On the performance of approximate equilibria in congestion games

George Christodoulou* Elias Koutsoupias† Paul G. Spirakis ‡

Abstract

We study the performance of approximate Nash equilibria for congestion games with polynomial latency functions. We consider how much the price of anarchy worsens and how much the price of stability improves as a function of the approximation factor ϵ. We give tight bounds for the price of anarchy of atomic and non-atomic congestion games and for the price of stability of non-atomic congestion games. For the price of stability of atomic congestion games we give non-tight bounds for linear latencies. Our results not only encompass and generalize the existing results of exact equilibria to ϵ-Nash equilibria, but they also provide a unified approach which reveals the common threads of the atomic and non-atomic price of anarchy results. By expanding the spectrum, we also cast the existing results in a new light.

1 Introduction

Algorithmic Game Theory has studied extensively and with remarkable success the computational issues of Nash equilibria. As a result, we understand almost completely the computational complexity of exact Nash equilibria (see for example [16, 7, 19, 33]). The results indicate a long-suspected drawback of Nash equilibria, namely that they cannot be computed effectively. Despite substantial recent progress [21, 19, 17, 34], it is still open whether approximate Nash equilibria share the same drawback. But in any case, they seem to provide a more reasonable equilibrium concept: It usually makes sense to assume that an agent is willing to accept a situation that is almost optimal to him.

In another direction, a large body of research in Algorithmic Game Theory concerns the degree of performance degradation of systems due to the selfish behavior of the users. Central to this area is the notion of the price of anarchy (PoA) [20, 25] and the price of stability (PoS) [3]. The first notion compares the social cost of the worst-case equilibrium to the social optimum, which could be obtained if every agent followed obediently a central authority. The second notion is very similar but it considers the best Nash equilibrium instead of the worst one.

A natural question then is how the performance of a system is affected when its users are approximately selfish: What is the *approximate price of anarchy* and the *approximate price of stability*? Clearly, by allowing the players to be almost rational (within an ϵ factor), we expand the equilibrium concept and we expect the PoA to get worse. On the other hand, the PoS should improve. The question is how they change as functions of the parameter ϵ. This is exactly the question that we address in this work.

We study two fundamental classes of games: the class of atomic congestion games [26, 23] and the class of non-atomic congestion (or selfish routing) games [5, 15, 22]. Both classes of games played central role in the development of the area of the PoA [20, 27, 28]. Although the PoA and PoS of these games for exact equilibria was established long ago [27, 11, 10, 4]—and actually some work [29, 32, 8] addressed partially the question for the PoA of approximate equilibria—our results add an unexpected understanding of the issues involved.

While the classes of atomic and non-atomic games are conceptually very similar, dissimilar techniques were employed in the past to answer the questions concerning the PoA and PoS. Moreover, the qualitative aspects of the answers were quite different. For instance, the maximum PoA for the non-atomic case is attained by the Pigou network (Figure 1), while for the atomic case it is attained [4, 11] by a completely different network (with the structure of Figure 2).

There are two main differences between the atomic and non-atomic games. The first difference is the *"uniqueness"* of equilibria: Non-atomic congestion games have a unique exact Nash (or Wardrop as it is called in these games) equilibrium, and therefore the PoS is not different than the PoA. On the other hand, atomic congestion games may have multiple exact equilibria. Perhaps, because of this, the problem

*University of Liverpool, Liverpool, UK. Email: gchristo@liv.ac.uk
†Department of Informatics, University of Athens. Email: elias@di.uoa.gr
‡Computer Engineering and Informatics Department, Patras University, Greece. Email: spirakis@cti.gr

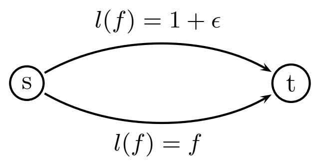

Figure 1: The Pigou network.

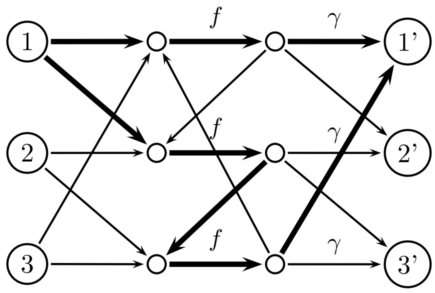

Figure 2: The figure illustrates the lower bound for the PoA for non-atomic selfish routing for the case $n = 3$. There are n distinct edge latency functions: $l(f) = f$, $l(f) = \gamma$ (a constant which depends on ϵ), $l(f) = 0$ (omitted in the picture). There are n commodities of rate 1 with source i and destination i' and 2 paths that connect them. The two paths for the first commodity are shown in bold lines. A similar construction works for atomic games as well.

of determining the PoS proved more challenging [10, 6] for this case. New techniques needed for upper bounding the PoS for linear latencies which exploited the potential of these games; for polynomial latencies, the problem is still open. Furthermore, the lower bound for linear latencies is quite complicated and, unlike the selfish routing case, it has a dependency on the number of players (it attains the maximum value at the limit).

The second main difference between the two classes of games is *"integrality"*: In the case of atomic games, when a player considers switching to another strategy, he has to take into account the extra cost that he will add to the edges (or facilities) of the new strategy. The number of players on the new edges increases by one and this changes the cost. On the other hand, in the selfish routing games the change of strategies has no additional cost. A simple—although not entirely rigorous—way to think about it, is to consider the effects of a tiny amount of flow that ponders whether to change path: it will not really affect the flow on the new edges (at least for continuous cost functions).

When we consider approximate equilibria, the uniqueness difference dissolves and only the integrality difference remains. But still, determining the PoS seems to be a harder problem than determining the PoA.

1.1 Our contribution and related work

Our work encompasses and generalizes some fundamental results in the area of the PoA [27, 11, 10, 4] (see also the recently published book [24] for background information). Our techniques not only provide a unifying approach but they cast the existing results in a new light. For instance, the Pigou network (Figure 1) is still the tight example for the PoS, but not for the PoA. For the latter, the network of Figure 2 is tight for $\epsilon \leq 1$ and the network of Figure 4 is tight for larger ϵ. We use the *multiplicative definition of approximate equilibria*: In atomic games, a player does not switch to a new strategy as long as his current cost is less than $1 + \epsilon$ times the new cost. In the selfish routing games, we use exactly the same definition [27]: the flow is at an ϵ-Nash equilibrium when the cost on its paths is less than $1 + \epsilon$ times the cost of every alternate path. There have been other definitions for approximate Nash equilibria in the literature. For example, for algorithmic issues of Nash equilibria, the most-studied one is the additive case [21, 16]. However, since the PoA is a ratio, the natural definition is the multiplicative one. A slightly different multiplicative approximate Nash equilibrium definition was used in [9], where they study convergence issues for congestion games.

There is a large body of work on the PoA in various models [24]. More relevant to our work are the following publications: For atomic congestion games, it was shown in [4, 11] that the pure PoA of linear latencies is 5/2. Later in [10], it was shown that the ratio 5/2 is tight even for mixed and correlated equilibria. For weighted congestion games, the PoA is $1 + \phi \approx 2.618$ [4]. For polynomial latencies of degree

p, [11] showed that the PoA of pure equilibria is $p^{\Theta(p)}$, and [4] showed $p^{\Theta(p)}$ bounds for mixed equilibria and for weighted congestion games. Aland et al. [1] gave exact bounds for weighted and unweighted congestion games. For the PoS of the atomic case, it was shown in [10, 6] that for linear congestion games it is $1+\sqrt{3}/3$. Also for the linear selfish routing games, [29, 32] give tight bounds for the PoA of ϵ-Nash equilibria when $\epsilon \leq 1$.

For the selfish routing paradigm and exact Nash equilibria, [27] gave the PoA (and consequently the PoS) for polynomial latencies of degree p with nonnegative coefficients and more general functions (see [13] or [24] for a simplified version of the proof); the results are extended to non-atomic games in [28].

In this work, we give tight upper and lower bounds for the PoA and PoS of atomic and non-atomic congestion games with polynomial latencies of degree p, with the exception of the PoS for atomic congestion games (Section 6) for which we only give non-tight bounds for linear latencies; Figure 3 shows the upper and lower bounds for this case. The bounds are functions of the approximation factor ϵ and the degree p. We assume throughout that the coefficients of the polynomials are nonnegative.

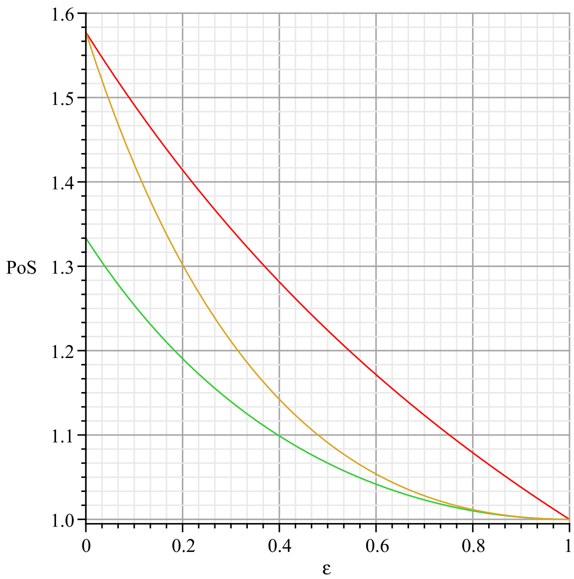

Figure 3: The upper and lower bound of the price of stability (Theorems 7 and 9) for atomic games and linear latencies. The lower line is the Pigou bound $\frac{4}{(1+\epsilon)(3-\epsilon)}$.

Our results reveal qualitative differences between exact and approximate Nash equilibria for non-atomic congestion games. These games have unique exact Nash equilibria, which makes the PoS to be equal to the PoA. When we consider the larger class of approximate equilibria the uniqueness of equilibria vanishes and the PoS and PoA diverge. Another interesting finding of our work is that for $\epsilon = p$, the PoS drops to 1 in both the atomic and non-atomic case (this was known in the non-atomic case [30, 14]).

For all cases we give appropriate *characterizations* of the ϵ-Nash equilibria, which are of independent interest and generalize known characterizations of exact equilibria. These characterizations apply to every nondecreasing and continuous latency functions, not just polynomials. Especially for the PoS, our characterization involves a notion which generalizes the potential function of Rosenthal [26, 24].

All upper bounds proofs have the following outline: Using an appropriate characterization for the ϵ-Nash equilibria and suitable algebraic inequalities—which for non-atomic games involve real numbers and for atomic games involve nonnegative integers—we bound the cost of the ϵ-Nash equilibria.

The cases of PoS however are more complicated. In these cases, the characterization of ϵ-Nash equilibria involves potential-like quantities. For exact Nash equilibria there is essentially a unique potential ; but for approximate equilibria, there are many choices for the potential and we have to figure out which potential gives the best results. This together with the technical challenges of the algebraic inequalities makes the proofs for the PoS harder. In fact, we don't have a direct way to establish a good upper bound on the PoS. Instead, we compute an upper bound on the PoA of a subset of strategies which is guaranteed to contain an ϵ-Nash equilibrium. Since this subset contains at least one ϵ-Nash equilibrium, it is also an upper bound on the PoS. The potential play the crucial role in selecting a useful subset of strategies: We focus on the subset of strategies that minimize the potential which is guaranteed to contain at least one ϵ-Nash equilibrium.

All the games that we consider in this work possess pure ϵ-Nash equilibria and our upper bound proofs are tuned to these, for simplicity. However, all our proofs can be directly extended to mixed and correlated equilibria—the difference will be an extra outer sum which corresponds to the expectation. The upper bounds of our theorems apply unmodified to the more general classes of *mixed* and *correlated* equilibria.

An unpublished preliminary version of this work [12], studied the same questions for linear latencies. Here we generalize the results to polynomial latencies which are technically much more challenging.

2 Definitions

A congestion game [26], also called an exact potential game [23], is a tuple $(N, E, (\mathcal{S}_i)_{i\in N}, (f_e)_{e\in E})$, where $N = \{1, \dots, n\}$ is a set of n players, E is a set of facilities, $\mathcal{S}_i \subseteq 2^E$ is a set of pure strategies for player i: a pure strategy $A_i \in \mathcal{S}_i$ is simply a subset of facilities and l_e is a cost (or latency) function, one for each facility $e \in E$. The cost of player i for the pure strategy profile $A = (A_1, \dots, A_n)$ is $c_i(A) = \sum_{e\in A_i} l_e(n_e(A))$, where $n_e(A)$ is the number of players who use facility e in the strategy profile A.

Definition 1. A pure strategy profile A is an ϵ-Nash equilibrium iff for every player $i \in N$

$$c_i(A) \leq (1+\epsilon)c_i(A_i, A_{-i}), \qquad \forall A_i \in \mathcal{S}_i \tag{1}$$

We believe that the multiplicative definition of approximate equilibria makes more sense in the framework that we consider: Given that the PoA is a ratio, we need a definition that is insensitive to scaling.

The social cost of a pure strategy profile A is the sum of the players' costs

$$\mathrm{C}(A) = \sum_{i\in N} c_i(A).$$

The approximate PoA and PoS, is the social cost of the worst-case and best-case ϵ-equilibrium over the optimal social cost

$$PoA = \max_{A \text{ is an } \epsilon\text{-N.E.}} \frac{\mathrm{C}(A)}{\mathrm{OPT}}, \quad PoS = \min_{A \text{ is an } \epsilon\text{-N.E.}} \frac{\mathrm{C}(A)}{\mathrm{OPT}}.$$

Instead of defining formally the class of non-atomic congestion games, we prefer to focus on the more illustrative–more restrictive though–class of selfish routing games. The difference in the two models is that in a non-atomic game, there does not exist any network and the strategies of the players are just subsets of facilities (as in the case of atomic congestion games) and they do not necessarily form a path in a network. Intuitively, one can view non-atomic congestion games as atomic congestion games where the number of players tends to infinity; in the limit, there are infinitely many players each controlling a negligible amount of flow. Since the amount of contribution of each player to the cost of an edge is negligible, when a player switches paths, the cost of every edge remains unaffected. This is not true for atomic congestion games; it turns out that this is the most important difference between atomic and non-atomic games in the analysis of the price of anarchy and stability.

Let $G = (V, E)$ be a directed graph, where V is a set of vertices and E is a set of edges. In this network we consider k commodities: source-node pairs (s_i, t_i) with $i = 1 \dots k$, that define the sources and destinations. The set of simple paths in every pair (s_i, t_i) is denoted by $\mathcal{P}_i$, while with $\mathcal{P} = \cup_{i=1}^k \mathcal{P}_i$ we denote their union. A flow f, is a mapping from the set of paths to the set of nonnegative reals $f : \mathcal{P} \to \mathbb{R}^+$. For a given flow f, the flow on an edge is defined as the sum of the flows of all the paths that use this edge $f_e = \sum_{P\in\mathcal{P}, e\in P} f_P$. We relate with every commodity (s_i, t_i) a traffic rate r_i, as the total traffic that needs to move from s_i to t_i. A flow f is feasible, if for every commodity $\{s_i, t_i\}$, the traffic rate equals the flow of every path in $\mathcal{P}_i$, $r_i = \sum_{P\in\mathcal{P}_i} f_P$. Every edge introduces a delay in the network. This delay depends on the load of the edge and is determined by a delay function, $l_e(\cdot)$. An instance of a routing game is denoted by the triple (G, r, l). The latency of a path P, for a given flow f, is defined as the sum of all the latencies of the edges that belong to P, $l_P(f) = \sum_{e\in P} l_e(f_e)$. The social cost that evaluates a given flow f, is the total delay due to f

$$C(f) = \sum_{P\in\mathcal{P}} l_P(f) f_P.$$

The total delay can also be expressed via edge flows $C(f) = \sum_{e\in E} l_e(f_e) f_e$.

From now on, when we refer to flows, we mean feasible flows. We define (as in [27]) the ϵ-Nash or ϵ-Wardrop equilibrium flows:

Definition 2. A feasible flow f, is an ϵ-Nash (or ϵ-Wardrop) equilibrium, if and only if for every commodity $i \in \{1, \dots, k\}$ and $P_1, P_2 \in \mathcal{P}_i$ with $f_{P_1} > 0$, $l_{P_1}(f) \leq (1+\epsilon) l_{P_2}(f)$.

3 Non-atomic PoA

In this section we give tight bounds for the approximate PoA for non-atomic congestion games. We present them for the special case of selfish routing (or network non-atomic congestion games), but they apply unchanged to the more general class of non-atomic congestion games. One of the technical difficulties here is that a different approach is required for small and large values of ϵ.

We start with a condition which relates the cost of ϵ-Nash equilibria to any other flow (and in particular the optimal flow). This is the generalization to approximate equilibria of the inequality established by Beckmann, McGuire, and Winston [5] for exact Wardrop equilibria.

Theorem 1. *If f is an ϵ-Nash flow and f^* is any feasible flow of a non-atomic congestion game:*

$$\sum_{e\in E} l_e(f_e) f_e \leq (1+\epsilon) \sum_{e\in E} l_e(f_e) f_e^*.$$

Proof. Let f be an ϵ-approximate Nash flow, and f^* be the optimum flow (or any other feasible flow). Given a flow f and for a particular commodity i, we denote by f_p^i and f_e^i the corresponding amount of flow that i routes through the path p and through edge e respectively. From the definition of ϵ-Nash equilibria (Inequality (1)), we get that for every commodity i and for every path p with non-zero flow in f and any other path p':

$$\sum_{e\in p} l_e(f_e) \leq (1+\epsilon) \sum_{e\in p'} l_e(f_e).$$

For every commodity i, we sum up these inequalities for all pairs of paths p and p' weighted with the amount of flow of f and f^* on these paths.

$$\sum_{p,p'} f_p^i {f^*_{p'}}^i \sum_{e\in p} l_e(f_e) \leq (1+\epsilon) \sum_{p,p'} f_p^i {f^*_{p'}}^i \sum_{e\in p'} l_e(f_e)$$
$$\sum_{p'} {f^*_{p'}}^i \sum_{e\in E} l_e(f_e) f_e^i \leq (1+\epsilon) \sum_{p} f_p^i \sum_{e\in E} l_e(f_e) {f^*_e}^i$$
$$(\sum_{p'} {f^*_{p'}}^i) \sum_{e\in E} l_e(f_e) f_e^i \leq (1+\epsilon) (\sum_{p} f_p^i) \sum_{e\in E} l_e(f_e) {f^*_e}^i$$

But $\sum_p f_p^i = \sum_{p'} {f^*_{p'}}^i$ is equal to the total rate r_i for the feasible flows f and f^*. Simplifying and summing up for all commodities i, we get

$$\sum_{e\in E} l_e(f_e) f_e \leq (1+\epsilon) \sum_{e\in E} l_e(f_e) f_e^*.$$

□ □

To bound the PoA, we will need the following lemma:

Lemma 1. *Let $g(x)$ be a polynomial with nonnegative coefficients of degree p. The inequality*

$$g(x)\, y \leq \alpha\, g(x)\, x + \beta\, g(y)\, y$$

holds for the following set of parameters:

$\alpha = \frac{p}{(p+1)(1+\epsilon)}$	$\beta = \frac{(1+\epsilon)^p}{p+1}$	*for every* $\epsilon \geq (1+p)^{1/p} - 1$
$\alpha = p/(p+1)^{1+1/p}$	$\beta = 1$	*for every* $\epsilon \geq 0$

Proof. We first consider the case of $\epsilon \geq (1+p)^{1/p} - 1$ and then we show how to adapt the proof for small values of ϵ.

To establish the lemma for every polynomial with nonnegative coefficients of degree p, it is sufficient to prove it for every one of its monomials. For $y = 0$ the lemma holds trivially. For $y > 0$, we divide by y^{k+1} and by letting $z = x/y$, it is sufficient to establish that $\alpha\, z^{k+1} - z^k + \beta \geq 0$ for every positive z and every $k \in [0, p]$.

The minimum value of $\alpha\, z^{k+1} - z^k + \beta$ is achieved when $z = k/((k+1)\alpha)$. This minimum value is

$$\beta - \frac{k^k}{(k+1)^{k+1}\, \alpha^k}. \quad (2)$$

Now, it is not hard to establish that, for every $\gamma \in (0,1)$ the expression $\frac{k^k}{(k+1)^{k+1}\gamma^k}$ as a function of k is convex. Since by its definition $\alpha < 1$, we get that the (2) achieves its maximum at the extreme points ($k = 0$ or $k = p$).

We now establish that the value of the expression is indeed nonnegative for the extreme values of k. For $k = p$, the minimum is achieved when $z = p/((p+1)\alpha) = 1+\epsilon$ and this minimum is 0. We also establish it for the other extreme value, $k = 0$: it is clearly true for every $\epsilon \geq (1+p)^{1/p} - 1$ (this is where we need to bound ϵ from below).

For small values of ϵ and $k = 0$, the above value of β does not satisfy the inequality. It is however sufficient to have $\beta = 1$. From the other extreme $k = p$, we need to have $\beta - \frac{p^p}{(p+1)^{p+1}\alpha^p} \geq 0$, which is satisfied for $\alpha = p/(p+1)^{1+1/p}$. □ □

We have now what we need to establish the upper bound of the main theorem of this section. The matching lower bounds are shown in Lemma 2 and Lemma 3. Since the lower bounds are based on network non-atomic games, we get that the theorem applies also to the special class of network non-atomic congestion games[1].

Theorem 2. *The PoA of non-atomic congestion games with latency functions polynomials of degree p with nonnegative coefficients is* $(1+\epsilon)^{p+1}$*, for every* $\epsilon \geq (p+1)^{1/p} - 1$*, and* $(1/(1+\epsilon) - p/(p+1)^{1+1/p})^{-1}$*, for every* $\epsilon \leq (p+1)^{1/p} - 1$.

Proof. The proof is based entirely on the inequality of Theorem 1:

$$\sum_{e \in E} l_e(f_e) f_e \leq (1+\epsilon) \sum_{e \in E} l_e(f_e) f_e^*.$$

We can bound the right hand side using Lemma 1, as follows

$$\sum_{e \in E} l_e(f_e) f_e^* \leq \sum_{e \in E} \alpha l_e(f_e) f_e + \beta l_e(f_e^*) f_e^*.$$

It follows that for large ϵ (i.e. $\epsilon \geq (p+1)^{1/p} - 1$), the PoA is at most

$$\frac{(1+\epsilon)\beta}{1-(1+\epsilon)\alpha} = (1+\epsilon)^{p+1},$$

and for small ϵ (i.e. $\epsilon \leq (p+1)^{1/p} - 1$), it is at most

$$\frac{(1+\epsilon)\beta}{1-(1+\epsilon)\alpha} = (1/(1+\epsilon) - p/(p+1)^{1+1/p})^{-1}.$$

The following two lemmas establish that these upper bounds are tight. □ □

Lemma 2 (Non-Atomic-PoA-Lower-Bound for $\epsilon \leq (p+1)^{1/p} - 1$)**.** *There are instances of non-atomic congestion games with polynomial latencies of degree p, with approximate PoA of at least* $(1/(1+\epsilon) - p/(p+1)^{1+1/p})^{-1}$*, for every* $\epsilon \leq (p+1)^{1/p} - 1$.

Proof. We construct an instance with $m+k$ commodities, each of them with unit flow, and $2(m+k)$ facilities. We choose m, k so that m/k gives an arbitrarily good approximation of $(p+1)^{1/p}$. There are two types of facilities (see Figure 2):

- $m+k$ facilities α_i, $i = 1, \ldots, m+k$, with latency $l_{\alpha_i}(x) = x^p$ and
- $m+k$ facilities β_i, $i = 1, \ldots, m+k$, with constant latency $l_{\beta_i}(x) = \gamma = m^{p+1}/(1+\epsilon) - km^p$.

For each commodity i there are two pure strategies, i.e.

$$A_i = \{\alpha_i, \ldots, \alpha_{i+m}\} \text{ and } P_i = \{\alpha_{i+m+1}, \ldots, \alpha_{i+m+k+1}\} \cup \{\beta_i\},$$

where the indices are taken cyclically (i.e., mod $m+k$).

[1] This means that one cannot use the special structure of network non-atomic congestion games to show an upper bound of a lower value.

First we prove that playing the first strategy $A = (A_1, \ldots, A_{m+k})$ is a ϵ-Nash equilibrium. The cost experienced by flow in commodity i is $c_i(A) = m^{p+1}$, as each commodity uses m facilities of type α with latency m^p each. The latency of strategy P_i is

$$c_i(P_i, A_{-i}) = km^p + \gamma = \frac{c_i(A)}{1+\epsilon}$$

and therefore A is an ϵ-Nash equilibrium.

The optimal cost is bounded from above by the cost of strategy P where commodity i has cost $c_i(P) = k^{p+1} + \gamma$ and so the PoA is at least

$$\frac{c_i(A)}{c_i(P)} = \frac{m^{p+1}}{k^{p+1} + \gamma} = (1/(1+\epsilon) - p/(p+1)^{1+1/p})^{-1}.$$

□ □

Lemma 3 (Non-Atomic-PoA-Lower-Bound for $\epsilon \geq (p+1)^{1/p} - 1$)**.** *There are instances of non-atomic congestion games with polynomial latencies of degree p, with approximate PoA of at least $(1+\epsilon)^{p+1}$, for every $\epsilon \geq (p+1)^{1/p} - 1$.*

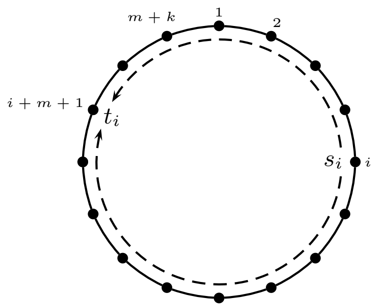

Figure 4: The lower bound for non-atomic linear latencies and large ϵ.

Proof. Consider an undirected cycle of $m + k$ vertices $C = (v_1, \ldots, v_{m+k})$ with edge latency $l_e(f_e) = f_e^p$ on every edge e (Figure 4). We choose m, k so that m/k gives an arbitrarily good approximation of $1 + \epsilon$. There are $m + k$ unit-demand commodities each one with source $s_i = v_i$ and destination $t_i = v_{i+m+1}$ (again indices are taken cyclically). Clearly, for each commodity i there are exactly two simple paths that connect s_i to t_i; i.e. the clockwise path using m edges and the counterclockwise path using k edges. If every commodity routes the traffic via the clockwise path, then the load on every edge is m. The latency experienced by players of each commodity is m^{p+1}–each commodity uses m edges each incurring cost m^p. The alternative path has length k and so the latency on that path is km^{p+1}, so this is an m/k-approximate flow. The optimal cost is upper-bounded by the cost of the flow that routes all the traffic counterclockwise having latency k^{p+1} per commodity; every edge has load k and every commodity uses k edges. This gives an approximate PoA of at least $(m/k)^{p+1} \approx (1+\epsilon)^{p+1}$. □ □

4 Atomic PoA

In this section we study the approximate PoA for the case of atomic congestion games for which we give a tight bound. It will be easier to present the proof if we first establish a few independent non-game-theoretic facts.

The following lemma was implicitly used in Aland et al. [1] and will play a central role in the proof:

Lemma 4. *Let $g(x)$ be any polynomial of degree at most p with nonnegative coefficients. For every $\alpha, \beta, z \in \mathbb{N}$:*

$$\beta \cdot g(\alpha + 1) \leq \kappa \cdot \alpha \cdot g(\alpha) + \lambda \cdot \beta \cdot g(\beta),$$

where

$$\kappa = \frac{(z+2)^p - (z+1)^p}{(z+1)^{p+1} - z^{p+1}}, \qquad \lambda = \frac{(z+1)^{2p+1} - (z+2)^p z^{p+1}}{(z+1)^{p+1} - z^{p+1}}$$

Also, we will need the following fact which follows by a standard straightforward argument when we take the derivative with respect to z: the expression $z^{p+1}/(z+1)^p$ is increasing in z, for nonnegative z.

We will also need the following lemma:

Lemma 5. *If* $1+\epsilon < (z+1)^{p+1}/(z+2)^p$ *then* $1-\kappa(1+\epsilon) > 0$, *where*

$$\kappa = \frac{(z+2)^p - (z+1)^p}{(z+1)^{p+1} - z^{p+1}},$$

is as in Lemma 4.

Proof. We observe that since $1+\epsilon < (z+1)^{p+1}/(z+2)^p$, it suffices to establish $\kappa(z+1)^{p+1}/(z+2)^p < 1$. We get

$$\begin{aligned}
\kappa(z+1)^{p+1}/(z+2)^p < 1 & \quad \Leftrightarrow \\
\frac{(z+2)^p - (z+1)^p}{(z+1)^{p+1} - z^{p+1}} \cdot \frac{(z+1)^{p+1}}{(z+2)^p} < 1 & \quad \Leftrightarrow \\
(z+2)^p(z+1)^{p+1} - (z+1)^p(z+1)^{p+1} < (z+1)^{p+1}(z+2)^p - z^{p+1}(z+2)^p & \quad \Leftrightarrow \\
z^{p+1}(z+2)^p < (z+1)^p(z+1)^{p+1} & \quad \Leftrightarrow \\
\frac{z^{p+1}}{(z+1)^p} < \frac{(z+1)^{p+1}}{(z+2)^p}. &
\end{aligned}$$

Since $z^{p+1}/(z+1)^p$ is increasing in z, the lemma holds. □ □

Define

$$\rho(z) = \frac{(1+\epsilon)\left((z+1)^{2p+1} - z^{p+1}(z+2)^p\right)}{(z+1)^{p+1} - z^{p+1} - (1+\epsilon)\left((z+2)^p - (z+1)^p\right)}, \tag{3}$$

We will show that the PoA is equal to $\rho(z)$ where z is the maximum integer which satisfies $\frac{z^{p+1}}{(z+1)^p} \le 1+\epsilon$.

By a straightforward calculation, when $\frac{z^{p+1}}{(z+1)^p} = 1+\epsilon$, the expression $\rho(z)$ satisfies $\rho(z) = \rho(z-1)$. This shows that in the expression $\rho(z)$ of the PoA we can also choose z to be the maximum integer which satisfies $\frac{z^{p+1}}{(z+1)^p} < 1+\epsilon$ (same as above but with a strict inequality).

We are now ready to state and prove the main theorem of the section.

Theorem 3 (Atomic-PoA)**.** *For any positive real* ϵ*, the approximate PoA of general congestion games with linear latencies is* $\rho(z)$ *as in* (3)*, where* $z \in \mathbb{N}$ *is the maximum integer which satisfies*

$$\frac{z^{p+1}}{(z+1)^p} < 1+\epsilon. \tag{4}$$

In particular, when $\frac{z^{p+1}}{(z+1)^p} = 1+\epsilon$*, then* $\rho(z)$ *is equal to the simpler expression* z^{p+1}*.*

Proof. Let $A = (A_1, \ldots, A_n)$ be an ϵ-Nash equilibrium, and $P = (P_1, \ldots, P_n)$ be the optimum allocation. From the definition of ϵ-Nash equilibria (Inequality (1)) we get

$$\sum_{e \in A_i} l_e(n_e(A)) \le (1+\epsilon) \sum_{e \in P_i} l_e(n_e(A)+1).$$

If we sum up for every player i and use Lemma 4, we get

$$\begin{aligned}
\mathrm{C}(A) &= \sum_{i \in N} c_i(A) \\
&= \sum_{i \in N} \sum_{e \in A_i} l_e(n_e(A)) \\
&= \sum_{e \in E} n_e(A) l_e(n_e(A)) \\
&\le (1+\epsilon) \sum_{e \in E} n_e(P) l_e(n_e(A)+1) \\
&\le (1+\epsilon) \cdot \kappa \cdot \sum_{e \in E} n_e(A) l_e(n_e(A)) + (1+\epsilon) \cdot \lambda \cdot \sum_{e \in E} n_e(P) l_e(n_e(P)) \\
&= (1+\epsilon) \cdot \kappa \cdot \mathrm{C}(A) + (1+\epsilon) \cdot \lambda \cdot \mathrm{OPT},
\end{aligned}$$

where the last inequality follows from Lemma 4. By reordering and substituting κ and λ as in Lemma 4, we get the bound ρ of the theorem.

We now discuss the constraints on z. Notice first that Lemma 4 requires z to be an integer. Furthermore, since in the last reordering step we divide by $1-\kappa(1+\epsilon)$, z must satisfy $1-\kappa(1+\epsilon)>0$.

Thus for every integer z which satisfies $1-\kappa(1+\epsilon)>0$, the PoA is bounded by $\rho(z)$. To get the best $\rho(z)$ we should use the optimal value of z: the one which satisfies these constraints and minimizes $\rho(z)$. To find the best value of z, we could argue directly[2], but there is a much simpler way: It suffices to find a matching lower bound. Indeed, the next lemma gives a matching lower of $\rho(z)$ where z is the maximum integer which satisfies $\frac{z^{p+1}}{(z+1)^p}\leq 1+\epsilon$. We need however to verify that this value of z is in the appropriate region; more precisely, we need to show that for this z we have $1-\kappa(1+\epsilon)>0$. But this z satisfies the premises of Lemma 5, and therefore it is in the desired region.

When $\frac{z^{p+1}}{(z+1)^p}=1+\epsilon$, it is straightforward to see that the expression of the theorem is simplified to z^{p+1}. □ □

To get a simpler expression of the PoA of the theorem for large ϵ, we observer that the expression $\frac{z^{p+1}}{(z+1)^p}$ is very close to $z-p$ (in the sense that their difference tends to 0). Thus for large values of ϵ, the PoA is approximately $z^{p+1}\approx(1+\epsilon+p)^{1+p}\approx(1+\epsilon)^{p+1}$.

We now present a matching lower bound in the next lemma.

Lemma 6 (Atomic-PoA-Lower-Bound)**.** *For any real positive ϵ, there are instances of congestion games with polynomial latencies of degree p, for which the approximate PoA is at least equal to $\rho(z)$ of the expression* (3)*, where z is the maximum integer which satisfies* (4)*.*

Proof. As we argued above, the bound of Theorem 3 is unaffected when we consider the constraint (4) to be tight. Since it is more convenient to prove the lower bound when the constraint is tight, we let $z\in\mathbb{N}$ be the maximum integer which satisfies the constraint (4) with strict inequality.

We will make use of the parameter

$$\gamma=\frac{(z+1)^{p+1}-(1+\epsilon)\,(z+2)^p}{(1+\epsilon)\,(z+1)^p-z^{p+1}},$$

which is a nonnegative real number for the selected z; in particular its numerator is nonnegative and its denominator is positive.

We now construct an instance with $z+2$ players and $2z+4$ facilities. There are two types of facilities:

- $z+2$ facilities of type α, with latency $l_e(x)=x^p$ and
- $z+2$ facilities of type β with latency $l_e(x)=\gamma x^p$.

Player i has two alternative pure strategies, i.e.

$$A_i=\{\alpha_1,\ldots,\alpha_{i-1},\alpha_{i+1},\ldots,\alpha_{z+2}\}\cup\{\beta_{i+1},\ldots,\beta_{i+1+z}\}\text{ and }P_i=\{\alpha_i,\beta_i\},$$

where the indices are taken cyclically (mod $z+2$).

First we prove that playing the first strategy $A=(A_1,\ldots,A_n)$ is a ϵ-Nash equilibrium. The cost of player i is

$$c_i(A)=(z+1)^{p+1}+\gamma z^{p+1},$$

as there are exactly $z+1$ players using facilities of type α and exactly z players using facilities of type β. If player i unilaterally switches to P_i he experiences cost

$$c_i(P_i,A_{-i})=(z+2)^p+\gamma(z+1)^p.$$

The parameter γ was selected to satisfy the equality

$$(z+2)^p+\gamma(z+1)^p=\frac{c_i(A)}{1+\epsilon}.$$

With this, $c_i(P_i,A_{-i})\leq c_i(A)/(1+\epsilon)$, which implies that A is an ϵ-Nash equilibrium.

The optimal cost is at most the cost of the strategy profile P, where every player has cost $c_i(P)=1+\gamma$ and so the PoA is at least

$$\frac{c_i(A)}{c_i(P)}=\frac{(z+1)^{p+1}+\gamma z^{p+1}}{1+\gamma}=\rho(z).$$

□ □

[2]Based on the observation that $\rho(z)=\rho(z-1)$ happens exactly when $\frac{z^{p+1}}{(z+1)^p}=1+\epsilon$ and by some monotonicity property of the derivative of $\rho(z)$, one can show that the minimum value of $\rho(z)$ is achieved for the z described in the theorem.

5 Non-atomic PoS

In this section we give tight upper and lower bounds for the non-atomic congestion games with polynomial latencies. We start with a theorem which characterizes the ϵ-Nash (or ϵ-Wardrop) equilibria for every non-atomic congestion game with arbitrary (not necessarily polynomial) latency functions. It involves a potential-like function and generalizes a well-known characterization of exact Nash (Wardrop) equilibria [27].

Theorem 4. *In any non-atomic congestion game with latency functions $l_e(f_e)$, let $\phi_e(f_e)$ be any integrable functions which satisfy*

$$\frac{l_e(f_e)}{(1+\epsilon)} \leq \phi_e(f_e) \leq l_e(f_e), \tag{5}$$

for every $f_e \geq 0$ and define $\Phi_e(f_e) = \int_0^{f_e} \phi_e(t)\,dt$. For a flow f, define $\Phi(f) = \sum_{e\in E} \Phi_e(f_e)$. If a flow f minimizes the potential function $\Phi(f)$, it is an ϵ-Nash equilibrium.

Furthermore, when the latency functions are nondecreasing, for any other flow f':

$$\sum_{e\in E} \phi_e(f_e) f_e \leq \sum_{e\in E} \phi_e(f_e) f'_e$$

Proof. Consider a flow f and two paths p and p' with common endpoints. Suppose that the flow f on path p is positive. We want to compute how much $\Phi(f)$ changes when we shift a small amount $\delta > 0$ of flow from path p to path p'. More precisely, if f' denotes the new flow, we compute the following limit

$$\lim_{\delta\to 0} \frac{\Phi(f') - \Phi(f)}{\delta} = \sum_{e\in p'} \phi_e(f_e) - \sum_{e\in p} \phi_e(f_e) \tag{6}$$

When f minimizes Φ, then the above quantity is nonnegative. But then we can bound the cost of paths p and p' with the two terms of this quantity as follows:

$$l_p(f) = \sum_{e\in p} l_e(f_e) \leq (1+\epsilon) \sum_{e\in p} \phi_e(f_e) \leq (1+\epsilon) \sum_{e\in p'} \phi_e(f_e) \leq (1+\epsilon) \sum_{e\in p'} l_e(f_e) = (1+\epsilon) l_{p'}(f),$$

which implies that f is an ϵ-Nash equilibrium.

For the second part, we observe that the expression (6), which is nonnegative when f minimizes Φ, implies that for every path p on which f is positive and every other path p' we must have

$$\sum_{e\in p} \phi_e(f_e) \leq \sum_{e\in p'} \phi_e(f_e).$$

Consider now another flow f' which satisfies the rate constraints for all commodities and let us sum the above inequalities for all paths p and p' of commodity i, weighted with the amount of flow in f and f'. More precisely, using the notation of the proof of Theorem 1:

$$\sum_{p,p'} f_p^i {f_{p'}^i}' \sum_{e\in p} \phi_e(f_e) \leq \sum_{p,p'} f_p^i {f_{p'}^i}' \sum_{e\in p'} \phi_e(f_e)$$
$$\sum_{p'} {f_{p'}^i}' \sum_{e\in E} \phi_e(f_e) f_e^i \leq \sum_{p} f_p^i \sum_{e\in E} \phi_e(f_e) {f_e^i}'.$$

But $\sum_{p'} {f_{p'}^i}' = \sum_p f_p^i$ is equal to the sum of the rates. If we remove from the expression this common factor, and sum up for all commodities i, the second part of the theorem follows. □ □

We now study the PoS of polynomial latency functions of the form $\ell_e(f) = \sum_{k=0}^p a_{e,k} f_e^k$. For these latency functions we define the potential function which has derivatives $\phi_e(f_e) = \sum_{k=0}^p \zeta_k a_{e,k} f_e^k$, for some ζ_k to be determined later and which satisfy:

$$\frac{1}{1+\epsilon} \leq \zeta_k \leq 1 \tag{7}$$

Theorem 5. *The PoS for non-atomic congestion games with polynomial latencies of degree p is exactly*

$$\left((1+\epsilon)\left(1 - \frac{p}{p+1}\left(\frac{1+\epsilon}{p+1}\right)^{1/p}\right)\right)^{-1},$$

for $\epsilon < p$, and it drops to 1 for $\epsilon \geq p$.

Proof. Our starting point is the second part of Theorem 4. Let f be a flow which minimizes the potential and let $\hat{f}$ be an optimal flow. We have that

$$\sum_{e\in E}\sum_{k=0}^{p} \zeta_k a_{e,k} f_e^{k+1} \le \sum_{e\in E}\sum_{k=0}^{p} \zeta_k a_{e,k} f_e^k \hat{f}_e.$$

We can bound from above the monomials $f_e^k \hat{f}_e$ as follows:

$$f_e^k \hat{f}_e \le \alpha_k f_e^{k+1} + \beta_k \hat{f}_e^{k+1},$$

where[3]

$$\alpha_k^k \beta_k = k^k/(k+1)^{k+1} \tag{8}$$

(and in particular $\alpha_0 = 0$ and $\beta_0 = 1$). We rearrange the terms to get:

$$\sum_{e\in E}\sum_{k=0}^{p} \zeta_k (1-\alpha_k) a_{e,k} f_e^{k+1} \le \sum_{e\in E}\sum_{k=0}^{p} \zeta_k \beta_k a_{e,k} \hat{f}_e^{k+1}.$$

The problem now is to select the parameters α_k, β_k, ζ_k to bound all the coefficients on the left side from below and all the coefficients on the right side from above:

$$\zeta_k\,(1-\alpha_k) \ge 1-\alpha_p \qquad\qquad \zeta_k\,\beta_k \le \beta_p \tag{9}$$

With these, the above inequality becomes

$$(1-\alpha_p)\sum_{e\in E}\sum_{k=0}^{p} a_{e,k} f_e^{k+1} \le \beta_p \sum_{e\in E}\sum_{k=0}^{p} a_{e,k} \hat{f}_e^{k+1},$$

which will immediately bound the PoS from above by $\beta_p/(1-\alpha_p)$.

There is a complication in selecting the parameters α_k, β_k, ζ_k as functions of ϵ, in that we need to have different choices for small and large values of ϵ:

For $k \le \epsilon$:	$\zeta_k = (k+1)/(1+\epsilon)$	$\beta_k = 1/(k+1)$
For $k > \epsilon$:	$\zeta_k = 1$	$\beta_k = 1/(1+\epsilon)$

The value of α_k is determined by Equation (8)[4]. We need to show that these values satisfy properties (7) and (9). We see immediately that property (7) is satisfied. For property (9), we need to do more work. We first observe that the second part of property (9) is satisfied always with equality. For the first part we get:

Case $\epsilon < k$: We first establish that $\alpha_k \le k/(k+1)$. Indeed, from $\alpha_k^k\beta_k = k^k/(k+1)^{k+1}$, we get $\alpha_k^k = (1+\epsilon)\,k^k/(k+1)^{k+1} \le (1+k)\,k^k/(k+1)^{k+1} = k^k/(k+1)^k$. We will also use the fact that the function $x^p/(x+1)^{p+1}$ is increasing in x when $x \le p$. We then have $\alpha_p^p = (1+\epsilon)\,p^p/(p+1)^{p+1} \ge (1+\epsilon)k^p/(k+1)^{p+1} = (1+\epsilon)k^k/(k+1)^{k+1}k^{p-k}/(k+1)^{p-k} \ge \alpha_k^k\alpha_k^{p-k} = \alpha_k^p$. Therefore, $\alpha_p \ge \alpha_k$, from which the first part of property (9) follows immediately.

Case $\epsilon \ge k$: In this case $\alpha_k = k/(k+1)$ and we need to show that $\zeta_k(1-\alpha_k) \ge 1-\alpha_p$ which is equivalent to $\alpha_p \ge \epsilon/(1+\epsilon)$. Since $\alpha_p^p = (1+\epsilon)\,p^p/(p+1)^{p+1}$, we use again the fact that the function $x^p/(x+1)^{p+1}$ is increasing in x when $x \le p$ to get that $\alpha_p^p \ge (1+\epsilon)\,\epsilon^p/(1+\epsilon)^{1+p} \ge \epsilon^p/(1+\epsilon)^p$. The claim follows. Notice also that this analysis holds also for the special case $k = 0$.

Therefore, the PoS is at most $\beta_p/(1-\alpha_p)$, where $\beta_p = 1/(1+\epsilon)$ and $\alpha_p^p\,\beta = p^p/(p+1)^{p+1}$. This establishes the upper bound of the theorem.

To establish the lower bound, we use the Pigou network with latencies $1+\epsilon$ and f_e^p and flow rate equal to 1. There is a unique ϵ-Nash equilibrium which uses the second edge with total latency 1. On the other hand, the optimal solution is to send flow x to the second edge and $1-x$ in the first edge and optimize for x. The social cost is $x^{p+1} + (1+\epsilon)(1-x)$. If we minimize this function, we match the upper bound. □ □

The fact that the PoS drops to 1 when $\epsilon = p$ was first shown in [30] and [14].

[3] We use the following fact: for every nonnegative x, y, k, and for every positive α_k, β_k which satisfy $\alpha_k^k\beta_k = k^k/(k+1)^{k+1}$: $x^k y \le \alpha_k x^{k+1} + \beta_k y^{k+1}$. Proof: Let $z = x/y$. We need to have $\alpha_k z^{k+1} - z^k + \beta_k \ge 0$. Taking the derivative, we see that the expression is minimized when $z = k/((k+1)\alpha_k)$ and the minimum value is $\beta_k - k^k/((k+1)^{k+1}\,\alpha_k^k)$, which is 0 for our choice of α_k and β_k.

[4] We can unify the above definitions by letting $q = \min\{k, \epsilon\}$ and

$$\zeta_k = (q+1)/(1+\epsilon) \qquad\qquad \beta_k = 1/(q+1),$$

but this is not very useful in the following analysis.

6 Atomic PoS

The PoS for atomic congestion games is by far the most difficult quantity that we treat in this work. Even the case of exact Nash equilibria is still open for polynomial latency functions. And the situation becomes even more complicated when we consider approximate equilibria. There are two sources of difficulty: one is to generalize appropriately the potential function which is essential for proving upper bounds on the PoS. The second difficulty, is that the technical depth of handling the situation increases considerably.

In general, it seems that the PoS is in many cases much more challenging than the PoA. For example, there are still open problems from the first paper to study the PoS [3] concerning some important cases of network design games (for example undirected networks [18, 2]).

We have a characterization theorem which generalizes the notion of potential function and parallels Theorem 4 for non-atomic games.

Theorem 6. *In an atomic congestion game with latency functions* $l_e(k)$*, let* $\phi_e(k)$ *be any functions which satisfy*

$$\frac{l_e(k)}{(1+\epsilon)} \leq \phi_e(k) \leq l_e(k), \tag{10}$$

for every $k \geq 1$ *and define* $\Phi_e(k) = \sum_{t=1}^{k} \phi_e(t)$*. For a strategy profile* A*, define* $\Phi(A) = \sum_{e \in E} \Phi_e(n_e(A))$*. If a strategy profile* A *is a local minimum of* Φ*, (i.e.,* $\Phi(A) \leq \Phi(P_i, A_{-i})$ *for every* i *and* P_i*), then* A *is an* ϵ*-Nash equilibrium.*

Furthermore, when the latency functions are nondecreasing, for any other strategy profile P*:*

$$\sum_{e \in E} \phi_e(n_e(A))\, n_e(A) \leq \sum_{e \in E} \phi_e(n_e(A)+1)\, n_e(P). \tag{11}$$

Proof. To show that A is an ϵ-Nash equilibrium, we need to show $c_i(A) \leq (1+\epsilon)\, c_i(P_i, A_{-i})$, for every i and P_i. We have $c_i(A) = \sum_{e \in A_i} l_e(n_e(A))$ and $c_i(P_i, A_{-i}) = \sum_{e \in P_i} l_e(n_e(P_i, A_{-i}))$. Since A is a local minimum, we have $\Phi(A) \leq \Phi(P_i, A_{-i})$ which is equivalent to

$$\sum_{e \in A_i} \phi_e(n_e(A)) \leq \sum_{e \in P_i} \phi_e(n_e(P_i, A_{-i})). \tag{12}$$

We use this to compare $c_i(A)$ and $c_i(P_i, A_{-i})$:

$$c_i(A) = \sum_{e \in A_i} l_e(n_e(A)) \leq (1+\epsilon) \sum_{e \in A_i} \phi_e(n_e(A)) \leq (1+\epsilon) \sum_{e \in P_i} \phi_e(n_e(P_i, A_{-i})) \leq (1+\epsilon)\, c_i(P_i, A_{-i}).$$

This shows that A is ϵ-Nash equilibrium. To show the useful inequality (11) of the theorem when l_e is nondecreasing, we use inequality (12). This inequality implies

$$\sum_{e \in A_i} \phi_e(n_e(A)) \leq \sum_{e \in P_i} \phi_e(n_e(A)+1),$$

because $n_e(P_i, A_{-i}) \leq n_e(A) + 1$. If we sum these inequalities for every i, we get the desired inequality (11). □ □

The above theorem gives us a way to bound the PoS as follows:

We first define an appropriate ϕ_e. We then combine two inequalities to bound the cost of an approximate Nash equilibrium A over the cost of some other strategy profile P: the first is Inequality (11) and the second is the inequality which expresses the fact that A is a *global* minimum: $\Phi(A) \leq \Phi(P)$, that is,

$$\sum_{e \in E} \Phi_e(n_e(A)) \leq \sum_{e \in E} \Phi_e(n_e(P)), \tag{13}$$

where

$$\Phi_e(k) = \sum_{t=1}^{k} \phi_e(t). \tag{14}$$

By multiplying Inequality (11) by a nonnegative constant δ and (13) by a nonnegative constant γ and rearranging the terms to get

$$0 \leq \sum_{e \in E} \left(\delta \left(\phi_e(n_e(A)+1)\, n_e(P) - \phi_e(n_e(A))\, n_e(A)\right) + \gamma \left(\Phi_e(n_e(P)) - \Phi_e(n_e(A))\right)\right).$$

The aim is to find appropriate γ and δ and as small as possible ρ that satisfy

$$\delta\left(\phi_e(n_e(A)+1)\,n_e(P) - \phi_e(n_e(A))\,n_e(A)\right) + \gamma\left(\Phi_e(n_e(P)) - \Phi_e(n_e(A))\right) \leq \rho\, l_e(n_e(P))\,n_e(P) - l_e(n_e(A))\,n_e(A), \quad (15)$$

for every nonnegative integers $n_e(A)$ and $n_e(P)$. This would show that the PoS is at most ρ, since $l_e(n_e(A))\,n_e(A)$ and $l_e(n_e(P))\,n_e(P)$ are the social cost of the A and P (for edge e).

The optimal selection for ϕ_e, δ and γ remains an open problem. Here, we consider only linear latencies. We need to consider values of ϵ between 0 and 1, as later in Theorem 8 we show that for larger values of ϵ the approximate PoS is equal to 1.

Theorem 7. *For linear latencies ($p = 1$), the PoS is at most*

$$\frac{1+\sqrt{3}}{\epsilon+\sqrt{3}},$$

for $0 \leq \epsilon \leq 1$.

Proof. We select

$$\begin{aligned}\phi_e(k) &= l_e(k) - \frac{\epsilon}{1+\epsilon} l_e(1)\\ \delta &= \rho - 1\\ \gamma &= \epsilon\rho + 1\end{aligned}$$

Observe that ϕ_e satisfies condition (10) and that both δ and γ are nonnegative.

The general bound follows from the above reasoning by solving (15) for ρ after plugging in the values for ϕ_e, γ and δ. To simplify the expressions, we can assume that $l_e(k) = k$; it is straightforward to check that the proof works also for latencies of the form $l_e(k) = c_1 k + c_0$, for nonnegative constants c_1 and c_0.

Since $\phi_e(k) = k - \epsilon/(1+\epsilon)$, we have that $\Phi_e(k) = k(k+1)/2 - k\,\epsilon/(1+\epsilon)$. To keep the length of the expressions short, let us define $a = n_e(A)$ and $b = n_e(B)$. By solving (15) we get

$$\rho \geq \frac{b^2 - 2ba + 3a^2 - b - a}{(2-\epsilon)b^2 - 2ba + (2+\epsilon)a^2 - (2-\epsilon)b - \epsilon a}.$$

An important technical point is that the denominator of the above expression is non-negative for every nonnegative integers a and b and $\epsilon \in [0,1]$ (otherwise, we could not solve the inequality for ρ): Since the denominator is a linear expression on ϵ, it suffices to show the claim only for $\epsilon = 0$ and $\epsilon = 1$. Indeed, for $\epsilon = 0$ and $\epsilon = 1$ the denominator is equal to $2b^2 - 2ba + 2a^2 - 2b = (b-a)^2 + (b-1)^2 + a^2 - 1$ and $b^2 - 2ba + 3a^2 - b - a = \left(b - a - \frac{1}{2}\right)^2 + 2\left(a - \frac{1}{2}\right)^2 - \frac{3}{4}$, respectively; both are nonnegative for integer a and b.

It follows from the above bound on ρ that an upper bound on the PoS is

$$\rho = \max_{a,b\in N} \frac{b^2 - 2ba + 3a^2 - b - a}{(2-\epsilon)b^2 - 2ba + (2+\epsilon)a^2 - (2-\epsilon)b - \epsilon a}.$$

We can rewrite this as $\rho = \max_{a,b\in N} \frac{A+B}{A\epsilon+B}$, where $A = a^2 - b^2 - a + b$ and $B = 2(a^2 - ab + b^2 - b)$. To show the bound, it suffices to show that $A/B \leq 1/\sqrt{3}$, for all nonnegative integers a and b; taking into account that B is nonnegative, this is equivalent to $B - A\sqrt{3} \geq 0$. Remarkably, this does not involve the parameter ϵ at all. To show the inequality, we find that

$$B - A\sqrt{3} = \frac{1}{2}\left(\sqrt{3}+2\right)\left(\left(b + \sqrt{3}a - 2\,a - \frac{1}{2}\right)^2 + (3\,\sqrt{3} - 5)\,a - \frac{1}{4}\right).$$

This is clearly positive for $a \geq 2$ because $(3\,\sqrt{3} - 5)\,2 > 1/4$. For $a = 0$, we have

$$B - A\sqrt{3} = \frac{1}{2}\left(\sqrt{3}+2\right) b\,(b-1)\,,$$

which is clearly nonnegative for integer b. Finally, for $a = 1$ we have

$$B - A\sqrt{3} = \frac{1}{2}\left(\sqrt{3}+2\right)(b-1)\left(b - 4 + 2\,\sqrt{3}\right),$$

which is also clearly nonnegative. □ □

We don't have a similar bound for general polynomials. We can however apply the proof technique to establish that the PoS drops to 1 exactly when ϵ becomes equal to the degree of the polynomial, i.e., when $\epsilon = p$.

Theorem 8. *The PoS for atomic congestion games with polynomial latencies of degree p is 1 when $\epsilon = p$. In other words, the strategy with optimal social cost is a p-Nash equilibrium. This is tight: for every $q < p$, there are congestion games for which the optimum is not a q-Nash equilibrium.*

For this case, we select

$$\begin{aligned}\phi_e(k) &= \frac{l_e(k)k}{p+1} - \frac{l_e(k-1)(k-1)}{p+1} \\ \delta &= 0 \\ \gamma &= \epsilon + 1\end{aligned}$$

With this selection of ϕ_e, we have that $\Phi_e(k) = \frac{l_e(k)k}{p+1}$. We now simply observe that for $\rho = 1$, the inequality(15) holds with equality for all $n_e(A)$ and $n_e(P)$.

We need to verify that ϕ_e satisfies condition (10), or equivalently that for every polynomial $l(k)$ of degree p with nonnegative coefficients we have that

$$\frac{l(k)}{p+1} \leq \frac{l(k)k}{p+1} - \frac{l(k-1)(k-1)}{p+1} \leq l(k).$$

The left inequality is straightforward, since it is equivalent to $l(k-1) \leq l(k)$. The right inequality is more involved and follows form the following technical lemma:

Lemma 7. *For every polynomial $Q(x)$ of degree p with nonnegative coefficients, $Q(x)x - Q(x+1)(x-p)$ is nonnegative for every $x \geq 0$.*

Proof. This clearly holds for $x \leq p$, so we can assume that $x > p$. We want to show that $\frac{Q(x+1)}{Q(x)} \leq \frac{x}{x-p}$. By a straightforward monotonicity argument, we get that $\frac{Q(x+1)}{Q(x)} \leq \frac{(x+1)^p}{x^p}$. It suffices to show that $\frac{(x+1)^p}{x^p} \leq \frac{x}{x-p}$ or equivalently $(x+1)^p(x-p) \leq x^{p+1}$, which follows easily from the Arithmetic Mean-Geometric Mean (AM-GM) Inequality (i. e., $x_1 \cdots x_n \leq (\frac{x_1+\cdots+x_n}{n})^n$). □ □

The fact that the PoS for $\epsilon < p$ is higher than 1 follows from Theorem 3 and the following general proposition.

Proposition 1. *The approximate PoS for the class of atomic congestion games with polynomial latencies is greater than or equal to the approximate PoS of non-atomic congestion games.*

To see that the proposition holds, consider a non-atomic game with latencies $l_e(f_e)$. We can turn it into an atomic game with n players, where n is very high, and latencies $l'_e(k) = l_e(k/n)$. Any ϵ-Nash equilibrium for the non-atomic game is an ϵ-Nash equilibrium for the atomic game. This in general is not sufficient to establish the proposition because the opposite may not hold. It suffices however to apply it to the Pigou construction at the end of the proof of Theorem 5, which gives the optimal value for the PoS of non-atomic games. The special structure of the Pigou construction guarantees that the resulting atomic game has trivially a unique ϵ-Nash equilibrium.

The discussion before Theorem 7 is the rough outline of all upper-bound proofs in this work. For the PoS of non-atomic congestion games, (Theorem 5), the situation is much simpler for two reasons: We don't need to use the global minimum inequality (i.e., we take $\gamma = 0$), and the parameters $n_e(A)$ and $n_e(P)$ are nonnegative reals, not integers; the hardest part is to find appropriate generalization of the potential. For the PoA of non-atomic (Theorem 2) and atomic (Theorem 3) games, the situation is substantially simpler, because we don't need a potential (or its generalization) and we have only one inequality. Among all 4 bounds, the atomic PoS is the only one which involves the combination of two types of inequalities (local and global constraints for the potential).

6.1 Lower bound for linear latencies

Theorem 5 and Proposition 1 show that the Pigou game gives a lower bound for non-atomic and atomic congestion games. While for the class of the non-atomic games this is tight, it is not tight for the class of atomic games. We construct here a game, which is much more complicated, and has PoS substantially higher than the Pigou game for linear latencies. The lower bound of this construction still does not match the upper bound of Theorem 7, except for the extreme values of $\epsilon = 0$ and $\epsilon = 1$. For $\epsilon = 0$, we get that the PoS is $1 + 1/\sqrt{3} \approx 1.57$, which matches the PoS of exact equilibria [10, 6]. It then decreases as a function of ϵ, and drops to 1 for $\epsilon = 1$.

Theorem 9 (Atomic-PoS-Lower-Bound-Linear)**.** *There are linear congestion games whose ϵ-Nash equilibria have PoS approaching*

$$2\,\frac{3+\epsilon+3\,\epsilon^3+2\,\epsilon^4+\left(1+\epsilon+\epsilon^2\right)\sqrt{2\,\epsilon^4+3\,\epsilon^3+\epsilon+3}}{6+2\,\epsilon+6\,\epsilon^3+4\,\epsilon^4+\epsilon\,(5+2\,\epsilon-\epsilon^2)\sqrt{2\,\epsilon^4+3\,\epsilon^3+\epsilon+3}}.$$

For small ϵ, this is approximately $1.5774 - 1.7955\epsilon$*.*

Proof. We construct a game, where the ratio of its unique *(and therefore best)* ϵ-Nash equilibrium, over the optimum profile, is high. There are $n = n_1 + n_2$ players divided into two sets G_1 and G_2, with $|G_1| = n_1$ and $|G_2| = n_2$. Each player i in G_1 has two strategies A_i and P_i. The players in G_2 have a unique strategy D. The strategy profile $A = (A_1, \ldots, A_{n_1}, D, \ldots, D)$ will be the unique ϵ-Nash equilibrium and $P = (P_1, \ldots, P_{n_1}, D, \ldots, D)$ will be the profile with optimal social cost.

There are 3 types of facilities:

- n_1 facilities α_i, $i = 1, \ldots, n_1$, each with cost function $l_{\alpha_i}(x) = \alpha x$. Facility α_i belongs only to strategy P_i.
- $n(n-1)$ facilities β_{ij}, $i, j = 1, \ldots, n_1$ with $i \neq j$, each with cost $l_{\beta_{ij}}(x) = \beta x$. Facility β_{ij} belongs only to strategies A_i and P_j.
- 1 facility γ with cost $l_\gamma(x) = x$. Facility γ belongs to A_i for $i = 1, \ldots, n_1$ and to D.

We first compute the cost of each player in G_1, for every possible strategy profile with k players playing strategies A_i and $n_1 - k$ players playing strategies P_i. By symmetry, we need only to consider the cost $cost_A(k)$ of player 1 and the cost $cost_P(k)$ of player n_1 of the strategy profile $S_k = (A_1, \ldots, A_k, P_{k+1}, \ldots, P_{n_1}, D, \ldots, D)$. Therefore,

$$cost_A(k) = (2n_1 - k - 1)\beta + (n_2 + k).$$

Similarly, we compute

$$cost_P(k) = \alpha + (n_1 + k - 1)\beta.$$

We now want to select the parameters α and β so that the strategy profile $A = S_{n_1}$ is the unique ϵ-Nash equilibrium of the game. In particular, we require that at every strategy profile S_k, player i, $i = 1, \ldots, k$, has no reason to switch to strategy P_i, because it is $(1+\epsilon)$ times more costly. This is expressed by the constraints

$$(1+\epsilon)\cdot cost_A(k) \leq cost_P(k-1), \quad \text{for every } k = 1, \ldots, n_1. \tag{16}$$

All these constraints are linear in k and they are satisfied by equality when

$$\alpha = (n_1 - 1)\frac{(1+\epsilon)(1+2\epsilon)}{2+\epsilon} + n_2 + 1 \qquad \beta = \frac{1+\epsilon}{2+\epsilon}$$

In summary, for the above values of the parameters α and β, we obtain the desired property that the strategy profile A is a ϵ-dominant equilibrium (in the strong sense where each player has a unique optimal strategy which is a factor ϵ factor better than any other). If we increase α by any small positive δ, inequality (16) becomes strict and the ϵ-dominant equilibrium is unique (and therefore unique ϵ-Nash equilibrium).

We now want to select the value of the parameter so that the ratio of the cost of A over the optimum is as high as possible. The PoS is

$$\frac{n_1 cost_A(n_1) + n_2(n_1 + n_2)}{n_1 cost_P(0) + n_2^2},$$

which gives the bound of the theorem when we take the limit for n_1 and n_2 and optimize their ratio. □ □

7 Conclusions

We considered the PoA and the PoS of approximate Nash equilibria for congestion games (atomic, selfish routing, and non-atomic) with polynomial latencies. We have used a unifying approach and obtained tight upper and lower bounds for all cases except for the PoS for atomic congestion games. This remains a challenging open problem even for the case of exact equilibria ($\epsilon = 0$).

Acknowledgments: The authors would like to thank Ioannis Caragiannis for many helpful discussions and Tim Roughgarden for useful pointers to literature. We would also like to thank an anonymous reviewer for helping us improving the presentation of the paper.

References

[1] S. Aland, D. Dumrauf, M. Gairing, B. Monien, and F. Schoppmann. Exact price of anarchy for polynomial congestion games. In *STACS*, pp 218–229, 2006.

[2] S. Albers. On the value of coordination in network design. In *SODA*, pp 294–303, 2008.

[3] E. Anshelevich, A. Dasgupta, J. M. Kleinberg, É. Tardos, T. Wexler, and T. Roughgarden. The price of stability for network design with fair cost allocation. In *FOCS*, pp 295–304, 2004.

[4] B. Awerbuch, Y. Azar, and A. Epstein. The price of routing unsplittable flow. In *STOC*, pp 57–66, 2005.

[5] M. Beckmann, C. B. McGuire, and C. B. Winston. *Studies in the Economics of Transportation.* Yale University Press, 1956.

[6] I. Caragiannis, M. Flammini, C. Kaklamanis, P. Kanellopoulos, and L. Moscardelli. Tight bounds for selfish and greedy load balancing. In *ICALP*, pp 311–322, 2006.

[7] X. Chen and X. Deng. Settling the complexity of two-player nash equilibrium. In *FOCS*, pp 261–272, 2006.

[8] Ho-Lin Chen and Tim Roughgarden. Network design with weighted players. In *SPAA*, pp 29–38, 2006.

[9] S. Chien and A. Sinclair. Convergence to approximate nash equilibria in congestion games. In *SODA*, pp 169–178, 2007.

[10] G. Christodoulou and E. Koutsoupias. On the price of anarchy and stability of correlated equilibria of linear congestion games. In *ESA*, pp 59–70, 2005.

[11] G. Christodoulou and E. Koutsoupias. The price of anarchy of finite congestion games. In *STOC*, pp 67–73, 2005.

[12] G. Christodoulou, E. Koutsoupias, and P. G. Spirakis. On the performance of approximate equilibria in congestion games. *CoRR*, abs/0804.3160, 2008.

[13] J. R. Correa, A. S. Schulz, and N. E. Stier Moses. Selfish routing in capacitated networks. *Math. Oper. Res.*, 29(4):961–976, 2004.

[14] José R. Correa, Andreas S. Schulz, and Nicolás E. Stier Moses. Fast, fair, and efficient flows in networks. *Operations Research*, 55:215–225, 2007.

[15] S. C. Dafermos and F. T. Sparrow. The traffic assignment problem for a general network. *Journal of Research of the National Bureau of Standards, Series B*, 73B(2):91–118, 1969.

[16] C. Daskalakis, P. W. Goldberg, and C. H. Papadimitriou. The complexity of computing a nash equilibrium. In *STOC*, pp 71–78, 2006.

[17] C. Daskalakis, A. Mehta, and C. H. Papadimitriou. A note on approximate nash equilibria. In *WINE*, 2006.

[18] A. Fiat, H. Kaplan, M. Levy, S. Olonetsky, and R. Shabo. On the price of stability for designing undirected networks with fair cost allocations. In *ICALP (1)*, pp 608–618, 2006.

[19] A. Fabrikant, C. Papadimitriou, and K. Talwar. On the complexity of pure equilibria. In *STOC*, 2004.

[20] E. Koutsoupias and C. H. Papadimitriou. Worst-case equilibria. In *STACS*, pp 404–413, 1999.

[21] R. J. Lipton, E. Markakis, and A. Mehta. Playing large games using simple strategies. In *EC*, pp 36–41, 2003.

[22] I. Milchtaich. Congestion games with player-specific payoff functions. *Games and Economic Behavior*, 13:111–124, 1996.

[23] D. Monderer and L. Shapley. Potential games. *Games and Economics Behavior*, 14:124–143, 1996.

[24] N. Nisan, T. Roughgarden, E. Tardos, and V.V. Vazirani. *Algorithmic Game Theory.* Cambridge University Press, 2007.

[25] C. H. Papadimitriou. Algorithms, games, and the internet. In *STOC*, pp 749–753, 2001.

[26] R. W. Rosenthal. A class of games possessing pure-strategy Nash equilibria. *International Journal of Game Theory*, 2:65–67, 1973.

[27] T. Roughgarden and E. Tardos. How bad is selfish routing? *J. ACM*, 49(2):236–259, 2002.

[28] T. Roughgarden and E. Tardos. Bounding the inefficiency of equilibria in nonatomic congestion games. *Games and Economic Behavior*, 47(2):389–403, 2004.

[29] T. Roughgarden. *Selfish Routing.* PhD. Thesis, Cornell University, May 2002.

[30] T. Roughgarden. How unfair is optimal routing? In *SODA*, pp 203–204, 2002.

[31] T. Roughgarden. The price of anarchy is independent of the network topology. *J. Comput. Syst. Sci.*, 67(2):341–364, 2003.

[32] T. Roughgarden. *Selfish Routing and the Price of Anarchy.* The MIT Press, 2005.

[33] A. Skopalik and B. Vöcking. Inapproximability of pure nash equilibria. In *STOC*, pp 355–364, 2008.

[34] H. Tsaknakis and P. G. Spirakis. An optimization approach for approximate nash equilibria. In *WINE*, pp 47–56, 2007.

[35] J. G. Wardrop. Some theoretical aspects of road traffic research. In *Proceedings of the Institute of Civil Engineers, Part II*, volume 1, pp 325–378, 1952.