Playing the Wrong Game: Bounding Externalities in Diverse Populations of Agents

RESHEF MEIR AND DAVID PARKES

The robustness of multiagent systems can be affected by mistakes or behavioral biases (e.g., risk-aversion, altruism, toll-sensitivity), with some agents playing the "wrong game." This can change the set of equilibria, and may in turn harm or improve the social welfare of agents in the system. We are interested in bounding what we call the *biased price of anarchy* (BPoA) in populations with diverse agent behaviors, which is the ratio between welfare in the "wrong" equilibrium and optimal welfare. We study nonatomic routing games, and derive an externality bound that depends on a key topological parameter of the underlying network. We then prove two general BPoA bounds for games with diverse populations: one that relies on the network structure and the *average bias* of all agents in the population, and one that is independent of the structure but depends on the *maximal bias*. Both types of bounds can be combined with known results to derive concrete BPoA bounds for a variety of specific behaviors (e.g., varied levels of risk-aversion).

ACM Reference format:
Reshef Meir and David Parkes. 2016. Playing the Wrong Game: Bounding Externalities in Diverse Populations of Agents. 1, 1, Article 1 (January 2016), 24 pages.
DOI: 10.1145/nnnnnnn.nnnnnnn

1 INTRODUCTION

Game theory is founded on the assumption that agents are rational decision makers, i.e. maximizing their utility, and that groups of agents reach an equilibrium outcome. In many games there is some objective measure of welfare that can be accurately measured in terms of money, time, and so on. Utility on the other hand, is subjective. It is not always easy to identify an explicit utility function that an agent is trying to maximize, if such exists. Even when there are objective factors that affect agents' utility, such as profit, effort, uncertainty, and temporal delays, various agents may weigh these factors differently or have subjective constraints and biases. Thus different agents may demonstrate different behavior even in the same situation.

As a concrete example, commuters may have some information on the expected congestion at each route via traffic reports or a cellphone app. However they also know that this information is inaccurate, and a risk-averse driver might take into account not just the expected congestion, but also the likelihood of an unexpected delay, a heuristic safety margin and so on. Moreover, different commuters may have different levels of risk-aversion, or act upon different heuristics.

The implications of these subjective differences and biases on a multiagent system are two-fold. First, from the perspective of an outside observer who cares about a particular objective (say, total latency), the agents are playing the "wrong game," either by applying some simple heuristics, or by optimizing a different utility function than the 'objective' one [18].

DOI: 10.1145/nnnnnnn.nnnnnnn

A second related issue, is that the behavior of one agent may exert significant negative externality on the utility of another. The extreme case is when some agents are adversarial, and act in a way that tries to *minimize* the utility of some other agents.

It is well known that even in games without biases or subjective utilities, there may be negative externalities, and that equilibrium outcomes may be suboptimal in terms of the utilitarian social welfare. This inefficiency is often quantified as the *Price of Anarchy* (PoA), which is the ratio between the worst equilibrium social welfare and the optimal social welfare. Biases may change the equilibria of the game, and thus have a large effect on the PoA. Externalities are widely studied in multiagent systems, but usually in the context of well-defined behavior such as maximizing utility in a game [28, 34] or finding an optimal or stable solution [7, 11, 12]. How can we hope to bound externalities in a diverse society of agents with assorted biases and behaviors? The answer will lie in a proper abstraction of such behaviors, but before that we will present our questions more formally.

Research goals. Suppose that in game $\mathbf{G}$, each agent i has some true cost function c^i (in this paper we consider negative utilities, i.e. costs). Now, each agent i sees her cost as some other function $\hat{c}^i$, and thus we are interested in the equilibria of the *biased game* $\hat{\mathbf{G}}$ comprised of modified utilities or costs $\{\hat{c}^i\}_{i=1,2,\dots}$. We ask the following:

- Is there a bound to the negative externality that type j exerts on type i, that applies *regardless* of the type $\hat{c}^j$? (Example: can we bound the equilibrium latency of type i commuters without any assumption on the risk-aversion level of others?)
- Is there a bound to the social cost (in the "objective game" $\mathbf{G}$) of $\hat{\mathbf{G}}$'s worst equilibrium? (Example: can we bound the total latency in equilibrium, given all commuters' subjective cost functions?)

We refer to the ratio between the latter measure and the optimal social cost in $\mathbf{G}$ as the *Biased Price of Anarchy* (BPoA), and note that it coincides with the PoA when $\hat{\mathbf{G}} = \mathbf{G}$.

While the first question is straight-forward, the second question may raise some conceptual debate. One may argue that since agents are acting so as to minimize their biased cost $\hat{c}^i$ rather than c^i, this is the cost we should take into account when computing the social cost or welfare. We justify using $\mathbf{G}$ as the baseline for welfare as follows. First, the social cost may be the objective used by the system designer or analyst, as in our example above. A different analyst might care about a different goal, resulting in a different objective game $\mathbf{G}$. Second, the social cost may be the sum of the agents' true utilities, while the agents are bounded-rational. For example they may be unaware of some roads [2] or the exact latency functions [41]; they may have uncertainty regarding actual congestion [42] or the overall amount of other agents [5]; and may assign wrong probabilities to rare outcomes [23].

Smoothness in routing games. Nonatomic routing games in the Wardrop model [43] are a good testbed for the questions above: they have very convenient theoretical properties, such as the uniqueness of equilibrium (up to identical utilities); equilibrium inefficiency and in particular the PoA is very well understood; and several biases have already been suggested and studied in the context of routing games (see Related Work). The *smoothness method* allows to leverage a property of the edge cost functions to obtain a tight upper bound on the Price of Anarchy that is independent of the network topology: if all cost functions are (λ, μ)-smooth for some parameters λ, μ, then the PoA is bounded by $\frac{\lambda}{1-\mu}$ [17, 36, 37]. Thus the smoothness of cost functions lets us *abstract away* the details of the game and prove PoA bounds on large classes of games.

1.1 Paper structure and contribution

After a short overview of nonatomic routing games, price of anarchy, and a measure of graph complexity called *parallel-width* (Section 2), we prove a tight bound on the negative externality in any routing game without any assumption on agents' behavior (Section 3). Our bound generalizes previous results from specific behaviors on series-parallel networks [2, 33] to arbitrary behaviors and networks (parametrized by their parallel-width).

Next, we adopt *smoothness* as an abstraction for general biases and behaviors. In Section 4 we extend the definition of a smooth cost function to $(\hat{\lambda}, \hat{\mu})$-biased-smoothness that takes into account both the "true" and the "biased" cost functions. This approach follows similar extensions for specific, modified costs [10, 13], and we review recent smoothness bounds for such specific settings.

In Section 5, we consider games where agents have diverse biases, and use our results from Section 3 along with biased-smoothness to derive several bounds on the Biased Price of Anarchy. For example, for symmetric games over series-parallel networks (which have parallel-width of 1), we derive a BPoA bound in terms of the "average" biased-smoothness: BPoA $\leq O(1) \sum_i \frac{r_i}{r} \frac{\hat{\lambda}^i}{1-\hat{\mu}^i}$, where $\frac{r_i}{r}$ is the fraction of type i agents in the population. For arbitrary networks, we get a bi-criteria result that depends both on the average bias as above, and on the parallel-width of the underlying network.

For the special case of polynomial cost functions, we leverage (Section 6) known results on the PoA in heterogeneous unbiased games to derive a structure-independent BPoA bound. In contrast to our main results, this bound depends on the *worst* bias rather than on the average bias.

1.2 Related Work

The most-well studied, modified cost function comes about as a result of tolls, where the travel time plus the imposed toll can be thought of as a modified cost function. In this context, most papers focus on the objective of minimizing total latency [15, 20, 26], and on the design optimal or practical toll schemes [10, 21, 25, 40]. Heterogeneous biases arise when different agents have different sensitivity to imposed tolls. We explain in the relevant sections how these papers technically relate to our work.

Chen et al. [13, 14] apply smoothness analysis to provide BPoA bounds for various games (including atomic congestion games) where agents are altruistic, i.e., part of their utility is derived from the social welfare. In the context of nonatomic routing games, their model is formally equivalent to toll-sensitivity (see Section 4.2).

Acemoglu et al. [2] study nonatomic congestion games where some agents are unaware of the existence of certain edges, which is equivalent to having a wrong cost function that assigns infinite costs to some edges. They prove that on directed series-parallel networks, such ignorance can only lead to a worse equilibrium than under true information, yet the *worst-case* PoA remains the same.

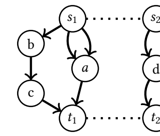
Another behavioral bias that has been studied in congestion games is risk aversion [3, 5, 30, 33, 35], which can often be written as biased cost functions (see Section 4.2 for details).

Finally, Babaioff et al. [6] consider congestion games where some of the agents are malicious. Babaioff et al. focused on the effect of a negligible amount of malicious agents, and showed examples where it can be either detrimental or (surprisingly) beneficial to the other agents, but without any upper bounds. Indeed, malicious behavior can be considered as another form of bias.

2 PRELIMINARIES

A *network* is a 2-terminal directed multigraph $G = \langle V, E, s, t \rangle$, where $s, t \in V$ are special vertices (*source* and *target*), and every edge $e \in E$ belongs to some simple $s - t$ path.

A network is *series-parallel* [19, 24] if it is either a single edge, or composed recursively by joining two series-parallel networks in series or in parallel. E.g., merging $\{s_1, s_2\}$ and $\{t_1, t_2\}$ in the following networks also results in a series-parallel network:



2.1 Nonatomic routing games

Following the definitions of Roughgarden [36] and Roughgarden and Tardos [38], a *nonatomic routing game* (NRG) is a tuple $\mathbf{G} = \langle G, m, \boldsymbol{c}, \boldsymbol{s}, \boldsymbol{t}, \boldsymbol{r} \rangle$, where

- $G = \langle V, E \rangle$ is a directed multigraph;
- $m \in \mathbb{N}$ is the number of agent types;
- $\boldsymbol{c} = (c_e)_{e \in E}$, where $c_e(x) \geq 0$ is the cost incurred when x agents use edge e;[1]
- $\boldsymbol{s}, \boldsymbol{t} \in V^m$, where (s_i, t_i) are the source and target nodes of type i agents;
- $\boldsymbol{r} = (r_i)_{i \leq m}$, where $r_i > 0$ is the total mass of type i agents. The total mass of agents of all types is $\sum_{i \leq m} r_i = r$, and we assume unless specified otherwise that $r = 1$.

We denote by $A_i \subseteq 2^E$ the set of all directed simple paths between the pair of nodes (s_i, t_i) in the graph. Thus A_i is the the set of *actions* available to agents of type i. We denote by $A = \cup_i A_i$ the set of all directed source-target simple paths. We assume that all cost functions mentioned in the paper (including biased costs mentioned later on) are non-decreasing, continuous, differentiable and semi-convex (i.e., $xc_e(x)$ is convex). Such cost functions are called *standard* [36].

Player-specific costs. A *nonatomic routing game with player-specific costs* (PNRG) is a tuple $\mathbf{G} = \left\langle G, m, (\boldsymbol{c}^i)_{i \leq m}, \boldsymbol{s}, \boldsymbol{t}, \boldsymbol{r} \right\rangle$. The difference from a NRG is that agents of each type i experience a cost of $c^i_e(x)$ when x agents use edge e. We can have multiple types with the same source and target nodes to allow diversity of behavior. To avoid confusion, we refer to (s_i, t_i) (or A_i) as the *demand type* and to $\boldsymbol{c}^i$ as the *cost type*. Thus the type i specifies both the demand type and the cost type.

A PNRG is *symmetric* if all agents have the same demand type, i.e., $A_i = A$ for all i. A PNRG is a *resource selection game* (RSG) if G is a network of parallel links. That is, if the action of every agent is to select a single (s, t) edge.

Flows. A *flow* (or action profile) of a PNRG is a vector $\boldsymbol{f} \in \mathbb{R}_+^{|A| \times m}$, where $f_{p,i}$ is the amount of agents of type i that use path $p \in A_i$. In a valid flow, $\sum_{p \in A_i} f_{p,i} = r_i$ for all i. The total traffic on path $p \in A$ is denoted by $f_p = \sum_{i=1}^m f_{p,i}$. Similarly, the total traffic on edge $e \in E$ is denoted by $f_e = \sum_{p: e \in p} f_p$. Denote the *support of type i strategy* in flow $\boldsymbol{f}$ by $P_i(\boldsymbol{f}) = \{p \in A : f_{i,p} > 0\}$. That is, all paths used by type i agents in flow $\boldsymbol{f}$.

The cost for an agent of type i in game $\mathbf{G}$, selecting a path $p \in A_i$ in flow $\boldsymbol{f}$, is $c^i(p, \boldsymbol{f}) = \sum_{e \in p} c^i_e(f_e)$.

Social cost. We denote by $SC_i(\mathbf{G}, \boldsymbol{f}) = \sum_{p \in A_i} f_{i,p} c(p, \boldsymbol{f})$ the cost experienced by type i agents in flow $\boldsymbol{f}$. By summing over all types, we get the social cost:

$$SC(\mathbf{G}, \boldsymbol{f}) = \sum_{i \leq m} SC_i(\mathbf{G}, \boldsymbol{f}) = \sum_{i \leq m} \sum_{p \in P_i(\boldsymbol{f})} f_{i,p} c(p, \boldsymbol{f}) = \sum_{e \in E} c_e(f_e) f_e$$

Thus the social cost only depends on the total traffic per edge. We denote by $\boldsymbol{f}^o(\mathbf{G}) \in \text{argmin}_{\boldsymbol{f}} SC(\mathbf{G}, \boldsymbol{f})$ some profile with minimal total cost, and $OPT(\mathbf{G}) = SC(\mathbf{G}, \boldsymbol{f}^o(\mathbf{G})) = \min_{\boldsymbol{f}} SC(\mathbf{G}, \boldsymbol{f})$.

[1] We use the term cost rather than latency, to reflect that agents may care about other factors.

Equilibrium. A flow f for an PNRG is an *equilibrium* in game $\mathbf{G}$ if for every agent type i, any used path $p \in P_i(f)$ and any $p' \in A_i$, we have $c^i(p,f) \leq c^i(p',f)$. That is, if no agent can switch to a path with a lower cost. This provides the analogy of a Nash equilibrium for nonatomic games.

It is known that in any NRG there is at least one equilibrium, and that this can be reached by a simple best-response dynamic. Further, all equilibria have the same social cost and in every equilibrium all agents of type i experience the same cost [1, 9, 31, 38]. Player-specific NRGs are also guaranteed to have at least one equilibrium [39], however, equilibrium costs may not be unique, and best-response dynamics may not converge, except in special cases [22].

Affine routing games. In an *affine* NRG, all cost functions take the form of a linear function. That is, $c_e(x) = a_e x + b_e$ for some constants $a_e \geq 0, b_e \geq 0$. The social cost can be written as $SC(\mathbf{G}, f) = \sum_{e \in E} a_e(f_e)^2 + b_e f_e$. *Pigou's example* is the special case of an affine RSG with two resources, where $c_1(x) = 1$ and $c_2(x) = ax$. We denote by $\mathbf{G}_P(a)$ the instance where $c_2(x) = ax$ (see Fig. 1a).

The price of anarchy. Let $EQ(\mathbf{G})$ be the set of equilibria in game $\mathbf{G}$. The *price of anarchy* (PoA) of a game is the ratio between the social cost in the worst equilibrium in $EQ(\mathbf{G})$ and the optimal social cost [27]. Formally, $\text{PoA}(\mathbf{G}) = \sup_{f^* \in EQ(G)} \frac{SC(\mathbf{G}, f^*)}{OPT(\mathbf{G})}$. For example in affine NRGs, it is known that $\text{PoA}(\mathbf{G}) \leq \frac{4}{3}$, and this bound is attained by the Pigou example of $\mathbf{G}_P(1)$ [38].

Smoothness. A cost function c is (λ, μ)-smooth for $\lambda \geq 0, \mu < 1$ if for any $x, x' \geq 0$, it holds that

$$c(x)x' \leq \lambda x' c(x') + \mu x c(x). \tag{1}$$

A NRG $\mathbf{G}$ is (λ, μ)-smooth if all cost functions in $\mathbf{G}$ are (λ, μ)-smooth. For any (λ, μ)-smooth NRG, $\text{PoA}(\mathbf{G}) \leq \frac{\lambda}{1-\mu}$ [17, 37].[2] Moreover, w.l.o.g. $\lambda = 1$ (that is, for any class of cost functions there is an optimal pair $(1, \mu)$ for some μ [16, 38]). For example, affine functions are $(1, \frac{1}{4})$-smooth, which again entails a PoA bound of $\frac{4}{3}$.

2.2 Parallel Width

Consider a network $G = \langle V, E, s, t \rangle$.

Definition 2.1. A set of edges $S \subseteq E$ is *parallel* if there is some $S' \subseteq E$ s.t. $S \subseteq S'$, and S' is a minimal cut between s and t in the network G.

Definition 2.2. A set of edges $S \subseteq E$ is *crossed* if there is a simple directed $s-t$ path p that contains S.

Definition 2.3 (Parallel Width). The *parallel-width* of a network, $\text{PW}(G)$, is the size of the largest set $S \subseteq E$ that is both parallel and crossed.

Intuitively, a parallel width of k means there are at least k non-intersecting source-target paths, and some additional path that edge-intersects all of them.

Example 2.4. Consider the Braess network in Fig. 1b (ignore the costs). The minimal $s-t$ cuts are: $\{sa, sb\}, \{at, bt\}$, and $\{sb, ab, at\}$. Thus the set $\{sa, bt\}$ is both parallel and crossed, which means $\text{PW}(G_B) \geq 2$. The set $\{sa, at\}$ is crossed but not parallel; and $\{sa, sb, ab\}$ is neither. In fact, the *only* parallel set of size greater than 2 is $\{sb, ab, at\}$, which is not crossed, thus $\text{PW}(G_B) < 3$. We conclude that the parallel-width of the Braess network is 2.

[2] The mentioned papers only require each c_e to be smooth around the equilibrium flow $x = f^*_e$. Our proofs require the cost functions to be smooth at several different places, and hence the slight difference in the definition.

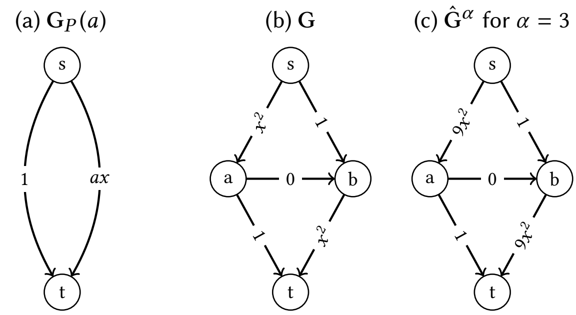


Figure 1. (1a) Pigou's example. (1b) An objective game G (a quadratic variation of the Braess paradox). (1c) The game $\hat{G}^\alpha$ is the same game played by pessimistic agents with parameter $\alpha = 3$. The biased costs of edges with fixed costs like sb do not change, but the biased cost on the edge sa for example is $\hat{c}^\alpha_{sa}(x) = c_{sa}(3x) = (3x)^2 = 9x^2$.

Definition 2.5. For any $k \geq 2$, we define the *k-crossed-parallel network* $G_{CP(k)} = \langle V, E, s, t\rangle$, where $V = \{s, t, a_2, \dots, a_k, b_1, \dots, b_{k-1}\}$, and $E = \bigcup_{i=2}^{k-1}\{(s, a_i), (a_i, b_i), (b_i, t), (b_i, a_{i+1})\} \cup \{(s, b_1), (a_k, t)\}$.

See Figure 2 for an example.

The 2-cross-parallel network is the Braess network G_B. The parallel-width of the k-cross-parallel network is exactly k, where the parallel edges are $\{(s, b_1), (a_2, b_2), \dots, (a_{k-1}, b_{k-1}), (a_k, t)\}$. The network $G_{CP(k)}$ was used in [6] to derive examples of games with high Price of Malice, and we will use it later in a similar way. Meir [29] proved that the parallel-width of acyclic networks can be characterized completely, where $PW(G) < k$ if and only if $G_{CP(k)}$ (or some small variants of which) is not embedded in G. For $k = 1$, this entails another characterization.

Proposition 2.6 (Meir [29]). *Let G be an acyclic network, then $PW(G) = 1$ if and only if G is series-parallel.*

3 BOUNDING EXTERNALITIES

Suppose we are given a game $\mathbf{G} = \langle G, m, (\mathbf{c}^j)_{j\leq m}, \mathbf{s}, \mathbf{t}, \mathbf{r}\rangle$, and agents play some equilibrium $\boldsymbol{f}^*$ of $\mathbf{G}$. For any single type i, all type i agents have the same cost in $\boldsymbol{f}^*$. We denote this cost by $C^{i,*} = c^i(p, \boldsymbol{f}^*)$, where $p \in P_i(\boldsymbol{f}^*)$ is an arbitrary path used by a type i agent. Note that $r_i \cdot C^{i,*} = SC_i(\mathbf{G}, \boldsymbol{f}^*)$.

Our goal in this section is to bound $C^{i,*}$. A-priori, this may seem difficult, as we do not assume *anything* about the types of the other agents. However, we will show that the negative externality of the other types can be bounded using the structural parameters of the network G (the parallel-width), and even adversarial agents cannot be much worse for i than more type i agents.

Given a game $\mathbf{G}$ and type i, we define a new game $\mathbf{G}^i$ by setting both the cost type and the demand type of all agents in $\mathbf{G}$ to i. That is, $\mathbf{G}^i = \langle G, 1, \mathbf{c}_i, s_i, t_i, r\rangle$. We also define a game $\mathbf{G}^{(k)} = \langle G, m, \mathbf{c}, \mathbf{s}, \mathbf{t}, k \cdot r\rangle$, where the demand in $\mathbf{G}$ is multiplied by k. Finally, we set $\mathbf{G}^{i,(k)} = (\mathbf{G}^i)^{(k)}$, i.e., $\mathbf{G}^{i,(k)}$ is a game where there all $k \cdot r$ agents are of type i. Note that $\mathbf{G}^{i,(1)} = \mathbf{G}^i$, and denote $C^i = C^{i,1}$.

Let $\boldsymbol{g}^*$ be an equilibrium of $\mathbf{G}^{i,(k)}$. As $\mathbf{G}^{i,(k)}$ is a symmetric game, all agents have the same cost in $\boldsymbol{g}^*$. We denote this cost by $C^{i,(k)} = c^i(p, \boldsymbol{g}^*)$, where $p \in \overline{P}(\boldsymbol{g}^*)$ is an arbitrary used path.

Finally, let $G|_i = \left\langle V^i, E^i, s_i, t_i\right\rangle$ be the network obtained from $G = \langle V, E\rangle$, by eliminating all edges and vertices that are not in A_i.

The following bound is our main result in this section. Prop. 3.2 shows that the bound is tight.

Theorem 3.1 (Externality theorem). *Let* $\mathbf{G}$ *be any PNRG played on a network* G. *If* $PW(G|_i) \leq k$, *then* $C^{i,*} \leq C^{i,(k)}$.

Before we prove the theorem, we explain its implications. For $k = 1$ (e.g., series-parallel networks), this means that there are no negative externalities due to type differences—the cost of type i agents may only increase when all others are of the same type. We note that this result (along with Prop. 3.2) strictly generalizes Theorem 5 from [2], which is attained as a special case for $k = 1$ and specific demand types (namely, agents that ignore certain edges); and Theorem 5.7 in [33] which is attained as a special case for $k = 1$ and symmetric games. It also implies that the Price of Malice [6] is 0 in series-parallel networks.

For larger values of k, Theorem 3.1 means that regardless of what the other agents are doing, the cost to the type i agents is never more than their equilibrium cost in a game where all agents are of type i, and their number is multiplied by k.

Proof. Denote $\overline{P} = P_i(\boldsymbol{g}^*)$ (the set of paths $p \in A_i$ s.t. $g^*_{p,i} > 0$). Assume towards a contradiction that $C^{i,*} > C^{i,(k)}$. This means that $c^i(p, \boldsymbol{f}^*) > C^{i,(k)} = c^i(p, \boldsymbol{g}^*)$ for any path $p \in \overline{P}$, as $c^i(p, \boldsymbol{f}^*)$ is either $C^{i,*}$ or higher. For any path $p \in \overline{P}$,

$$\sum_{e \in p} c^i_e(f^*_e) = c^i(p, \boldsymbol{f}^*) > c^i(p, \boldsymbol{g}^*) = \sum_{e \in p} c^i_e(g^*_e),$$

thus there is an edge $e = e_p \in p$ s.t. $c^i_e(f^*_e) > c^i_e(g^*_e)$. Consider the set $E_{\overline{P}} = \{e_p | p \in \overline{P}\}$. Since c^i_e is monotone, this means $f^*_e > g^*_e$ for every edge $e \in E_{\overline{P}}$.

Consider the weighted directed graph $H = \langle V, F \rangle$, where $F = \{e \in p | p \in \overline{P}\}$, and the capacity (weight) of every edge $e \in F$ is g^*_e.

By construction, $E_{\overline{P}}$ is a cut between s_i and t_i in H. Let $\overline{E} \subseteq E_{\overline{P}}$ s.t. $\overline{E}$ is a minimal cut (not necessarily of minimum size or minimum weight), then

$$\sum_{e \in \overline{E}} f^*_e > \sum_{e \in \overline{E}} g^*_e \geq k \cdot r. \tag{2}$$

The first inequality is since $f^*_e > g^*_e$ holds for every $e \in E_{\overline{P}}$ and thus for every $e \in \overline{E}$. The second inequality follows from the min-cut-max-flow theorem, since $\overline{E}$ is some $s_i - t_i$ cut in H, and the flow between s_i, t_i is exactly $k \cdot r$ in any valid flow of $\mathbf{G}^{i,(k)}$ (which by definition has a mass of $k \cdot r$ type i agents).

Now, as there are only r agents in $\mathbf{G} = (G, m, (\boldsymbol{c}^i)_{i \leq m}, \boldsymbol{s}, \boldsymbol{t}, \boldsymbol{r})$, then by Eq. (2) and the pigeon-hole principle there must be some agents in profile $\boldsymbol{f}^*$ using strictly more than k edges from the set $\overline{E}$. Choose some path used by such an agent and denote it by $p' \in A_i$. We define $S = p' \cap \overline{E}$, thus by our selection of p', $|S| > k$.

Finally, all edges of S are contained in the minimal cut $\overline{E}$ between s_i, t_i (thus S is parallel), and S is also contained by the simple path p' (thus S is crossed). By definition of parallel-width, $\text{PW}(G|_i) \geq |S| > k$, which is a contradiction to our assumption. □

3.1 Lower bound

Theorem 3.1 is tight in a very strong sense: the bound cannot be improved for *any network* and *any class* of cost functions. We assume that a class of cost functions contains all constant functions, as well as at least one non-constant function.

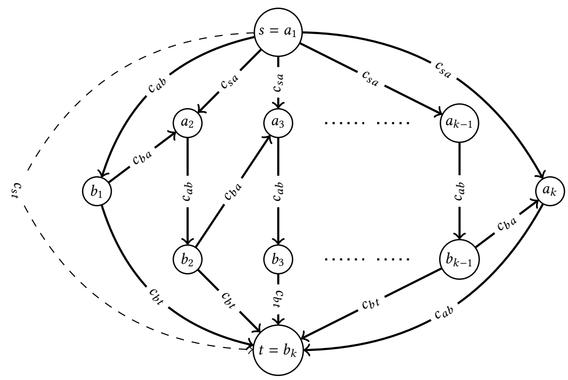


Figure 2. The solid edges compose the network $G_{CP(k)}$. The edge labels are used in Prop. 3.2 and 5.4, where edges with the same label (e.g. c_{ab}) have the same cost function. The dashed (s,t) edge is not part of $G_{CP(k)}$ and is used only in Prop. 5.4.

Proposition 3.2. *Let any $k \geq 2$, network G with $PW(G) = k$, any class of cost functions C, and any $\delta > 0$. There is a game with two types $\mathbf{G} = \left\langle G, (\boldsymbol{c}^1, \boldsymbol{c}^2), s, t, \boldsymbol{r}\right\rangle$, such that $c^1_e, c^2_e \in C$ for all $e \in E$, and $C^{1,*} > C^{1,(k)} - \delta$.*

Proof. We first prove for the k-Cross-Parallel network $G_{CP(k)}$. The demand type of both agents is (s,t), thus all paths from s to t are valid strategies. We define edge costs as follows. Recall that C contains all constant functions, and at least one strictly increasing function c. For type $i = 1$, we set $c^i_{ab}(x) = c(x)$. The cost of all other edges is 0. For type $j = 2$, we set $c^j_{ab}(x) = c^j_{ba}(x) = 0, c^j_{sa}(x) = c^j_{bt}(x) = c(1)$. We set $r_i = \epsilon$ (ϵ will be determined later) and $r_j = 1 - \epsilon$.

Thus in the unique equilibrium $\boldsymbol{g}^*$ of game $\mathbf{G}^{i,(k)}$, the agents (k in total) split evenly over all k short paths, and $C^{i,(k)} = c^i_{ab}(\frac{k}{k}) = c^i_{ab}(1) = c(1)$.

In the unique equilibrium $\boldsymbol{f}^*$ of game $\mathbf{G}$, all type j agents take the long path $p = (s, b_1, a_2, b_2, \ldots, b_{k-1}, a_k, t)$ since $c^j(p, \boldsymbol{f}^*) = 0$. The type i agents split evenly over all k short paths , and thus each experiences a cost of $C^{i,*} = c^i_{ab}(r_j + \frac{r_i}{k}) = c(r_j + \frac{r_i}{k}) > c(r_j)$.

Let $\epsilon > 0$ such that $c(1-\epsilon) > c(1) - \delta$. Since c is continuous and strictly increasing such ϵ must exist. Finally,
$C^{i,*} > c(r_j) = c(1-\epsilon) > c(1) - \delta = c(1) = C^{i,(k)} - \delta$, as required.

Now, consider an arbitrary network G with $PW(G) \geq k$. Consider some cross-parallel set S of size k and the path p that contains it, and S will play the role of the ab edges above. Specifically, For type i agents we set $c^i_e \equiv c^i_{ab} \equiv c$ for any $e \in S$, and $c^i_e \equiv c \equiv 0$ for all other edges. For type j, we set $c^j_e \equiv 0$ for all $e \in p$, and $c^j_e \equiv c(1)$ for all other edges. It is easy to verify that $C^{i,*}, C^{i,(k)}$ are the same as for $G_{CP(k)}$. □

4 BIASED PRICE OF ANARCHY

Given an NRG $\mathbf{G} = \langle G, m, \boldsymbol{c}, \boldsymbol{s}, \boldsymbol{t}, \boldsymbol{r}\rangle$ and modified cost functions $\hat{c}^i_e$ for every type $i \leq m$ and edge $e \in E$, we obtain a PNRG $\hat{\mathbf{G}} = \left\langle G, m, (\hat{\boldsymbol{c}}^i)_{i \leq m}, \boldsymbol{s}, \boldsymbol{t}, \boldsymbol{r}\right\rangle$. We assume that agents act based on their modified cost functions, irrespective of whether this is a rational behavior or not. We refer to $\hat{\mathbf{G}}$ as the *biased game*, where every agent of type i experiences a cost of $\hat{c}^i_e$ rather than c_e. Both games $\mathbf{G}$

and $\hat{G}$ have a role in our model, and we often denote the overall setting as $\mathcal{G} = \langle G, \hat{G} \rangle$. The way we interpret $\mathcal{G}$ is that agents play the biased game $\hat{G}$ (and thus, it is the equilibria of $\hat{G}$ that matter), whereas their true costs are according to game G. We say that $\mathcal{G}$ is *uniform* if all agents in $\hat{G}$ have the same cost type (same bias).

Biased Price of Anarchy. We measure the price of anarchy in a game with biased costs by comparing the equilibria of $\hat{G}$ (denoted by $\hat{\boldsymbol{f}}^*$) to the optimum of G. Formally:

$$\mathrm{BPoA}(\mathcal{G}) = \mathrm{BPoA}(G, \hat{G}) = \sup_{\hat{\boldsymbol{f}}^* \in EQ(\hat{G})} \frac{SC(G, \hat{\boldsymbol{f}}^*)}{OPT(G)}. \tag{3}$$

In the uniform bias case where $\hat{\boldsymbol{c}}^i = \hat{\boldsymbol{c}}$ for all i, the game $\hat{G}$ is just another NRG. Chen et al. [14] referred to the BPoA (when applied to altruism) as the *Robust PoA*. The *Price of Risk Aversion* [33] and the *Deviation Ratio* [26] are similar concepts to the BPoA, except they compare $\hat{\boldsymbol{f}}^*$ to the equilibrium of the unbiased game G.

A simple example of a biased cost is induced by *pessimism*, which is one way to model risk-aversion [30]. Suppose that whenever faced with some expected load f_e on edge e, an agent takes a safety margin by playing as if the actual load is $\alpha \cdot f_e$ (for some fixed private parameter $\alpha \geq 1$). Such an agent will play *as if* every cost function c_e is replaced with a new cost function $\hat{c}^\alpha_e$, where $\hat{c}^\alpha_e(x) = c_e(\alpha x)$ (see Fig. 1). We denote the (uniform) game where all agents play according to $(\hat{c}^\alpha_e)_{e \in E}$ by $\hat{G}^\alpha$.

Example 4.1. For example, the equilibrium $\boldsymbol{f}^*$ of the objective game G in Fig. 1 is suboptimal, as all agents take the long path $s - a - b - t$. Thus $\mathrm{PoA}(G) \cong \frac{2}{1.23} = 1.625$. In contrast, in the equilibrium $\hat{\boldsymbol{f}}^*$ of $\hat{G}^\alpha$ for $\alpha = 3$, agents divide equally among the two short paths, which leads to $\mathrm{BPoA}(G, \hat{G}^\alpha) \cong \frac{1.25}{1.23} = 1.016$.

4.1 Smoothness for Biased Costs

Our goal is to provide bounds on the biased Price of Anarchy for a given game with biased costs $\langle G, \hat{G} \rangle$. That each of c_e and $\hat{c}^i_e$ are smooth is insufficient to provide such a bound.

Example 4.2. Consider a Pigou game with pessimistic agents $\mathcal{G}_\alpha = \langle G_\alpha, \hat{G}^\alpha_\alpha \rangle$ where $G_\alpha = G_P(2/\alpha)$ and thus $\hat{G}^\alpha_\alpha = G_P(2)$. For any α the equilibrium of $\hat{G}^\alpha_\alpha$ (and thus of $\mathcal{G}_\alpha$) is the same: 1/2 of all agents use each resource. However as α increases, the optimal flow of G_α shifts more agents to the resource with variable cost, and the optimal social cost decreases to 0. Thus the gap between the equilibrium cost and the optimal cost (the BPoA) goes to infinity with α even though both of $G_\alpha, \hat{G}^\alpha_\alpha$ are affine.

It is crucial, then, to extend the definition of smoothness to games with biased costs in a way that takes into account both c and $\hat{c}$. This technique has been applied before for specific modified costs, for example nonatomic games with restricted tolls [10] and atomic games with altruistic agents [13]. We provide a general extension.

Definition 4.3. Let $\lambda \geq 0, \mu < 1$. The function c is (λ, μ)-*biased-smooth* w.r.t. biased cost function $\hat{c}$, if for any $x, x' \in \mathbb{R}_+$,

$$c(x)x + \hat{c}(x)(x' - x) \leq \lambda c(x')x' + \mu c(x)x. \tag{4}$$

It is instructive to check the familiar case where there is no bias. Indeed, if $\hat{c} = c$, and c is (λ, μ)-smooth, then

$$c(x)x + \hat{c}(x)(x' - x) = c(x)x + c(x)(x' - x) = c(x)x' \leq \lambda c(x')x' + \mu c(x)x,$$

and Eq. (4) collapses to "standard" smoothness (Eq. (1)).

Recall that the PoA of a (λ, μ)-smooth game is bounded by $\frac{\lambda}{1-\mu}$. This bound extends to games with biased costs that are biased-smooth and when all agents have the *same bias*. This was explicitly shown for specific biases, but we write down the general formulation for completeness.

Theorem 4.4 (Bonifaci et al. [10]). *Consider a uniform game with biased costs $\mathcal{G} = \langle \mathbf{G}, \hat{\mathbf{G}} \rangle$ where every cost function c_e is $(\hat{\lambda}, \hat{\mu})$-biased smooth w.r.t. biased cost function $\hat{c}_e$. Let $\hat{\boldsymbol{f}}^* \in EQ(\hat{\mathbf{G}})$, and $\boldsymbol{f}$ any valid flow. Then $SC(\mathbf{G}, \hat{\boldsymbol{f}}^*) \leq \frac{\hat{\lambda}}{1-\hat{\mu}} SC(\mathbf{G}, \boldsymbol{f})$.*
In particular, $BPoA(\mathbf{G}, \hat{\mathbf{G}}) \leq \frac{\hat{\lambda}}{1-\hat{\mu}}$.

Proof. The proof extends the standard proof of PoA bounds for nonatomic congestion games via smoothness arguments. In any equilibrium $\hat{\boldsymbol{f}}^*$, the *variational inequality* $\sum_e \hat{c}_e(\hat{f}^*_e)\hat{f}^*_e \leq \sum_e \hat{c}_e(\hat{f}^*_e) f_e$ holds (see [17, 37]). Thus,

$$SC(\mathbf{G}, \hat{\boldsymbol{f}}^*) \leq SC(\mathbf{G}, \hat{\boldsymbol{f}}^*) + \sum_e \hat{c}_e(\hat{f}^*_e) f_e - \sum_e \hat{c}_e(\hat{f}^*_e)\hat{f}^*_e$$
$$\leq \sum_e [\hat{\lambda} c_e(f_e) f_e + \hat{\mu} c_e(\hat{f}^*_e)\hat{f}^*_e] = \hat{\lambda} SC(\mathbf{G}, \boldsymbol{f}) + \hat{\mu} SC(\mathbf{G}, \hat{\boldsymbol{f}}^*).$$

We get the bound by reorganizing the terms. The only part that differs from the standard smoothness is the first inequality. □

An alternative derivation of optimal tolls. We can also check that the extension provides the right result in regard to modified costs $\tilde{c}(x) = c(x) + c'(x)x$ (where $c'(x) = \frac{\partial c(x)}{\partial x}$) that represent optimal tolls and should lead to optimal play [8]. Let's confirm this result via a biased-smoothness argument.

Observation 4.5. *If we set modified cost $\tilde{c}(x) = c(x) + c'(x)x$, then cost function c is $(1, 0)$-biased smooth w.r.t. to $\tilde{c}$.*

This follows immediately from the convexity of $xc(x)$, and affirms via Theorem 4.4 that $\text{BPoA}(\mathbf{G}, \tilde{\mathbf{G}}) = 1$ for any game. For example, adding optimal tolls to the game in Fig. 1c will set $\tilde{c}_{sa}(x) = 3x^2$, and in the only equilibrium, $SC(\mathbf{G}, \boldsymbol{f}^*) = 1.23 = OPT(\mathbf{G})$.

4.2 Bounds for Specific Biases

Theorem 4.4 lets us bound the BPoA of a population of agents with *uniform bias*, and in the next sections we will prove theorems that bound the BPoA of populations with *heterogeneous biases*. However, all these theorems require some explicit bounds on the biased smoothness parameters $\hat{\lambda}$ and $\hat{\mu}$ in a given scenario (i.e., for the specific bias and specific class of cost functions).

There are by now many such studied biases. We already introduced pessimism, which depends on a single parameter α, and where the biased cost is $\hat{c}^\alpha_e(x) = c_e(\alpha x)$. Toll-sensitive agents [15] have a single parameter β, and their biased cost is $\hat{c}^\beta_e(x) = c_e(x) + \beta x c'_e(x)$ (meaning that marginal tolls are imposed, but agents discount or over-weigh them by a factor of β). Altruist agents [14] have exactly the same biased cost function as tolls, except that now β should be interpreted as how much they care about hurting others. In the Mean-Var risk-aversion model [32, 33], $c_e(x)$ is a distribution over costs, and the biased cost is $\hat{c}^\gamma_e(x) = \mathbb{E}[c_e(x)] + \gamma VAR[c_e(x)]$, i.e., γ represents the sensitivity of agents to variance. We assume that $VAR[c_e(x)] \leq \tau$ for some constant τ. *Small error* means that c and $\hat{c}$ are within a small multiplicative factor of $1 + \delta$ from one another (see Sec. 6 for details).

(a) Maximal values of $\hat{\lambda}, \hat{\mu}$ under various biases

Bias	Pessimism		Toll-sensitivity / Altruism		Risk-Aversion	Small error
Params.	$\alpha \leq 2$	$\alpha \geq 2$	$\beta \leq 1$	$\beta \geq 1$	any γ, τ	any $\delta, \hat{\delta}$
$\hat{\lambda}$	1	$\frac{\alpha^2}{4\alpha-4}$	1	$\frac{(\beta+1)^2}{4\beta}$	$1+\gamma\tau$	$1+\delta$
$\hat{\mu}$	$1+\frac{\alpha^2}{4}-\alpha$	0	$\frac{(\beta+1)^2}{4}-\beta$	0	$\frac{1}{4}$	$1-\frac{3}{4(1+\hat{\delta})}$

(b) BPoA($\mathcal{G}^{\alpha}$)

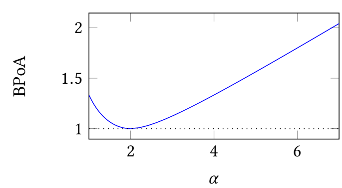


Figure 3. Biased-smoothness bounds for affine cost functions. For a trivial bias of $\alpha = 1, \beta = 0, \gamma\tau = 0$, or $\delta = \hat{\delta} = 0$, we always get the familiar affine smoothness of $(\hat{\lambda} = 1, \hat{\mu} = \frac{1}{4})$. The bound on Altruism with $\beta \leq 1$ is due to Chen et al. [14]. The bound on small error is due to Prop. 6.2 applied to affine cost functions. All other bounds are from a companion workshop paper [4].

Figure 3 summarizes known biased smoothness bounds for routing games with *affine* cost functions (there are also known bounds for more general classes of cost functions, not shown here). Note that from every column in the table we can derive a BPoA bound via Theorem 4.4. For example, for any biased game $\left\langle \mathbf{G}, \hat{\mathbf{G}}^{\alpha} \right\rangle$ where all cost functions in $\mathbf{G}$ are affine, and all agents in $\hat{\mathbf{G}}$ are pessimistic with parameter $\alpha \leq 2$, we have $BPoA(\mathbf{G}, \hat{\mathbf{G}}^{\alpha}) \leq \frac{4}{4\alpha-\alpha^2}$. This bound, which is tight, is illustrated in Fig. 3b. Note that for certain values of α, the BPoA is lower than the PoA, i.e., bias steers the society to better outcomes (we also saw this phenomenon in Example 4.1).

We note that BPoA bounds have also been derived independently using different techniques by Nikolova and Stier-Moses [33], by Meir and Parkes [30] and by Kleer at al. [26].

5 STRUCTURE-DEPENDENT BOUNDS

In this section, we derive a bound, tight up to a constant, on the equilibrium cost for agents participating in a game with heterogeneous biases. This bound depends on their own biased-smoothness parameters, as well as on the structure of the network. In some cases, this also allows us to derive BPoA bounds that depend on the *average bias*. Most importantly, our bounds in this section do not assume any specific class of cost functions, and work for any function and any bias, as long as biased smoothness holds.

5.1 Upper bound

Our primary question is whether we can get a bound on the social cost of any equilibrium of $\mathcal{G}$ in terms of the smoothness parameters of all types, and the optimal social cost.

In the next theorem and corollaries, each c_e is (λ, μ)-smooth (as per Eq. (1)), each $\hat{c}^i_e$ is (λ^i, μ^i)-smooth , and c_e is $(\hat{\lambda}^i, \hat{\mu}^i)$-biased-smooth w.r.t. $\hat{c}^i_e$ (as per Def. 4.3). Recall that $\frac{r_i}{r}$ is the fraction of agents of type i.

Theorem 5.1 (Smoothness theorem). *Consider any game with biased costs $\mathcal{G} = \left\langle \mathbf{G}, \hat{\mathbf{G}} \right\rangle$ where $PW(G) \leq k$. Let $\hat{f}^*$ be an equilibrium of $\hat{\mathbf{G}}$. Then for any type i,*

$$SC_i(\mathbf{G}, \hat{\boldsymbol{f}}^*) \leq \frac{r_i}{r} \frac{\lambda \lambda^i \hat{\lambda}^i}{(1-\sqrt{\mu})^2 (1-\mu^i)(1-\hat{\mu}^i)} \frac{1}{k} OPT(\mathbf{G}^{i,(k)}).$$

In simple words, the theorem says that for every agent type, the equilibrium cost to all r_i agents of this type may increase (compared to the optimal cost) by a factor that only depends on the smoothness of cost functions, their own bias, and the structure of the network, but not on the biases of other types. This result is bi-criteria, as we compare to the optimum of a game with higher demand. Since costs are increasing functions, $\frac{1}{k} OPT(\mathbf{G}^{i,(k)})$ is *increasing* in k.

Lemma 5.2. *For every $x > 0$, and every type i, it holds that:*
(a) $\hat{c}^i(x) \geq (1-\hat{\mu}^i)c(x)$;[3] *and (b) $\hat{c}^i(x) \leq \frac{\lambda \hat{\lambda}^i}{(1-\sqrt{\mu})^2} c(x)$.*

Proof. For (a), define quantity $x' = 0$, so that by $(\hat{\lambda}^i, \hat{\mu}^i)$-biased-smoothness,

$$x(c(x) - \hat{c}^i(x)) = c(x)x + \hat{c}^i(x)(x' - x) \leq \hat{\lambda}^i c(x')x' + \hat{\mu}^i c(x)x = \hat{\mu}^i c(x)x.$$

Hence $c(x) - \hat{c}^i(x) \leq \hat{\mu}^i c(x)$.

For (b), set $q = \sqrt{1/\mu}$ and consider $x' = qx$.

$$c(x)x + \hat{c}^i(x)(x' - x) \leq \hat{\lambda}^i c(x')x' + \hat{\mu}^i c(x)x \qquad \Rightarrow$$
$$c(x)x + \hat{c}^i(x)(q-1)x \leq \hat{\lambda}^i c(qx)qx + \hat{\mu}^i c(x)x \qquad \Rightarrow$$
$$\hat{c}^i(x) \leq \frac{q}{q-1} \hat{\lambda}^i c(qx) + (\hat{\mu}^i - 1)c(x) \leq \frac{q}{q-1} \hat{\lambda}^i c(qx) \qquad (\hat{\mu}^i \leq 1)$$

Since c is (λ, μ)-smooth, $c(qx)x \leq \lambda c(x)x + \mu c(qx)qx$ for every $x > 0$, and thus $c(qx) \leq \frac{\lambda}{1-q\mu} c(x)$. Finally,

$$\hat{c}^i(x) \leq \hat{\lambda}^i \lambda \frac{q}{(q-1)(1-q\mu)} c(x) = \frac{\lambda \hat{\lambda}^i}{(1-\sqrt{\mu})^2} c(x),$$

as required. □

Proof sketch of Theorem 5.1. Let $\hat{\boldsymbol{g}}^*$ be an equilibrium of $\hat{\mathbf{G}}^{i,(k)}$. Define $\hat{C}^{i,*} = \hat{c}^i(p, \hat{\boldsymbol{f}}^*)$ for some used path $p \in P_i(\hat{\boldsymbol{f}}^*)$, and $\hat{C}^{i,(k)} = \hat{c}^i(p, \hat{\boldsymbol{g}}^*)$ for some used path $p \in P_i(\hat{\boldsymbol{g}}^*)$. In words, $\hat{C}^{i,*}, \hat{C}^{i,(k)}$ are the *perceived* costs to agents of type i in equilibria $\hat{f}^*$ and $\hat{g}^*$, respectively. Note that they are not affected by which used path we choose.

The next steps are: to upper bound the ratio between $rk \cdot \hat{C}^{i,(k)} = SC(\hat{\mathbf{G}}^{i,(k)}, \hat{\boldsymbol{g}}^*)$ and $OPT(\mathbf{G}^{i,(k)})$, using Lemma 5.2(b) on every used edge; and to lower bound $\hat{C}^{i,*}$ by the *true cost* $c(p, \hat{\boldsymbol{f}}^*)$ of path p, using Lemma 5.2(a) on all edges of p. The most important part is the inequality $\hat{C}^{i,*} \leq \hat{C}^{i,(k)}$, which is due to Theorem 3.1 applied to the biased game $\hat{\mathbf{G}}$.

Finally, recall that what we need to bound is social cost for all type i agents in the "real game" $\mathbf{G}$, which is $SC_i(\mathbf{G}, \hat{\boldsymbol{f}}^*)$. Since each such type i agent has a cost of exactly $c(p, \hat{\boldsymbol{f}}^*)$, we now chain all the previous inequalities to get the desired upper bound. □

5.2 Implications

Theorem 5.1 has a number of useful corollaries. Some use the fact that for series-parallel networks $k = 1$ (Prop. 2.6).

Corollary 5.3. *Consider any game with biased costs $\mathcal{G} = \langle \mathbf{G}, \hat{\mathbf{G}} \rangle$ where $PW(G) \leq k$. Let $\hat{\boldsymbol{f}}^*$ be an equilibrium of $\mathcal{G}$. For any type i:*

[3] For some biases (e.g. tolls), we have a tighter inequality $\hat{c}^i(x) \geq c(x)$ from the definition. In such cases we can eliminate the factor $1 - \hat{\mu}^i$ in Theorem 5.1 and all following corollaries.

(1) *if* $\mathbf{G}$ *is affine,* $SC_i(\mathbf{G}, \hat{f}^*) \leq 4\frac{r_i}{r}\frac{\lambda^i\hat{\lambda}^i}{(1-\mu^i)(1-\hat{\mu}^i)}k \cdot OPT(\mathbf{G}^i)$*;*

(2) *if type i is unbiased,* $SC_i(\mathbf{G}, \hat{f}^*) \leq \frac{r_i}{r}\frac{\lambda}{1-\mu}\frac{1}{k}OPT(\mathbf{G}^{i,(k)})$.

Moreover, if $\mathbf{G}$ *is symmetric, then:*

(3) $SC(\mathbf{G}, \hat{f}^*) \leq \left(\sum_{i\leq m}\frac{r_i}{r}\frac{\lambda\lambda^i\hat{\lambda}^i}{(1-\sqrt{\mu})^2(1-\mu^i)(1-\hat{\mu}^i)}\right)\frac{1}{k}OPT(\mathbf{G}^{(k)})$*;*

(4) *if* $\mathbf{G}$ *is affine,* $BPoA(\mathcal{G}) \leq O(1)k\sum_{i\leq m}\frac{r_i}{r}\frac{\hat{\lambda}^i}{1-\hat{\mu}^i}$*;*

(5) *If G is series-parallel, then* $BPoA(\mathcal{G}) \leq O(1)\sum_{i\leq m}\frac{r_i}{r}\frac{\hat{\lambda}^i}{1-\hat{\mu}^i}$.

These corollaries provide us with more explicit bounds. In particular: (1) for affine cost functions we can get rid of the bi-criteria result since the optimal social cost is linear in the demand (more generally, for polynomials of degree d the factor becomes $O(k^d)$). (2) entails that in a setting where most agents are unbiased and there is only a small fraction of adversarial agents (as in [6]), the damage to the unbiased agents is limited. (3)-(5) show that the equilibrium social cost in symmetric games depends on the *average bias* over all agents: each type of biased agents can only increase the social cost by a factor that is proportional to their mass (and also affected by the parallel-width of the network). Kleer at al. [26] proved a result similar in spirit to our Corollary 5.3(5), but restricted to toll-sensitivity.

5.3 Lower Bounds

Our main theorems both rely on the restricted structure of the underlying network. Prop. 3.2 already shows the necessity of this restriction when some agents behave adversarially.

We next show that the neither smoothness parameters $(\hat{\lambda}_i, \hat{\mu}_i)$ nor the structural parameter $k = PW(G)$ can be relaxed in Theorem 5.1 and in its corollaries.

Proposition 5.4. *Let* $k \geq 2$*, any smoothness parameters* $(\hat{\lambda}^i, \hat{\mu}^i)$*, and any* $r_i < \frac{1}{2}$*. There is an affine symmetric game with biases* $\mathcal{G} = \langle \mathbf{G}, \hat{\mathbf{G}}\rangle$ *with* $PW(G) = k$*, such that in the unique equilibrium* $\hat{f}^*$ *of* $\hat{\mathbf{G}}$*, the cost for all type i agents increases by a factor of* $\Omega(1)\frac{\hat{\lambda}^i}{1-\hat{\mu}^i}k$ *compared to* $OPT(\mathbf{G})$.

Proof sketch. The proof uses the k-Cross-Parallel network $G_{CP(k)}$ with one additional edge e_{st} (see Fig. 2). We define the real costs as follows: $c_{sa} = c_{bt} = c_{ba} \equiv 0, c_{st} \equiv \frac{q}{2}, c_{ab}(x) = x$, where $q = \frac{\hat{\lambda}^i}{1-\hat{\mu}^i}$. In the flow of $\mathbf{G}$ where all agents split evenly, the cost is $C^o = \frac{1}{k}$. In the biased game, $\hat{c}^i_{ab}(x) = qx$ (thus $\frac{\hat{\lambda}^i}{1-\hat{\mu}^i} \leq \frac{q}{1-0} = q$), and we modify all other agents so they prefer the long path that intersects all short paths. Thus in equilibrium $\hat{f}^*$ all type i agents use edge e_{st} for a cost of $\frac{q}{2} \geq \frac{\hat{\lambda}^i}{2(1-\hat{\mu}^i)} \geq \frac{\hat{\lambda}^i}{2(1-\hat{\mu}^i)}k \cdot C^o$. □

Since we only used affine cost functions, the bounds in both Theorem 5.1 and Cor. 5.3(1) are tight up to a constant. Also, since the network used in the proof is embedded in $G_{CP(k+1)}$, and thus embedded in any acyclic network with $\text{PW}(G) > k$ (due to [29]), the example in Prop. 5.4 can be constructed for *any acyclic network G* with $\text{PW}(G) \geq k + 1$.

6 STRUCTURE-INDEPENDENT BOUNDS

In this section, we leverage known PoA bounds in routing games with player-specific costs [22] to obtain BPoA bounds in biased games with heterogeneous agents. These are bounds that do not depend on the network structure, but only on the cost functions and agents' biases.

The bounds in this section are typically worse than those in Section 5, as they depend on the worst bias rather than the average bias, and this dependency is polynomial rather than linear. Yet, these bounds hold regardless of network structure.

For two cost functions $c, \hat{c}$ and $r > 0$, denote by $\Delta(c, \hat{c}, r) = \sup_{x \in [0,r]} \frac{c(x)}{\hat{c}(x)}$. Let $\Psi(\mathrm{G}) = \max_{i,j \le m} \max_{e \in E} \Delta(c^i_e, c^j_e, r)$.

Theorem 6.1 (Gairing et al. [22]). *Consider any PNRG* $\mathrm{G} = \left\langle G, (\mathbf{c}^i)_{i \le m}, \mathbf{r} \right\rangle$ *where* $\mathbf{c}^i$ *are polynomials of degree d for all i. Then* $PoA(\mathrm{G}) \le (d+1) \cdot \Psi(\mathrm{G})^d$.[4]

Intuitively, the PoA is low if for each edge e and any x, all functions $(c^i_e(x))_{i \le m}$ attain similar values. We can use this PoA bound to prove a similar BPoA bound that is independent of the network structure. Given a game with biased costs $\mathcal{G} = \left\langle \mathrm{G}, \hat{\mathrm{G}} \right\rangle$, let:

- $\overline{\Phi}(\mathcal{G}) = \max_{i \le m} \max_{e \in E} \Delta(\hat{c}^i_e, c_e, r)$;
- $\underline{\Phi}(\mathcal{G}) = \max_{i \le m} \max_{e \in E} \Delta(c_e, \hat{c}^i_e, r)$;
- and $\Psi(\mathcal{G}) = \Psi(\hat{\mathrm{G}})$.

We first observe that Δ is tightly related to biased-smoothness, as biased-smoothness bounds entail bounds on Δ, and vice versa.

Proposition 6.2. *Consider* $c, \hat{c}$ *such that* c *is* (λ, μ)*-smooth,* $\Delta(c, \hat{c}, r) \le 1 + \delta$, *and* $\Delta(\hat{c}, c, r) \le 1 + \hat{\delta}$ *for some* $\delta, \hat{\delta} \ge 0$ *and any* $r > 0$. *Then* c *is* $\left((1+\delta)\lambda, \frac{\mu + \hat{\delta}}{1 + \hat{\delta}}\right)$*-biased-smooth w.r.t.* $\hat{c}$.

Proposition 6.3. *Consider* $c, \hat{c}$ *such that* c *is* (λ, μ)*-smooth, and* $(\hat{\lambda}, \hat{\mu})$*-biased smooth w.r.t.* $\hat{c}$. *Then for any* $r > 0$ $\Delta(c, \hat{c}, r) \le \frac{1}{1 - \hat{\mu}}$; $\Delta(\hat{c}, c, r) \le \frac{\lambda}{(1 - \sqrt{\mu})^2} \hat{\lambda}$. *Also, if* c *is a polynomial of degree at most d,* $\Delta(\hat{c}, c, r) \le (d+1) \cdot \mathrm{e} \cdot \hat{\lambda}$, *where* e *is the natural logarithm base.*

Proposition 6.3 follows from Lemma 5.2 and from the smoothness of polynomial functions. Next, we derive a BPoA bound. Due to tightness results on the PoA [22], we cannot hope to significantly improve the bound in this approach.

Theorem 6.4. *Consider any game* $\mathcal{G} = \left\langle \mathrm{G}, \hat{\mathrm{G}} \right\rangle$ *where all of* $\hat{c}^i_e$ *are polynomials of degree at most d. Then* $BPoA(\mathcal{G}) \le (d+1)(\underline{\Phi}(\mathcal{G})\overline{\Phi}(\mathcal{G}))^{d+1}$.

Proof sketch. Consider an equilibrium flow $\hat{\boldsymbol{f}}^* \in EQ(\hat{\mathrm{G}})$. We use the definitions of $\underline{\Phi}, \overline{\Phi}$ to show that $SC(\mathrm{G}, \hat{\boldsymbol{f}}^*) \le \underline{\Phi}(\mathcal{G})\overline{\Phi}(\mathcal{G})\mathrm{PoA}(\hat{\mathrm{G}})$. We then apply the bound $\mathrm{PoA}(\hat{\mathrm{G}}) \le (d+1) \cdot \Psi(\mathrm{G})^d$ from Theorem 6.1, and show that $\Psi(\hat{\mathrm{G}}) = \Psi(\mathcal{G}) \le \underline{\Phi}(\mathcal{G})\overline{\Phi}(\mathcal{G})$. □

Finally, we can combine Proposition 6.3 and Theorem 6.4 to obtain a BPoA bound that depends only on the smoothness parameters of the "worst" type:

Corollary 6.5. *Consider any game with biased costs* $\mathcal{G} = \langle \mathrm{G}, \hat{\mathrm{G}} \rangle$ *where for all* $i \le m$ *and* $e \in E$: *(a)* $c_e, \hat{c}^i_e$ *are polynomials of degree at most d; and (b)* c_e *is* $(\hat{\lambda}, \hat{\mu})$*-biased smooth w.r.t.* $\hat{c}^i_e$.
Then $BPoA(\mathcal{G}) \le (d+1)^{d+2} \mathrm{e}^{d+1} \left(\frac{\hat{\lambda}}{1 - \hat{\mu}}\right)^{d+1}$.

[4] The tight bound given by Gairing et al. [22] is better by a factor of up to $(d+1)$, however we are mainly interested in the fact that it is independent of the network structure, and so we use a simplified form.

7 CONCLUSION AND FUTURE WORK

We have considered strategic settings in which participants are making routing decisions based on individually perceived costs. From the perspective of a system analyst who considers the objective costs, the agents' behavior deviates from optimal self-interested play. Whether these deviations come from a cognitive limitation, subjective preferences, a behavioral bias, or external influence such as tolls, it is important to understand how the equilibria of the game are affected.

Our work is the first to provide (Biased) PoA bounds for populations with heterogeneous arbitrary biases, and the first to consider heterogeneous biases in general networks. Our main results (Sec. 3 and 5) bound the equilibrium cost of each agent type separately, based only on their own biases and on the structure of the network. When considering the entire population, this entails a bi-criteria bound on the social cost (in some cases can be written as a Biased PoA bound) that depends on the *average* bias in the population.

Since in nonatomic routing games all mixed and correlated equilibria have the same social cost anyway [9], smoothness does not add more "robustness," and is thus perhaps considered less interesting in such games. Our results show that smoothness is useful for a *different kind of robustness*, namely to heterogeneous biases.

An additional result (Sec. 6) relies on the PoA of the heterogeneous "wrong" game, and while it does not require the network to have a particular structure, it provides bounds stated in terms of the *worst* bias in the population, and only applies for polynomial cost functions. We can think of these results as bounding the negative externalities of each type *on all others*, whereas the results in Sections 3 and 5 bound the externalities that all others may inflict *on agents of type* i.

We hope that our results and techniques will inspire further progress in understanding both the effect of network topology and the effect of bounded rationality on equilibria.

Some challenges are to derive BPoA bounds that only depend on the average bias (i.e., not on the network structure); and bounds on the negative externality that become negligible for a small amount of adversarial agents.

Biased smoothness can also be applied to obtain robust PoA bounds in *atomic* congestion games and other normal form games (i.e. for mixed and correlated equilibria), by extending the standard smoothness definition as in [14]. An important open question is whether atomic games with heterogeneous biases can be similarly analyzed, by showing biased-smoothness independently for each type of agent.

REFERENCES

[1] Heayat Zakaie Aashtiani and Thomas L Magnanti. 1981. Equilibria on a congested transportation network. *SIAM Journal on Algebraic Discrete Methods* 2, 3 (1981), 213–226.

[2] Daron Acemoglu, Ali Makhdoumi, Azarakhsh Malekian, and Asuman Ozdaglar. 2016. Informational Braess' Paradox: The Effect of Information on Traffic Congestion. *arXiv preprint arXiv:1601.02039* (2016).

[3] Haris Angelidakis, Dimitris Fotakis, and Thanasis Lianeas. 2013. Stochastic Congestion Games with Risk-Averse Players. In *Proceedings of the 6th International Symposium on Algorithmic Game Theory*. Springer, 86–97.

[4] [Anonimized]. 2015. Playing the Wrong Game: Smoothness Bounds for Congestion Games with Heterogeneous Biases. (2015).

[5] Itai Ashlagi, Dov Monderer, and Moshe Tennenholtz. 2009. Two-terminal routing games with unknown active players. *Artificial Intelligence* 173, 15 (2009), 1441–1455.

[6] Moshe Babaioff, Robert Kleinberg, and Christos H Papadimitriou. 2007. Congestion games with malicious players. In *Proceedings of the 8th ACM conference on Electronic commerce*. ACM, 103–112.

[7] Bikramjit Banerjee and Landon Kraemer. 2010. Coalition structure generation in multi-agent systems with mixed externalities. In *Proceedings of the 9th International Conference on Autonomous Agents and Multiagent Systems: volume 1-Volume 1*. International Foundation for Autonomous Agents and Multiagent Systems, 175–182.

[8] M. Beckmann, B. McGuire, and C. Winsten. 1956. *Studies in the Economics of Transportation*. Yale University Press, New Haven.

[9] Avrim Blum, Eyal Even-Dar, and Katrina Ligett. 2006. Routing without regret: On convergence to Nash equilibria of regret-minimizing algorithms in routing games. In *Proceedings of the twenty-fifth annual ACM symposium on Principles of distributed computing*. ACM, 45–52.

[10] Vincenzo Bonifaci, Mahyar Salek, and Guido Schäfer. 2011. Efficiency of Restricted Tolls in Non-atomic Network Routing Games. In *Proceedings of the 4th International Symposium on Algorithmic Game Theory*. 302–313.

[11] Simina Brânzei, Tomasz Michalak, Talal Rahwan, Kate Larson, and Nicholas R Jennings. 2013. Matchings with externalities and attitudes. In *Proceedings of the 2013 international conference on Autonomous agents and multi-agent systems*. International Foundation for Autonomous Agents and Multiagent Systems, 295–302.

[12] Patrice Caire, Baptiste Alcalde, Leendert van der Torre, and Chattrakul Sombattheera. 2011. Conviviality measures. In *The 10th International Conference on Autonomous Agents and Multiagent Systems-Volume 2*. International Foundation for Autonomous Agents and Multiagent Systems, 895–902.

[13] Po-An Chen, Bart De Keijzer, David Kempe, and Guido Schäfer. 2011. The robust price of anarchy of altruistic games. In *Proceedings of the 7th Workshop on Internet and Network Economics*. Springer, 383–390.

[14] Po-An Chen, Bart De Keijzer, David Kempe, and Guido Schäfer. 2014. Altruism and its impact on the price of anarchy. *ACM Transactions on Economics and Computation* 2, 4 (2014), 17.

[15] Richard Cole, Yevgeniy Dodis, and Tim Roughgarden. 2003. Pricing network edges for heterogeneous selfish users. In *Proceedings of the thirty-fifth annual ACM symposium on Theory of computing*. ACM, 521–530.

[16] José R Correa, Andreas S Schulz, and Nicolás E Stier-Moses. 2004. Selfish routing in capacitated networks. *Mathematics of Operations Research* 29, 4 (2004), 961–976.

[17] José R Correa, Andreas S Schulz, and Nicolás E Stier-Moses. 2008. A geometric approach to the price of anarchy in nonatomic congestion games. *Games and Economic Behavior* 64, 2 (2008), 457–469.

[18] Giovanna Devetag and Massimo Warglien. 2008. Playing the wrong game: An experimental analysis of relational complexity and strategic misrepresentation. *Games and Economic Behavior* 62, 2 (2008), 364 – 382.

[19] David Eppstein. 1992. Parallel recognition of series-parallel graphs. *Information and Computation* 98, 1 (1992), 41–55.

[20] Lisa Fleischer, Kamal Jain, and Mohammad Mahdian. 2004. Tolls for heterogeneous selfish users in multicommodity networks and generalized congestion games. In *Foundations of Computer Science, 2004. Proceedings. 45th Annual IEEE Symposium on*. IEEE, 277–285.

[21] Dimitris Fotakis, Dimitris Kalimeris, and Thanasis Lianeas. 2015. Improving Selfish Routing for Risk-Averse Players. In *International Conference on Web and Internet Economics*. Springer, 328–342.

[22] Martin Gairing, Burkhard Monien, and Karsten Tiemann. 2011. Routing (un-) splittable flow in games with player-specific affine latency functions. *ACM Transactions on Algorithms (TALG)* 7, 3 (2011), 31.

[23] Ralph Hertwig, Greg Barron, Elke U Weber, and Ido Erev. 2004. Decisions from experience and the effect of rare events in risky choice. *Psychological science* 15, 8 (2004), 534–539.

[24] Andreas Jakoby, Maciej Liśkiewicz, and Rüdiger Reischuk. 2006. Space efficient algorithms for directed series–parallel graphs. *Journal of Algorithms* 60, 2 (2006), 85–114.

[25] Tomas Jelinek, Marcus Klaas, and Guido Schäfer. 2014. Computing Optimal Tolls with Arc Restrictions and Heterogeneous Players.. In *STACS*. 433–444.

[26] Pieter Kleer and Guido Schäfer. 2016. The impact of worst-case deviations in non-atomic network routing games. In *International Symposium on Algorithmic Game Theory*. Springer, 129–140.

[27] Elias Koutsoupias and Christos Papadimitriou. 1999. Worst-case equilibria. In *Proceedings of the 16th annual conference on Theoretical aspects of computer science*. Springer-Verlag, 404–413.

[28] Piotr Krysta, Tomasz Michalak, Tuomas Sandholm, and Michael Wooldridge. 2010. Combinatorial auctions with externalities. In *AAMAS'10*. 1471–1472.

[29] Reshef Meir. 2017. Directed graph embeddings. *arXiv preprint arXiv:1711.01806* (2017).

[30] Reshef Meir and David Parkes. 2015. Congestion Games with Distance-Based Strict Uncertainty. In *Proceedings of the 29th AAAI Conference on Artificial Intelligence*.

[31] Igal Milchtaich. 2000. Generic uniqueness of equilibrium in large crowding games. *Mathematics of Operations Research* 25, 3 (2000), 349–364.

[32] Evdokia Nikolova and Nicolás E Stier-Moses. 2014. A mean-risk model for the traffic assignment problem with stochastic travel times. *Operations Research* 62, 2 (2014), 366–382.

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

[34] Jean Oh and Stephen F Smith. 2008. A few good agents: multi-agent social learning. In *Proceedings of the 7th international joint conference on Autonomous agents and multiagent systems-Volume 1*. International Foundation for Autonomous Agents and Multiagent Systems, 339–346.

[35] Georgios Piliouras, Evdokia Nikolova, and Jeff S Shamma. 2016. Risk sensitivity of price of anarchy under uncertainty. *ACM Transactions on Economics and Computation (TEAC)* 5, 1 (2016), 5.

[36] Tim Roughgarden. 2003. The price of anarchy is independent of the network topology. *J. Comput. System Sci.* 67, 2 (2003), 341–364.
[37] Tim Roughgarden. 2009. Intrinsic robustness of the price of anarchy. In *Proceedings of the 41st annual ACM symposium on Theory of computing*. ACM, 513–522.
[38] Tim Roughgarden and Éva Tardos. 2004. Bounding the inefficiency of equilibria in nonatomic congestion games. *Games and Economic Behavior* 47, 2 (2004), 389–403.
[39] David Schmeidler. 1973. Equilibrium points of nonatomic games. *Journal of Statistical Physics* 7, 4 (1973), 295–300.
[40] Guni Sharon, Josiah Hanna, Tarun Rambha, Michael Albert, Peter Stone, and Stephen D Boyles. 2016. Delta-Tolling: Adaptive Tolling for Optimizing Traffic Throughput.. In *ATT@ IJCAI*.
[41] Moshe Tennenholtz and Aviv Zohar. 2009. Learning equilibria in repeated congestion games. In *Proceedings of The 8th International Conference on Autonomous Agents and Multiagent Systems-Volume 1*. International Foundation for Autonomous Agents and Multiagent Systems, 233–240.
[42] Shoshana Vasserman, Michal Feldman, and Avinatan Hassidim. 2015. Implementing the Wisdom of Waze.. In *IJCAI*. 660–666.
[43] John Glen Wardrop. 1952. Road paper. Some theoretical aspects of road traffic research. In *ICE Proceedings: Engineering Divisions*, Vol. 1. 325–362.

A STRUCTURE-DEPENDENT BOUNDS

Lemma A.1. $SC(\hat{\mathbf{G}}^{i,(k)}, \hat{\boldsymbol{g}}^*) \leq \frac{\lambda\lambda^i\hat{\lambda}^i}{(1-\sqrt{\mu})^2(1-\mu^i)} OPT(\mathbf{G}^{i,(k)})$.

Proof. Since by definition $\hat{\mathbf{G}}^i$ (and thus also $\hat{\mathbf{G}}^{i,(k)}$) is (λ^i, μ^i)-smooth, for any flow $\boldsymbol{f}$ of $\hat{\mathbf{G}}^{i,(k)}$,

$$SC(\hat{\mathbf{G}}^{i,(k)}, \hat{\boldsymbol{g}}^*) \leq \frac{\lambda^i}{1-\mu^i} SC(\hat{\mathbf{G}}^{i,(k)}, \boldsymbol{f}). \tag{5}$$

We denote the optimal flow in $\mathbf{G}^{i,(k)}$ by $\boldsymbol{g}^o = \boldsymbol{f}^o(\mathbf{G}^{i,(k)})$.

$$\begin{aligned} SC(\hat{\mathbf{G}}^{i,(k)}, \hat{\boldsymbol{g}}^*) &\leq \frac{\lambda^i}{1-\mu^i} SC(\hat{\mathbf{G}}^{i,(k)}, \boldsymbol{g}^o) = \frac{\lambda^i}{1-\mu^i} \sum_{e\in E} g^o_e \hat{c}^i_e(g^o_e) && \text{(By (5) with } \boldsymbol{f} = \boldsymbol{g}^o) \\ &\leq \frac{\lambda^i}{1-\mu^i} \sum_{e\in E} g^o_e \left(\frac{\lambda\hat{\lambda}^i}{(1-\sqrt{\mu})^2} c_e(g^o_e) \right) && \text{(from Lemma 5.2(b))} \\ &= \frac{\lambda\lambda^i\hat{\lambda}^i}{(1-\sqrt{\mu})^2(1-\mu^i)} \sum_{e\in E} g^o_e c_e(g^o_e) = \frac{\lambda\lambda^i\hat{\lambda}^i}{(1-\sqrt{\mu})^2(1-\mu^i)} OPT(\mathbf{G}^{i,(k)}), \end{aligned}$$

as required. □

Theorem 5.1. *Consider any game with biased costs $\mathcal{G} = \langle \mathbf{G}, \hat{\mathbf{G}} \rangle$ where $PW(G) \leq k$. Let $\hat{\boldsymbol{f}}^*$ be an equilibrium of $\hat{\mathbf{G}}$. Then for any type i,*

$$SC_i(\mathbf{G}, \hat{\boldsymbol{f}}^*) \leq \frac{r_i}{r} \frac{\lambda\lambda^i\hat{\lambda}^i}{(1-\sqrt{\mu})^2(1-\mu^i)(1-\hat{\mu}^i)} \frac{1}{k} OPT(\mathbf{G}^{i,(k)}).$$

Proof. Similarly to the definitions before Theorem 3.1, we define $\hat{\boldsymbol{f}}^*$ as an equilibrium of $\hat{\mathbf{G}}$, $\hat{C}^{i,*} = \hat{c}^i(p, \hat{\boldsymbol{f}}^*)$ for some used path $p \in P_i(\hat{\boldsymbol{f}}^*)$, and $\hat{C}^{i,(k)} = \hat{c}^i(p, \hat{\boldsymbol{g}}^*)$ for some used path $p \in P_i(\hat{\boldsymbol{g}}^*)$. Note that $\hat{C}^{i,*}, \hat{C}^{i,(k)}$ are not effected by which used path we choose.

By Theorem 3.1 applied to games $\hat{\mathbf{G}}, \hat{\mathbf{G}}^{i,(k)}$,

$$r \cdot \hat{C}^{i,*} \leq r \cdot \hat{C}^{i,(k)} = \frac{1}{k}(r \cdot k)\hat{C}^{i,(k)} = \frac{1}{k}(\sum_{p\in A} \hat{g}^*_p)\hat{C}^{i,(k)} = \frac{1}{k} \sum_{p\in P^i(\hat{\boldsymbol{g}}^*)} \hat{g}^*_p \hat{c}^i(p, \hat{\boldsymbol{g}}^*) = \frac{1}{k} SC(\hat{\mathbf{G}}^{i,(k)}, \hat{\boldsymbol{g}}^*).$$

On the other hand, for any used path $p \in P^i(\hat{f}^*)$, by Lemma 5.2(a)

$$\hat{C}^{i,*} = \hat{c}^i(p, \hat{f}^*) = \sum_{e \in p} \hat{c}^i_e(\hat{f}^*_e) \geq \sum_{e \in p}(1 - \hat{\mu}^i)c_e(\hat{f}^*_e) = (1 - \hat{\mu}^i)c(p, \hat{f}^*).$$

Combining the above bounds, we get an upper bound on the total cost for all type i agents:

$$SC_i(\mathbf{G}, \hat{f}^*) = \sum_{p \in A} \hat{f}^*_{i,p} c(p, \hat{f}^*) \leq \sum_{p \in A} \hat{f}^*_{i,p} \frac{\hat{C}^{i,*}}{1 - \hat{\mu}^i} = \frac{r_i \hat{C}^{i,*}}{1 - \hat{\mu}^i} = \frac{r_i (r \cdot \hat{C}^{i,*})}{r(1 - \hat{\mu}^i)} \leq \frac{r_i}{r} \frac{1}{(1 - \hat{\mu}^i)} \frac{1}{k} SC(\hat{\mathbf{G}}^{i,(k)}, \hat{\boldsymbol{g}}^*).$$

Finally, we get the theorem by plugging in the bound of Lemma A.1. □

Corollary 5.3. *Consider any game with biased costs $\mathcal{G} = \left\langle \mathbf{G}, \hat{\mathbf{G}} \right\rangle$ where $PW(G) \leq k$. Let $\hat{f}^*$ be an equilibrium of $\mathcal{G}$. For any type i:*

1. *if* $\mathbf{G}$ *is affine,* $SC_i(\mathbf{G}, \hat{f}^*) \leq 4\frac{r_i}{r}\frac{\lambda^i \hat{\lambda}^i}{(1-\mu^i)(1-\hat{\mu}^i)} k \cdot OPT(\mathbf{G}^i)$;
2. *if type i is unbiased,* $SC_i(\mathbf{G}, \hat{f}^*) \leq \frac{r_i}{r}\frac{\lambda}{1-\mu}\frac{1}{k} OPT(\mathbf{G}^{i,(k)})$.

Moreover, if $\mathbf{G}$ *is symmetric, then:*

3. $SC(\mathbf{G}, \hat{f}^*) \leq \left(\sum_{i \leq m} \frac{r_i}{r} \frac{\lambda \lambda^i \hat{\lambda}^i}{(1-\sqrt{\mu})^2(1-\mu^i)(1-\hat{\mu}^i)}\right) \frac{1}{k} OPT(\mathbf{G}^{(k)})$;
4. *if* $\mathbf{G}$ *is affine,* $BPoA(\mathcal{G}) \leq O(1)k \sum_{i \leq m} \frac{r_i}{r} \frac{\hat{\lambda}^i}{1-\hat{\mu}^i}$;
5. *If G is series-parallel, then* $BPoA(\mathcal{G}) \leq O(1) \sum_{i \leq m} \frac{r_i}{r} \frac{\hat{\lambda}^i}{1-\hat{\mu}^i}$.

Proof.

(1) First, for affine functions, $\lambda = 1, \mu = \frac{1}{4}$, thus $\frac{\lambda}{(1-\sqrt{\mu})^2} = \frac{1}{1/4} = 4$ (by Lemma. A.2 there is no better pair (λ, μ)). Second, for any $k \geq 1$ and any flow $\hat{f}$ in G,

$$SC(\mathbf{G}, k \cdot \hat{f}) = \sum_{e \in E}(k \cdot f_e)c_e(k \cdot f_e) = \sum_{e \in E}(k \cdot f_e)(a_e \cdot k \cdot f_e + b_e) \leq k^2 \sum_{e \in E} f_e(a_e f_e + b_e) = k^2 SC(\mathbf{G}, \hat{f}),$$

thus $OPT(\mathbf{G}^{i,(k)}) = SC(\mathbf{G}, \boldsymbol{f}^o(\mathbf{G}^{i,(k)})) \leq SC(\mathbf{G}, k \cdot \boldsymbol{f}^o(\mathbf{G}^i)) \leq k^2 SC(\mathbf{G}, \boldsymbol{f}^o(\mathbf{G}^i)) = k^2 OPT(\mathbf{G}^i)$. We get the corollary by replacing the respective terms in Theorem 5.1. More generally, for degree-d polynomials, the approximation factor is $O(k^d)$.

(2) Strictly speaking, Cor. 5.3(2) is not entailed by Theorem 5.1. Proof is by following the same steps of the proof of Theorem 5.1, except for Lemma 5.2, which becomes redundant for unbiased agents (see Footnote 3). Thus $\frac{\lambda \hat{\lambda}^i}{(1-\sqrt{\mu})^2(1-\hat{\mu}^i)}$ can be omitted from the bound, which leaves us with $\frac{\lambda^i}{1-\mu^i} = \frac{\lambda}{1-\mu}$ since $\hat{\boldsymbol{c}}^i = \boldsymbol{c}^i$.

(3) This is since if $\mathbf{G}$ is symmetric, then $\mathbf{G}^i = \mathbf{G}$ and $\mathbf{G}^{i,(k)} = \mathbf{G}^{(k)}$ for all i and k. Thus

$$SC(\mathbf{G}, \hat{f}^*) = \sum_{i \leq m} SC_i(\mathbf{G}, \hat{f}^*) \leq \left(\sum_{i \leq k} \frac{r_i}{r} \frac{\lambda \lambda^i \hat{\lambda}^i}{(1 - \sqrt{\mu})^2(1 - \mu^i)(1 - \hat{\mu}^i)}\right) \frac{1}{k} OPT(\mathbf{G}^{(k)}).$$

(4) From (1),

$$SC(\mathbf{G}, \hat{f}^*) = \sum_{i \leq m} SC_i(\mathbf{G}, \hat{f}^*) \leq \sum_{i \leq m} 4 \frac{r_i}{r} \frac{\lambda^i \hat{\lambda}^i}{(1 - \mu^i)(1 - \hat{\mu}^i)} k \cdot OPT(\mathbf{G}^i) = 4kOPT(\mathbf{G}) \sum_{i \leq m} \frac{r_i}{r} \frac{\lambda^i \hat{\lambda}^i}{(1 - \mu^i)(1 - \hat{\mu}^i)}.$$

(5) Follows immediately from (3) with $k = 1$.

□

A.1 A single smoothness parameter

For a given class of cost functions, we denote their smoothness parameters by (λ, μ_λ), i.e. by fixing $\lambda \geq 1$ (necessary for any class that includes constant functions), and finding the minimal $\mu = \mu_\lambda$ such that all functions in the class are (λ, μ) smooth. Recall that $\frac{\lambda}{1-\mu_\lambda}$ is always minimized for $\lambda = 1$ [16].

Lemma A.2. *For polynomial functions, $\frac{\lambda}{(1-\sqrt{\mu}_\lambda)^2}$ is minimized for $\lambda = 1$.*

We suspect that this is true for any function class that includes all constant functions.

Proof. Consider any class of polynomial functions of degree at most d. It is not hard to show that $\mu_\lambda = \frac{d\lambda}{(\lambda(1+d))^{1+\frac{1}{d}}} = \lambda^{-\frac{1}{d}} \frac{d}{(1+d)^{1+\frac{1}{d}}}$ (similar to the smoothness bounds in [16]).

Thus

$$\frac{\lambda}{(1-\sqrt{\mu})^2} = \frac{\lambda}{\left(1-\sqrt{\lambda^{-\frac{1}{d}}\frac{d}{(1+d)^{1+\frac{1}{d}}}}\right)^2} = \frac{\lambda}{\left(1-\lambda^{-\frac{1}{2d}}\frac{d}{(1+d)^{\frac{d+1}{2d}}}\right)^2}$$

$$= \frac{1}{\left(\lambda^{-\frac{1}{2}} - \lambda^{-\frac{1}{2}}\lambda^{-\frac{1}{2d}}\frac{d}{(1+d)^{\frac{d+1}{2d}}}\right)^2}$$

$$= \frac{1}{\left(\lambda^{-\frac{1}{2}} - \lambda^{-\frac{d+1}{2d}}\frac{d}{(1+d)^{\frac{d+1}{2d}}}\right)^2} = \frac{1}{y(\lambda)^2}$$

Now, $\frac{\partial \lambda^{-\frac{1}{2}}}{\partial \lambda} = -\frac{1}{2}\lambda^{-\frac{3}{2}}$, whereas

$$\frac{\partial\left(\lambda^{-\frac{d+1}{2d}}\frac{d}{(1+d)^{\frac{d+1}{2d}}}\right)}{\partial \lambda} = -\frac{d+1}{2d}\frac{d}{(1+d)^{\frac{d+1}{2d}}}\lambda^{-\frac{d+3}{2d}}$$

$$= -\frac{1}{2(1+d)^{\frac{d+1}{2d}-1}}\lambda^{-\frac{d+3}{2d}}$$

$$= -\frac{1}{2(1+d)^{-\frac{d-1}{2d}}}\lambda^{-\frac{d+3}{2d}}$$

$$= -\frac{1}{2}\lambda^{\frac{d+3}{2d}}(1+d)^{\frac{d-1}{2d}}$$

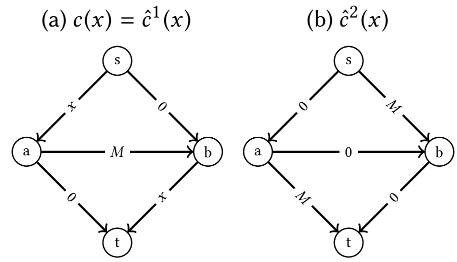


Figure 4. All agents need to select an $s-t$ path. Some fraction of type 2 agents adopt the modified cost function $\hat{c}^2$, where $M \gg 1$ is some large constant.

It thus holds that for the *derivative* of $y(\lambda)$:

$$
\begin{aligned}
y'(\lambda) = -\frac{1}{2}\left(\lambda^{\frac{d+3}{2d}}(1+d)^{\frac{d-1}{2d}} - \lambda^{-\frac{3}{2}}\right) \leq 0 && \iff \\
\left(\lambda^{\frac{d+3}{2d}}(1+d)^{\frac{d-1}{2d}} - \lambda^{-\frac{3}{2}}\right) \geq 0 && \iff \\
\lambda^{\frac{d+3}{2d}}(1+d)^{\frac{d-1}{2d}} \geq \lambda^{-\frac{3}{2}} && \iff \\
\lambda^{\frac{d+3}{2d}}\lambda^{\frac{3}{2}} \geq (1+d)^{-\frac{d-1}{2d}} && \iff \\
\lambda^{\frac{4d+3}{2d}} \geq (1+d)^{-\frac{d-1}{2d}} && \iff \\
\lambda^{4d+3} \geq (1+d)^{-(d-1)} && \iff \\
\lambda \geq (1+d)^{-\frac{d-1}{4d+3}}
\end{aligned}
$$

Note that the last right-hand expression is upper-bounded by 1 for all $d \geq 1$. In particular for all $\lambda \geq 1$, $y'(\lambda)$ is non-positive, $y(\lambda)$ is non-increasing, and $\frac{\lambda}{(1-\sqrt{\mu}_\lambda)^2} = \frac{1}{y(\lambda)^2}$ is non-decreasing. Hence $\frac{1}{(1-\sqrt{\mu}_1)^2} \leq \frac{\lambda}{(1-\sqrt{\mu}_\lambda)^2}$ for all $\lambda \geq 1$. □

A.2 Lower bounds

Theorem A.3 ([29]). *Let G be a 2-terminal directed acyclic graph, and let $k \geq 2$. The following conditions coincide:*

(1) G is a directed series-parallel graph;
(2) The directed Braess graph G_B is not d-embedded in G;
(3) PW(G) = 1;

Proposition A.4. *Consider any graph such that $PW(G) > 1$. For any $\epsilon > 0$ and any $M' > 0$, there is a symmetric game with biases $\mathcal{G} = (\mathrm{G}, \hat{\mathrm{G}})$ and equilibria $\boldsymbol{f}^* \in EQ(\mathrm{G})$, $\hat{\boldsymbol{f}}^* \in EQ(\hat{\mathrm{G}})$, such that:*

- *Only a fraction $r_2 = \epsilon$ of the agents in $\mathcal{G}$ are biased;*
- *For the unbiased type 1, $SC_1(\mathrm{G}, \hat{\boldsymbol{f}}^*) \geq M' \cdot SC_1(\mathrm{G}, \boldsymbol{f}^*)$;*
- *In particular, $BPoA(\mathcal{G}) > (1-\epsilon)M'$.*

Proof. By theorem A.3 it is sufficient to construct a bad example on the Braess graph to prove the proposition.

Indeed, consider the network in Figure 4, where the value M is some large number that will be set later on. Suppose the total mass is one unit, and that all agents need to flow from s to t. In the game $\mathbf{G}$, costs are as in Figure 4 (a). Thus in the equilibrium $\boldsymbol{f}^*$, the agents will split evenly among the paths $s-a-t$ and $s-b-t$, so that each agent incurs the same cost of $(2\cdot 1/2)^M = 1^M = 1$, and thus $SC_1(\boldsymbol{f}^*, \mathbf{G}) = r_1 = 1-\epsilon$. Now, denote the unbiased agents as type 1, and suppose that some fraction $r_2 = \epsilon > 0$ of agents is of a type 2, with biased costs $\hat{c}^2$. See Figure 4 (b).

In the equilibrium $\hat{\boldsymbol{f}}^*$, all type 2 agents will select the path $s-a-b-t$, whereas all type 1 agents will split evenly. Thus there will be $\hat{f}^*_e = \frac{1}{2}+\frac{\epsilon}{2} > f^*_e$ agents on each edge $s-a$ and $b-t$.

We now compute the cost for the type 1 agents in $\hat{\boldsymbol{f}}^*$. Set $M > \frac{M'}{\epsilon}$. For every such agent taking the path $p=(s-a-t)$, the cost is

$$c(p, \hat{\boldsymbol{f}}^*) = c_{ua}(\hat{f}^*_{ua}) = c_{ua}(2(\frac{1}{2}+\frac{\epsilon}{2}))^M = (1+\epsilon)^M > M'.$$

Thus the total cost for type 1 agents is $SC_1(\mathbf{G}, \hat{\boldsymbol{f}}^*) > (1-\epsilon)M' = M' SC_1(\mathbf{G}, \boldsymbol{f}^*)$.

Finally, since $\boldsymbol{f}^*$ is a valid flow in $\mathbf{G}$,

$$\text{BPoA}(\mathcal{G}) \geq \frac{SC(\hat{\boldsymbol{f}}^*, \mathbf{G})}{SC(\boldsymbol{f}^*, \mathbf{G})} \geq \frac{SC_1(\hat{\boldsymbol{f}}^*, \mathbf{G})}{SC(\boldsymbol{f}^*, \mathbf{G})} > \frac{(1-\epsilon)M'}{1} = (1-\epsilon)M',$$

as required. □

Proposition 5.4. *Let any $k \geq$, any $q > 2$, and any $r_i < \frac{1}{2}$. There is a symmetric game with biases $\mathcal{G} = \langle \mathbf{G}, \hat{\mathbf{G}} \rangle$, and a flow $\hat{\boldsymbol{f}}^*$ s.t.:*

1. *$PW(\mathbf{G}) = k$;*
2. *in $\hat{\mathbf{G}}$ there are r_i agents of type i, and $\sum_{j\leq m} r_j = 1$;*
3. *$\langle \mathbf{G}, \hat{\mathbf{G}}^i \rangle$ is $(\hat{\lambda}^i, \hat{\mu}^i)$-biased smooth, where $\frac{\hat{\lambda}^i}{1-\hat{\mu}^i} = q$;*
4. *$\hat{\boldsymbol{f}}^*$ is the unique equilibrium of $\hat{\mathbf{G}}$;*
5. *$SC_i(\mathbf{G}, \hat{\boldsymbol{f}}^*) \geq \Omega(1)\frac{r_i}{rk}\frac{\hat{\lambda}^i}{1-\hat{\mu}^i} OPT(\mathbf{G}^{i,(k)}) = \Omega(1)\frac{r_i}{r}\frac{\hat{\lambda}^i}{1-\hat{\mu}^i} kOPT(\mathbf{G})$.*

Proof. Consider the k-Cross-Parallel graph $G_{CP(k)}$ (see Fig. 2), with one additional edge e_{st}. We define the real costs as follows: $c_{sa} = c_{bt} = c_{ba} \equiv 0, c_{st} \equiv \frac{q}{2}, c_{ab}(x) = x$. There is a mass of $r=1$ agents, and all paths from s to t are allowed (the game is symmetric). In the equilibrium flow of $\mathbf{G}^{i,(k)}$ the mass $rk=k$ splits evenly among all k parallel paths (excluding e_{st}), and the cost for each agent is $c_{ab}(1) = 1$. In particular, $OPT(\mathbf{G}^{i,(k)}) = k$. Similarly, in the flow of $\mathbf{G}$ where all agents split evenly, $OPT(\mathbf{G}) = r\cdot c_{ab}(\frac{1}{k}) = \frac{1}{k}$.

We now define the modified costs. For type i, $\hat{c}^i_{ab}(x) = qx$, whereas all other costs remain unchanged. For all other agents (denote as type j), we set $\hat{c}^j_{st} = \hat{c}^j_{sa} = \hat{c}^j_{bt} \equiv 10qk$ and $\hat{c}^j_{ab} = \hat{c}^j_{ba} \equiv 0$. We argue that $(\mathbf{G}, \hat{\mathbf{G}}^i)$ is $(q,0)$-biased smooth: We know (see Table 3)) that any affine function $c(x)$ is $(\frac{q^2}{4(q-1)}, 0)$-biased smooth w.r.t. $\hat{c}(x) = c(qx)$ for any $q\geq 1$. Note that $\frac{q^2}{4(q-1)} \leq q$ for all $q > \frac{4}{3}$, and thus $\frac{\hat{\lambda}^i}{1-\hat{\mu}^i} \leq \frac{q}{1-0} = q$.

In any equilibrium of $\hat{\mathbf{G}}$, all type j agents follow the long path through all vertices $(s, b_1, a_2, b_2, \ldots, b_{k-1}, a_k, t)$. Hence for any path that includes an $a-b$ edge, the cost for i is at least $\hat{c}^i_{ab}(r_j) = q\cdot r_j > \frac{q}{2}$. Thus in the only equilibrium $\hat{\boldsymbol{f}}^*$, all type i agents will select the direct path e_{st} and experience a cost of $\frac{q}{2}$.

Putting everything together,

$$SC_i(\mathbf{G}, \hat{\boldsymbol{f}}^*) = r_i\frac{q}{2} = \frac{1}{2}\frac{r_i}{rk}qk = \Omega(1)\frac{r_i}{rk}\frac{\hat{\lambda}^i}{1-\hat{\mu}^i} OPT(\mathbf{G}^{i,(k)}) = \Omega(1)\frac{r_i}{r}\frac{\hat{\lambda}^i}{1-\hat{\mu}^i} kOPT(\mathbf{G}),$$

as required. □

Since we only used affine cost functions, the bounds in both Theorem 5.1 and Cor. 5.3(1) are tight up to a constant. Also, since the graph used in the proof is a minor of $G_{CP(k+1)}$, and thus by Theorem A.3 a minor of any graph with $\mathrm{PW}(G) > k$, the example in Prop. 5.4 can be constructed for *any graph* G with $\mathrm{PW}(G) > k$.

B STRUCTURE-INDEPENDENT BOUNDS

Proposition 6.2. *Consider $c, \hat{c}$ such that c is (λ, μ)-smooth, $\Delta(c, \hat{c}, r) \leq 1 + \delta$, and $\Delta(\hat{c}, c, r) \leq 1 + \hat{\delta}$ for some $\delta, \hat{\delta} \geq 0$. Then c is $\left((1+\delta)\lambda, \frac{\mu+\hat{\delta}}{1+\hat{\delta}}\right)$-biased-smooth w.r.t. $\hat{c}$ (in the range $[0, r]$).*

Proof. Consider some $x, x' \in [0, r]$. If $x' \geq x$, then

$$\begin{aligned}
c(x)x + \hat{c}(x)(x'-x) &\leq c(x)x + (1+\delta)c(x)(x'-x)\\
&= c(x)x + (1+\delta)c(x)x' - c(x)x - \delta c(x)x = (1+\delta)c(x)x' - \delta c(x)x\\
&\leq (1+\delta)(\lambda c(x')x' + \mu c(x)x) - \delta c(x)x = (1+\delta)\lambda c(x')x' + (\mu + \delta(\mu-1))c(x)x\\
&\leq (1+\delta)\lambda c(x')x' + \mu c(x)x \leq (1+\delta)\lambda c(x')x' + \frac{\mu+\hat{\delta}}{1+\hat{\delta}}c(x)x \qquad (\mu < 1)
\end{aligned}$$

If $x > x'$, then

$$\begin{aligned}
c(x)x + \hat{c}(x)(x'-x) &= c(x)x - \hat{c}(x)(x-x') \leq c(x)x - \frac{1}{1+\hat{\delta}}c(x)(x-x')\\
&= (1 - \frac{1}{1+\hat{\delta}})c(x)x + \frac{1}{1+\hat{\delta}}c(x)x'\\
&= \frac{\hat{\delta}}{1+\hat{\delta}}c(x)x + \frac{1}{1+\hat{\delta}}c(x)x' \leq \frac{\hat{\delta}}{1+\hat{\delta}}c(x)x + \frac{1}{1+\hat{\delta}}(\lambda c(x')x' + \mu c(x)x)\\
&= \frac{\lambda}{1+\hat{\delta}}c(x')x' + \frac{\hat{\delta}+\mu}{1+\hat{\delta}}c(x)x\\
&\leq \lambda c(x')x' + \frac{\mu+\hat{\delta}}{1+\hat{\delta}}c(x)x \leq (1+\delta)\lambda c(x')x' + \frac{\mu+\hat{\delta}}{1+\hat{\delta}}c(x)x
\end{aligned}$$

We therefore get that for any $x, x' \geq 0$, $c(x)x + \hat{c}(x)(x'-x) \leq (1+\delta)\lambda c(x')x' + (\mu + \hat{\delta}(1-\mu))c(x)x$, meaning that c is $\left((1+\delta)\lambda, \mu + \hat{\delta}(1-\mu)\right)$-biased smooth w.r.t $\hat{c}$. □

Proposition 6.3. *Let $r > 0$. Suppose that c is (λ, μ)-smooth, and $(\hat{\lambda}, \hat{\mu})$-biased smooth w.r.t. $\hat{c}$. Then $\Delta(c, \hat{c}, r) \leq \frac{1}{1-\hat{\mu}}$; $\Delta(\hat{c}, c, r) \leq \frac{\lambda}{(1-\sqrt{\mu})^2}\hat{\lambda}$. Also, for polynomials of degree at most d, $\Delta(\hat{c}, c, r) \leq (d+1)\mathrm{e}\hat{\lambda}$, where e is the natural logarithm base.*

Proof. By definition, $\Delta(c, \hat{c}, r) \leq \sup_{x>0} \frac{c(x)}{\hat{c}(x)}$, which is upper bounded by $\frac{1}{1-\hat{\mu}}$ due to Lemma 5.2(a). Similarly, $\Delta(c, \hat{c}, r) \leq \sup_{x>0} \frac{\hat{c}(x)}{c(x)} \leq \frac{\lambda}{(1-\sqrt{\mu})^2}\hat{\lambda}$ due to Lemma 5.2(b).

Due to biased-smoothness, for all $x, x' \geq 0$, $c(x) + \hat{c}(x)(x'-x) \leq \hat{\lambda}c(x')x' + \hat{\mu}c(x)x$. Thus for polynomials of degree d,

$$
\begin{aligned}
&\hat{c}(x)(x'-x) \le \hat{\lambda}c(x')x' + (\hat{\mu}-1)c(x)x && \Rightarrow \\
&\hat{c}(x)(x'-x) \le \hat{\lambda}c(x')x' && \Rightarrow \\
&\hat{c}(x)\epsilon x \le \hat{\lambda}c((1+\epsilon)x)(1+\epsilon)x && \Rightarrow && \text{(For } x' = (1+\epsilon)x\text{)} \\
&\hat{c}(x)\epsilon \le \hat{\lambda}(1+\epsilon)^d c(x)(1+\epsilon) && \Rightarrow && (c \text{ is degree } d \text{ polynomial}) \\
&\hat{c}(x) \le \hat{\lambda}\frac{1}{\epsilon}(1+\epsilon)^{d+1}c(x) \\
&\le \hat{\lambda}\frac{1}{1/(d+1)}\left(1+\frac{1}{d+1}\right)^{d+1} c(x) &&&& \text{(For } \epsilon = \tfrac{1}{d+1}\text{)} \\
&\le (d+1)\cdot \mathrm{e} \cdot \hat{\lambda}c(x),
\end{aligned}
$$

as required. □

Theorem 6.4. *Consider any game with biased costs $\mathcal{G} = \left\langle \mathbf{G}, \hat{\mathbf{G}} \right\rangle$ where $\hat{\boldsymbol{c}}^i$ are polynomial functions for all i. Then $BPoA(\mathcal{G}) \le \underline{\Phi}(\mathcal{G})\overline{\Phi}(\mathcal{G})PoA(\hat{\mathbf{G}}) \le (d+1)(\underline{\Phi}(\mathcal{G})\overline{\Phi}(\mathcal{G}))^{d+1}$.*

Proof. Consider games $\mathbf{G}$ and $\hat{\mathbf{G}}$, both over the network $G = (V, E)$. For any flow $\boldsymbol{f}$, we have

$$
\begin{aligned}
SC(\mathbf{G}, \boldsymbol{f}) &= \sum_{e\in E} f_e c_e(f_e) = \sum_{e\in E}\sum_{i\le m} f_{e,i}c_e(f_e) \le \sum_{e\in E}\sum_{i\le m} f_{e,i}\underline{\Phi}(\mathcal{G})\hat{c}^i_e(f_e) \\
&= \underline{\Phi}(\mathcal{G})\sum_{i\le m}\sum_{e\in E} f_{e,i}\hat{c}^i_e(f_e) = \underline{\Phi}(\mathcal{G})\sum_{i\le m} SC_i(\hat{\mathbf{G}}, \boldsymbol{f}) \le \underline{\Phi}(\mathcal{G})SC(\hat{\mathbf{G}}, \boldsymbol{f}).
\end{aligned}
$$

Similarly, $SC(\hat{\mathbf{G}}, \boldsymbol{f}) \le \overline{\Phi}(\mathcal{G})SC(\mathbf{G}, \boldsymbol{f})$ for every flow $\boldsymbol{f}$.

For $\hat{\boldsymbol{f}}^* \in EQ(\hat{\mathbf{G}})$, $\hat{\boldsymbol{f}}^o = \hat{\boldsymbol{f}}^o(\mathbf{G})$, we have

$$
\begin{aligned}
SC(\mathbf{G}, \hat{\boldsymbol{f}}^*) &\le \underline{\Phi}(\mathcal{G})SC(\hat{\mathbf{G}}, \hat{\boldsymbol{f}}^*) \le \underline{\Phi}(\mathcal{G})\mathrm{PoA}(\hat{\mathbf{G}})SC(\hat{\mathbf{G}}, \hat{\boldsymbol{f}}^o) \\
&\le \underline{\Phi}(\mathcal{G})\mathrm{PoA}(\hat{\mathbf{G}})\overline{\Phi}(\mathcal{G})SC(\mathbf{G}, \hat{\boldsymbol{f}}^o) \\
&= \underline{\Phi}(\mathcal{G})\overline{\Phi}(\mathcal{G})\mathrm{PoA}(\hat{\mathbf{G}}),
\end{aligned}
$$

Due to Theorem 6.1, $\mathrm{BPoA}(\mathcal{G})\overline{\Phi}(\mathcal{G})\underline{\Phi}(\mathcal{G})\mathrm{PoA}(\hat{\mathbf{G}}) \le (d+1)\underline{\Phi}(\mathcal{G})\Psi(\mathcal{G})^d\overline{\Phi}(\mathcal{G})$, which is a constant independent of the network G.

Finally,

$$
\begin{aligned}
\Psi(\mathcal{G}) = \Psi(\hat{\mathbf{G}}) &= \max_{i,j\le m}\Delta(\hat{\boldsymbol{c}}^i, \hat{\boldsymbol{c}}^j, r) = \max_{i,j\le m}\max_{e\in E}\sup_{x\in[0,r]}\frac{\hat{c}^i_e(x)}{\hat{c}^j_e(x)} \\
&= \max_{i,j\le m}\max_{e\in E}\sup_{x\in[0,r]}\frac{\hat{c}^i_e(x)}{c_e(x)}\frac{c_e(x)}{\hat{c}^j_e(x)} \\
&\le \max_{i\le m}\max_{e\in E}\sup_{x\in[0,r]}\frac{\hat{c}^i_e(x)}{c_e(x)}\max_{i\le m}\max_{e\in E}\sup_{x\in[0,r]}\frac{c_e(x)}{\hat{c}^j_e(x)} \\
&= \max_{i\le m}\Delta(\hat{\boldsymbol{c}}^i, \boldsymbol{c}, r)\max_{i\le m}\Delta(\boldsymbol{c}, \hat{\boldsymbol{c}}^j, r) = \overline{\Phi}(\mathcal{G})\underline{\Phi}(\mathcal{G}),
\end{aligned}
$$

which entails the theorem. □

Corollary 6.5. *Consider any game with biased costs* $\mathcal{G} = \left\langle \text{G}, \hat{\text{G}} \right\rangle$ *where for all* $i \leq m$ *and* $e \in E$*: (a)* $c_e, \hat{c}^i_e$ *are polynomials of degree at most d; and (b)* c_e *is* $(\hat{\lambda}, \hat{\mu})$*-biased smooth w.r.t.* $\hat{c}^i_e$*. Then* $BPoA(\mathcal{G}) \leq (d+1)\text{e}\frac{\hat{\lambda}}{1-\hat{\mu}} PoA(\hat{\text{G}}) \leq (d+1)^{d+2}\text{e}^{d+1}\left(\frac{\hat{\lambda}}{1-\hat{\mu}}\right)^{d+1}$.

Proof. By Prop. 6.3,

$$\underline{\Phi}(\mathcal{G}) = \max_{i \leq m} \max_{e \in E} \Delta(c, \hat{c}^i, r) = \max_{i \leq m} \max_{e \in E} \sup_{x \in [0,r]} \frac{c_e(x)}{\hat{c}^i_e(x)} \leq \max_{i \leq m} \frac{1}{1-\hat{\mu}^i} \leq \frac{1}{1-\hat{\mu}};$$

$$\overline{\Phi}(\mathcal{G}) = \max_{i \leq m} \max_{e \in E} \Delta(\hat{c}^i, c, r) = \max_{i \leq m} \max_{e \in E} \sup_{x \in [0,r]} \frac{\hat{c}^i_e(x)}{c_e(x)} \leq \max_{i \leq m}((d+1) \cdot \text{e} \cdot \hat{\lambda}^i) = (d+1) \cdot \text{e} \cdot \hat{\lambda}.$$

By combining the above bounds with Theorem 6.4, the proof is complete. □