrm{s}+} = 0$ otherwise, so that $\sum_{p \in \mathcal{P}_{\mathrm{c}}^{\mathrm{a}}} \delta_p^{\mathrm{s}} \ell_p^{\mathrm{mc}}(x) \leq \sum_{p \in \mathcal{P}_{\mathrm{c}}^{\mathrm{a}}} \delta_p^{\mathrm{s}+} \ell_p^{\mathrm{mc}}(x)$, and $\sum_{p \in \mathcal{P}_{\mathrm{c}}^{\mathrm{s}}} \delta_p^{\mathrm{s}+} \ell_p^{\mathrm{mc}}(x) \geq 0$. Now, for each $p \in \mathcal{P}_{\mathrm{c}}^{\mathrm{a}}$ such that $\delta_p^{\mathrm{s}} > 0$, notice that $\ell_p(x) = \ell_p(\bar{x})$, thus ℓ_p is constant. So, (21) continues as

$$\begin{aligned}
\sum_{p \in \mathcal{P}^{\mathrm{a}}} \delta_p^{\mathrm{s}} \ell_p^{\mathrm{mc}}(x) + \sum_{p \in \mathcal{P}_{\mathrm{c}}^{\mathrm{a}}} \delta_p^{\mathrm{s}} \ell_p^{\mathrm{mc}}(x) &\leq \sum_{p \in \mathcal{P}^{\mathrm{a}}} \delta_p^{\mathrm{s}} \ell_p^{\mathrm{mc}}(x) + \sum_{p \in \mathcal{P}_{\mathrm{c}}^{\mathrm{a}}} \delta_p^{\mathrm{s}+} \ell_p^{\mathrm{mc}}(x) \\
&= \sum_{p \in \mathcal{P}^{\mathrm{a}}} \delta_p^{\mathrm{s}} \ell_p^{\mathrm{mc}}(x) + \sum_{p \in \mathcal{P}_{\mathrm{c}}^{\mathrm{a}}} \delta_p^{\mathrm{s}+} \ell_p(x) \\
&\leq 0,
\end{aligned} \tag{22}$$

where (22) follows from the fact that for any $p \in \mathcal{P}_{\mathrm{c}}^{\mathrm{a}}$ and any $q \in \mathcal{P}^{\mathrm{a}}$, $\ell_p(x) \leq \ell_q(x) \leq \ell_q^{\mathrm{mc}}(x)$. Thus we have that $\mathcal{L}(x) - \mathcal{L}(\bar{x}) \leq 0$, and thus the perversity index is less than 1 when $\overline{p}$ does not exist.

Hence we may assume $\overline{p}$ exists, and continue by supposing that the bound in (20) is false, then we have:

$$\begin{aligned}
(1 + \alpha d)\Lambda_{\mathrm{s}}(\bar{x}) &< \ell_{\overline{p}}^{\mathrm{mc}_\alpha}(x) \\
&= \sum_{e \in \overline{p}} \ell_e^{\mathrm{mc}_\alpha}(x_e) \\
&= \sum_{e \in \overline{p}} \left[a_{e,0} + \sum_{i=1}^{d} (1 + \alpha i) a_{e,i} x_e^i\right] \\
&\leq \sum_{e \in \overline{p}} \left[a_{e,0} + (1 + \alpha d) \sum_{i=1}^{d} a_{e,i} x_e^i\right].
\end{aligned} \tag{23}$$

It is clear that $1 + \alpha d > 0$, so we have that

$$\begin{aligned}\Lambda_{\mathrm{s}}(\bar{x}) &< \sum_{e \in \overline{p}} \left[\frac{a_{e,0}}{(1+\alpha d)} + \sum_{i=1}^{d} a_{e,i} x_e^i \right] \\ &\leq \sum_{e \in \overline{p}} \left[a_{e,0} + \sum_{i=1}^{d} a_{e,i} x_e^i \right] \\ &= \ell_{\overline{p}}(x),\end{aligned}$$

contradicting [16, Proposition 2]. Hence we have obt the bound in (20), and Nash flow conditions give u bound in (19) by extension.

Proof of Lemma 3.2: We first derive an upper bour the cost of autonomous traffic in x. Here, let path $\overline{p}$ d $x^{\mathrm{a}}_{\overline{p}} > 0$, that is, $\overline{p}$ is any path used by altruistic traffic so that for any $p \in \mathcal{P}^{\mathrm{a}}$, if $\ell_p^{\mathrm{mc}_\alpha}(x) > \ell_{\overline{p}}^{\mathrm{mc}_\alpha}(x)$, then x^{a}_p

$$\begin{aligned}\sum_{p \in \mathcal{P}} x_p^{\mathrm{a}} \ell_p(x) &\leq \sum_{p \in \mathcal{P}} x_p^{\mathrm{a}} \ell_p^{\mathrm{mc}_\alpha}(x) \\ &= \ell_{\overline{p}}^{\mathrm{mc}_\alpha}(x) \sum_{p \in \mathcal{P}} x_p^{\mathrm{a}} \\ &\leq r^{\mathrm{a}}(1+\alpha d)\Lambda_{\mathrm{s}}(\bar{x}), \qquad (25)\end{aligned}$$

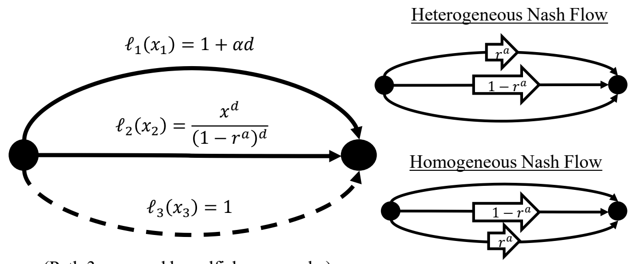
where (25) follows from Lemma 4.1. Now, we are ready to compute the bound in (14):

$$\begin{aligned}\mathcal{L}(x) &= \sum_{p \in \mathcal{P}} x_p^{\mathrm{s}} \ell_p(x) + \sum_{p \in \mathcal{P}} x_p^{\mathrm{a}} \ell_p(x) \\ &\leq (1 - r^{\mathrm{a}})\Lambda_{\mathrm{s}}(\bar{x}) + r^{\mathrm{a}}(1+\alpha d)\Lambda_{\mathrm{s}}(\bar{x}) \qquad (26) \\ &\leq (1 + \alpha d r^{\mathrm{a}})\mathcal{L}(\bar{x}), \qquad (27)\end{aligned}$$

where (26) follows from [16, Proposition 2], and (27) follows from $\bar{x}$ being an all-selfish Nash flow. ■

Proof of Lemma 3.3: We proceed by construction (depicted in Figure 3): consider a parallel network consisting of 3 edges, so that $E = \mathcal{P} = \{1, 2, 3\}$, $\mathcal{P}^{\mathrm{a}} = \{1, 2\}$ and $\mathcal{P}^{\mathrm{s}} = \mathcal{P}$. Let the latency functions for $\mathcal{P}$ be $\ell_1(x_1) = 1 + \alpha d$, $\ell_2(x_2) = \frac{x^d}{(1-r^{\mathrm{a}})^d}$, and $\ell_3(x_3) = 1$, respectively.

Now, $\bar{x} = \{0, 1 - r^{\mathrm{a}}, r^{\mathrm{a}}\}$ is a Nash flow for $(G, 0)$, where all autonomous traffic routes on path 2, the mass of selfish traffic routing on path 3 is equal to the autonomous fraction of traffic, and the remaining selfish traffic routes on path 2. This flow is feasible, since $r^{\mathrm{a}} \leq r^{\mathrm{s}}$, and Nash conditions are satisfied as $\ell_2(1 - r^{\mathrm{a}}) = \ell_3(r^{\mathrm{a}}) \leq \ell_1(0)$. Next, we show that $x = \{r^{\mathrm{a}}, 1 - r^{\mathrm{a}}, 0\}$ is a Nash flow for (G, α). It is feasible since $1 - r^{\mathrm{a}} \leq r^{\mathrm{s}}$, and the latency autonomous traffic experiences is a Nash flow as $\ell_1^{\mathrm{mc}_\alpha}(r^{\mathrm{a}}) \leq \ell_2^{\mathrm{mc}_\alpha}(1 - r^{\mathrm{a}})$. Now, selfish traffic will never use path 1 in any Nash flow, since $\ell_1(0) = 1 + \alpha d \geq 1 = \ell_2(r^{\mathrm{s}})$. Finally, straightforward computation shows that $\mathcal{L}(\bar{x}) = 1$, and $\mathcal{L}(x) = 1 + \alpha d r^{\mathrm{a}} = (1 + \alpha d r^{\mathrm{a}})\mathcal{L}(\bar{x})$. Thus, the bound in (15) is obtained. ■



(Path 3 accessed by selfish users only.)

Fig. 3: Network used to construct matching lower bound in Lemma 3.3.

REFERENCES

[1] L. J. Ratliff, R. Dong, S. Sekar, and T. Fiez, "A Perspective on Incentive Design: Challenges and Opportunities," *Annual Review of Control, Robotics, and Autonomous Systems*, vol. 2, no. 1, pp. 305–338, 2019.

[2] T. Roughgarden, *Selfish Routing and the Price of Anarchy*. MIT press, 2005.

[3] T. Roughgarden, "Intrinsic Robustness of the Price of Anarchy," *Journal of the ACM (JACM)*, vol. 62, no. 5, pp. 32:1–32:42, 2015-11-02.

[4] T. Roughgarden, "Stackelberg scheduling strategies," in *Proceedings of the Thirty-Third Annual ACM Symposium on Theory of Computing*, pp. 104–113, 2001.

[5] Y. Cheng and C. Langbort, "On informational nudging for boundedly rational decision makers," in *2018 IEEE Conference on Decision and Control (CDC)*, pp. 4791–4796, IEEE, 2018.

[6] B. L. Ferguson, P. N. Brown, and J. R. Marden, "Carrots or sticks? the effectiveness of subsidies and tolls in congestion games," in *2020 American Control Conference (ACC)*, pp. 1853–1858, IEEE, 2020.

[7] P.-A. Chen, B. D. Keijzer, D. Kempe, and G. Schäfer, "Altruism and its impact on the price of anarchy," *ACM Transactions on Economics and Computation (TEAC)*, vol. 2, no. 4, pp. 1–45, 2014.

[8] R. Li, P. N. Brown, and R. Horowitz, "Employing altruistic vehicles at on-ramps to improve the social traffic conditions," in *2021 American Control Conference (ACC)*, pp. 4547–4552, IEEE, 2021.

[9] P. N. Brown and J. R. Marden, "Can taxes improve congestion on all networks?," *IEEE Transactions on Control of Network Systems*, vol. 7, no. 4, pp. 1643–1653, 2020.

[10] P. N. Brown, "When altruism is worse than anarchy in nonatomic congestion games," in *2021 American Control Conference (ACC)*, pp. 4503–4508, IEEE, 2021.

[11] I. Milchtaich, "Network topology and the efficiency of equilibrium," *Games and Economic Behavior*, vol. 57, no. 2, pp. 321–346, 2006.

[12] M. J. Beckmann, C. B. McGuire, C. B. Winsten, and T. C. Koopmans, "Studies in the economics of transportation," *(No Title)*, 1955.

[13] A. Mas-Colell, "On a theorem of Schmeidler," *Journal of Mathematical Economics*, vol. 13, no. 3, pp. 201–206, 1984.

[14] P. Kleer and G. Schäfer, "Path deviations outperform approximate stability in heterogeneous congestion games," in *Algorithmic Game Theory: 10th International Symposium, SAGT 2017, L'Aquila, Italy, September 12–14, 2017, Proceedings 10*, pp. 212–224, Springer, 2017.

[15] E. Nikolova and N. E. Stier-Moses, "The burden of risk aversion in mean-risk selfish routing," in *Proceedings of the Sixteenth ACM Conference on Economics and Computation*, pp. 489–506, 2015.

[16] S. Sekar, L. Zheng, L. J. Ratliff, and B. Zhang, "Uncertainty in multi-commodity routing networks: When does it help?," *IEEE Transactions on Automatic Control*, vol. 65, no. 11, pp. 4600–4615, 2019.

[17] E. Bıyık, D. A. Lazar, R. Pedarsani, and D. Sadigh, "Altruistic autonomy: Beating congestion on shared roads," in *Algorithmic Foundations of Robotics XIII: Proceedings of the 13th Workshop on the Algorithmic Foundations of Robotics 13*, pp. 887–904, Springer, 2020.

[18] P. N. Brown and J. R. Marden, "Fundamental limits of locally-computed incentives in network routing," in *2017 American Control Conference (ACC)*, pp. 5263–5268, IEEE, 2017.

[19] T. Lianeas, E. Nikolova, and N. E. Stier-Moses, "Asymptotically tight bounds for inefficiency in risk-averse selfish routing," *arXiv preprint arXiv:1510.02067*, 2015.

[20] R. Meir and D. Parkes, "Playing the wrong game: Bounding externalities in diverse populations of agents," *arXiv preprint arXiv:1411.1751*, 2014.

[21] W. H. Sandholm, "Evolutionary implementation and congestion pricing," *The Review of Economic Studies*, vol. 69, no. 3, pp. 667–689, 2002.