The Price of Routing Unsplittable Flow

Baruch Awerbuch* Yossi Azar† Amir Epstein‡

Abstract

The essence of the routing problem in real networks is that the traffic demand from a source to destination must be satisfied by choosing a single path between source and destination. The splittable version of this problem is when demand can be satisfied by many paths, namely a flow from source to destination. The unsplittable, or discrete version of the problem is more realistic yet is more complex from the algorithmic point of view; in some settings optimizing such unsplittable traffic flow is computationally intractable.

In this paper, we assume this more realistic unsplittable model, and investigate the "price of anarchy", or deterioration of network performance measured in total traffic latency under the selfish user behavior. We show that for linear edge latency functions the price of anarchy is exactly 2.618 for weighted demand and exactly 2.5 for unweighted demand. These results are easily extended to (weighted or unweighted) atomic "congestion games", where paths are replaced by general subsets. We also show that for polynomials of degree d edge latency functions the price of anarchy is $d^{\Theta(d)}$. Our results hold also for mixed strategies.

Previous results of Roughgarden and Tardos showed that for linear edge latency functions the price of anarchy is exactly $\frac{4}{3}$ under the assumption that each user controls only a negligible fraction of the overall traffic (this result also holds for the splittable case). Note that under the assumption of negligible traffic pure and mixed strategies are equivalent and also splittable and unsplittable models are equivalent.

1 Introduction

1.1 The model

A major component of any large-scale network system is the routing mechanism, namely choosing a communication path between a sender and a receiver of traffic. In most cases, such as the Internet, wireless networks, or overlay networks built on top of the Internet, traffic from a sender to a receiver is sent over a *single* path; splitting the traffic causes the problem of packet reassembly at the receiver and thus is generally avoided. When choosing routing paths, the typical objective is to minimize the total latency. In most of these network systems it is infeasible to maintain one centralized authority that imposes efficient routing strategies on the network traffic. As a result users act independently and "selfishly": each user tries to minimize her traffic cost based on current network traffic.

This problem can be mathematically formalized using classical game theory as follows. The network users are viewed as independent agents participating in a noncooperative game. Each agent wishes to use the minimum latency path from its source to its destination, given the link congestion

*Dept. of Computer Science, Johns Hopkins University. E-Mail: baruch@cs.jhu.edu. Research supported by NSF grants ANIR-0240551 and CCR-0311795.

†School of Computer Science, Tel-Aviv University. E-Mail: azar@tau.ac.il. Research supported in part by the German-Israeli Foundation.

‡School of Computer Science, Tel Aviv University. E-Mail: amirep@tau.ac.il. Research supported in part by the German-Israeli Foundation.

caused by the rest of the agents. This system is said to be in Nash Equilibrium if no agent has an incentive to change his path from its source to its destination. It is well known that Nash Equilibria do not in general optimize the social welfare (see, e.g, ”The Prisoner’s Dilemma” [9, 18]) and can be far from the global optimum.

Equilibria can be defined for pure strategies, where a single path is chosen by each user and for mixed strategies, where a probability distribution over the paths is used instead of a single path. In this paper we consider both pure and mixed strategies. This is in contrast to the case where each user controls only a negligible fraction of the overall traffic in which pure and mixed strategies are the same [24].

An additional complication that raised when considering arbitrary unsplittable traffic is that in this general case there is not necessarily pure Nash Equilibrium. Nash [17] proved that every game has a randomized Nash Equilibrium. In a classical paper Rosenthal [20] proved that every congestion game (i.e. when all demands are of unit size) has a pure Nash Equilibrium. Hence when the problem can be represented as a congestion game a pure Nash Equilibrium always exists, e.g, when all network users control the same amount of traffic. For general demands the existence of pure Nash Equilibrium is still open. Recently for linear edge latency functions Fotakis et. al [11] proved that pure Nash Equilibrium always exists.

The degradation of network performance caused by the lack of a centralized authority can be measured using the worst-case coordination ratio (price of anarchy) suggested by Koutsoupias and Papadimitriou [14, 19] which is the ratio between the worst possible Nash Equilibrium and the social optimum.

Previous results (Roughgarden and Tardos [24]) showed that under the assumption that each user controls only a negligible fraction of the overall traffic the price of anarchy is exactly $\frac{4}{3}$ for linear edge latency functions and they also showed bicriteria results for continuous and nondecreasing edge latency functions. The above results were extended to the splittable case by Roughgarden [21].

1.2 Our Results

We prove the following results for general networks with unsplittable flow:

- For linear latency functions we prove that the worst-case coordination ratio for weighted demand is exactly 2.618 for pure and mixed strategies. For unweighted demand we show that the worst-case coordination ratio is exactly 2.5 for pure strategies.

- For polynomials of degree d latency functions we prove that the worst-case coordination ratio is $d^{\Theta(d)}$ for pure and mixed strategies. More precisely the worst-case coordination ratio is at most $O(2^d d^{d+1})$ and at least $\Omega(d^{d/2})$.

- All the above results can be extended to weighted and unweighted atomic congestion games, where arbitrary subsets are used instead of paths (see [20, 10] for definition of congestion games).

Our results for the price of anarchy for bounded degree polynomials is a constant (independent of the network size), which stands in contrast to the maximum load in the Koutsoupias/Papadimitriou model. On the other hand we note that the lower bound for bounded degree polynomials implies that there is no constant $c > 0$ that applies for all polynomials such that the total latency of selfish users is bounded from above by the total latency incurred by optimally routing c times as much traffic. We note that considering mixed strategies is essential (and a complication that does not exist in most previous work), since Nash equilibrium for pure strategies need not exist in this model (see [20]). We also note that for unsplittable (atomic) network congestion games only special cases were previously considered. Recently and independently Christodoulou and Koutsoupias [4] showed similar results to our results for the special case of unweighted demands.

Techniques: The lower bounds imply that the bounds of Roughgarden and Tardos for the splittable model cannot carry over to this model, so new techniques must be needed. In our proofs for evaluating the price of anarchy for pure strategies we compare the delay encountered by each agent to the delay it would encounter if it changes to the optimal route. We combine these bounds in a weighted fashion and transform it to a relation between the total delay of the Nash Equilibrium and the total delay of the optimal routes. For mixed strategies we separate the problem into two logical steps. In the first step we consider the latency of the expected load of each edge and in the second step we augment it to the expectation of the total latency. The first step turns out to be equivalent to the pure strategies case where requests are allowed to be split. The second step adds technical complications that does not occur in the proofs for pure strategies. Although the results for pure strategies follows from the results of mixed strategies for simplifying the presentation we start with proving the results for pure strategies. We note that in our proofs we also use techniques and inequalities appear in [1, 25].

1.3 Related Work

Splittable or negligible flows. Unregulated traffic routing for general networks has been modeled as network flow model since the 1950's [3, 26] (see [24] for further historical references). Roughgarden and Tardos [24] initiated the study of price of anarchy in this model (dubbed worst-case coordination ratio by Papadimitriou [14]). Note that under the assumption of negligible traffic as in Roughgarden and Tardos model, splittable and unsplittable models are equivalent and also pure and mixed strategies are equivalent. They proved that for linear latency functions the worst-case coordination ratio is exactly $\frac{4}{3}$. They also proved that for general continuous and nondecreasing latency functions the total latency of the routes chosen by selfish network users is no more than the total latency incurred by optimally routing twice as much traffic. Specifically, for polynomial of degree d they showed that the coordination ratio is linear in d. When network users can control a significant portion of the overall traffic, but are permitted to route their flow fractionally Roughgarden [21] showed that all known bounds on the price of anarchy for nonatomic selfish routing games carry over to the atomic splittable case. Roughgarden [23] also showed that the cost of unregulated traffic does not depend on the complexity of network topology. He also studied the impact of latency functions belonging to specific classes. Roughgarden et al. [5, 6, 22] studied various ways to construct and price networks such that the cost incurred in unregulated traffic is minimized.

Coordination ratio for parallel links. Most of the work on the parallel links model was done with the maximum load measure. Koutsoupias and Papadimitriou [14] initiated the study of worst-case coordination ratio in networks composed of m parallel links with possibly different speeds. They showed that for two links the worst-case coordination ratio is exactly $\frac{3}{2}$ for identical links and $\phi = \frac{1+\sqrt{5}}{2}$ for links with possibly different speeds. They also obtained non-tight bounds for the general case of m identical links. Mavronicolas and Spirakis [16] obtained tight results for the general case of any number of links. Their results are for the special case of fully-mixed strategies in which the probability of assigning any task to any link is non-zero. They proved that for the identical links model the worst-case coordination ratio is $\Theta(\frac{\log m}{\log \log m})$. They proved the same result for the general related links model where all tasks have equal weights and $m \leq n$. Czumaj and Vöcking [8] improved these results by proving tight bounds for the m parallel links model. They showed that for the identical links model the worst-case coordination ratio is $\Theta(\frac{\log m}{\log \log m})$ and for the general related links model the worst-case coordination ratio is $\Theta(\frac{\log m}{\log \log \log m})$. In [2] additional models were considered. Czumaj et. al [7] continued to study this problem and characterized the coordination and bicriteria ratios for different families of cost functions. For other results in this model with other social costs see, e.g, [12, 13, 15, 25].

1.4 Paper structure

The paper is organized as follows. Section 2 includes formal definitions and notations. In section 3 we prove the results for networks with linear latency functions. In section 4 we prove the results for networks with polynomials of degree d latency functions.

2 Definitions and preliminaries

2.1 The Model

We consider the following model: there is a directed graph $G = (V, E)$. Each edge $e \in E$ is given a load-dependent latency function $f_e : \mathcal{R}^+ \to \mathcal{R}^+$. There are n users, where user j ($j = 1, \ldots, n$) has a bandwidth request defined by a tuple (s_j, t_j, w_j), where $s_j, t_j \in V$ are the source/destination pair, and $w_j \in \mathcal{R}^+$ corresponds to the required bandwidth. We denote the set of (simple) $s_j - t_j$ paths by $\mathcal{Q}_j$. Request j can be assigned to any path Q from the set of paths $\mathcal{Q}_j$, such that the required bandwidth w_j has to be reserved along the path Q.

We assume that the users are non-cooperative and each one wishes to minimize its own cost with no regard to the global optimum. We consider two types of users strategy systems:

1. **Pure strategies:** user j selects a single path $Q \in \mathcal{Q}_j$ and assigns his request to it. Each user is aware of the choices made by all other user when making his decision.

2. **Mixed strategies:** user j selects a probability distribution $\{p_j\}$ $(Q \in \mathcal{Q}_j)$ over all the set of paths from the source s_j to the destination t_j. Each user is aware of the probability distributions selected by all other users.

2.2 Pure strategies definition

First, we give some simpler notations we use for system $\mathcal{S}$ of pure strategies. Let Q_j be the path associated with request j and let $\mathcal{Q} = \cup_j Q_j$ be the set of paths associated with all the requests. We define $J(e) = \{j | e \in \mathcal{Q}\}$ the set of requests assigned to a path containing the edge e. The load on edge e is defined by: $l_e = \sum_{j \in J(e)} w_j$.

For the optimal routes let Q_j^* be the path associated with request j and let $\mathcal{Q}^* = \cup_j Q_j^*$ be the set of paths associated with all the requests. We define $J^*(e) = \{j | e \in \mathcal{Q}^*\}$ the set of requests assigned to a path containing the edge e. We denote the load on edge e by l_e^*.

Definition 2.1 *The latency of user j for assigning his request in system $\mathcal{S}$ to path Q (instead of path Q_j) is defined as:*

$$c_{Q,j} = \sum_{(e \in Q) \wedge (e \in Q_j)} f_e(l_e) + \sum_{(e \in Q) \wedge (e \notin Q_j)} f_e(l_e + w_j). \tag{1}$$

2.3 Mixed strategies definition

Given a system $\mathcal{S}$ of mixed strategies with probability distribution $\{p_j\}$, we denote the probability of assigning request j to edge e by $p_{e,j}$. We define the following random variables:

- A set of indicator random variables $\{X_{Q,j}\}$, where $X_{Q,j}$ indicates whether request j is assigned to path Q. By definition: $Pr[X_{Q,j} = 1] = p_{Q,j}$.
- A set of indicator random variables $\{X_{e,j}\}$, where $X_{e,j}$ indicates whether request j is assigned to edge e. By definition: $X_{e,j} = \sum_{Q|e \in Q} X_{Q,j}$ and $Pr[X_{e,j} = 1] = p_{e,j}$.
- For each edge e ($e \in E$) we define a random variable l_e, indicating the total load on the edge: $l_e = \sum_{j=1}^{n} X_{e,j} w_j$.

Definition 2.2 *The expected latency of user j for assigning his request in system $\mathcal{S}$ to path Q is defined as:* $c_{Q,j} = E[\sum_{e \in Q} f_e(l_e) | X_{Q,j} = 1] = \sum_{e \in Q} E[f_e(l_e + (1 - X_{Q,j}) w_j)$

Notice that for a system $\mathcal{S}$ of pure strategies we have $X_{Q,j} = \{0, 1\}$, matching equation 1.

2.4 Nash equilibrium and Coordination ratio

Nash equilibrium is characterized by the property that there is no incentive for any user to change its strategy and defined as follows

Definition 2.3 (Nash Equilibrium) *A system $\mathcal{S}$ is said to be in Nash Equilibrium if and only if for every $j \in \{1, \ldots, n\}$ and $Q, Q' \in \mathcal{Q}_j$, with $p_{Q,j} > 0$, $c_{Q,j} \leq c_{Q',j}$.*

Definition 2.4 *The expected cost $C(\mathcal{S})$ for a given system $\mathcal{S}$ of pure or mixed strategies is defined as the expected total latency incurred by $\mathcal{S}$, that is $C(\mathcal{S}) = E\left[\sum_{e \in E} f_e(l_e) l_e\right]$.*

We are interested in estimating the worst-case coordination ratio when Nash equilibrium exists. We denote the optimal system of pure strategies by $\mathcal{S}^*$.

Definition 2.5 (Coordination Ratio) *The coordination ratio is defined as $R = max_{\mathcal{S}} \frac{C(\mathcal{S})}{C(\mathcal{S}^*)}$, where the maximum is taken over all strategies $\mathcal{S}$ in Nash equilibrium.*

3 Nash Equilibrium for Linear Latency Functions

In this section we consider the case where the latency of each edge is linear in the edge congestion. Specifically $f_e(x) = a_e x + b_e$ for each edge $e \in E$, where a_e and b_e are nonnegative reals. We show that for linear latency functions the worst-case coordination ratio is exactly 2.618 for pure and mixed strategies.

3.1 Upper Bounds for Linear Latency Functions

We begin by proving an upper bound on the worst-case coordination ratio in the case of linear latency functions.

3.1.1 Pure Strategies

We start with the following Lemma.

Lemma 3.1 *Let $\vec{l}$, $\vec{l'}$ be any load vectors of the edges. Then*

$$\sum_{e\in E} a_e l_e l_e^{'} \leq \sqrt{\sum_{e\in E}(a_e l_e + b_e)l_e \sum_{e\in E}(a_e l'_e + b_e)l'_e}.$$

Proof: We use Cauchy-Schwartz inequality and get

$$\sum_{e\in E} a_e l_e l_e^{'} \leq \sqrt{\sum_{e\in E} a_e l_e^2 \sum_{e\in E} a_e {l'_e}^2} \leq \sqrt{\sum_{e\in E}(a_e l_e + b_e)l_e \sum_{e\in E}(a_e l'_e + b_e)l'_e}.$$

This completes the proof. ■

Theorem 3.1 *For linear latency functions and pure strategies the worst-case coordination ratio R is at most $\frac{3+\sqrt{5}}{2} \approx 2.618$.*

Proof: Let $\mathcal{Q}$ be the routes of system $\mathcal{S}$ of pure strategies in Nash equilibrium and let $\mathcal{Q}^*$ be the optimal routes for these linear latency functions. Let Q_j be the path of request j in $\mathcal{Q}$ and let Q_j^* be the path of request j in $\mathcal{Q}^*$. Recall that we denote by l the load vector of the system $\mathcal{S}$ in Nash Equilibrium and we denote by l^* the load vector of the global optimum. According to the definition of Nash Equilibrium 2.3 we have

$$\sum_{e\in Q_j} a_e l_e + b_e \leq \sum_{(e\in Q_j^*)\wedge(e\in Q_j)} a_e l_e + b_e + \sum_{(e\in Q_j^*)\wedge(e\notin Q_j)} a_e(l_e + w_j) + b_e \leq \sum_{e\in Q_j^*} a_e(l_e + w_j) + b_e.$$

We multiply the above inequality by w_j and get

$$\sum_{e\in Q_j}(a_e l_e + b_e)w_j \leq \sum_{e\in Q_j^*}(a_e l_e + b_e)w_j + a_e w_j^2.$$

We sum all the above inequalities for all j and get

$$\sum_j \sum_{e\in Q_j}(a_e l_e + b_e)w_j \leq \sum_j \sum_{e\in Q_j^*}(a_e l_e + b_e)w_j + a_e w_j^2.$$

Classifying the above sums according to the edges indices $J(e)$ for the left hand-side and $J^*(e)$ for the right hand-side yields

$$\sum_{e\in E}\sum_{j\in J(e)}(a_e l_e + b_e)w_j \leq \sum_{e\in E}\sum_{j\in J^*(e)}(a_e l_e + b_e)w_j + a_e w_j^2. \quad (2)$$

Note that for any $d \geq 1$ we have

$$\sum_{j\in J(e)} w_j = l_e, \qquad \sum_{\mathrm{j}\in \mathrm{J}^*(\mathrm{e})} \mathrm{w_j} = \mathrm{l_e^*}, \qquad \sum_{\mathrm{j}\in \mathrm{J}^*(\mathrm{e})} \mathrm{w_j^d} \leq \mathrm{l_e^{*d}}. \quad (3)$$

Substituting (3) in (2), we get

$$\sum_{e\in E}(a_e l_e + b_e)l_e \le \sum_{e\in E}(a_e l_e + b_e)l_e^* + a_e l_e^{*2} = \sum_{e\in E} a_e l_e l_e^* + \sum_{e\in E}(a_e l_e^* + b_e)l_e^*.$$

Applying Lemma 3.1 to the first term of the right-hand side of the inequality, we obtain

$$\sum_{e\in E}(a_e l_e + b_e)l_e \le \sqrt{\sum_{e\in E}(a_e l_e + b_e)l_e \sum_{e\in E}(a_e l_e^* + b_e)l_e^*} + \sum_{e\in E}(a_e l_e^* + b_e)l_e^*.$$

We denote the square root of the ratio of the Nash routes cost and the optimal routes cost by

$$x = \sqrt{\frac{\sum_{e\in E}(a_e l_e + b_e)l_e}{\sum_{e\in E}(a_e l_e^* + b_e)l_e^*}}.$$

Then, we divide the above inequality by $\sum_{e\in E}(a_e l_e^* + b_e)l_e^*$ and express the result in terms of x. Thus $x^2 \le x+1$ and hence $x^2 \le \frac{3+\sqrt{5}}{2}$, which completes the proof. ∎

The result of Theorem 3.1 is tight as we show in the lower bounds section. However for unweighted demand the result can be improved as follows.

Theorem 3.2 *For linear latency functions, unweighted demand and pure strategies the worst-case coordination ratio R is at most* 2.5.

Proof: According to inequality (2) obtained in the proof of Theorem 3.1 for weighted network congestion games with linear latency functions and pure strategies we have

$$\sum_{e\in E}\sum_{j\in J(e)}(a_e l_e + b_e)w_j \le \sum_{e\in E}\sum_{j\in J^*(e)}(a_e l_e + b_e)w_j + a_e w_j^2. \tag{4}$$

Substituting (3) in (4) and using the fact that $w_j = w_j^2 = 1$, we get

$$\sum_{e\in E}(a_e l_e + b_e)l_e \le \sum_{e\in E}(a_e l_e + b_e)l_e^* + a_e l_e^* = \sum_{e\in E} a_e l_e l_e^* + a_e l_e^* + \sum_{e\in E} b_e l_e^*. \tag{5}$$

The proof requires the following Lemma which properties appear in [25] and have simple proofs.

Lemma 3.2 *Let $i \ge 0, j \ge 0$ be integers. Then*

1. $ij = \frac{1}{3}j^2 + \frac{3}{4}i^2 - \frac{1}{3}(j - \frac{3}{2}i)^2$
2. $\frac{9}{8}i^2 + \frac{3}{2}i - \frac{1}{2}(j - \frac{3}{2}i)^2 \le \frac{5}{2}i^2$.

Applying property 1 of Lemma 3.2 to the first term of the right-hand side of inequality (5), we obtain

$$\sum_{e\in E}(a_e l_e + b_e)l_e \le \sum_{e\in E} a_e\left(\frac{1}{3}l_e^2 + \frac{3}{4}l_e^{*2} - \frac{1}{3}(l_e - \frac{3}{2}l_e^*)^2 + l_e^*\right) + b_e l_e^*,$$

and this is equivalent to

$$\sum_{e\in E}(a_el_e+\frac{3}{2}b_e)l_e \le \sum_{e\in E} a_e\left(\frac{9}{8}l_e^{*2}+\frac{3}{2}l_e^*-\frac{1}{2}(l_e-\frac{3}{2}l_e^*)^2\right)+\frac{3}{2}b_el_e^*.$$

Hence

$$\sum_{e\in E}(a_el_e+b_e)l_e \le \sum_{e\in E} a_e\left(\frac{9}{8}l_e^{*2}+\frac{3}{2}l_e^*-\frac{1}{2}(l_e-\frac{3}{2}l_e^*)^2\right)+\frac{3}{2}b_el_e^*.$$

Applying property 2 of Lemma 3.2 to the first term of the right-hand side of the inequality, we obtain

$$\sum_{e\in E}(a_el_e+b_e)l_e \le \sum_{e\in E}\frac{5}{2}a_el_e^{*2}+\frac{3}{2}b_el_e^* \le \frac{5}{2}(a_el_e^{*2}+b_el_e^*).$$

This completes the proof. ■

3.1.2 Mixed Strategies

Theorem 3.3 *For linear latency functions and mixed strategies the worst-case coordination ratio R is at most $\frac{3+\sqrt{5}}{2}\approx 2.618$.*

Proof: Let $\{p_{Q,j}\}$ be the probability distribution of the system $\mathcal{S}$ of mixed strategies. Let $\mathcal{Q}^*$ be the optimal routes for these linear latency functions. Let Q_j^* be the path of request j in $\mathcal{Q}^*$. According to definition 2.2, the expected latency of user j for assigning his request to path Q in system $\mathcal{S}$ is

$$c_{Q,j}=E[\sum_{e\in Q}a_e(l_e+(1-X_{Q,j})w_j)+b_e]=\sum_{e\in Q}a_e(E[l_e]+(1-p_{Q,j})w_j)+b_e.$$

According to definition 2.3, at Nash Equilibrium we have for any path Q with $p_{Q,j}>0$, the inequality $c_{Q,j}\le c_{Q_j^*,j}$. By substituting $c_{Q,j}$ and $c_{Q_j^*,j}$ in the inequality we have

$$\sum_{e\in Q}a_e(E[l_e]+(1-p_{Q,j})w_j)+b_e \le \sum_{e\in Q_j^*}a_e(E[l_e]+(1-p_{Q_j^*,j})w_j)+b_e \le \sum_{e\in Q_j^*}a_e(E[l_e]+w_j)+b_e.$$

We multiply the above inequality by $p_{Q,j}\cdot w_j$ and get

$$\sum_{e\in Q}(a_eE[l_e]+b_e)p_{Q,j}w_j+a_e(1-p_{Q,j})p_{Q,j}w_j^2 \le \sum_{e\in Q_j^*}a_e(E[l_e]p_{Q,j}w_j+p_{Q,j}w_j^2)+b_ep_{Q,j}w_j.$$

We sum all the above inequalities for all the paths Q for request j and classifying them according to the edges paths. This yields

$$\begin{aligned}\sum_{e\in E}(a_eE[l_e]+b_e)\sum_{Q|e\in Q}p_{Q,j}w_j+\sum_{e\in E}a_e\sum_{Q|e\in Q}(1-p_{Q,j})p_{Q,j}w_j^2&\\ \le \sum_{e\in Q_j^*}a_e(E[l_e]w_j+w_j^2)+b_ew_j&\qquad,\end{aligned}$$

where in the right-hand side we use the fact that $\sum_Q p_{Q,j} = 1$. We sum all the above inequalities for all j, exchange the order of summation in the left hand-side of the inequality and get

$$\sum_{e \in E}(a_e E[l_e] + b_e) \sum_j \sum_{Q|e \in Q} p_{Q,j} w_j + \sum_{e \in E} a_e \sum_j \sum_{Q|e \in Q} (p_{Q,j} - p_{Q,j}^2) w_j^2 \\ \leq \sum_j \sum_{e \in Q_j^*} a_e(E[l_e] w_j + w_j^2) + b_e w_j. \quad (6)$$

The proof requires the following Lemmas.

Lemma 3.3 *For any system of mixed strategies:* $E[l_e^2] - (E[l_e])^2 = \sum_j p_{e,j}(1 - p_{e,j}) w_j^2$.

Proof: We have

$$E[l_e^2] - (E[l_e])^2 = Var[l_e] = \sum_j p_{e,j}(1 - p_{e,j}) w_j^2.$$

The first equality follows from the definition of the variance. The second equality follows from the linearity of expectation and the independence of the indicator random variables $X_{e,j}$ and $X_{e,k}$ for $j \neq k$. This completes the proof of the Lemma. ■

The proof of the following Lemma appears in Appendix A.1.

Lemma 3.4 *For any system of mixed strategies:*

$$\sum_{e \in E} \sum_j (a_e E[l_e] + b_e) \sum_{Q|e \in Q} p_{Q,j} w_j + \sum_{e \in E} a_e \sum_j \sum_{Q|e \in Q} (p_{Q,j} - p_{Q,j}^2) w_j^2 \geq \sum_{e \in E} a_e E[l_e^2] + \sum_{e \in E} b_e E[l_e].$$

The proof of the following Lemma appears in Appendix A.2.

Lemma 3.5 *For any system of mixed strategies and for the global optimal solution:*

$$\sum_j \sum_{e \in Q_j^*} a_e(E[l_e] w_j + w_j^2) + b_e w_j \leq \sqrt{\sum_{e \in E} a_e E[l_e^2] + b_e E[l_e]} \sqrt{\sum_{e \in E} a_e l_e^{*2} + b_e l_e^*} + \sum_{e \in E}(a_e l_e^* + b_e) l_e^*.$$

To complete the proof of Theorem 3.3 we apply Lemma 3.4 and Lemma 3.5 to inequality (6) and obtain

$$\sum_{e \in E} a_e E[l_e^2] + \sum_{e \in E} b_e E[l_e] \leq \sqrt{\sum_{e \in E} a_e E[l_e^2] + b_e E[l_e]} \sqrt{\sum_{e \in E} a_e l_e^{*2} + b_e l_e^*} + \sum_{e \in E}(a_e l_e^* + b_e) l_e^*.$$

We denote the square root of the ratio of the Nash routes cost and the optimal routes cost by

$$x = \sqrt{\frac{\sum_{e \in E} a_e E[l_e^2] + b_e E[l_e]}{\sum_{e \in E}(a_e l_e^* + b_e) l_e^*}}.$$

Then we divide the above inequality by $\sum_{e \in E}(a_e l_e^* + b_e) l_e^*$ and express the result in terms of x. Thus, $x^2 \leq x + 1$ and therefore $x^2 \leq \frac{3+\sqrt{5}}{2}$, which completes the proof of the Theorem. ■

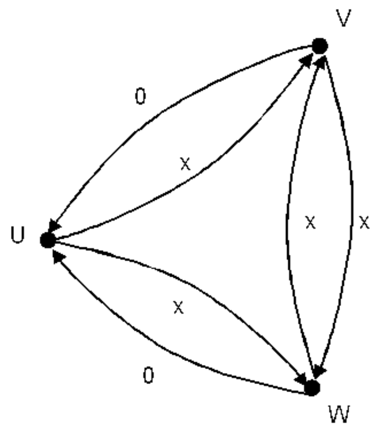

Figure 1: A network congestion game

3.2 Lower Bounds for Linear Latency Functions

In this section we prove a lower bound on the worst-case coordination ratio in the case of linear latency functions.

Theorem 3.4 *For pure strategies and linear latency functions,* $R \geq 2.618$.

Proof: Let denote the golden ratio by $\phi = \frac{1+\sqrt{5}}{2}$. We construct the following example which uses the network shown in Figure 1. We consider an atomic weighted network congestion game with four players. Player 1 has a bandwidth request (U, V, ϕ) (player 1 has to move ϕ unit of bandwidth from U to V), player 2 has a bandwidth request (U, W, ϕ), player 3 has a bandwidth request $(V, W, 1)$ and player 4 has a bandwidth request $(W, V, 1)$. In the optimal solution player 1 uses the path UV, player 2 uses the path UW, player 3 uses the path VW and player 4 uses the path WV. The costs of the players are $\phi^2, \phi^2, 1$ and 1 respectively. The total cost is $2\phi^2 + 2$. If player 1 chooses the path UWV, player 2 chooses the path UVW, player 3 uses the path VUW and player 4 uses the path WUV, then this is a Nash equilibrium and the costs of the players are $\phi(2\phi + 1), \phi(2\phi + 1), (\phi + 1)$ and $(\phi + 1)$ respectively. The total cost is $4\phi^2 + 4\phi + 2$. Thus $R = (4\phi^2 + 4\phi + 2)/(2\phi^2 + 2) = (8\phi + 6)/(2\phi + 4) = 1 + \phi \approx 2.618$, by the fact that $\phi^2 = \phi + 1$. This completes the proof. ■

Theorem 3.5 *For pure strategies, unweighted demand and linear latency functions,* $R \geq 2.5$.

Proof: We construct the same example as in the proof of Theorem 3.4 with the modification that here we set $\phi = 1$. We obtain the same optimal solution and the same Nash equilibrium as in the proof of Theorem 3.4. Thus $R = (4\phi^2 + 4\phi + 2)/(2\phi^2 + 2) = 10/4 = 2.5$. This completes the proof. ■

4 Nash Equilibrium for Polynomial Latency Functions

In this section we consider the case where the latency of each edge is a polynomial of degree d in the edge congestion. Specifically $f_e(x) = \sum_i a_{e,i} x^i$ for each edge $e \in E$, where $a_{e,i}$ are nonnegative reals. We show that for polynomials of degree d latency functions the worst-case coordination ratio is at most $O(2^d d^{d+1})$ for pure and mixed strategies and we show that for polynomials of degree d latency functions the worst-case coordination ratio is lower bounded by $\Omega(d^{d/2})$.

4.1 Upper Bounds for Polynomial Latency Functions

We begin by proving an upper bound on the worst-case coordination ratio in the case of polynomial of degree d latency functions.

4.1.1 Pure Strategies

The proof of the following Theorem appears in Appendix A.3.

Theorem 4.1 *For polynomial latency functions of degree* d *and pure strategies, we have* $R = O(2^d d^{d+1})$.

4.1.2 Mixed Strategies

The proof of the following Theorem appears in Appendix A.4.

Theorem 4.2 *For polynomial latency functions of degree* d *and mixed strategies, we have* $R = O(2^d d^{d+1})$.

4.2 Lower Bounds for Polynomial Latency Functions

In this section we prove a lower bound on the worst-case coordination ratio in the case of polynomial of degree d latency functions. The proof of the following Theorem appears in Appendix A.5.

Theorem 4.3 *For pure strategies and polynomial latency functions of degree* d, $R = \Omega(d^{d/2})$.

References

[1] B. Awerbuch, Y. Azar, E. Grove, M. Kao, P. Krishnan, and J. Vitter. Load balancing in the l_p norm. In *Proc. of the 36th Annual Symposium on Foundations of Computer Science (FOCS95)*, pages 383–391, 1995.

[2] B. Awerbuch, Y. Azar, Y. Richter, and D. Tsur. Tradeoffs in worst-case equilibria. In *Proceedings of 1st WAOA*, pages 41–52, 2003.

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

[4] G. Christodoulou and E. Koutsoupias. The price of anarchy of finite congestion games. In *Proc. 37th ACM Symp. on Theory of Computing*, 2005. To appear.

[5] R. Cole, Y. Dodis, and T. Roughgarden. How much can taxes help selfish routing? In *Proceedings of the 4th ACM conference on Electronic commerce*, pages 98–107, 2003.

[6] R. Cole, Y. Dodis, and T. Roughgarden. Pricing network edges for heterogeneous selfish users. In *Proc. 35th ACM Symp. on Theory of Computing*, pages 521–530, 2003.

[7] A. Czumaj, P. Krysta, and B. Vöcking. Selfish traffic allocation for server farms. In *Proc. 34th ACM Symp. on Theory of Computing*, pages 287–296, 2002.

[8] A. Czumaj and B. Vöcking. Tight bounds for worst-case equilibria. In *Proc. 13rd ACM-SIAM Symp. on Discrete Algorithms*, pages 413–420, 2002.

[9] P. Dubey. Inefficiency of nash equilibria. *Mathematics of Operations Research*, 11(1):1–8, 1986.

[10] A. Fabrikant, Christos Papadimitriou, and Kunal Talwar. The complexity of pure equilibria. In *Proc. 36th ACM Symp. on Theory of Computing*, pages 604–612, 2004.

[11] D. Fotakis, S. Kontogiannis, and P. Spirakis. Selfish unsplittable flows. In *International Colloquium on Automata, Languages and Programming - ICALP '04*. to appear.

[12] M. Gairing, T. Lücking, M. Mavronicolas, and B. Monien. The price of anarchy for polynomial social cost. In *Proceedings of the 29th International Symposium on Mathematical Foundations of Computer Science (MFCS 2004)*, pages 574–585, 2004.

[13] M. Gairing, T. Lücking, M. Mavronicolas, B. Monien, and M. Rode. Nash equilibria in discrete routing games with convex latency functions. In *Proc. of the 31st International Colloquium on Automata, Languages and Programming (ICALP 2004)*, pages 645–657, 2004.

[14] E. Koutsoupias and C.H. Papadimitriou. Worst-case equilibria. In *Proc. 16th Symp. on Theoretical Aspects of Comp. Science*, pages 404–413, 1999.

[15] T. Lücking, M. Mavronicolas, B. Monien, and M. Rode. A new model for selfish routing. In *Proc. of the 21st International Symposium on Theoretical Aspects of Computer Science (STACS 2004)*, pages 547–558, 2004.

[16] M. Mavronicolas and P. Spirakis. The price of selfish routing. In *Proc. 33rd ACM Symp. on Theory of Computing*, pages 510–519, 2001.

[17] J. F. Nash. Equilibrium points in n-person games. In *Proceedings of National Academy of Sciences*, volume 36, pages 48–49, 1950.

[18] G. Owen. *Game Theory*. Academic Press, third edition, 1995.

[19] C.H. Papadimitriou. Algorithms, games and the internet. In *Proc. 33rd ACM Symp. on Theory of Computing*, pages 749–753, 2001.

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

[21] T. Roughgarden. Selfish routing with atomic players. In *SODA 2005*. to appear.

[22] T. Roughgarden. Designing networks for selfish users is hard. In *Proc. 42nd IEEE Symp. on Found. of Comp. Science*, pages 472–481, 2001.

[23] T. Roughgarden. The price of anarchy is independent of the network topology. In *Proc. 34th ACM Symp. on Theory of Computing*, pages 428–437, 2002.

[24] T. Roughgarden and É. Tardos. How bad is selfish routing. In *Proc. 41st IEEE Symp. on Found. of Comp. Science*, pages 93–102, 2000.

[25] S. Suri, C. Toth, and Y. Zhou. Selfish load balancing and atomic congestion games. In *ACM SPAA '04 (Symp. on Parallelism in Algorithms and Architectures)*, 2004.

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

Appendix

A Omitted proofs

A.1 Proof of Lemma 3.4

We have

$$
\begin{aligned}
&\sum_{e\in E}\sum_{j}(a_eE[l_e]+b_e)\sum_{Q|e\in Q}p_{Q,j}w_j+\sum_{e\in E}a_e\sum_{j}\sum_{Q|e\in Q}(p_{Q,j}-p_{Q,j}^2)w_j^2\\
&\quad\geq\ \sum_{e\in E}\sum_{j}(a_eE[l_e]+b_e)p_{e,j}w_j+\sum_{e\in E}a_e\sum_{j}(p_{e,j}-p_{e,j}^2)w_j^2\\
&\quad=\ \sum_{e\in E}a_e\sum_{j}E[l_e]p_{e,j}w_j+(p_{e,j}-p_{e,j}^2)w_j^2+\sum_{e\in E}b_e\sum_{j}p_{e,j}w_j\\
&\quad=\ \sum_{e\in E}a_e\left((E[l_e])^2+\sum_{j}p_{e,j}(1-p_{e,j})w_j^2\right)+\sum_{e\in E}b_eE[l_e]\\
&\quad=\ \sum_{e\in E}a_eE[l_e^2]+\sum_{e\in E}b_eE[l_e],
\end{aligned}
$$

where the first inequality follows from the fact that $p_{e,j}=\sum_{Q|e\in Q}p_{Q,j}$. The second equality follows from the fact that $E(l_e)=\sum_j p_{e,j}w_j$. The last equality follows from Lemma 3.3. This completes the proof of the Lemma.

A.2 Proof of Lemma 3.5

Classifying the sum of the optimal paths according to the edges indices $J^*(e)$, we get

$$
\sum_{j}\sum_{e\in Q_j^*}a_e(E[l_e]w_j+w_j^2)+b_ew_j=\sum_{e\in E}a_e\sum_{j\in J^*(e)}(E[l_e]w_j+w_j^2)+b_ew_j. \tag{7}
$$

Substituting (3) in (7), we obtain

$$
\begin{aligned}
&\sum_{e\in E}a_e\sum_{j\in J^*(e)}(E[l_e]w_j+w_j^2)+b_ew_j\\
&\quad\leq\ \sum_{e\in E}a_e(E[l_e]l_e^*+l_e^{*2})+b_el_e^*\\
&\quad=\ \sum_{e\in E}a_eE[l_e]l_e^*+\sum_{e\in E}(a_el_e^*+b_e)l_e^*\\
&\quad\leq\ \sqrt{\sum_{e\in E}a_e(E[l_e])^2\sum_{e\in E}a_el_e^{*2}}+\sum_{e\in E}(a_el_e^*+b_e)l_e^*\\
&\quad\leq\ \sqrt{\sum_{e\in E}a_eE[l_e^2]+b_eE[l_e]}\sqrt{\sum_{e\in E}a_el_e^{*2}+b_el_e^*}+\sum_{e\in E}(a_el_e^*+b_e)l_e^*,
\end{aligned}
$$

where the second inequality follows from Lemma 3.1. The last inequality follows from the fact that $E[l_e^2]\geq(E[l_e])^2$. This completes the proof of the Lemma.

A.3 Proof of Theorem 4.1

Let $\mathcal{Q}$ be the routes of system $\mathcal{S}$ of pure strategies in Nash equilibrium and let $\mathcal{Q}^*$ be the optimal routes for these polynomial latency functions. Let Q_j be the path of request j in $\mathcal{Q}$ and let Q_j^* be the path of request j in $\mathcal{Q}^*$. According to definition 2.3 of Nash Equilibrium we have [1]

$$\begin{aligned}\sum_{e\in Q_j}\sum_i a_{e,i}l_e^i &\le \sum_{(e\in Q_j^*)\wedge(e\in Q_j)}\sum_i a_{e,i}l_e^i + \sum_{(e\in Q_j^*)\wedge(e\notin Q_j)}\sum_i a_{e,i}(l_e+w_j)^i \\ &\le \sum_{e\in Q_j^*}\sum_i a_{e,i}(l_e+w_j)^i.\end{aligned}$$

We multiply the above inequality by w_j and get

$$\sum_{e\in Q_j}\sum_i a_{e,i}l_e^i w_j \le \sum_{e\in Q_j^*}\sum_i a_{e,i}(l_e+w_j)^i w_j.$$

Next, we sum all the above inequalities for all j to get

$$\sum_j\sum_{e\in Q_j}\sum_i a_{e,i}l_e^i w_j \le \sum_j\sum_{e\in Q_j^*}\sum_i a_{e,i}(l_e+w_j)^i w_j.$$

Classifying the above sums according to the edges indices $J(e)$ for the left hand-side and $J^*(e)$ for the right hand-side yields

$$\sum_{e\in E}\sum_i\sum_{j\in J(e)} a_{e,i}l_e^i w_j \le \sum_{e\in E}\sum_i\sum_{j\in J^*(e)} a_{e,i}(l_e+w_j)^i w_j. \tag{8}$$

The proof requires the following Lemma which appears in [1] and has simple proof.

Lemma A.1 *The function $f(x,y)=(x+y)^d$ is bounded by: $cx^d+\left(y\left(\frac{d}{\ln c}+1\right)\right)^d$ for any $c>1$.*

We also need the following Lemmas.

Lemma A.2 *Let $f_e(x)=\sum_i a_{e,i}x^i$ be the polynomial of degree d. Then for any system of pure strategy:*

$$\sum_{e\in E}\sum_i\sum_{j\in J(e)} a_{e,i}l_e^i w_j = \sum_{e\in E} f_e(l_e)l_e.$$

Proof: We have

$$\sum_{e\in E}\sum_i\sum_{j\in J(e)} a_{e,i}l_e^i w_j = \sum_{e\in E}\sum_i a_{e,i}l_e^{i+1} = \sum_{e\in E} f_e(l_e)l_e,$$

where the first equality follows from the fact that $l_e=\sum_{j\in J(e)}w_j$ and the second equality follows from the definition of the polynomial $f_e(x)$. This completes the proof of the Lemma. ■

[1] To simplify the notation throughout the entire paper we assume $0^0=0$.

Lemma A.3 *Let $f_e(x) = \sum_i a_{e,i} x^i$ be the polynomial of degree d and let $c > 1$. Then for any system of pure strategies and for the global optimal solution:*

$$\sum_{e \in E} \sum_i \sum_{j \in J^*(e)} a_{e,i}(l_e + w_j)^i w_j \leq c(d+1)(\sum_{e \in E} f_e(l_e) l_e)^{d/(d+1)} (\sum_{e \in E} f_e(l_e^*) l_e^*)^{1/(d+1)}.$$

Proof: We have

$$\begin{aligned}
\sum_{e \in E} \sum_i \sum_{j \in J^*(e)} a_{e,i}(l_e + w_j)^i w_j &= \sum_{e \in E} \sum_i a_{e,i} \sum_{j \in J^*(e)} (l_e + w_j)^i w_j \\
&\leq \sum_{e \in E} \sum_i a_{e,i} \sum_{j \in J^*(e)} \left(c l_e^i w_j + \left(\frac{i}{\ln c} + 1 \right)^i w_j^{i+1} \right) \\
&\leq \sum_{e \in E} \sum_i a_{e,i} \left(c l_e^i l_e^* + \left(\frac{d}{\ln c} + 1 \right)^d l_e^{*i+1} \right) \\
&= c \sum_{e \in E} \sum_i a_{e,i} l_e^i l_e^* + \left(\frac{d}{\ln c} + 1 \right)^d \sum_{e \in E} \sum_i a_{e,i} l_e^{*i+1},
\end{aligned}$$

where the first inequality follows from Lemma A.1. The second inequality follows from (3) and the fact that the function $\left(\frac{x}{\ln c} + 1\right)^x$ is an increasing function for $x \geq 0$. We apply Holder's inequality:

$$\sum_i a_i^\alpha b_i^\beta \leq \left(\sum_i a_i \right)^\alpha \left(\sum_i b_i \right)^\beta, \tag{9}$$

for $\alpha + \beta = 1$. We use $a_i = a_{e,i} l_e^{i+1}$, $b_i = a_{e,i} l_e^{*i+1}$, $\alpha = i/(i+1)$, and $\beta = 1/(i+1)$. This yields

$$\begin{aligned}
&\sum_{e \in E} \sum_i \sum_{j \in J^*(e)} a_{e,i}(l_e + w_j)^i w_j \\
&\leq c \sum_i (\sum_{e \in E} a_{e,i} l_e^{i+1})^{i/(i+1)} (\sum_{e \in E} a_{e,i} l_e^{*i+1})^{1/(i+1)} + \left(\frac{d}{\ln c} + 1 \right)^d \sum_{e \in E} \sum_i a_{e,i} l_e^{*i+1} \\
&\leq c \sum_i (\sum_{e \in E} \sum_k a_{e,k} l_e^{k+1})^{i/(i+1)} (\sum_{e \in E} \sum_k a_{e,k} l_e^{*k+1})^{1/(i+1)} + \left(\frac{d}{\ln c} + 1 \right)^d \sum_{e \in E} \sum_i a_{e,i} l_e^{*i+1} \\
&= c \sum_i (\sum_{e \in E} f_e(l_e) l_e)^{i/(i+1)} (\sum_{e \in E} f_e(l_e^*) l_e^*)^{1/(i+1)} + \left(\frac{d}{\ln c} + 1 \right)^d \sum_{e \in E} f_e(l_e^*) l_e^*
\end{aligned}$$

where the equality follows from the definition of the polynomial $f_e(x)$. Next, we use the fact that for $x \geq y > 0$ and $1 \geq \alpha \geq \alpha' \geq 0$ we have

$$x^\alpha y^{1-\alpha} \geq x^{\alpha'} y^{1-\alpha'}.$$

We apply it for $x = C(\mathcal{S}) = \sum_{e \in E} f_e(l_e) l_e$ and $y = C(\mathcal{S}^*) = \sum_{e \in E} f_e(l_e^*) l_e^*$ to get

$$
\begin{aligned}
&\sum_{e\in E}\sum_i \sum_{j\in J^*(e)} a_{e,i}(l_e+w_j)^i w_j \\
&\leq \quad c\sum_i (\sum_{e\in E} f_e(l_e)l_e)^{d/(d+1)}(\sum_{e\in E} f_e(l_e^*)l_e^*)^{1/(d+1)} + \left(\frac{d}{\ln c}+1\right)^d \sum_{e\in E} f_e(l_e^*)l_e^* \\
&= \quad c(d+1)(\sum_{e\in E} f_e(l_e)l_e)^{d/(d+1)}(\sum_{e\in E} f_e(l_e^*)l_e^*)^{1/(d+1)} + \left(\frac{d}{\ln c}+1\right)^d \sum_{e\in E} f_e(l_e^*)l_e^*.
\end{aligned}
$$

This completes the proof of the Lemma. ■

To complete the proof of Theorem 4.1 we Apply Lemma A.2 to the left-hand side of (8) and we apply Lemma A.3 to the right-hand side of (8). This yields

$$
\sum_{e\in E} f_e(l_e)l_e \leq c(d+1)(\sum_{e\in E} f_e(l_e)l_e)^{d/(d+1)}(\sum_{e\in E} f_e(l_e^*)l_e^*)^{1/(d+1)} + \left(\frac{d}{\ln c}+1\right)^d \sum_{e\in E} f_e(l_e^*)l_e^*.
$$

Let

$$
x = \frac{(\sum_{e\in E} f_e(l_e)l_e))^{1/d}}{(\sum_{e\in E} f_e(l_e^*)l_e^*))^{1/d}}.
$$

Then we divide the above inequality by $\sum_{e\in E} f_e(l_e^*)l_e^*$ and express the result in terms of x. Thus,

$$
x^{d+1} \leq c(d+1)x^d + \left(\frac{d}{\ln c}+1\right)^d.
$$

Next we divide by x^d to get

$$
x \leq c(d+1) + \left(\frac{\frac{d}{\ln c}+1}{x}\right)^d.
$$

It is easy to see that for $c = 2-\epsilon$ we get $x \leq 2(d+1)$. Hence $\frac{C(\mathcal{Q})}{C(\mathcal{Q}^*)} = x^{d+1} = O(2^d d^{d+1})$. This completes the proof of the Theorem.

A.4 Proof of Theorem 4.2

Let $\{p_{Q,j}\}$ be the probability distribution of the system $\mathcal{S}$ of mixed strategies. Let $\mathcal{Q}^*$ be the optimal routes for these polynomial latency functions. Let Q_j^* be the path of request j in $\mathcal{Q}^*$.

According to definition (2.2), the expected latency of user j for assigning his request to path Q in system $\mathcal{S}$ is

$$
c_{Q,j} = \sum_{e\in Q} E[f_e(l_e + (1-X_{Q,j})w_j)]
$$

According to definition 2.3, at Nash Equilibrium we have for any path Q with $p_{Q,j} > 0$, the inequality $c_{Q,j} \leq c_{Q_j^*,j}$. By substituting $c_{Q,j}$ and $c_{Q_j^*,j}$ in the inequality we have

$$
\begin{aligned}
\sum_{e\in Q} E[f_e(l_e + (1-X_{Q,j})w_j)] \quad &\leq \quad \sum_{e\in Q_j^*} E[f_e(l_e + (1-X_{Q_j^*,j})w_j)] \\
&\leq \quad \sum_{e\in Q_j^*} E[f_e(l_e + w_j)].
\end{aligned}
$$

We multiply the above inequality by $p_{Q,j} \cdot w_j$ and get

$$\sum_{e \in Q} E[f_e(l_e + (1 - X_{Q,j})w_j)]p_{Q,j}w_j \leq \sum_{e \in Q_j^*} E[f_e(l_e + w_j)]p_{Q,j}w_j.$$

Then, we sum all the above inequalities for all the paths Q and classify them according to the edges paths. This Yields

$$\sum_{e \in E} \sum_{Q | e \in Q} E[f_e(l_e + (1 - X_{Q,j})w_j)]p_{Q,j}w_j \quad \leq \quad \sum_{e \in Q_j^*} E[f_e(l_e + w_j)]w_j,$$

where in the right-hand side of the inequality we use the fact that $\sum_Q p_{Q,j} = 1$. We sum all the above inequalities for all j and get

$$\sum_j \sum_{e \in E} \sum_{Q | e \in Q} E[f_e(l_e + (1 - X_{Q,j})w_j)]p_{Q,j}w_j \quad \leq \quad \sum_j \sum_{e \in Q_j^*} E[f_e(l_e + w_j)]w_j. \tag{10}$$

The proof requires the following Lemmas.

Lemma A.4 *For any system of mixed strategies:*

$$E[f_e(l_e)l_e] = \sum_j E[f_e(l_e + (1 - X_{e,j})w_j)]p_{e,j}w_j.$$

Proof: We have

$$\begin{aligned}
E[f_e(l_e)l_e] &= E[(\sum_j X_{e,j}w_j)f_e(l_e)] = \sum_j w_j E[X_{e,j}f_e(l_e)] \\
&= \sum_j w_j Pr[X_{e,j} = 0]E[X_{e,j}f_e(l_e)|X_{e,j} = 0] + w_j Pr[X_{e,j} = 1]E[X_{e,j}f_e(l_e)|X_{e,j} = 1] \\
&= \sum_j p_{e,j}w_j E[X_{e,j}f_e(l_e)|X_{e,j} = 1] \\
&= \sum_j E[f_e(l_e + (1 - X_{e,j})w_j)]p_{e,j}w_j,
\end{aligned}$$

where the second equality follows from the linearity of expectation and the fourth equality follows from the fact that $p_{e,j} = Pr[X_{e,j} = 1]$. This completes the proof of the Lemma. ■

Lemma A.5 *For any system of mixed strategies:*

$$\sum_j \sum_{e \in E} \sum_{Q | e \in Q} E[f_e(l_e + (1 - X_{Q,j})w_j)]p_{Q,j}w_j \geq \sum_{e \in E} E[f_e(l_e)l_e]$$

Proof: We have

$$\begin{aligned}
\sum_j \sum_{e\in E} \sum_{Q|e\in Q} E[f_e(l_e+(1-X_{Q,j})w_j)]p_{Q,j}w_j &\geq \sum_j \sum_{e\in E} \sum_{Q|e\in Q} E[f_e(l_e+(1-X_{e,j})w_j)]p_{Q,j}w_j \\
&= \sum_j \sum_{e\in E} E[f_e(l_e+(1-X_{e,j})w_j)]p_{e,j}w_j \\
&= \sum_{e\in E} \sum_j E[f_e(l_e+(1-X_{e,j})w_j)]p_{e,j}w_j \\
&= \sum_{e\in E} E[f_e(l_e)l_e].
\end{aligned}$$

The inequality follows from the fact that for every edge $e \in Q$, it holds $X_{e,j} \geq X_{Q,j}$. The first equality follows from the fact that $p_{e,j} = \sum_{Q|e\in Q} p_{Q,j}$. The last equality follows from Lemma A.4. This completes the proof of the Lemma. ∎

Lemma A.6 *Let $c > 1$. Then for any system of mixed strategies and for the global optimal solution:*

$$\sum_j \sum_{e\in Q_j^*} E[f_e(l_e+w_j)]w_j \leq c(d+1)(\sum_{e\in E} E[f_e(l_e)l_e])^{d/(d+1)}(\sum_{e\in E} f_e(l_e^*)l_e^*)^{1/(d+1)} + \left(\frac{d}{\ln c}+1\right)^d \sum_{e\in E} f_e(l_e^*)l_e^*.$$

Proof: To simplify the proof we first consider only the polynomial latency functions of degree d $f_e(x) = x^d$. Then we extend it to the general case. We have

$$\begin{aligned}
\sum_j \sum_{e\in Q_j^*} E[(l_e+w_j)^d]w_j &= \sum_{e\in E} \sum_{j\in J^*(e)} E[(l_e+w_j)^d]w_j \\
&\leq \sum_{e\in E} \sum_{j\in J^*(e)} \left(cE[l_e^d]w_j + \left(\frac{d}{\ln c}+1\right)^d w_j^{d+1}\right) \\
&\leq \sum_{e\in E} \left(cE[l_e^d]l_e^* + \left(\frac{d}{\ln c}+1\right)^d l_e^{*d+1}\right) \\
&= c\sum_{e\in E} E[l_e^d]l_e^* + \left(\frac{d}{\ln c}+1\right)^d \sum_{e\in E} l_e^{*d+1}.
\end{aligned}$$

The first equality is obtained by classifying the sums according to the edges indices $J^*(e)$. The first inequality follows from Lemma A.1 and from the linearity of expectation. The second inequality follows from (3). We apply Holder's inequality (9), where we use $a_i = E[l_e^d]^{(d+1)/d}$, $b_i = l_e^{*d+1}$, $\alpha = d/(d+1)$, and $\beta = 1/(d+1)$. This yields

$$\begin{aligned}
&\sum_j \sum_{e\in Q_j^*} E[(l_e+w_j)^d]w_j \\
&\leq c(\sum_{e\in E} E[l_e^d]^{(d+1)/d})^{d/(d+1)}(\sum_{e\in E} l_e^{*d+1})^{1/(d+1)} + \left(\frac{d}{\ln c}+1\right)^d \sum_{e\in E} l_e^{*d+1} \\
&\leq c(\sum_{e\in E} E[l_e^{d+1}])^{d/(d+1)}(\sum_{e\in E} l_e^{*d+1})^{1/(d+1)} + \left(\frac{d}{\ln c}+1\right)^d \sum_{e\in E} l_e^{*d+1},
\end{aligned}$$

where the last inequality follows from the fact that $E[l_e^{d+1}] \geq E[l_e^d]^{(d+1)/d}$. Since we assumed that $f_e(l_e) = l_e^d$ then the last inequality is

$$\sum_j \sum_{e \in Q_j^*} E[f_e(l_e{+}w_j)]w_j \leq c(\sum_{e\in E} E[f_e(l_e)l_e])^{d/(d+1)}(\sum_{e\in E} f_e(l_e^*)l_e^*)^{1/(d+1)}{+}\left(\frac{d}{\ln c}+1\right)^d \sum_{e\in E} f_e(l_e^*)l_e^*.$$

Now we return to the case of general polynomials $f_e(x) = \sum_i a_{e,i}x^i$ of degree d for the latency functions. By the linearity of expectation and by the same considerations we used in Theorem 4.1, we obtain

$$\sum_j \sum_{e \in Q_j^*} E[f_e(l_e{+}w_j)]w_j \leq c(d{+}1)(\sum_{e\in E} E[f_e(l_e)l_e])^{d/(d+1)}(\sum_{e\in E} f_e(l_e^*)l_e^*)^{1/(d+1)}{+}\left(\frac{d}{\ln c}+1\right)^d \sum_{e\in E} f_e(l_e^*)l_e^*.$$

This completes the proof of the Lemma. ■

To complete the proof of Theorem 4.2 we apply Lemma A.5 to the left-hand side of (10) and we apply Lemma A.6 to the right-hand side of (10). This yields

$$\sum_{e\in E} E[f_e(l_e)l_e] \leq c(d{+}1)(\sum_{e\in E} E[f_e(l_e)l_e])^{d/(d+1)}(\sum_{e\in E} f_e(l_e^*)l_e^*)^{1/(d+1)} + \left(\frac{d}{\ln c}+1\right)^d \sum_{e\in E} f_e(l_e^*)l_e^*.$$

Let

$$x = \frac{(\sum_{e\in E} E[f_e(l_e)l_e])^{1/d}}{(\sum_{e\in E} f_e(l_e^*)l_e^*))^{1/d}}.$$

Then we divide the above inequality by $\sum_{e\in E} f_e(l_e^*)l_e^*$ and express the result in terms of x. Thus,

$$x^{d+1} \leq c(d+1)x^d + \left(\frac{d}{\ln c}+1\right)^d$$

and hence as in Theorem 4.1 we obtain $R = O(2^d d^{d+1})$. This completes the proof of the Theorem.

A.5 Proof of Theorem 4.3

We use the construction given for the maximum latency in [2] and apply it for the average latency. This construction can be easily modelled as a directed graph. We construct the following problem instance for the restricted assignment model. Let $T > 0$ and let $l > 0$ be large enough. We consider $l+1$ groups of links such that in group $k = 0, \ldots, l$ there are $\frac{T}{k!}$ links. Denote the number of links in group k by n_k. For each link we consider the latency function $f(x) = x^d$. We partition the tasks to l groups. In group $k = 1, \ldots, l$ there are $k \cdot n_k$ unit jobs, each can be assigned to any link from groups $[k-1, \ldots, l]$. Observe that the optimal solution assigns the jobs of group k ($k = 1, \ldots, l$) to the links in group $k-1$, one job per link. We define the following system of pure strategies, denote it by S: all jobs from group k ($k = 1, \ldots, l$) are assigned to links from group k, k jobs per link. We choose $T = l!$ to maintain integrality.

Lemma A.7 *The system $\mathcal{S}$ is in Nash Equilibrium.*

Proof: Denote by S_k $(k = 1, \ldots, l)$ the set of links to which jobs from job group k can be assigned. Let j be a job from group k and consider the assignment of j to link i from link group k. Clearly, $c_{ij} = k^d$, and for each link $t \in S_k$ we have $c_{tj} \geq ((k-1)+1)^d = k^d = c_{ij}$. Hence the system $\mathcal{S}$ is in Nash Equilibrium. This completes the proof. ■

We assume that $l > d$. We denote the optimal solution cost by OPT. For large enough l we have

$$OPT = \sum_{k=0}^{l-1} \frac{T}{k!} \cdot 1^d = T \sum_{k=0}^{l-1} \frac{1}{k!} \approx T \cdot e.$$

and

$$C(S) = \sum_{k=1}^{l} \frac{T}{k!} \cdot k^d \geq \frac{T}{\ulcorner d/2 \urcorner !} \cdot (\ulcorner d/2 \urcorner)^d = T \cdot \Omega(d^{d/2}),$$

where the second equality follows from Stirling's formula. Hence

$$R \geq \frac{C(\mathcal{S})}{OPT} = \Omega(d^{d/2}).$$

This completes the proof.