

# Atomic Routing Games on Maximum Congestion

Costas Busch  
Computer Science Department  
Rensselaer Polytechnic Institute  
buschc@cs.rpi.edu

Malik Magdon-Ismail  
Computer Science Department  
Rensselaer Polytechnic Institute  
magdon@cs.rpi.edu

## Abstract

We study atomic routing congestion games in which each player chooses a path in the network from its strategy set (a collection of paths) with the objective to minimize the *maximum congestion* along any edge on its selected path. The social cost is the global maximum congestion on any edge in the network. We show that for arbitrary routing games, the *price of stability* is 1, and the *price of anarchy*,  $PoA$ , is bounded by  $\kappa - 1 \leq PoA \leq c(\kappa^2 + \log^2 n)$ , where  $\kappa$  is the length of the longest cycle in the network,  $n$  is the size of the network and  $c$  is a constant. Further, any best response dynamic converges to a Nash equilibrium. Our bounds show that for maximum congestion games, the topology of the network, in particular the length of cycles, plays an important role in determining the quality of the Nash equilibria.

**Classification:** Algorithms; Current Challenges: game theory (routing congestion games).

# 1 Introduction

A fundamental issue in the management of large scale communication networks is to route the packet traffic so as to optimize the network performance. Our measure of network performance is the worst bottleneck (most used link) in the system. The model we use for network traffic is that of finite, unsplittable packets (atomic flow), and each packet's path is controlled independently by a selfish player. Ideally, when these players are left to their own devices, i.e., when they act selfishly in selecting their paths so as to minimize the worst bottleneck *from their point of view*, the resulting network performance should not be significantly worse than it would have been had the players coordinated their actions to optimize the network performance.

The Nash equilibrium (NE) is a natural outcome for a game with selfish players – a stable state in which no player can unilaterally improve her situation. In the recent literature, the *price of anarchy* [19, 29] and the *price of stability* [1, 2] have become prevalent measures of the quality of the equilibria of uncoordinated selfish behavior relative to coordinated optimal behavior. The former quantifies the worst possible outcome with selfish agents, and the latter measures the minimum penalty in performance required to ensure a stable equilibrium outcome.

We study atomic routing games (unsplittable traffic) [27, 32]. An atomic routing game is specified by  $N$  players corresponding to  $N$  source-destination node pairs on a network  $G$ . The strategy set available to each player is a set of edge-simple paths (typically all edge-simple paths in  $G$ ) from the player's source to the destination. A strategy of a player is a selection of one of the paths in the player's strategy set.

The *congestion* on an edge in the network is the number of paths that use this edge; the congestion of a path is the maximum congestion over all edges on the path; the congestion of the network is the maximum congestion over all edges in the network. We will use the network congestion as the measure of social cost. The congestion governs the delivery time of the packets in the network which makes it a natural measure of the social cost for the outcome of the routing game. In the literature this choice for the social cost is often referred to as the maximum social cost [7, 8, 19, 33]. Similarly, we choose for the player cost function the congestion on her path, i.e., the worst bottleneck along her chosen path, often referred to as the maximum player cost. The network congestion is an important metric for formally analyzing the performance of packet scheduling algorithms [9, 21, 22, 28, 30].

We only consider pure Nash equilibria. At the outset, we will show that there always exist optimal pure equilibria, therefore, the price of stability is  $PoS = 1$ . We then study the worst case equilibria, the price of anarchy ( $PoA$ ). We give a characterization of the price of anarchy (worst equilibria) in terms of the length of cycles in the network. Thus, the price of anarchy depends only on the network topology. In order to obtain this result we introduce a new technique called *edge-expansion* which bounds the social cost when path lengths are restricted. The edge-expansion technique may be of independent interest. For example, edge-expansion immediately gives bounds for the price of anarchy in networks where the path lengths are naturally short, such as the Hypercube and Butterfly-like networks.

## 1.1 Definitions and Statement of Results

An instance  $\mathcal{R}$  of a *routing (congestion) game* is a tuple  $(\mathbf{N}, G, \{\mathcal{P}_i\}_{i \in \mathbf{N}})$ , where  $\mathbf{N} = \{1, 2, \dots, N\}$  are the players,  $G = (V, E)$  is an undirected connected graph with  $|V| = n$ , and  $\mathcal{P}_i$  is a collection of *edge-simple* paths. Each path in  $\mathcal{P}_i$  is a path in  $G$  that has the same source  $s_i \in V$  and

destination  $t_i \in V$ ; each path in  $\mathcal{P}_i$  is a pure strategy available to player  $i$ . A pure strategy profile  $\mathbf{p} = [p_1, p_2, \dots, p_N]$  is a collection of pure strategies (paths), one for each player, where  $p_i \in \mathcal{P}_i$ . We refer to a pure strategy profile as a *routing*. On a finite network, a routing game is necessarily a finite game.

For any routing  $\mathbf{p}$  and any edge  $e \in E$ , the *edge-congestion*  $C_e(\mathbf{p})$  is the number of paths in  $\mathbf{p}$  that use edge  $e$ . For any path  $p$ , the *path-congestion*  $C_p(\mathbf{p})$  is the maximum edge congestion over all edges in  $p$ ,  $C_p(\mathbf{p}) = \max_{e \in p} C_e(\mathbf{p})$ . The *network congestion* is the maximum edge-congestion over all edges in  $E$ ,  $C(\mathbf{p}) = \max_{e \in E} C_e(\mathbf{p})$ . The *social or global cost*  $SC(\mathbf{p})$  is the network congestion,  $SC(\mathbf{p}) = C(\mathbf{p})$ . The *player or local cost*  $pc_i(\mathbf{p})$  for player  $i$  is her path-congestion,  $pc_i(\mathbf{p}) = C_{p_i}(\mathbf{p})$ . When the context is clear, we will drop the dependence on  $\mathbf{p}$  and use  $C_e, C_p, C, SC, pc_i$ .

We use the standard notation  $\mathbf{p}_{-i}$  to refer to the collection of paths  $\{p_1, \dots, p_{i-1}, p_{i+1}, \dots, p_N\}$ , and  $(p_i; \mathbf{p}_{-i})$  as an alternative notation for  $\mathbf{p}$  which emphasizes the dependence on  $p_i$ . Player  $i$  is *locally optimal* in routing  $\mathbf{p}$  if  $pc_i(\mathbf{p}) \leq pc_i(p'_i; \mathbf{p}_{-i})$  for all paths  $p'_i \in \mathcal{P}_i$ . A routing  $\mathbf{p}$  is in a Nash Equilibrium ( $\mathbf{p}$  is a *Nash-routing*) if every player is locally optimal. Nash-routings quantify the notion of a stable selfish outcome. A routing  $\mathbf{p}^*$  is an optimal pure strategy profile if it has minimum attainable social cost: for any other pure strategy profile  $\mathbf{p}$ ,  $SC(\mathbf{p}^*) \leq SC(\mathbf{p})$ .

We quantify the quality and diversity of the Nash-routings by the *price of stability* (*PoS*) and the *price of anarchy* (*PoA*) (sometimes referred to as the coordination ratio). Let  $\mathbf{P}$  denote the set of distinct Nash-routings, and let  $SC^*$  denote the social cost of an optimal routing  $\mathbf{p}^*$ . Then,

$$PoS = \inf_{\mathbf{p} \in \mathbf{P}} \frac{SC(\mathbf{p})}{SC^*}, \quad PoA = \sup_{\mathbf{p} \in \mathbf{P}} \frac{SC(\mathbf{p})}{SC^*}. \quad (1)$$

Given a routing  $\mathbf{p}$ , a *best response move* by player  $i$  is to change its path selection from  $p_i$  (if it is not locally optimal) to a path  $p'_i$  which is locally optimal given  $\mathbf{p}_{-i}$ , thus,  $pc_i(p'_i; \mathbf{p}_{-i}) < pc_i(p_i; \mathbf{p}_{-i})$ . A *best response dynamic* is an arbitrary sequence of best response moves. If there are no infinite best response dynamics, then every best response strategy converges. The (worst case) convergence time is the length of the longest best response dynamic. A best response dynamic is maximal if it cannot be extended by any best response move. Any maximal best response dynamic must end at a Nash-routing. Our first result is on the existence and quality of pure Nash-routings for arbitrary routing game instances.

**Theorem 1** *An instance  $\mathcal{R}$  of a routing game is specified by a tuple  $(\mathbf{N}, G, \{\mathcal{P}_i\}_{i \in \mathbf{N}})$ .*

(i) *Every instance of a routing game, has a pure Nash-routing which is optimal, hence  $PoS = 1$ .*

(ii) *For every routing game, there are no infinite best response dynamics.*

(iii) *For any routing game, the network congestion is non-increasing along a best response dynamic.*

Part (i) in Theorem 1 indicates the existence of an optimal Nash-routing. Part (ii) indicates that one way to find a Nash-routing is to set the initial paths of each player arbitrarily from which any greedy best response dynamic will converge to a Nash-routing. Part (iii) shows that if the initial paths are chosen well (with near optimal network congestion), then the Nash-routing which results after any best response dynamic will have near optimal congestion. There exist routing algorithms for finding such a good initial set of paths, for arbitrary networks [3, 4, 17, 31], as well as special classes of networks [5, 6, 25].

Our next result characterizes the worst case Nash-routing in terms of the topology of the network. In particular, we show that cycles in the network play an important role in determining whether or not there can exist bad Nash-routings. For a graph  $G$ , the *edge-cycle number*  $\kappa_e(G)$  is

the length of the longest edge simple cycle in  $G$  (similarly one could define the node-cycle number  $\kappa_n$ ). We will drop the dependence on  $G$  when the context is clear.

**Theorem 2** *For any undirected graph  $G$  with edge-cycle number  $\kappa_e$ ,*

- (i) *there exists a routing game  $(\mathbf{N}, G, \{\mathcal{P}_i\}_{i \in \mathbf{N}})$  for which  $PoA \geq \kappa_e - 1$ ;*
- (ii) *for any routing game  $(\mathbf{N}, G, \{\mathcal{P}_i\}_{i \in \mathbf{N}})$ ,  $PoA \leq c(\kappa_e^2 + \log^2 n)$ .*

(We will use  $c$  to represent a generic constant, whose specific value may vary even from line to line in the same set of equations.) Theorem 2 characterizes the Nash-routings in terms of the network topology, the key parameter being the edge-cycle number. Thus, for networks with small cycles, the price of anarchy is small. Let  $m$  denote the number of edges in the network. Since  $\kappa_e \leq m$ , we have that  $PoA \leq c \cdot m^2$ . In graphs with Euler cycles,  $\kappa_e = m$ , thus,  $m - 1 \leq PoA \leq c \cdot m^2$ .

The lower bound in part (i) of Theorem 2 is obtained by constructing a game instance where the players have their sources and destination in the largest cycle. A useful result needed to prove the upper bound in part (ii) of Theorem 2 is the *edge-expansion theorem*, which bounds the price of anarchy for arbitrary instances of routing games in terms of the maximum length of the paths in the strategy sets. Specifically, for a routing game  $(\mathbf{N}, G, \{\mathcal{P}_i\}_{i \in \mathbf{N}})$  in which the length of any path in a strategy set is at most  $\ell$ , the edge-expansion theorem states that  $SC(\mathbf{p}) < 2(\ell \cdot SC^* + \log n)$ , where  $\mathbf{p}$  is any Nash-routing, and  $SC^*$  is the social cost of an optimal pure strategy profile. The edge-expansion result implies that  $PoA < 2(\ell + \log n)$ .

To illustrate the main idea in the proof of Theorem 2 part (ii), first consider 2-connected graphs. Using Menger's theorem [38], we establish that  $\ell \leq c \cdot \kappa_e^2$ , from which Theorem 2 follows immediately by using the edge-expansion theorem. However, if the graph  $G$  is not 2-connected, then the relation  $\ell \leq c \cdot \kappa_e^2$  will generally not hold, and  $\ell$  could be much larger than  $c \cdot \kappa_e^2$ . In order to obtain the result, however, we will decompose an arbitrary graph  $G$  into a tree consisting of 2-connected components. We show that if in  $G$  the Nash-routing has network congestion  $C$ , then there is some 2-connected component  $H$  which has congestion  $C' \approx C$ . If the subpaths in  $H$  of the set of players contributing to this congestion  $C'$  were all locally optimal in  $H$ , then we would have a Nash-routing in  $H$ , and we could apply the 2-connected result. However, not all the players contributing to the congestion  $C'$  in  $H$  will be locally optimal in  $H$ , and so to convert this basic idea into a proof, we generalize the edge-expansion theorem so that it applies to situations where not all players are locally optimal (we refer to such situations as partial Nash-routings).

The edge-expansion result may also be of independent interest, as it can be used to give good bounds on the price of anarchy for networks where it is natural to use paths with short length. For example in the Hypercube and Butterfly [20], if we choose bit-fixing paths, then  $\ell = O(\log n)$ , which implies that  $PoA \leq c \log n$ .

## 1.2 Related Work

General congestion games were introduced and studied in [27, 32]. The application of game theory in computer science, specifically the introduction of the price of anarchy was introduced in [19]. Since then, there has been significant activity. Briefly, the models studied can be categorized by the topology of the network; the nature of the player and social costs; the nature of the traffic (atomic or splittable); the nature of the strategy sets; the nature of the equilibria studied (pure or mixed). A brief taxonomy of some relevant existing results, specifically in the context of congestion games and routing is given in Table 1, where part (a) gives a breakdown according to the kind of

	Atomic	Splittable
Pure	[7], [23], [32] [37]* <b>Our Work</b>	[33], [34], [35], [36]
Mixed	[10]*, [11]*, [12]*, [15]*, [14]*, [16]*, [18]*, [19]*, [24]*, [26]*, [29]*	[8], [13]*

(a)

	Maximum $SC$	Sum $SC$	Other $SC$	**
Max. $pc$	<b>Our Work</b>	–	–	[23]
Sum $pc$	[7], [8], [10]*, [11]*, [12]*, [13]*, [14]*, [18]*, [19]*, [26]*, [29]* [33]	[7], [16]*, [34], [35], [36], [37]*	[15]*, [24]*	[23], [32]

(b)

\* A specific network model is used (eg. parallel links) or specific player strategy sets (eg. singleton sets).

\*\* Results on existence or convergence to equilibrium, as opposed to quality of equilibria.

Table 1: Existing work in routing congestion games.

flow (atomic or splittable) and equilibria (mixed or pure), and part (b) according to the social and player cost (sum or maximum).

Typically, the research in the literature has focused on computing upper and lower bounds on the price of anarchy. The vast majority of the work on maximum social cost has been for parallel link networks, with only a few recent results on general topologies [7, 8, 33]. Essentially, all of the work has focused on the sum player cost, which corresponds to the sum of the edge congestions on a path (as opposed to the maximum edge congestion on the path, which we consider here).

Our results are for the maximum player cost, and the only result which has a brief discussion of the maximum player cost is [23] where the authors focus on parallel link networks, but also give some results for general topologies. In [23], the main content is to establish the existence of pure Nash-routings. We present a systematic study of pure Nash-routings in atomic routing games with maximum player cost. Pure equilibria with atomic players and maximum player cost introduces essentially combinatoric conditions for the equilibria, in contrast to infinitely splittable flow, or mixed equilibria, which can be characterised by Wardrop-type equilibrium conditions.

We note that our results assume that the network latency function is the identity. Our results can easily be extended to any monotonic latency function, as the combinatorial structure of the maximum player and social cost are invariant to any monotonic transformation of the latency function.

**Outline of Paper.** We start with the proof of Theorem 1 in Section 2. We then continue by presenting the edge expansion theorem in Section 3. In section 4, we give the lower bound in Theorem 2. In Section 5, we give the upper bound in Theorem 2 for 2-connected graphs, and in Section 6 we give our general result for arbitrary networks. We conclude in Section 7. Most of the technical proofs have been placed in a technical appendix.

## 2 Existence of Optimal Nash-routings

The goal in this section is to establish main Theorem 1. First, we establish the existence of pure Nash-routings and compute the price of stability and anarchy; we also show that every best response strategy converges, and the social cost is non-increasing along a best response strategy. Note that the strategy set for any player may be *any* collection of edge simple paths from her source to destination, typically all the possible paths.

For any routing  $\mathbf{p}$ , define the *congestion vector*  $\mathbf{C}(\mathbf{p}) = [m_0(\mathbf{p}), m_1(\mathbf{p}), m_2(\mathbf{p}), \dots]$ , where each component  $m_k(\mathbf{p})$  is the number of edges with congestion  $k$ . Note that  $\sum_k m_k(\mathbf{p}) = m$ , where  $m$  is the number of edges in the network. The social cost (network congestion)  $SC(\mathbf{p})$  is the maximum  $k$  for which  $m_k > 0$ . We define a lexicographic total order on routings as follows. Let  $\mathbf{p}$  and  $\mathbf{p}'$  be two routings, with  $\mathbf{C}(\mathbf{p}) = [m_0, m_1, m_2, \dots]$ , and  $\mathbf{C}(\mathbf{p}') = [m'_0, m'_1, m'_2, \dots]$ . Two routings are equal, written  $\mathbf{p} =_c \mathbf{p}'$ , if and only if  $m_k = m'_k$  for all  $k \geq 0$ ;  $\mathbf{p} <_c \mathbf{p}'$  if and only if there is some  $k^*$  such that  $m_{k^*} < m'_{k^*}$  and  $\forall k > k^*, m_k \leq m'_k$ .

Let  $(\mathbf{N}, G, \{\mathcal{P}_i\}_{i \in \mathbf{N}})$  be an instance of a routing game. Since there are only finitely many routings (as a player's path may use any edge at most once), there exists at least one minimum routing. There may be many distinct routings all of which are minimum (and equal to each other) in our total order. Let  $\mathbf{p}^*$  be a minimum routing in this total order (which we know exists); then, for all routings  $\mathbf{p}$ ,  $\mathbf{p}^* \leq_c \mathbf{p}$ . Any minimum routing is an optimal routing; indeed, if  $SC(\mathbf{p}) < SC(\mathbf{p}^*)$  for some other routing  $\mathbf{p}$ , then the maximum  $k$  for which  $m_k(\mathbf{p}) > 0$  is smaller than the corresponding  $k$  for  $\mathbf{p}^*$ , contradicting the fact that  $\mathbf{p}^* \leq_c \mathbf{p}$ .

**Lemma 2.1** *At least one minimum routing  $\mathbf{p}^*$  exists. All minimum routings are optimal routings.*

We now show that greedy moves can only decrease a routing according to  $<_c$ . A greedy move is available to player  $i$  if she can obtain a lower path congestion by changing her current path from  $p_i$  to a new path  $p'_i$ . We say that a greedy move takes the original routing  $(p_i; \mathbf{p}_{-i})$  to a new routing  $(p'_i; \mathbf{p}'_{-i})$  in which  $p_i$  is replaced by  $p'_i$ . A greedy move decreases the number of high congestion edges, by transferring the congestion to lower congestion edges. Thus, we can show:

**Lemma 2.2** *If a greedy move by any player takes  $\mathbf{p}$  to  $\mathbf{p}'$ , then  $\mathbf{p}' <_c \mathbf{p}$ .*

Since there is a finite number of routings, by Lemma 2.2, all best response dynamics are finite.

**Corollary 2.3** *Starting from an arbitrary initial routing, every best response dynamic is finite.*

A consequence of Lemma 2.2 is that every minimum routing is a Nash-routing. Indeed, suppose that, in a minimum routing  $\mathbf{p}^*$ , some player has an available greedy move which takes  $\mathbf{p}^*$  to  $\mathbf{p}'$ . Then, by Lemma 2.2,  $\mathbf{p}' <_c \mathbf{p}^*$  which contradicts the minimality of  $\mathbf{p}^*$ , therefore every player must be locally optimal in a minimum routing. Since at least one minimum routing exists and all minimum routings are optimal (Lemma 2.1), we have

**Lemma 2.4** *Every minimum routing  $\mathbf{p}^*$  (at least one exists) is an optimal Nash-routing.*

**Corollary 2.5** *The price of stability is  $PoS = 1$ .*

Our main Theorem 1 follows from Lemmas 2.2, 2.4 and Corollaries 2.3, 2.5.

### 3 The Edge-Expansion Theorem

The edge-expansion result will be essential to proving the the second part of Theorem 2, and is also of independent interested in path-constrained routing games where the paths have bounded length. The edge-expansion theorem can be used to bound the price of anarchy for an arbitrary routing game in terms of the maximum path length in the strategy set of the players. In fact, we will prove a more general result which bounds the quality of *partial* Nash-routings, in which only some of the players are locally optimal.

Let  $\mathcal{R} = (\mathbf{N}, G, \{\mathcal{P}_i\}_{i \in \mathbf{N}})$  be an instance of a routing game. Let  $\mathcal{P} = \bigcup_{i \in \mathbf{N}} \mathcal{P}_i$ . The *path-length* of  $\mathcal{R}$  is  $\ell = \max_{p \in \mathcal{P}} |p|$ . A *path-cut* for player  $i$  is a set of edges  $E_i$  such that every path in  $\mathcal{P}_i$  must use at least one of the edges in  $E_i$ . The congestion of a path-cut  $C(E_i)$  is the minimum congestion of any edge in  $E_i$ ,  $C(E_i) = \min_{e \in E_i} C_e$ . If player  $i$  is locally optimal with congestion  $pc_i$ , then every alternative path for that player must have congestion at least  $pc_i - 1$ .

**Lemma 3.1** *Consider an arbitrary routing  $\mathbf{p} = [p_1, p_2, \dots, p_N]$  for game  $\mathcal{R}$ . If player  $i$  is locally optimal, then there is a path-cut  $E_i$  for player  $i$  with congestion  $C(E_i) \geq pc_i - 1$ .*

### 3.1 Edge-Expansion Process

In general, an arbitrary routing  $\mathbf{p}$  with social cost (network congestion)  $C$  will be a partial Nash-routing with some players locally optimal and some not. If all players are locally optimal, then  $\mathbf{p}$  is a Nash-routing, but this is not a requirement for the following discussion.

In order to initiate the edge-expansion process, assume that at least one player is locally optimal and has player cost  $C$ . Let  $\mathcal{E}_0$  be the set of edges with congestion  $C_0 = C$  that are used by at least one locally optimal player. Let  $\Pi_0$  be set of locally optimal players that use at least one edge in  $\mathcal{E}_0$ . By Lemma 3.1, each player in  $\Pi_0$  has a path-cut with congestion at least  $C_0 - 1$ . Let  $\mathcal{E}_1$  denote the union of  $\mathcal{E}_0$  with all these path-cuts of every player in  $\Pi_0$ . Thus,  $\mathcal{E}_0 \subseteq \mathcal{E}_1$  and each edge in  $\mathcal{E}_1$  has congestion at least  $C_1 = C_0 - 1$ . Let  $\Pi_1$  denote the set of locally optimal players whose paths in  $\mathbf{p}$  use at least one edge in  $\mathcal{E}_1$ . Note that  $\Pi_0 \subseteq \Pi_1$ . Each player in  $\Pi_1$  has player cost at least  $C_1$ , since every edge in  $\mathcal{E}_1$  has congestion at least  $C_1$ .

We continue this process inductively as follows. Suppose that for  $i \geq 1$ , edge set  $\mathcal{E}_i$  has been constructed as the union of  $\mathcal{E}_{i-1}$  with path cuts for the players in  $\Pi_{i-1}$ , thus every edge in  $\mathcal{E}_i$  has congestion at least  $C_i = C_{i-1} - 1 = C - i$ . We now construct  $\Pi_i$ , the set of locally optimal players whose paths use at least one edge in  $\mathcal{E}_i$ ; every player in  $\Pi_i$  has player cost at least  $C_i$ . By Lemma 3.1, each player in  $\Pi_i$  has a path-cut with congestion  $C_i - 1$ , and we construct  $\mathcal{E}_{i+1}$  to be the union of  $\mathcal{E}_i$  with all these path-cuts of the players in  $\Pi_i$ . Using this inductive construction, we obtain a sequence of edge sets,  $\mathcal{E}_0 \subseteq \mathcal{E}_1 \subseteq \mathcal{E}_2, \dots$ , with  $C(\mathcal{E}_j) \geq C_j = C - j$ , and corresponding to each edge set, a set of locally optimal players  $\Pi_0 \subseteq \Pi_1 \subseteq \Pi_2 \dots$ . We continue this inductive construction up to edge set  $\mathcal{E}_s$  which is the first set for which  $|\mathcal{E}_s| \leq 2|\mathcal{E}_{s-1}|$ . We will refer to this inductive process as the *edge-expansion process*, since by construction  $\mathcal{E}_{i-1} \subseteq \mathcal{E}_i$ , for  $1 \leq i \leq s$ , and the edge set size doubles from step  $i - 1$  to step  $i$ , for  $i < s$ .

### 3.2 Edge-Expansion Properties

Since the edge sets double in size at each expansion step, and there are at most  $\frac{1}{2}n^2$  edges,

**Lemma 3.2**  $|\mathcal{E}_s| \geq 2^{s-1}$  and  $1 \leq s < 2 \log n$ .

In routing  $\mathbf{p}$ , let  $F(C') \subseteq \mathbf{N}$  denote the set of non-locally optimal players with player cost at least  $C'$ , i.e.  $pc_i \geq C' \forall i \in F(C')$ . We now give a bound for  $C$  in terms of  $C^*$ , establishing a relationship between the congestion of a routing and the optimal routing.

**Theorem 3.3 (Edge-expansion)**  $C < 2\ell \cdot (C^* + F(C - 2 \log n)) + 2 \log n$ .

**Proof:** From the edge-expansion process and Lemma 3.2, each edge in set  $\mathcal{E}_{s-1}$  has congestion at least  $C_{s-1}$ . Let  $M$  be the number of times edges in  $\mathcal{E}_{s-1}$  are used by the paths in  $\mathbf{p}$ . Then,  $M > C_{s-1} \cdot |\mathcal{E}_{s-1}|$ , since each edge has congestion at least  $C_{s-1}$ . By construction, in  $\mathbf{p}$ , the congestion in each of the edges of  $\mathcal{E}_{s-1}$  is caused only by the players in  $A = \Pi_{s-1} \cup B$ , where  $B \subseteq F(C_{s-1})$  contains the non locally optimal players that use edges in  $\mathcal{E}_{s-1}$ . Each packet in  $A$  uses paths of length at most  $\ell$ . Therefore, each player in  $A$  can use at most  $\ell$  edges in  $\mathcal{E}_{s-1}$ , hence,  $M \leq \ell \cdot |A|$ . Thus,  $\ell \cdot |A| > C_{s-1} \cdot |\mathcal{E}_{s-1}|$ . Since,  $|A| \leq |\Pi_{s-1}| + |F(C_{s-1})|$ , we obtain,

$$C_{s-1} < \frac{\ell}{|\mathcal{E}_{s-1}|} \cdot (|\Pi_{s-1}| + |F(C_{s-1})|).$$

We now give an upper bound on  $|\Pi_{s-1}|$ .  $\mathcal{E}_s$  contains a path-cut for every player in  $\Pi_{s-1}$ , every such player must use at least one edge in  $\mathcal{E}_s$  in any routing, including the optimal routing  $\mathbf{p}^*$ . Thus, edges in  $\mathcal{E}_s$  are used at least  $|\Pi_{s-1}|$  times, hence some edge is used at least  $|\Pi_{s-1}|/|\mathcal{E}_s|$  times, by the pigeonhole principle. Hence  $C^* \geq |\Pi_{s-1}|/|\mathcal{E}_s|$  (note that  $|\mathcal{E}_s| > 0$ ). By the definition of  $s$ ,  $|\mathcal{E}_s| \leq 2|\mathcal{E}_{s-1}|$ , hence  $|\Pi_{s-1}| \leq 2|\mathcal{E}_{s-1}|C^*$ . Using this upper bound for  $|\Pi_{s-1}|$ , we obtain

$$C_{s-1} < 2\ell \cdot \left( C^* + \frac{|F(C_{s-1})|}{2|\mathcal{E}_{s-1}|} \right).$$

Since  $C_{s-1} = C - (s - 1)$  and  $2|\mathcal{E}_{s-1}| \geq 2^s$  (Lemma 3.2), we obtain

$$C < 2\ell \cdot \left( C^* + \frac{|F(C - s + 1)|}{2^s} \right) + s - 1.$$

For  $C'' < C'$ ,  $F(C') \subseteq F(C'')$  which means that  $|F(C')|$  is non-increasing in  $C'$ . Thus  $|F(C - s + 1)| \leq |F(C - 2\log n)|$ . Since  $2^s \geq 2$ , the claim follows.  $\blacksquare$

Theorem 3.3 depends on the number of players which are not locally optimal in the partial Nash-routing. We can obtain a tighter result for Nash-routings. In a Nash-routing,  $F(C') = 0$ , for any  $C' > 0$ , so Theorem 3.3 gives

**Theorem 3.4 (Edge-expansion for Nash-routings)** *If  $\mathbf{p}$  is a Nash-routing, then  $C < 2\ell \cdot C^* + 2\log n$ , and  $PoA < 2(\ell + \log n)$ .*

## 4 Lower Bound on Price of Anarchy

In this section, we give a lower bound on the price of anarchy with respect to the cycle number of a graph, establishing part (i) of Theorem 2. We consider an arbitrary connected graph  $G = (V, E)$ . We will construct a routing game instance on the longest cycle. For each player, the source and destination are adjacent pairs in the cycle. In the optimal solution each player follows the edge connecting the source and destination, resulting to network congestion 1. However, there is a Nash-routing where each player uses the long path in the cycle, giving congestion proportional to the size of the cycle. The details are in the following result.

**Theorem 4.1 (PoA lower bound)** *Let  $G$  be a graph with edge cycle number  $\kappa_e$ . Then, there is an instance of a routing game  $(\mathbf{N}, G, \{\mathcal{P}_i\}_{i \in \mathbf{N}})$  on  $G$  with  $PoA = \kappa_e - 1$ .*



**Proof:** Let  $Q = e_1, \dots, e_{\kappa_e}$  be an edge simple cycle with length  $\kappa_e$ . We construct a routing game with  $\kappa_e$  players,  $\mathbf{N} = \{1, 2, \dots, \kappa_e\}$ . There is a player  $i$  corresponding to every edge  $e_i = (u_i, v_i)$  in  $Q$ . We set the source  $s_i = u_i$  and the destination  $t_i = v_i$ . The strategy set  $\mathcal{P}_i$  is the collection of edge simple paths from  $s_i$  to  $t_i$ . The forward path  $f_i$  is composed solely of the edge  $(u_i, v_i)$ . The backward path  $b_i$  is obtained by traversing the edges of cycle  $Q$  in reverse until edge  $e_i$  is reached ( $e_i$  is not used in the backward path).  $\{f_i, b_i\} \subseteq \mathcal{P}_i$ .

Since  $Q$  is edge simple, if every player uses his forward path (the forward routing,  $\mathbf{p} = [f_1, \dots, f_{\kappa_e}]$ ),  $C = 1$ . Thus, the optimal social cost is 1. If on the other hand, all the players use their backward paths (the backward routing  $\bar{\mathbf{p}} = [b_1, \dots, b_{\kappa_e}]$ ), then player  $i$  uses every edge in  $Q$  except  $e_i$  exactly once. Thus, the congestion on every edge in  $Q$  is  $N - 1 = \kappa_e - 1$ .

We will show that  $\bar{\mathbf{p}}$  is a Nash-routing. Suppose for contradiction that some player  $k$  is not locally optimal in  $\bar{\mathbf{p}}$ . So for some other path  $p$ , player  $k$  has lower path congestion,  $pc_k(p; \bar{\mathbf{p}}_{-k}) < pc_k(b_k; \bar{\mathbf{p}}_{-k}) = \kappa_e - 1$ . Since every edge on  $Q$  has congestion  $\kappa_e - 1$  in routing  $\bar{\mathbf{p}}$ , at least  $\kappa_e - 2$  players other than player  $k$  use every edge on  $Q$ . Thus, if  $p$  uses any edge on  $Q$ , then  $pc_k(p; \bar{\mathbf{p}}_{-k}) = \kappa_e - 1$ , which does not improve its cost, so we conclude that  $p$  does not use any edge on  $Q$ . Therefore,  $p$  has length at least 2 (since  $p \neq e_k$ ). Thus, replacing  $e_k \in Q$  by  $p$  results in a new edge simple cycle  $Q'$  that is strictly longer than  $Q$ , a contradiction. Thus,  $\bar{\mathbf{p}}$  is a Nash-routing. ■

## 5 Upper Bound on Price of Anarchy for 2-Connected Graphs

We will show that the price of anarchy is upper bounded by the square of the edge-cycle number. First, however, we prove this result for the special case of 2-connected graphs. The purpose of this section is to demonstrate how the edge-expansion theorem can be used to bound the price of anarchy using topological properties of the graph. This section will develop the intuition behind the result, before we consider the general result in Section 6.

A graph  $G$  is  $k$ -connected if the minimum edge-cut of  $G'$  has size at least  $k$ . By Menger's theorem [38],  $G'$  is  $k$ -connected if and only if for every pair of nodes  $u, v$  in  $G'$  there are at least  $k$  edge-disjoint  $(u, v)$ -paths.

For 2-connected subgraphs, using Menger's theorem, we can show that the maximum path length in the graph,  $\ell$ , is bounded by  $c \cdot \kappa_e^2$ . The proof relies on the observation that the longest path  $p$  must have at least  $\sqrt{\ell}$  edges in common with the largest cycle  $q$ , since otherwise, we would be able to construct a larger cycle by combing pieces of  $p$  and  $q$ . We have:

**Lemma 5.1** *If  $G$  is 2-connected, then  $\kappa_e(G) \geq \sqrt{2\ell} - \frac{3}{2}$ , where  $\ell$  is the longest path length in  $G$ .*

Lemma 5.1 bounds the longest path length in  $G$  with respect to the edge-cycle number in the graph. The edge-expansion Theorem 3.4 bounds the price of anarchy in terms of the longest path in the players' strategy sets. By combining these two results, we obtain:

**Theorem 5.2 (PoA upper bound for 2-connected graphs)** *For an arbitrary routing game  $(\mathbf{N}, G, \{\mathcal{P}_i\}_{i \in \mathbf{N}})$  on an arbitrary 2-connected graph  $G$ ,  $PoA \leq c(\kappa_e^2 + \log n)$ .*

## 6 Upper Bound on Price of Anarchy for Arbitrary Graphs

We now give the more general version of Theorem 5.2, establishing part (ii) of Theorem 2, which states that the price of anarchy of any routing game on an arbitrary graph  $G$  is upper bounded by

the square of the edge-cycle number. The main idea behind the result is that any Nash-routing in  $G$  can be mapped to a partial Nash-routing on some 2-connected subgraph of  $G$ . In this partial Nash-routing, many players are locally optimal, and so we can then apply the general edge-expansion theorem (Theorem 3.3) to this partial Nash-routing to obtain the upper bound.

## 6.1 Canonical Subgraphs

Consider an arbitrary connected graph  $G = (V, E)$ . A *subgraph*  $G' = (V', E')$  of  $G$  contains a subset of the nodes,  $V' \subseteq V$ , and a subset of the edges  $E' \subseteq E$ , where each edge in  $E'$  is incident with two nodes in  $V'$ . We say that  $G'$  is an *induced* subgraph by the node set  $V'$  if  $E'$  contains all the edges in  $E$  that are incident with a pair of vertices in  $V'$ . We say that two subgraphs are *adjacent* if the intersection of their node sets is non-empty. The *union* of two subgraphs  $G' = (V', E')$  and  $G'' = (V'', E'')$  is  $\overline{G} = (V' \cup V'', E' \cup E'')$ . We will focus on 2-connected subgraphs. The following result can be easily verified:

**Lemma 6.1**  *$G$  contains a 2-connected subgraph if and only if it is not a tree.*

If  $G$  is not a tree, then  $G$  contains at least one 2-connected subgraph (Lemma 6.1). A 2-connected subgraph  $G'$  is *maximal* if there is no larger 2-connected subgraph  $G'' = (V'', E'')$  that contains  $G'$ , so if  $G''$  is 2-connected, then  $E' \not\subseteq E''$ . Let  $A_1, \dots, A_\alpha$  be all the maximal 2-connected subgraphs of  $G$ , where  $\alpha \geq 1$ , and  $A_i = (V_{A_i}, E_{A_i})$ . Any two subgraphs  $A_i$  and  $A_j$ ,  $i \neq j$ , are node-disjoint since otherwise their union would be 2-connected, which contradicts their maximality.

From  $G$  we can construct two subgraphs  $A$  and  $B$ , where  $A$  consists of  $A_1, \dots, A_\alpha$ , while  $B$  consists of the remaining edges in  $G$ :  $A = (V_A, E_A)$  and  $B = (V_B, E_B)$ , where  $E_A = \bigcup_{i=1}^{\alpha} E_{A_i}$ ,  $E_B = E - E_A$ , and  $V_A$  and  $V_B$  are the nodes adjacent to the edges in  $E_A$  and  $E_B$ , respectively. Note that graphs  $A$  and  $B$  are edge-disjoint, however, they may have common nodes. Subgraph  $B$  consists of one or more disjoint maximal connected components (each containing at least two nodes), which we will denote  $B_1, \dots, B_\beta$ . (Graph  $A$  consists of connected components  $A_1, \dots, A_\alpha$ .) We refer to the  $A_i$  as the *type-a canonical subgraphs* of  $G$  and the  $B_i$  as the *type-b canonical subgraphs* of  $G$ . We now give some useful results relating to canonical subgraphs.

**Lemma 6.2** *Every type-b canonical subgraph is a tree and can have at most one node in common with any type-a canonical subgraph.*

We now define a simple bipartite graph  $H = (V_H, E_H)$  that represents the structure of  $G$ . We define  $V_H = \{a_1, \dots, a_\alpha, b_1, \dots, b_\beta\}$ , where node  $a_i$  corresponds to the type-a canonical subgraph  $A_i$ , and node  $b_j$  corresponds to the type-b canonical subgraph  $B_j$ . The edge  $(a_i, b_j)$  is in  $E_H$  if and only if the canonical subgraphs  $A_i$  and  $B_j$  are adjacent (have a common node). A natural bipartition for  $H$  is  $(\mathcal{A}, \mathcal{B})$ , where  $\mathcal{A} = \{a_1, \dots, a_\alpha\}$  and  $\mathcal{B} = \{b_1, \dots, b_\beta\}$ . This is a valid bipartition because only canonical subgraphs of different types can be adjacent. The nodes in  $H$  inherit the same type as their corresponding canonical subgraph in  $G$ . Since  $G$  is connected, it follows immediately that  $H$  is connected. Further, we have:

**Lemma 6.3** *Graph  $H$  is a tree.*

## 6.2 Canonical Subpaths

A node in  $G$  can belong to at most one type-a subgraph and one type-b subgraph, since no two canonical subgraphs of the same type are adjacent. If a node is a member of one canonical subgraph, then its type is the type of the subgraph. If the node belongs to two canonical subgraphs then it is of type-a (we assign it to the type-a canonical subgraph).

Let  $p = v_1, v_2, \dots, v_k, k > 1$ , be an edge-simple path in  $G$ . We can write  $p$  as a concatenation of subpaths  $p = q_1 q_2 \dots q_k$ , where  $|q_i| > 0, \forall i$ , with the following properties: (i) the subpaths are edge disjoint; (ii) all the nodes of a subpath  $q_i$  are in the same subgraph and have the same type, which will also be the type and subgraph of  $q_i$ ; (iii) the types of the subpaths alternate, i.e. the types of  $q_i$  and  $q_{i+1}$  are different; (iv) There is no type-a subpath with one node (any type-a subpath with one node can be merged with two adjacent type-b subpaths in the same type-b subgraph). We refer to the  $q_i$  as the *canonical subpaths* of  $p$ . Note that there is a unique canonical subpath decomposition for path  $p$ .

Since type-b subgraphs are trees and graph  $H$  is a tree, an arbitrary path in  $G$  can form cycles only inside type-a canonical subgraphs (in the respective type-a canonical subpaths). As a consequence, a path from a source node to a destination node follows a unique sequence of type-b edges (the union of all the edges in the type-b subpaths). Thus, we can obtain the following crucial result on paths that connect the same endpoints in  $G$ .

**Lemma 6.4** *Any two edge-simple paths from nodes  $s$  to  $t$  in  $G$  use the same sequence of type-b edges.*

## 6.3 Subgames in Canonical Subgraphs

Consider a routing game  $\mathcal{R} = (\mathbf{N}, G, \{\mathcal{P}_i\}_{i \in \mathbf{N}})$  in  $G$ . Let  $\mathbf{p}$  be a routing with network congestion  $C$ . Let  $\mathbf{p}^*$  denote an optimal routing for  $\mathcal{R}$  with congestion  $C^*$ . An immediate consequence of Lemma 6.4 is that every path in  $\mathbf{p}$  uses the same edges as its corresponding path in  $\mathbf{p}^*$ , hence

**Lemma 6.5** *Any type-b edge  $e$  has the same congestion in  $\mathbf{p}$  and  $\mathbf{p}^*$ , i.e.  $C_e(\mathbf{p}) = C_e(\mathbf{p}^*) \leq C^*$ .*

As a consequence of Lemma 6.5, all the high congestion edges in  $\mathbf{p}$  must occur in type-a subpaths.

**Lemma 6.6** *For path  $p$ , if  $C_p(\mathbf{p}) > C^*$ , then  $p$  must have a type-a subpath  $q$  with  $C_q(\mathbf{p}) = C_p(\mathbf{p})$ .*

Suppose now that  $\mathbf{p}$  is an arbitrary Nash-routing which has network congestion  $C$ . For a type-a subgraph  $A$ , let  $\mathbf{p}_A = \{p_1, \dots, p_\gamma\}$  denote the paths in  $\mathbf{p}$  that use edges in  $A$ , and denote the respective users as  $\mathbf{N}_A$ , where  $|\mathbf{N}_A| = \gamma$ . Let  $Q_A = \{q_1, \dots, q_\gamma\}$  denote the type-a canonical subpaths of the paths in  $\mathbf{p}_A$  that are in  $A$  ( $q_i$  is a subpath of  $p_i$ ). Each subpath  $q_i$  has a first node  $s_i^A$  in  $A$  and a last node  $t_i^A$  in  $A$ .

In subgraph  $A$ , we define a new routing game  $\mathcal{R}_A = (\mathbf{N}_A, A, \{\mathcal{P}_i^A\}_{i \in \mathbf{N}_A})$ , where  $\mathcal{P}_i^A$  contains all the type-a subpaths of  $\mathcal{P}_i$  that are in  $A$  and have the same source and destination as  $q_i$ . We refer to  $\mathcal{R}_A$  as the *subgame* of  $\mathcal{R}$  for subgraph  $A$ .  $Q_A$  is a possible routing for  $\mathcal{R}_A$ . If  $q_i$  is locally optimal for player  $i$  in  $A$ , we say that its corresponding path  $p_i$  in  $G$  is *satisfied* in subgame  $\mathcal{R}_A$ . In other words, if path  $p_i$  is satisfied in  $\mathcal{R}_A$ , player  $i$  does not wish to change the choice  $q_i$  in  $A$ . Every player with high player cost (higher than  $C^*$ ) must be satisfied in a type-a subgraph, since otherwise it would violate Lemma 6.6. Thus, we have:

**Lemma 6.7** *If player  $i$  has path  $p_i$  and  $pc_i > C^*$ , then player  $i$  is satisfied in some subgame  $\mathcal{R}_A$  in a type-a subgraph  $A$ , and player  $i$  has congestion  $pc_i$  in  $A$ .*

## 6.4 Main Result

Consider routing game  $\mathcal{R} = (\mathbf{N}, G, \{\mathcal{P}_i\}_{i \in \mathbf{N}})$  in  $G$  and a Nash-routing  $\mathbf{p}$  with congestion  $C(\mathbf{p}) = C$ . Lemma 6.7, implies that each user is satisfied in some type-a subgraph (not necessarily the same). In any type-a subgraph, the resulting routing in the subgame may be a partial Nash-routing, since some users may not be satisfied in it. We first show that there is a subgraph with high congestion where the number of unsatisfied players is bounded. For a canonical type-a subgraph  $A$ , let  $F_A(C')$  denote the set of non-locally optimal players in the subgame  $\mathcal{R}_A$  whose congestion in  $\mathcal{R}$  is at least  $C'$ . We will use  $C_A$  to denote the congestion in a subgraph  $A$ . We have:

**Lemma 6.8** *Suppose that  $C(\mathbf{p}) > C^* + x(1 + \log n)$  for some  $x > 0$ . Then, there is a type-a canonical subgraph  $A$  with congestion  $C_A \geq C - x \log n$  and  $|F_A(C_A - x)| \leq 2C^*$ .*

Lemma 6.8 will allow us to apply the edge-expansion theorem for partial Nash-routing, to obtain our upper bound on the price of anarchy, completing the proof of Theorem 2.

**Theorem 6.9 (PoA upper bound for arbitrary graphs)**  $PoA \leq c \cdot (\kappa_e^2(G) + \log^2 n)$ .

**Proof:** Let  $x = 2 \log n$ . If  $C \leq C^* + x(1 + \log n)$ , then there is nothing to prove because  $C/C^* \leq 1 + 2 \log n(1 + \log n)/C^* \leq c \log^2 n$ , for some generic constant  $c$ . So, suppose that  $C > C^* + x(1 + \log n)$ . By Lemma 6.8, there exists a type-a subgraph  $A$  such that  $C_A \geq C - 2 \log^2 n$  and  $|F_A(C_A - 2 \log n)| \leq 2C^*$ . By applying Theorem 3.3 to the subgame  $\mathcal{R}_A$  we obtain,

$$C_A < 2\ell \cdot (C_A^* + F_A(C_A - 2 \log n)) + 2 \log n,$$

where  $\ell$  is the length of the longest edge-simple path in the player's strategies in  $\mathcal{R}_A$ ,  $n'$  is the number of nodes in  $A$  and  $C_A^*$  is the optimal congestion for the subgame  $\mathcal{R}_A$ . Note that  $n' \leq n$ , and the subgame  $\mathcal{R}_A$  cannot have a higher optimal congestion than the full game  $\mathcal{R}$ , hence  $C^* \geq C_A^*$ . Since  $|F_A|$  is monotonically non-increasing ( $F_A(C') \subseteq F_A(C'')$  for  $C'' < C'$ ), we have that

$$\begin{aligned} C - 2 \log^2 n &< 2\ell \cdot (C^* + F(C_A - 2 \log n)) + 2 \log n, \\ &\leq 2\ell \cdot (C^* + 2C^*) + 2 \log n. \end{aligned}$$

From Lemma 5.1, we have that  $\ell \leq c\kappa_e^2(A) \leq c\kappa_e^2(G)$ , therefore,

$$C \leq c \cdot (\kappa_e^2(G)C^* + \log^2 n).$$

After dividing by  $C^*$ , we obtain the desired result. ■

## 7 Discussion

In Theorem 2, we believe that the result in (i) is tight, i.e. that the bound in (ii) can be considerably improved. Specifically, we leave open the following conjecture:

*Conjecture 1.* For any routing game  $(\mathbf{N}, G, \{\mathcal{P}_i\}_{i \in \mathbf{N}})$ ,  $PoA \leq \kappa_e - 1$ .

An interesting future direction is to obtain similar results when the latency functions at each link are more general and not necessarily the same. We conclude by noting that all our results have been stated for paths that are edge-simple. Analogous results could be obtained for node-simple paths, with social and player costs defined with respect to node-congestion.

## References

- [1] E. Anshelevich, A. Dasgupta, J. Kleinberg, , E. Tardos, T. Wexler, and T. Roughgarden. The price of stability for network design with fair cost allocation. In *Proc. FOCS*, 2004.
- [2] E. Anshelevich, A. Dasgupta, E. Tardos, and T. Wexler. Near optimal network design with selfish agents. In *Proc. STOC*, 2003.
- [3] Y. Azar, E. Cohen, A. Fiat, H. Kaplan, and H. Räcke. Optimal oblivious routing in polynomial time. In *Proceedings of the 35th Annual ACM Symposium on Theory of Computing (STOC)*, pages 383–388, San Diego, CA, June 2003. ACM Press.
- [4] M. Bienkowski, M. Korzeniowski, and H. Räcke. A practical algorithm for constructing oblivious routing schemes. In *Proceedings of the 15th Annual ACM Symposium on Parallelism in Algorithms and Architectures*, pages 24–33, Jun. 2003.
- [5] C. Busch, M. Magdon-Ismail, and J. Xi. Oblivious routing on geometric networks. In *17th ACM Symp. on Parallelism in Alg. and Arch. (SPAA)*, Las Vegas, Nevada, July 17-20 2005.
- [6] C. Busch, M. Magdon-Ismail, and J. Xi. Optimal oblivious path selection on the mesh. In *Proceedings of the International Parallel and Distributed Processing Symposium (IPDPS 2005)*, Denver, Colorado, USA, April 2005.
- [7] G. Christodoulou and E. Koutsoupias. The price of anarchy of finite congestion games. In *Proc. STOC*, 2005.
- [8] J. R. Correa, A. S. Schulz, and N. E. Stier Moses. Computational complexity, fairness, and the price of anarchy of the maximum latency problem. In *Proc. 10th Conf. on Integer Programming and Combinatorial Optimization (IPCO)*, 2004.
- [9] R. Cypher, F. M. auf der Heide, C. Scheideler, and B. Vöcking. Universal algorithms for store-and-forward and wormhole routing. In *In Proc. of the 28th ACM Symp. on Theory of Computing*, pages 356–365, 1996.
- [10] A. Czumaj, P. Krysta, and B. Vöcking. Selfish traffic allocation for server farms. In *Proc. STOC*, 2002.
- [11] A. Czumaj and B. Vöcking. Tight bounds for worst-case equilibria. In *Proc. SODA*, 2002.
- [12] D. Fotakis, S. Kontogiannis, E. Koutsoupias, M. Mavronicolas, and P. Spirakis. The structure and complexity of Nash equilibria for a selfish routing game. In *Proc. ICALP*, 2002.
- [13] D. Fotakis, S. Kontogiannis, and P. Spirakis. Selfish unsplittable flows. In *Proc. ICALP*, 2004.
- [14] M. Garing, T. Lücking, M. Mavronicolas, and B. Monien. Computing nash equilibria for scheduling on restricted parallel links. In *Proc. STOC*, 2004.
- [15] M. Garing, T. Lücking, M. Mavronicolas, and B. Monien. The price of anarchy for polynomial social cost. In *Proc. MFCS*, 2004.
- [16] M. Garing, T. Lücking, M. Mavronicolas, B. Monien, and M. Rode. Nash equilibria in discrete routing games with convex latency functions. In *Proc. ICALP*, 2004.
- [17] C. Harrelson, K. Hildrum, and S. Rao. A polynomial-time tree decomposition to minimize congestion. In *Proceedings of the 15th Annual ACM Symposium on Parallelism in Algorithms and Architectures*, pages 34–43, Jun. 2003.
- [18] E. Koutsoupias, M. Mavronicolas, and P. Spirakis. Approximate equilibria and ball fusion. In *Proc. SIROCCO*, 2002.
- [19] E. Koutsoupias and C. Papadimitriou. Worst-case equilibria. In *Proc. STACS*, 1999.

- [20] F. T. Leighton. *Introduction to Parallel Algorithms and Architectures: Arrays - Trees - Hypercubes*. Morgan Kaufmann, San Mateo, 1992.
- [21] F. T. Leighton, B. M. Maggs, and S. B. Rao. Packet routing and job-scheduling in  $O(\text{congestion} + \text{dilation})$  steps. *Combinatorica*, 14:167–186, 1994.
- [22] T. Leighton, B. Maggs, and A. W. Richa. Fast algorithms for finding  $O(\text{congestion} + \text{dilation})$  packet routing schedules. *Combinatorica*, 19:375–401, 1999.
- [23] L. Libman and A. Orda. Atomic resource sharing in noncooperative networks. *Telecommunication Systems*, 17(4):385–409, 2001.
- [24] T. Lücking, M. Mavronicolas, B. Monien, and M. Rode. A new model for selfish routing. In *Proc. STACS*, 2004.
- [25] B. M. Maggs, F. Meyer auf der Heide, B. Vöcking, and M. Westerman. Exploiting locality in data management in systems of limited bandwidth. In *Proceedings of the 38th Annual Symposium on the Foundations of Computer Science*, pages 284–293, 1997.
- [26] M. Mavronicolas and P. Spirakis. The price of selfish routing. In *Proc. STOC*, 2001.
- [27] D. Monderer and L. S. Shapely. Potential games. *Games and Economic Behavior*, 1996.
- [28] R. Ostrovsky and Y. Rabani. Universal  $O(\text{congestion} + \text{dilation} + \log^{1+\epsilon} N)$  local control packet switching algorithms. In *Proceedings of the 29th Annual ACM Symposium on the Theory of Computing*, pages 644–653, New York, May 1997.
- [29] C. H. Papadimitriou. Algorithms, games and the internet. In *Proc. STOC*, pages 749–753, 2001.
- [30] Y. Rabani and É. Tardos. Distributed packet switching in arbitrary networks. In *Proceedings of the Twenty-Eighth Annual ACM Symposium on the Theory of Computing*, pages 366–375, Philadelphia, Pennsylvania, 22–24 May 1996.
- [31] H. Räcke. Minimizing congestion in general networks. In *Proceedings of the 43rd Annual Symposium on the Foundations of Computer Science*, pages 43–52, Nov. 2002.
- [32] R. W. Rosenthal. A class of games possessing pure-strategy Nash equilibria. *International Journal of Game Theory*, 1973.
- [33] T. Roughgarden. The maximum latency of selfish routing. In *Proc. SODA*, 2004.
- [34] T. Roughgarden. Selfish routing with atomic players. In *Proc. SODA*, 2005.
- [35] T. Roughgarden and Éva Tardos. How bad is selfish routing. *Journal of the ACM*, 49(2):236–259, March 2002.
- [36] T. Roughgarden and Éva Tardos. Bounding the inefficiency of equilibria in nonatomic congestion games. *Games and Economic Behavior*, 47(2):389–403, 2004.
- [37] S. Suri, C. D. Tóth, and Y. Zhou. Selfish load balancing and atomic congestion games. In *Proc. SPAA*, 2004.
- [38] D. B. West. *Introduction to Graph Theory*. Prentice Hall, Upper Saddle River, NJ, U.S.A., 2001.

## A Proofs of Section 2

**Proof of Lemma 2.2.** Suppose that a greedy move by player  $i$  takes  $\mathbf{p}$  to  $\mathbf{p}'$ , and so  $C_{p'_i}(\mathbf{p}') < C_{p_i}(\mathbf{p})$ . Let  $k = C_{p_i}(\mathbf{p})$ . Since only player  $i$  has changed his path, the only edges with higher congestion in  $\mathbf{p}'$  than in  $\mathbf{p}$  are edges on the path  $p'_i$ . Some edges on  $p_i$  decreased in congestion by 1 as a result of the greedy move. In particular, all edges of congestion  $k$  on  $p_i$  have decreased in congestion by 1, since all edges on  $p'_i$  have final congestion less than  $k$ . Thus,  $m_k(\mathbf{p}') \leq m_k(\mathbf{p}) - 1$ , since at least one edge of congestion  $k$  dropped in congestion and no new edges reached congestion  $k$ . To conclude that  $\mathbf{p}' <_c \mathbf{p}$ , we note that no edge with congestion greater than  $k$  has been affected by the greedy move, hence  $m_j(\mathbf{p}') = m_j(\mathbf{p})$  for all  $j > k$ . ■

## B Proofs of Section 3

**Proof of Lemma 3.1.** Since player  $i$  is locally optimal, every path in  $\mathcal{P}_i$  must have path-congestion at least  $pc_i - 1$ . Indeed, if not, then there is a path  $p'_i \in \mathcal{P}_i$  with path-congestion at most  $pc_i - 2$ . If player  $i$  switches from  $p_i$  to  $p'_i$ , his cost is at most  $pc_i - 1$ , which contradicts  $p_i$  being locally optimal for  $i$ . For every path  $p \in \mathcal{P}_i$ , let edge  $e(p) \in p$  be an edge with maximum congestion on  $p$  ( $C(e) \geq pc_i - 1$ ). Let  $E_i = \cup_{p \in \mathcal{P}_i} e(p)$ . Since  $E_i$  contains at least one edge from every path in  $\mathcal{P}_i$ , it is a path-cut for player  $i$  and every edge in  $E_i$  has congestion at least  $pc_i - 1$ . Thus,  $C(E_i) \geq pc_i - 1$ . ■

**Proof of Lemma 3.2.** If  $s = 1$  there is nothing to prove, so assume that  $s > 1$ . Since  $|\mathcal{E}_k| > 2|\mathcal{E}_{k-1}|$  for  $k = 1, \dots, s-1$ ,  $|\mathcal{E}_k| > 2^k|\mathcal{E}_0|$ . Since  $|\mathcal{E}_0| \geq 1$ , we have  $|\mathcal{E}_k| > 2^k$ . By construction,  $\mathcal{E}_{i-1} \subseteq \mathcal{E}_i$ , for  $1 \leq i \leq s$ ; thus,  $|\mathcal{E}_s| \geq |\mathcal{E}_{s-1}| \geq 2^{s-1}$ . Since  $|\mathcal{E}_s| \leq |E| < \frac{1}{2}n^2$ ,  $2^{s-1} < \frac{1}{2}n^2$  implying  $s < 2 \log n$ . ■

## C Proofs of Section 5

**Proof of Lemma 5.1.** Let  $u$  and  $v$  be the respective starting and ending nodes in the longest path in  $G$ . Since the min  $(u, v)$ -cut has size at least two, by Menger's theorem [38], there is a pair of edge-disjoint paths  $p_1, p_2$  from  $u$  to  $v$ ; let  $l_1 \leq l_2$  be the lengths of these paths respectively. Let  $p$  be an edge-simple  $(u, v)$ -path with length  $\ell$ . Path  $p$  can be decomposed into  $2k$  path segments as follows,

$$p = \lambda_0 \kappa_1 \lambda_1 \kappa_2 \lambda_2 \cdots \kappa_k \lambda_k,$$

where each  $\kappa_i$  has length at least one and consists only of edges on  $p_1$ , and each “excursion”  $\lambda_i$  does not contain any edges on  $p_1$ . Since each excursion  $\lambda_i$  connects two (not necessarily distinct) nodes on  $p_1$ , it follows that there is an edge simple cycle composed of  $\lambda_i$  together with the segment of  $p_1$  between these two nodes. The length of this cycle is at least  $|\lambda_i|$ , so we have that  $\kappa_e(G) \geq |\lambda_i|$  for all  $i \in [0, k]$ . Since path  $p$  is edge simple, there can be at most  $l_1 + 1$  excursions (as each of the  $\kappa_i$

must contain distinct edges), i.e.  $k \leq l_1$ . We now compute the length of  $p$  as follows,

$$\begin{aligned} \ell &= \sum_{i=0}^k |\lambda_i| + \sum_{i=1}^k |\kappa_i|, \\ &\leq \sum_{i=0}^k \kappa_e(G) + \sum_{i=1}^k |\kappa_i|, \\ &\leq \kappa_e(G) \cdot (l_1 + 1) + l_1. \end{aligned}$$

Solving for  $\kappa_e(G)$ , we have that  $\kappa_e(G) \geq (\ell + 1)/(l_1 + 1) - 1$ . Since  $p_1$  and  $p_2$  form an edge-simple cycle,  $\kappa_e(G) \geq l_1 + l_2 \geq 2l_1$ . Combining these inequalities, we have

$$\kappa_e(G) \geq \max \left\{ 2l_1, \frac{\ell + 1}{l_1 + 1} - 1 \right\}.$$

Suppose that  $\kappa_e(G) < \sqrt{2\ell} - \frac{3}{2}$ . Since  $\kappa_e(G) \geq 2l_1$ , we have  $l_1 < \sqrt{\ell/2} - \frac{3}{4}$ . Therefore,

$$\begin{aligned} \kappa_e(G) &\geq \frac{\ell + 1}{l_1 + 1} - 1, \\ &> \frac{\ell + 1}{\sqrt{\ell/2} + \frac{1}{4}} - 1, \\ &= \sqrt{2\ell} - \frac{3}{2} + \epsilon, \end{aligned}$$

where  $\epsilon = 9/(\sqrt{32\ell} + 2) > 0$ . This contradiction concludes the proof. ■

## D Proofs of Section 6.1

The following result follows from the definition of a 2-connected graph.

**Lemma D.1**  *$G$  is 2-connected if and only if every pair of nodes are on some edge-simple cycle.*

Note that only canonical subgraphs of different types can be adjacent (since both the type-a and type-b subgraphs have a maximality property). We obtain the following basic result.

**Lemma D.2** *Any edge-simple path  $p$  that leaves a type-a subgraph does not re-enter it.*

**Proof:** To the contrary, suppose that such a path  $p$  exists which leaves a type-a subgraph  $A$  at node  $v$  and re-enters it for the first time at node  $u$  ( $u$  may be the same node as  $v$ ). Let  $p'$  be the edge-simple subpath of  $p$  from  $v$  to  $u$ . By construction, none of the edges in  $p'$  are in  $A$ , and the same is true for all nodes in  $p'$  except  $v, u$ . We will show that the union of  $A$  with  $p'$  is 2-connected, contradicting the maximality of  $A$ . Let  $A' = A \cup p'$ , and let  $z, w$  be any two nodes in  $A'$ . If both  $z$  and  $w$  are in  $A$ , then they lie on an edge-simple cycle (from Lemma D.1, as  $A$  is 2-connected). Let  $q$  be an edge-simple  $(v, u)$ -path in  $A$ , possibly of length zero. The cycle  $r = q \cup p'$  is an edge-simple cycle in  $A'$  that contains any two nodes on  $r$  so if  $z, w \in p'$ , then they are on an edge-simple cycle. Suppose that  $z \in A$  and  $w \notin A$ . There is edge-simple cycle  $s$  containing  $z$  and  $u$  (since  $z, u \in A$ ). Starting at  $z$ , this cycle must therefore enter cycle  $r$ , leave  $r$  and return to  $z$ . We can splice  $r$  onto  $s$  at the first node at which  $s$  enters  $r$  and the last node at which  $s$  leaves  $r$ , choosing the splice of  $r$  that contains  $w$ . In this way, we construct an edge-simple cycle containing  $z, w$ . Thus every pair of nodes in  $A'$  lies on an edge simple cycle, so  $A'$  is 2-connected (Lemma D.1). ■



**Proof of Lemma 6.2.** Suppose some  $B_i$  is not a tree. Then it contains a 2-connected subgraph (Lemma 6.1), a contradiction since every 2-connected subgraph is a subgraph of some  $A_i$  which must be edge disjoint from  $B_i$  by construction.

If some  $B_i$  has at least two distinct nodes  $u, v$  in common with some  $A_i$ , then an edge-simple path exists in  $B_i$  from  $u$  to  $v$  as  $B_i$  is connected. This path leaves  $A_i$  and re-enters it, which contradicts Lemma D.2. ■

**Proof of Lemma 6.3.** Since  $H$  is connected, it suffices to show that  $H$  is acyclic. To the contrary, suppose that  $H$  contains a node-simple cycle  $c_H = h_0, h_1, \dots, h_k, h_{k+1}$ , where  $h_{k+1} = h_0$ . Note that since  $H$  is bipartite,  $c_H$  is an even cycle. Any pair of adjacent canonical subgraphs identifies a unique (by Lemma 6.2) node of intersection. Thus corresponding to  $c_H$  is a unique sequence of nodes  $v_0, v_1, \dots, v_k$  in  $G$ , where  $v_i$  is the node that corresponds to the adjacent canonical subgraphs  $h_i, h_{i+1}$ . The nodes  $v_i$  are not necessarily all different, however every consecutive pair is in a single canonical component, including the pair  $(v_k, v_0)$ . Thus, in  $G$ , there is an edge-simple path  $p_i, i = 0, \dots, k$ , from  $v_i$  to  $v_{i+1}$  (where  $v_{k+1} = v_0$ ) which is contained in the canonical subgraph  $h_{k+1}$ , since the  $h$ 's are connected. Since the canonical subgraphs form an edge partition of  $E$ , the concatenation of these paths,  $q = p_0 p_1 \dots p_k$  is an edge-simple cycle in  $G$ . At least one of the  $h_i$  is of type-a, hence  $q$  must leave a type-a canonical subgraph and re-enter it, which contradicts Lemma D.2. ■

## E Proofs of Section 6.2

Here we prove Lemma 6.4 after first giving some elementary properties of canonical subgraph decompositions. Consider the canonical subpath decomposition of a path  $p = q_1 \dots q_k$ . Let  $S(p) = G_1, \dots, G_k$  denote the *subgraph sequence* of the respective canonical subgraphs that contain the canonical subpaths. We have,

**Lemma E.1** *For an edge-simple path  $p$ , no type-a canonical subgraph repeats in  $S(p)$ .*

**Proof:** Let  $A$  be a type-a repeated canonical subgraph. Some other canonical subgraph separates two occurrences of  $A$  in  $S(p)$ . Thus,  $p$  must leave and re-enter  $A$ , contradicting Lemma D.2. ■

For any edge-simple path  $p$  let  $r(p)$  denote the *reduced* node-simple path that we obtain from  $p$  after removing any cycles. Note that cycles in  $p$  exist only in type-a canonical subpaths. Thus,  $r(p)$  is similar to  $p$  with the difference that some type-a subpaths are removed (and the adjacent type-b subpaths are merged). Therefore, we have,

**Lemma E.2** *Let  $p$  be an edge-simple path with reduced node-simple path  $r(p)$ , then: (i) no canonical subgraph (type-a or type-b) repeats in  $S(r(p))$ , and (ii) paths  $p$  and  $r(p)$  visit the same sequence of type-b edges.*

**Proof:** (ii) follows immediately because none of the cycles removed contain type-b edges. For (i), we need only consider repeated type-b canonical subgraphs. let  $B$  be the first such subgraph that is repeated for the first time. If there is more than one subgraph between these two occurrences of  $B$ , then there is an edge-simple cycle in  $H$ , which contradicts  $H$  being a tree (Lemma 6.3). Therefore the only possibility is that a single type-a canonical subgraph  $A$  occurs between the two

occurrences of  $B$  in the decomposition, i.e.  $S(r(p)) = \cdots BAB \cdots$ . Since  $A$  and  $B$  have exactly one common node (Lemma 6.2), this means the subpath in  $A$  has the same first and last node, and hence  $A$  should have been removed, a contradiction. ■

For subgraph sequence  $S(p) = G_1, \dots, G_k$ , let  $h_1, \dots, h_k$  be the nodes in  $H$  corresponding to  $G_1, \dots, G_k$ . The *projection*  $H(p)$  of  $p$  into  $H$  is the walk in  $H$  given by  $H(p) = h_1, \dots, h_k$ . Note that  $H(p)$  may not be a path, since  $p$  may visit the same type-b subgraph more than once. However, from Lemma E.2 the projection  $H(r(p))$  is a node-simple path. Using projections on  $H$ , and the fact that  $H$  is a tree (Lemma 6.3), we can obtain an equivalence result between reduced paths.

**Lemma E.3** *For any two edge-simple paths  $p$  and  $q$  from node  $s$  to  $t$  in  $G$ ,  $S(r(p)) = S(r(q))$ .*

**Proof:** Suppose that  $S(r(p)) \neq S(r(q))$ , then  $H(r(p)) \neq H(r(q))$ . Let  $H(r(p)) = u_1, \dots, u_k$  and  $H(r(q)) = w_1, \dots, w_l$ , and assume that  $k \leq l$ . If  $l = 1$ , then  $u_1 \neq w_1$ , and  $s, t$  are both in  $u_1, w_1$ , which means that two nodes are common to two different canonical subgraphs, contradicting Lemma 6.2. Therefore  $l \geq 2$ . Since  $s$  is a node common to the subgraphs corresponding to  $u_1$  and  $w_1$ , either  $u_1 = w_1$  or  $u_1$  and  $w_1$  are adjacent in  $H$ . Similarly,  $u_k = w_l$  or  $u_k$  and  $w_k$  are adjacent in  $H$ , as  $t$  is a common node.

Let  $j$  be the first index for which  $u_j \neq w_j$ . First suppose that  $1 < j \leq k$ , so  $u_1 \cdots u_{j-1} = w_1 \cdots w_{j-1}$ , and  $u_{j-1}$  is adjacent to both  $u_j, w_j$ . Now consider the sequence of nodes  $w_{j-1}w_jw_{j+1} \cdots w_l u_k u_{k-1} \cdots u_j u_{j-1}$ . This is a valid walk (with possibly repeated nodes), since either  $u_k = w_l$  or  $u_k$  is adjacent to  $w_l$ . Further, this is a cycle, since  $w_{j-1} = u_{j-1}$ . This cycle contains at least 3 different nodes, because  $w_{j-1} \neq w_j$  and  $u_{j-1} \neq u_j$  (Lemma E.2), and  $u_j \neq w_j$ . This cycle can be reduced to a node simple cycle with at least 3 different nodes, contradicting the fact that  $H$  is a tree (Lemma 6.3).

If  $j = 1$ , then we construct a cycle  $w_1 \cdots w_l u_k \cdots u_1 w_1$ , which we know is valid since  $u_1$  is adjacent to  $w_1$ . This cycle contains at least 3 different nodes since  $u_1 \neq w_2$ . If  $j = k + 1$  (it must be that  $l > k$ ), then we construct a cycle  $w_k \cdots w_l u_k$ , which is valid since  $u_k = w_k$ . This cycle also contains at least 3 different nodes because since  $r(q)$  is node simple, it cannot use  $t$  to get into  $w_{k+1}$ . Hence,  $w_{k+1} \neq w_l$  because this would imply that  $w_{k+1}$  and  $u_k$  have at least two nodes in common. Therefore either  $l > k + 1$  or  $w_l \neq u_k$ . ■

We are now ready to give the proof of Lemma 6.4.

**Proof of Lemma 6.4.** Let  $p$  and  $q$  be two edge-simple paths from  $s$  to  $t$ . By Lemma E.2, we only need to show that  $r(p)$  and  $r(q)$  use the same set of type-b edges. The result follows essentially because each type-b subgraph is a tree (Lemma 6.2).

From Lemma E.3,  $S(r(p)) = S(r(q))$ . Type-b edges are used only in the type-b subgraphs which appear in the same order in  $S(r(p))$  and  $S(r(q))$ . Consider a type-b subgraph occurring in this subgraph decomposition. Either it is the first, or the last, or type-a subgraphs occur before and after. In all cases the subpath in this type-b subgraph is from the same node  $u$  to the same node  $v$  in both of the subgraph decompositions. This is because these nodes are either the unique nodes of intersection between the same type-a and type-b subgraph, or they are the source or destination, which are the same for both paths. Since each type-b subgraph is a tree, there is a unique path from  $u$  to  $v$ , which must be the same in both subpath decompositions. ■

## F Proofs of Section 6.3

**Proof of Lemma 6.7.** We show that if the claim is false, then path  $p_i$  is not locally optimal for player  $i$  in  $\mathcal{R}$ , contradicting the fact that  $\mathbf{p}$  is a Nash-routing. Indeed, we know from Lemma 6.6 that  $p_i$  uses type-a canonical subgraphs. If none of these type-a canonical subpaths are not locally optimal for their respective subgames, then they can all be switched in favor of paths with strictly lower congestion than  $C$ . This will give a valid path for player  $i$  with strictly lower congestion than  $C$ , hence  $p_i$  is not locally optimal for player  $i$ . ■

## G Proofs of Section 6.4

**Proof of Lemma 6.8.** Let  $f_A = |F_A(C_A - x)|$ , and suppose that  $C > C^* + x(1 + \log n)$ . Assume, that every type-a subgraph  $A$  with congestion  $C(A) \geq C - x \log n$  has  $f_A > 2C^*$ . we will obtain a contradiction by showing that  $H$  has a cycle. Since  $\mathbf{p}$  is a Nash-routing, every player with congestion  $C > C^*$  is locally optimal in at least one type-a subgame of  $\mathcal{R}$  (Lemma 6.7). Thus, there is at least one type-a canonical subgraph  $A_1$  with  $C_{A_1} = C$ . We will now root  $H$  at the its type-a node  $a_1$  which corresponds to  $A_1$  and define a type-a tree  $H_a$  composed only of the type-a nodes in  $H$ . The root of  $H_a$  is also  $A_1$ . By assumption,  $f_{A_1} > 2C^*$ . Since  $C - x > C^* + x \log n$ , these  $f_{A_1}$  players which are not locally optimal in subgame  $\mathcal{R}_{A_1}$  have congestion at least  $C - x$  and are locally optimal in some other subgame. Therefore, their paths leave  $A_1$  and enter some other type-a canonical subgraph.

**Claim G.1** *If  $K$  paths leave a type-a canonical subgraph  $A$ , they must use at least  $\lceil K/C^* \rceil$  distinct edges out of  $A$*

**Proof:** If not, then one of the exit edges (which is a type-b edge) will have congestion greater than  $C^*$ , contradicting Lemma 6.5. □

We now build the rooted tree  $H_a$  inductively as follows. The root node is  $a_1$ . Suppose that  $\alpha$  is a node in  $H_a$  corresponding to type-a subgraph  $A$ , with the following two properties:

- (i)  $C_A - x > C^*$ ;
- (ii)  $f_A = |F_A(C_A - x)| > 2C^*$ .

Then, we define three potential children for  $\alpha$  as follows. Since there are  $f_A > 2C^*$  players with congestion at least  $C_A - x > C^*$  which are not locally optimal in subgame  $\mathcal{R}_A$ , these  $f_A$  players must be locally optimal in some other subgame. Therefore all these  $f_A$  paths leave  $A$  and proceed to their respective subgames where they are locally optimal with congestion at least  $C_A - x$ . By Claim G.1, they use at least three distinct type-b edges  $e_1, e_2, e_3$  in leaving  $A$  (note that these three edges may be in the same type-b canonical subgraph, but this will not affect the argument). Let  $p_1, p_2, p_3$  be three paths with congestion at least  $C_A - x$  that exit  $A$  on the edges  $e_1, e_2, e_3$  respectively and continue on to their respective canonical subgraphs  $A_1, A_2, A_3$  in which they are locally optimal. At least two of these subgraphs correspond to nodes that are not the parent (if it exists) of  $\alpha$  in  $H_a$ ; these two nodes are two children  $c_1(\alpha)$  and  $c_2(\alpha)$  of  $\alpha$  in  $H_a$  (if more than two of these children are different from the parent, we arbitrarily pick two). The depth of a child is one greater than the depth of its parent (the depth of the root is 0). The next few lemmas give some properties of  $H_a$  that will be needed to complete the proof of the theorem.

**Claim G.2**  $H_a$  is a tree.

**Proof:**  $H_a$  is connected, by construction. Suppose that  $H_a$  contains a node simple cycle. By construction, an edge between nodes  $\alpha_1, \alpha_2$  in  $H_a$  implies the existence of an edge simple path leaves one type-a subgraph and enters the second. Hence, there is a path that leaves a type-a node and re-enters it. This path can be made edge simple by removing all cycles, which contradicts Lemma D.2.  $\square$

The nodes in  $H_a$  can be viewed as constructed level by level. Each node in  $H_a$  that satisfies the two conditions above has exactly two children. Note that  $a_1$  satisfies these two conditions, initiating the construction of  $H_a$ . The nodes in  $H_a$  which do not satisfy the conditions (i) and (ii) are leaves. Thus all nodes in  $H_a$  are either leaves or have two children.

**Claim G.3** A node at depth  $d \leq \log n$  cannot be a leaf.

**Proof:** Let  $\alpha$  be a node at depth  $d$ , corresponding to type-a canonical subgraph  $A$ . We show that  $C_A \geq C - d \cdot x$  by induction on  $d$ . Certainly when  $d = 0$ , the claim holds since  $C_{A_1} = C$ . Consider  $d > 0$ . The parent of  $A$ ,  $Par_A$ , has depth  $d - 1$ , so  $C_{Par_A} \geq C - (d - 1)x$ , by the induction hypothesis. Since  $d - 1 \leq \log n$ , by assumption  $f_{Par_A} > 2C^*$  and by construction of the children in  $H_a$ ,  $A$  is a subgraph in which some player is locally optimal in the subgame  $\mathcal{R}_A$  and has congestion at least  $C_{Par_A} - x \geq C - (d - 1) \cdot x + x$ . Therefore,  $C_A \geq C - d \cdot x$ .

Since  $d \leq \log n$ , we conclude that  $C_A - x \geq C - x(1 + \log n) > C^*$  by assumption in the statement of the theorem. Thus, condition (i) is satisfied for  $\alpha$  to have children. Since  $C_A \geq C - x \log n$ , by assumption  $f_A > 2C^*$ , hence condition (ii) is satisfied for  $\alpha$  to have children. Since both conditions are satisfied,  $\alpha$  cannot be a leaf node.  $\square$

We are now ready to conclude the proof of the theorem by obtaining a contradiction. Since  $H_a$  must have a leaf node, we conclude that the depth of  $H_a$  is at least  $1 + \log n$ . Since every node at depth at most  $\log n$  has 2 children, we conclude that  $H_a$  has  $\sum_{i=1}^{\log n} 2^i$  nodes. Evaluating this sum, we have that  $H_a$  contains  $2n - 1$  nodes, and since  $n > 1$ , we have our contradiction since  $H_a$  cannot possibly contain more nodes than  $G$ .  $\blacksquare$