Alcom-FT Technical Report Series
ALCOMFT-TR-03-141

The Structure and Complexity of Extreme Nash Equilibria*

Martin Gairing† *Thomas Lücking*† *Marios Mavronicolas*‡ *Burkhard Monien*†
Paul Spirakis§

(June 18, 2003)

*This work has been partially supported by the IST Program of the European Union under contract numbers IST-1999-14186 (ALCOM-FT) and IST-2001-33116 (FLAGS). by funds from the Joint Program of Scientific and Technological Collaboration between Greece and Cyprus, and by research funds at University of Cyprus.

†Faculty of Computer Science, Electrical Engineering and Mathematics, University of Paderborn, Fürstenallee 11, 33102 Paderborn, Germany. Email: {gairing,luck,bm}@uni-paderborn.de

‡Department of Computer Science, University of Cyprus, 1678 Nicosia, Cyprus. Email: mavronic@ucy.ac.cy

§Computer Technology Institute, P. O. Box 1122, 261 10 Patras, Greece, & Department of Computer Engineering and Informatics, University of Patras, Rion, 265 00 Patras, Greece. Email: spirakis@cti.gr

Abstract

We study the combinatorial structure and computational complexity of *extreme* Nash equilibria, ones that maximize or minimize a certain objective function, in the context of a *selfish routing* game. Specifically, we assume a collection of n *users,* each employing a *mixed strategy,* which is a probability distribution over m parallel *links,* to control the routing of its own assigned *traffic.* In a *Nash equilibrium,* each user selfishly routes its traffic on those links that minimize its *expected latency cost.* The *social cost* of a Nash equilibrium is the expectation, over all random choices of the users, of the maximum, over all links, *latency* through a link. We are interested in computing (exactly or approximately) the social cost of the *worst* and the *best* Nash equilibrium – the ones attaining the maximum (*worst*) and the minimum (*best*) social cost, respectively. We provide a comprehensive collection of structural and complexity results, where some of the latter benefit from the former.

Our structural results provide substantial evidence for the *Fully Mixed Nash Equilibrium Conjecture,* which states that the worst Nash equilibrium is the *fully mixed Nash equilibrium,* where each user chooses each link with positive probability. Specifically, we prove:

- The Fully Mixed Nash Equilibrium Conjecture is valid for pure Nash equilibria.
- Under a certain condition, the social cost of any Nash equilibrium is within a factor of $6+\varepsilon$, for any constant $\varepsilon > 0$, of that of the fully mixed Nash equilibrium, assuming that link *capacities* are identical. This is shown through applying and extending techniques from the theory of *stochastic orders.* The factor reduces to 1 if one traffic is at least as large as the sum of the other traffics.

Our complexity results include hardness, approximability and inapproximability ones.

- For identical link capacities and under a certain condition, there is a randomized, polynomial-time algorithm to approximate the worst social cost within a factor arbitrarily close to $6+\varepsilon$, for any constant $\varepsilon > 0$. This builds on the structural results and resorts to a known randomized, polynomial-time approximation scheme (RPTAS) for the social cost of the fully mixed Nash equilibrium.
- For any arbitrary integer $k > 0$, it is $\mathcal{NP}$-hard to decide whether or not any given allocation of users to links can be transformed into a pure Nash equilibrium using at most k selfish steps, even if the number of links is 2.
- Assuming identical link capacities, there exists a polynomial-time approximation scheme (PTAS) to approximate the best social cost, over all pure Nash equilibria, within a factor of $1+\varepsilon$, for any arbitrary constant $\varepsilon > 0$. (This resorts to sorting traffics and a sequence of n selfish steps.)
- For any arbitrary constant $\varepsilon > 0$, it is $\mathcal{NP}$-hard to approximate the worst social cost within a multiplicative factor $2 - \frac{2}{m+1} - \varepsilon$. The quantity $2 - \frac{2}{m+1}$ is the tight upper bound on the ratio of the worst social cost and the optimal cost in the model of identical capacities. Moreover, for any fixed number of links, there exists a pseudopolynomial-time algorithm to compute the worst pure Nash equilibrium.

1 Introduction

1.1 Motivation and Framework

A *Nash equilibrium* [22, 23] represents a stable state of the play of a *strategic game,* in which each player holds an accurate opinion about the (expected) behavior of other players and acts rationally. An issue that arises naturally in this context concerns the computational complexity of Nash equilibria of any given strategic game. Due to the ultimate significance of Nash equilibrium as a prime solution concept in contemporary *Game Theory* [24], this issue has become a fundamental algorithmic problem that is being intensively studied in the Theory of Computing community today (see, e.g., [4, 7, 31]); in fact, it is arguably one of the few, most important algorithmic problems for which no *general* polynomial-time algorithms are known today (cf. [26]).

The problem of computing arbitrary Nash equilibria becomes even more challenging when one considers *extreme* Nash equilibria, ones that maximize or minimize a certain *objective function.* So, understanding the combinatorial structure of extreme Nash equilibria is a necessary prerequisite to either designing efficient algorithms to compute them or establishing corresponding hardness and thereby designing efficient approximation algorithms. In this work, we embark on a systematic study of the combinatorial structure and the computational complexity of extreme Nash equilibria; our study is carried out within the context of a simple *selfish routing* game, originally introduced in a pioneering work by Koutsoupias and Papadimitriou [17], that we describe next.

We assume a collection of n users, each employing a *mixed strategy,* which is a probability distribution over m parallel *links,* to control the shipping of its own assigned *traffic.* For each link, a *capacity* specifies the rate at which the link processes traffic. In a Nash equilibrium, each user selfishly routes its traffic on those links that minimize its *expected latency cost,* given the network congestion caused by the other users. A user's *support* is the set of those links on which it may ship its traffic with non-zero probability. The *social cost* of a Nash equilibrium is the expectation, over all random choices of the users, of the maximum, over all links, *latency* through a link.

Our study distinguishes between *pure* Nash equilibria, where each user chooses exactly one link (with probability one), and *mixed* Nash equilibria, where the choices of each user are modeled by a probability distribution over links. We also distinguish in some cases between models of *identical capacities,* where all link capacities are equal, and of *arbitrary capacities.*

1.2 The Fully Mixed Nash Equilibrium Conjecture.

In this work, we formulate and study a natural conjecture asserting that the fully mixed Nash equilibrium $\mathbf{F}$ is the *worst* Nash equilibrium with respect to social cost. Formally, we conjecture:

Conjecture 1.1 (Fully Mixed Nash Equilibrium Conjecture) *For any traffic vector* $\mathbf{w}$ *such that the fully mixed Nash equilibrium* $\mathbf{F}$ *exists, and for any Nash equilibrium* $\mathbf{P}$*,* $\mathsf{SC}(\mathbf{w}, \mathbf{P}) \leq \mathsf{SC}(\mathbf{w}, \mathbf{F})$.

Clearly, the Fully Mixed Nash Equilibrium Conjecture is intuitive and natural: the fully mixed Nash equilibrium favors "collisions" between different users (since each user assigns its traffic with positive probability to *every* link); thus, this increased probability of "collisions" favors a corresponding increase to the (expected) maximum total traffic through a link, which is, precisely, the social cost. More importantly, the Fully Mixed Nash Equilibrium Conjecture is also significant since it precisely identifies the *worst* possible Nash equilibrium for the selfish routing game we consider; this will enable designers of Internet protocols not only to avoid choosing the worst-case Nash equilibrium, but also to calculate the worst-case loss to the system at *any* Nash equilibrium due to its deliberate lack of coordination, and to evaluate the Nash equilibrium of choice against the (provably) worst-case one.

1.3 Contribution and Significance

Our study provides quite strong evidence in support of the Fully Mixed Nash Equilibrium Conjecture by either establishing or near establishing the conjecture in a number of interesting instances of the problem.

We start with the model of arbitrary capacities, where traffics are allowed to vary arbitrarily. There we prove that the Fully Mixed Nash Equilibrium Conjecture holds for *pure* Nash equilibria. We next turn to the case of identical capacities. Through a delicate probabilistic analysis, we establish that in the special case, that the number of links is equal to the number of users and for a suitable large number of users, the social cost of *any* Nash equilibrium is less than $6 + \varepsilon$ (for any $\varepsilon > 0$) times the social cost of the fully mixed Nash equilibrium. Our proof employs concepts and techniques from *majorization theory* [18] and *stochastic orders* [30], such as comparing two random variables according to their *stochastic variability* (cf. [28, Section 9.5]).

For pure Nash equilibria we show that it is $\mathcal{NP}$-hard to decide whether or not any given allocation of users to links can be transformed into a pure Nash equilibrium using at most *k selfish steps,* even if the number of links is 2. Furthermore, we prove that there exists a polynomial-time approximation scheme (PTAS) to approximate the social cost of the best pure Nash equilibrium to any arbitrary accuracy. The proof involves an algorithm that transforms any pure strategy profile into a pure Nash equilibrium with at most the same social cost, using at most n reassignments of users. We call this technique *Nashification,* and it may apply to other instances of the problem as well.

Still for pure Nash equilibria, we give a tight upper bound on the ratio between $\mathsf{SC}(\mathbf{w}, \mathbf{L})$ and $\mathsf{OPT}(\mathbf{w})$ for any Nash equilibrium $\mathbf{L}$. Then we show that it is $\mathcal{NP}$-hard to approximate the worst-case Nash equilibrium with a ratio that is better than this upper bound. We close our section about pure Nash equilibria with a pseudopolynomial algorithm for computing the worst-case Nash equilibrium for any fixed number of links.

1.4 Related Work and Comparison

The selfish routing game considered in this paper was first introduced by Koutsoupias and Papadimitriou [17] as a vehicle for the study of the price of selfishness for routing over non-cooperative networks, like the Internet. This game was subsequently studied in the work of Mavronicolas and Spirakis [19], where fully mixed Nash equilibria were introduced and analyzed. In both works, the aim had been to quantify the amount of performance loss in routing due to selfish behavior of the users. (Later studies of the selfish routing game from the same point of view, that of performance, include the works by Koutsoupias *et al.* [16], and by Czumaj and Vöcking [2].)

The closest to our work is the one by Fotakis *et al.* [7], which focuses on the combinatorial structure and the computational complexity of Nash equilibria for the selfish routing game we consider. The Fully Mixed Nash Equilibrium Conjecture formulated and systematically studied in this paper has been inspired by two results due to Fotakis *et al.* [7] that confirm or support the conjecture. First, Fotakis *et al.* [7, Theorem 4.2] establish the Fully Mixed Nash Equilibrium Conjecture for the model of identical capacities and assuming that $n = 2$. Second, Fotakis *et al.* [7, Theorem 4.3] establish that, for the model of arbitrary capacities, the social cost of any Nash equilibrium is no more than 49.02 times the social cost of the (generalized) fully mixed Nash equilibrium; Theorem 5.1 shows that the constant 49.02 can be reduced to $6 + \varepsilon$ when restricted to the model of identical capacities.

The routing problem considered in this paper is equivalent to the multiprocessor scheduling problem. Here, pure Nash equilibria and Nashification translate to local optima and sequences of local improvements. A schedule is said to be *jump optimal* if no job on a processor with maximum load can improve by moving to another processor [29].

Obviously, the set of pure Nash equilibria is a subset of the set of jump optimal schedules. Moreover, in the model of identical processors every jump optimal schedule can be transformed into a pure Nash equilibrium without altering the makespan. Thus, for this model the strict upper bound $2 - 2/(m+1)$ on the ratio between best and worst makespan of jump optimal schedules [6, 29] also holds for pure Nash equilibria.

Algorithms for computing a jump optimal schedule from any given schedule have been proposed in [1, 6, 29]. The fastest algorithm is given by Schuurman and Vredeveld [29]. It always moves the job with maximum weight from a makespan processor to a processor with minimum load, using $O(n)$ moves. However, in all algorithms the resulting jump optimal schedule is not necessarily a Nash equilibrium.

1.5 Road Map.

The rest of this paper is organized as follows. Section 2 presents some preliminaries. Stochastic orders are treated in Section 3. Pure Nash equilibria are contrasted to the fully mixed Nash equilibrium in Section 4. Worst mixed Nash equilibria are contrasted to the fully mixed Nash equilibrium in Section 5. Sections 6 and 7 consider best and worst pure Nash equilibria, respectively. We conclude, in Section 8, with a discussion of our results and some open problems.

2 Framework

Most of our definitions are patterned after those in [19, Section 2] and [7, Section 2], which, in turn, were based on those in [17, Sections 1 & 2].

2.1 Mathematical Preliminaries and Notation.

For any integer $m \geq 1$, denote $[m] = \{1, \ldots, m\}$. Denote Γ the *Gamma function*; that is, for any natural number N, $\Gamma(N+1) = N!$, while for any arbitrary real number $x > 0$, $\Gamma(x) = \int_0^\infty t^{x-1}e^{-t}dt$. The Gamma function is invertible; both Γ and its inverse Γ^{-1} are increasing. It is well known that $\Gamma^{-1}(N) = \frac{\lg N}{\lg \lg N}(1 + o(1))$ (see, e.g., [11]). For our purposes, we shall use the fact that for any pair of an arbitrary real number α and an arbitrary natural number N, $\left(\frac{\alpha}{e}\right)^\alpha = N$ if and only if $\alpha = \Gamma^{-1}(N) + \Theta(1)$. For an event E in a sample space, denote $\mathbf{Pr}(E)$ the probability of event E happening.

For a random variable X, denote $\mathcal{E}(X)$ the *expectation* of X. In the *balls-and-bins* problem, m balls are thrown into m bins uniformly at random. (See [15] for a classical introduction to this problem.) It is known that the expected maximum number of balls thrown over a bin equals the quantity $R(m) = \Gamma^{-1}(m) - \frac{3}{2} + o(1)$ [11].

In the paper, we make use of the following inequality, that holds due to Hoeffding.

Theorem 2.1 ([20], Theorem 2.3.) *Let the random variables $X_1, X_2, ..., X_n$ be independent, with $0 \leq X_k \leq 1$ for each k. Let $S_n = \sum X_k$, let $\mu = \mathcal{E}(S_n)$. Then, for any $\beta > 0$,*

$$\mathbf{Pr}(S_n \geq (1+\beta)\mu) \leq e^{-((1+\beta)ln(1+\beta)-\beta)\mu}.$$

Note that Theorem 2.1 also holds if $0 \leq X_k \leq \kappa$ for some constant $\kappa > 0$.

2.2 General

We consider a *network* consisting of a set of m parallel *links* $1, 2, \ldots, m$ from a *source* node to a *destination* node. Each of n *network users* $1, 2, \ldots, n$, or *users* for short, wishes to route a particular amount of traffic along a (non-fixed) link from source to destination. Denote w_i the *traffic* of user $i \in [n]$. Define the $n \times 1$ *traffic vector* $\mathbf{w}$ in the natural way. Assume throughout that $m > 1$ and $n > 1$. Assume also, without loss of generality, that $w_1 \geq w_2 \geq \ldots \geq w_n$. For a traffic vector $\mathbf{w}$, denote $W = \sum_1^n w_i$.

A *pure strategy* for user $i \in [n]$ is some specific link. A *mixed strategy* for user $i \in [n]$ is a probability distribution over pure strategies; thus, a mixed strategy is a probability distribution over the set of links. The *support* of the mixed strategy for user $i \in [n]$, denoted *support*(i), is the set of those pure strategies (links) to which i assigns positive probability.

A *pure strategy profile* is represented by an n-tuple $\langle \ell_1, \ell_2, \ldots, \ell_n \rangle \in [m]^n$; a *mixed strategy profile* is represented by an $n \times m$ *probability matrix* $\mathbf{P}$ of nm probabilities p_i^j, $i \in [n]$ and $j \in [m]$, where p_i^j is the probability that user i chooses link j. For a probability matrix $\mathbf{P}$, define *indicator variables* $I_i^\ell \in \{0, 1\}$, $i \in [n]$ and $\ell \in [m]$, such that $I_i^\ell = 1$ if and only if $p_i^\ell > 0$. Thus, the support of the mixed strategy for user $i \in [n]$ is the set $\{\ell \in [m] \mid I_i^\ell = 1\}$.

For each link $\ell \in [m]$, define the *view* of link ℓ, denoted *view*(ℓ), as the set of users $i \in [n]$ that potentially assign their traffics to link ℓ; so, *view*$(\ell) = \{i \in [n] \mid I_i^\ell = 1\}$. For each link $\ell \in [m]$, denote $V^\ell = |view(\ell)|$. A mixed strategy profile $\mathbf{P}$ is *fully mixed* [19, Section 2.2] if for all users $i \in [n]$ and links $j \in [m]$, $I_i^j = 1$.*

*An earlier treatment of fully mixed strategies in the context of *bimatrix games* has been found in [27], called

2.3 System, Models and Cost Measures

Denote $c^\ell > 0$ the *capacity* of link $\ell \in [m]$, representing the rate at which the link processes traffic. So, the *latency* for traffic w through link ℓ equals w/c^ℓ. In the model of *identical capacities,* all link capacities are equal to 1; link capacities may vary arbitrarily in the model of *arbitrary capacities.* For a pure strategy profile $\langle \ell_1, \ell_2, \ldots, \ell_n \rangle$, the *latency cost for user i,* denoted λ_i, is $(\sum_{k:\ell_k=\ell_i} w_k)/c^{\ell_i}$; that is, the latency cost for user i is the latency of the link it chooses. For a mixed strategy profile $\mathbf{P}$, denote δ^ℓ the *actual traffic* on link $\ell \in [m]$; so, δ^ℓ is a random variable for each link $\ell \in [m]$, denote θ^ℓ the *expected traffic* on link $\ell \in [m]$; thus, $\theta^\ell = \mathcal{E}(\delta^\ell) = \sum_{i=1}^n p_i^\ell w_i$. Given $\mathbf{P}$, define the $m \times 1$ *expected traffic vector* $\mathbf{\Theta}$ induced by $\mathbf{P}$ in the natural way. Given $\mathbf{P}$, denote Λ^ℓ the *expected latency* on link $\ell \in [m]$; clearly, $\Lambda^\ell = \frac{\theta^\ell}{c^\ell}$. Define the $m \times 1$ *expected latency vector* $\mathbf{\Lambda}$ in the natural way. For a mixed strategy profile $\mathbf{P}$, the *expected latency cost* for user $i \in [n]$ on link $\ell \in [m]$, denoted λ_i^ℓ, is the expectation, over all random choices of the remaining users, of the latency cost for user i had its traffic been assigned to link ℓ; thus,

$$\lambda_i^\ell = \frac{w_i + \sum_{k=1,k\neq i} p_k^\ell w_k}{c^\ell} = \frac{(1-p_i^\ell)w_i + \theta^\ell}{c^\ell}.$$

For each user $i \in [n]$, the *minimum expected latency cost,* denoted λ_i, is the minimum, over all links $\ell \in [m]$, of the expected latency cost for user i on link ℓ; thus, $\lambda_i = \min_{\ell\in[m]} \lambda_i^\ell$. For a probability matrix $\mathbf{P}$, define the $n \times 1$ *minimum expected latency cost vector* λ induced by $\mathbf{P}$ in the natural way.

Associated with a traffic vector $\mathbf{w}$ and a mixed strategy profile $\mathbf{P}$ is the *social cost* [17, Section 2], denoted $\mathsf{SC}(\mathbf{w}, \mathbf{P})$, which is the expectation, over all random choices of the users, of the maximum (over all links) latency of traffic through a link; thus,

$$\mathsf{SC}(\mathbf{w}, \mathbf{P}) = \mathcal{E}\left(\max_{\ell\in[m]} \frac{\sum_{k:\ell_k=\ell} w_k}{c^\ell}\right) = \sum_{\langle \ell_1,\ell_2,\ldots,\ell_n\rangle \in [m]^n} \left(\prod_{k=1}^n p_k^{\ell_k} \cdot \max_{\ell\in[m]} \frac{\sum_{k:\ell_k=\ell} w_k}{c^\ell}\right).$$

Note that $\mathsf{SC}(\mathbf{w}, \mathbf{P})$ reduces to the maximum latency through a link in the case of pure strategies. On the other hand, the *social optimum* [17, Section 2] associated with a traffic vector $\mathbf{w}$, denoted $\mathsf{OPT}(\mathbf{w})$, is the *least possible* maximum (over all links) latency of traffic through a link; thus,

$$\mathsf{OPT}(\mathbf{w}) = \min_{\langle \ell_1,\ell_2,\ldots,\ell_n\rangle \in [m]^n} \max_{\ell\in[m]} \frac{\sum_{k:\ell_k=\ell} w_k}{c^\ell}.$$

---

there *completely mixed strategies.* See also [21] for a subsequent treatment in the context of *strategically zero-sum games.* Datta [3] studied recently some universality properties of fully mixed Nash equilibria (calling them *totally mixed*).

2.4 Nash Equilibria

We are interested in a special class of mixed strategies called Nash equilibria [22, 23] that we describe below. Say that a user $i \in [n]$ is *satisfied for the probability matrix* $\mathbf{P}$ if for all links $\ell \in [m]$, $\lambda_i^\ell = \lambda_i$ if $I_i^\ell = 1$, and $\lambda_i^\ell > \lambda_i$ if $I_i^\ell = 0$; thus, a satisfied user has no incentive to unilaterally deviate from its mixed strategy. A user $i \in [n]$ is *unsatisfied for the probability matrix* $\mathbf{P}$ if i is not satisfied for the probability matrix $\mathbf{P}$. The probability matrix $\mathbf{P}$ is a *Nash equilibrium* [17, Section 2] if for all users $i \in [n]$ and links $\ell \in [m]$, $\lambda_i^\ell = \lambda_i$ if $I_i^\ell = 1$, and $\lambda_i^\ell > \lambda_i$ if $I_i^\ell = 0$. Thus, each user assigns its traffic with positive probability only on links (possibly more than one of them) for which its expected latency cost is minimized. The *fully mixed Nash equilibrium* [19], denoted $\mathbf{F}$, is a Nash equilibrium that is a fully mixed strategy. Mavronicolas and Spirakis [19, Lemma 15] show that all links are *equiprobable* in a fully mixed Nash equilibrium, which is unique (for the model of identical capacities).

Fix any traffic vector $\mathbf{w}$. The *worst Nash equilibrium* is the Nash equilibrium $\mathbf{P}$ that maximizes $\mathsf{SC}(\mathbf{w}, \mathbf{P})$; the *best Nash equilibrium* is the Nash equilibrium that minimizes $\mathsf{SC}(\mathbf{w}, \mathbf{P})$. The *worst social cost,* denoted $\mathsf{WC}(\mathbf{w})$, is the social cost of the worst Nash equilibrium; correspondingly, the *best social cost,* denoted $\mathsf{BC}(\mathbf{w})$, is the social cost of the best Nash equilibrium.

Fotakis *et al.* [7, Theorem 1] consider starting from any arbitrary pure strategy profile and following a particular sequence of selfish steps, where in a *selfish step,* exactly one unsatisfied user is allowed to change its pure strategy. A selfish step is a *greedy selfish step* if the unsatisfied user chooses its best link. A (greedy) selfish step does not increase the social cost of the initial pure strategy profile. Fotakis *et al.* [7, Theorem 1] show that this sequence of selfish steps eventually converges to a Nash equilibrium, which proves its existence; however, it may take a large number of steps. It follows that if the initial pure strategy profile has minimum social cost, then the resulting (pure) Nash equilibrium will have minimum social cost as well. This implies that there exists a pure Nash equilibrium with minimum social cost. Thus, we have $\mathsf{BC}(\mathbf{w}) = \mathsf{OPT}(\mathbf{w})$.

2.5 Algorithmic Problems.

We list a few algorithmic problems related to Nash equilibria that will be considered in this work. The definitions are given in the style of Garey and Johnson [10]. A problem instance is a tuple (n, m, w, c) where n is the number of users, m is the number of links, $w = (w_i)$ is a vector of n user traffics and $c = (c^j)$ is a vector of m link capacities.

$\mathbf{\Pi_1}$: NASH EQUILIBRIUM SUPPORTS

INSTANCE: A problem instance (n, m, w, c).

OUTPUT: Indicator variables $I_i^j \in \{0,1\}$, where $i \in [n]$ and $j \in [m]$, that support a Nash equilibrium for the system of the users and the links.

Fotakis *et al.* [7, Theorem 2] establish that NASH EQUILIBRIUM SUPPORTS is in $\mathcal{P}$ when restricted to pure equilibria. We continue with two complementary to each other optimization problems (with respect to social cost).

Π_2: BEST NASH EQUILIBRIUM SUPPORTS
INSTANCE: A problem instance (n, m, w, c).
OUTPUT: Indicator variables $I_i^j \in \{0,1\}$, where $i \in [n]$ and $j \in [m]$, that support the best Nash equilibrium for the system of the users and the links.

Π_3: WORST NASH EQUILIBRIUM SUPPORTS
INSTANCE: A problem instance (n, m, w, c).
OUTPUT: Indicator variables $I_i^j \in \{0,1\}$, where $i \in [n]$ and $j \in [m]$, that support the worst Nash equilibrium for the system of the users and the links.

Fotakis *et al.* [7, Theorems 3 and 4] establish that both BEST NASH EQUILIBRIUM SUPPORTS and WORST NASH EQUILIBRIUM SUPPORTS are $\mathcal{NP}$-hard. Since both problems can be formulated as an integer program, it follows that they are $\mathcal{NP}$-complete.

Π_4: NASH EQUILIBRIUM SOCIAL COST
INSTANCE: A problem instance (n, m, w, c); a *Nash equilibrium* $\mathbf{P}$ for the system of the users and the links.
OUTPUT: The social cost of the Nash equilibrium $\mathbf{P}$.

Fotakis *et al.* [7, Theorem 8] establish that NASH EQUILIBRIUM SOCIAL COST is $\#\mathcal{P}$-complete. Furthermore, Fotakis *et al.* [7, Theorem 9] show that there exists a fully polynomial, randomized approximation scheme for NASH EQUILIBRIUM SOCIAL COST.

The following two problems, inspired by NASH EQUILIBRIUM SOCIAL COST are introduced for the first time in this work.

Π_5: WORST NASH EQUILIBRIUM SOCIAL COST
INSTANCE: A problem instance (n, m, w, c).
OUTPUT: The worst social cost $\mathsf{WSC}(\mathbf{w})$.

Π_6: BEST NASH EQUILIBRIUM SOCIAL COST
INSTANCE: A problem instance (n, m, w, c).
OUTPUT: The best social cost $\mathsf{BSC}(\mathbf{w})$.

Π_7: k-NASHIFY
INSTANCE: A problem instance (n, m, w, c); a pure strategy profile $\mathbf{L}$ for the system of the users

and the links.
QUESTION: Is there a sequence of at most k selfish steps that transform $\mathbf{L}$ to a (pure) Nash equilibrium?

The following problem is a variant of k-NASHIFY in which k is part of the input.

$\mathbf{\Pi_8}$: NASHIFY
INSTANCE: A problem instance (n, m, w, c); a pure strategy profile $\mathbf{L}$ for the system of the users and the links; an integer $k > 0$.
QUESTION: Is there a sequence of at most k selfish steps that transform $\mathbf{L}$ to a (pure) Nash equilibrium?

In our hardness and completeness proofs, we will employ the following $\mathcal{NP}$-complete problems [14]:

$\mathbf{\Pi_9}$: BIN PACKING
INSTANCE: A finite set $\mathcal{U}$ of items, a size $s(u) \in \mathbf{N}$ for each $u \in \mathcal{U}$, a positive integer bin capacity B, and a positive integer K.
QUESTION: Is there a partition of $\mathcal{U}$ into disjoint sets $\mathcal{U}_1, \ldots, \mathcal{U}_K$ such that for each set $\mathcal{U}_i$, $1 \leq i \leq K$, $\sum_{u \in \mathcal{U}_i} s(u) \leq B$?

$\mathbf{\Pi_{10}}$: PARTITION
INSTANCE: A finite set $\mathcal{U}$ and a size $s(u) \in \mathbf{N}$ for each element $u \in \mathcal{U}$.
QUESTION: Is there a subset $\mathcal{U}' \subseteq \mathcal{U}$ such that $\sum_{u \in \mathcal{U}'} s(u) = \sum_{u \in \mathcal{U} \setminus \mathcal{U}'} s(u)$?

We note that BIN PACKING is *strongly $\mathcal{NP}$-complete* [9, 25].†

3 Stochastic Order Relations

In this section, we treat stochastic order relations; we establish a certain stochastic order relation for the expected maxima of certain sums of Bernoulli random variables.

Recall that a function $f : \Re \rightarrow \Re$ is *convex* if for all numbers λ such that $0 < \lambda < 1$, $f(\lambda x_1 + (1-\lambda)x_2) \leq \lambda f(x_1) + (1-\lambda) f(x_2)$. We proceed to describe a stochastic order relation between two random variables.

Definition 3.1 *For any pair of arbitrary random variables X and Y, say that X* is stochastically more variable than Y *if for all increasing and convex functions $f : \Re \rightarrow \Re$, $\mathcal{E}(f(X)) \geq \mathcal{E}(f(Y))$.*

†A problem is *strongly $\mathcal{NP}$-complete* if it remains $\mathcal{NP}$-complete even if any instance of length n is restricted to contain integers of size polynomial in n. So, strongly $\mathcal{NP}$-complete problems admit no pseudopolynomial-time algorithms unless $\mathcal{P} = \mathcal{NP}$.

Call *stochastically more variability* the corresponding stochastic order relation on the set of random variables. (See [28, Section 9.5] for a more complete treatment of the notion of stochastically more variable and [18, 30] for more on majorization theory and stochastic orders.) The following lemma [28, Proposition 9.5.1] provides an alternative, analytic characterization of stochastically more variability.

Lemma 3.1 *Consider any pair of non-negative random variables X and $\widehat{X}$. Then, X is stochastically more variable than $\widehat{X}$ if and only if for all numbers $\alpha \geq 0$, $\int_{x=\alpha}^{\infty} \mathbf{Pr}(X > x)dx \geq \int_{x=\alpha}^{\infty} \mathbf{Pr}(\widehat{X} > x)dx$.*

Consider now a setting of the balls-and-bins problem where n balls $1, \ldots, n$ with *traffics* $w_1, \ldots, w_n$ are allocated into m bins $1, \ldots, m$ uniformly at random. So, for each pair of a ball $i \in [n]$ and a link $j \in [m]$, define Bernoulli random variables $Y_i^j = w_i$ with probability $\frac{1}{m}$ and 0 with probability $1 - \frac{1}{m}$, and $\widetilde{Y_i^j} = \frac{W}{n}$ with probability $\frac{1}{m}$ and 0 with probability $1 - \frac{1}{m}$. For each link $j \in [m]$, define the random variables $\delta^j = \sum_{i \in [n]} Y_i^j$ and $\widetilde{\delta^j} = \sum_{i \in [n]} \widetilde{Y_i^j}$; thus, each of δ^j and $\widetilde{\delta^j}$, $j \in [m]$, is a sum of Bernoulli random variables; denote $\theta^j = \mathcal{E}(\delta^j)$ and $\widetilde{\theta^j} = \mathcal{E}(\widetilde{\delta^j})$ the expectations of δ^j and $\widetilde{\delta^j}$, respectively. Note that

$$\theta^j = \mathcal{E}\left(\sum_{i \in [n]} Y_i^j\right) = \sum_{i \in [n]} \mathcal{E}\left(Y_i^j\right) = \sum_{i \in [n]} \left(w_i \frac{1}{m} + 0\left(1 - \frac{1}{m}\right)\right) = \frac{\sum_{i \in [n]} w_i}{m} = \frac{W}{m},$$

while

$$\widetilde{\theta^j} = \mathcal{E}(\widetilde{\delta^j}) = \mathcal{E}\left(\sum_{i \in [n]} \widetilde{Y_i^j}\right) = \sum_{i \in [n]} \mathcal{E}\left(\widetilde{Y_i^j}\right) = \sum_{i \in [n]} \left(\frac{W}{n}\frac{1}{m} + 0\left(1 - \frac{1}{m}\right)\right) = n\frac{W}{n}\frac{1}{m} = \frac{W}{m}.$$

So, $\theta^j = \widetilde{\theta^j}$ for each bin $j \in [m]$.

For two numbers $x, y \in \Re^+$ define

$$[x - y] = \begin{cases} x - y & : \text{ if } x > y \\ 0 & : \text{ else.} \end{cases}$$

We can then show the following preliminary lemma:

Lemma 3.2 *Let $b_i \in \Re^+$ for $i \in [n]$ and let $d = \frac{1}{n}\sum_{i=1}^{n} b_i$. Then for all $x \geq 0$*

$$\sum_{i=1}^{n} [b_i - x] \geq n \cdot [d - x]$$

Proof: Without loss of generality, assume that $b_1 \leq b_2 \leq \ldots \leq b_n$. The claim is true if $x > d$. If $x \leq b_1$, then $x \leq d$ and

$$\sum_{i=1}^{n} [b_i - x] = \sum_{i=1}^{n} (b_i - x) = n \cdot (d - x).$$

Now let $b_j < x \leq b_{j+1}$ and $d > x$. It follows that

$$\begin{aligned}
\sum_{i=1}^{n} [b_i - x] &= \sum_{i=j+1}^{n} (b_i - x) = \sum_{i=j+1}^{n} b_i - (n-j)x = \sum_{i=j+1}^{n} b_i - n \cdot x + j \cdot x \\
&\geq \sum_{i=j+1}^{n} b_i - n \cdot x + \sum_{i=1}^{j} b_i = \sum_{i=1}^{n} b_i - n \cdot x \\
&= n \cdot (d - x)
\end{aligned}$$

■

We finally prove:

Lemma 3.3 (Stochastically More Variability Lemma) *For any traffic vector* $\mathbf{w}$, $\max\left\{\delta^1, \ldots, \delta^m\right\}$ *is stochastically more variable than* $\max\left\{\widetilde{\delta^1}, \ldots, \widetilde{\delta^m}\right\}$.

Proof: Define the discrete random variables $X = \max\{\delta^1, \ldots, \delta^m\}$ and $\widetilde{X} = \max\{\widetilde{\delta^1}, \ldots, \widetilde{\delta^m}\}$. We then have to show that for all $\alpha \geq 0$,

$$\int_{x=\alpha}^{\infty} \mathbf{Pr}\left(X > x\right) \, dx \geq \int_{x=\alpha}^{\infty} \mathbf{Pr}\left(\widetilde{X} > x\right) \, dx.$$

Let S_k be the collection of all pure strategy profiles, where the maximum number of traffics on any link $j \in [m]$ is exactly k. If $i \neq j$, then $S_i \cap S_j = \emptyset$. Furthermore

$$\bigcup_{i=\lceil \frac{n}{m} \rceil}^{n} S_i = [m]^n.$$

For any pure strategy profile $L \in S_k$, define $\text{Link}(L)$ to be the smallest index of a link, holding k traffics. Furthermore, for any pure strategy profile L, let $I(L)$ be the collection of users that are assigned to $\text{Link}(L)$. Every set of k traffics is equal to some $I(L), L \in S_k$ with the same probability, say p_k. Define the actual traffic on $Link(L)$ as

$$b(L) = \sum_{i \in I(L)} w_i.$$

If all traffics are identical the actual traffic on $Link(L)$ for a pure strategy profile $L \in S_k$ is simply $\widetilde{b}(L) = k \cdot \frac{W}{n}$.

Every pure strategy profile $L \in [m]^n$ occurs with the same probability $\frac{1}{m^n}$ and defines together with $b(L)$ a discrete random variable Z. Z is a discrete random variable that can take every possible value $b(L), L \in [m]^n$.

It is easy to see, that X is stochastically more variable than Z, since for any pure strategy profile L, Z refers to the actual traffic on $Link(L)$, whereas X refers to the maximum actual traffic over all links.

We will complete our proof by showing, that Z is stochastically more variable than $\widetilde{X}$.

Since Z and $\widetilde{X}$ are discrete random variables

$$\int_{x=\alpha}^{\infty} \mathbf{Pr}(Z > x)dx = \sum_{k=\lceil \frac{n}{m} \rceil}^{n} (p_k \cdot A_k), \quad \text{where } A_k = \sum_{L \in S_k} [b(L) - \alpha]$$

and

$$\int_{x=\alpha}^{\infty} \mathbf{Pr}(\widetilde{X} > x)dx = \sum_{k=\lceil \frac{n}{m} \rceil}^{n} \left(p_k \cdot \widetilde{A_k}\right), \quad \text{where } \widetilde{A_k} = |S_k| \cdot [k \cdot \frac{W}{n} - \alpha]$$

Since for a fixed k each traffic contributes with the same probability to $b(L)$

$$\sum_{L \in S_k} b(L) = |S_k| \cdot k \cdot \frac{W}{n}.$$

It follows from Lemma 3.2 that $A_k \geq \widetilde{A_k}$ for each k. Therefore Z is stochastically more variable than $\widetilde{X}$, which completes the proof of the lemma. ■

By definition of stochastically more variability, Lemma 3.3 immediately implies:

Corollary 3.4 *For any traffic vector* $\mathbf{w}$, $\mathcal{E}\left(\max\left\{\delta^1, \ldots, \delta^m\right\}\right) \geq \mathcal{E}\left(\max\left\{\widetilde{\delta^1}, \ldots, \widetilde{\delta^m}\right\}\right)$.

In the balls-and-bins game in which m balls are thrown uniformly at random into m bins, Corollary 3.4 shows that, if the sum of the ball weights is the same, the expected maximum load over all bins is larger when the balls have different weights in comparison to all balls having the same weight.

4 Pure versus Fully Mixed Nash Equilibria

In this section, we establish the Fully Mixed Nash Equilibrium Conjecture for the case of pure Nash equilibria. This result holds also for the model of arbitrary capacities. We start by proving:

Lemma 4.1 *Fix any traffic vector* $\mathbf{w}$*, mixed Nash equilibrium* $\mathbf{P}$ *and user* i*. Then,* $\lambda_i(\mathbf{w}, \mathbf{P}) \leq \lambda_i(\mathbf{w}, \mathbf{F})$.

Proof: Let $\mathbf{P} = \left(p_k^j\right), \mathbf{F} = \left(f_k^j\right)$ for $k \in [n]$ and $j \in [m]$. Then

$$\begin{aligned}
\sum_{j\in[m]}\left(\sum_{k\in[n],k\neq i} p_k^j w_k\right) &= \sum_{k\in[n],k\neq i} w_k \left(\sum_{j\in[m]} p_k^j\right) \\
&= \sum_{k\in[n],k\neq i} w_k,
\end{aligned}$$

and

$$\begin{aligned}
\sum_{j\in[m]}\left(\sum_{k\in[n],k\neq i} f_k^j w_k\right) = \sum_{k\in[n],k\neq i} w_k \left(\sum_{j\in[m]} f_k^j\right) \\
= \sum_{k\in[n],k\neq i} w_k.
\end{aligned}$$

It follows that

$$\sum_{j\in[m]}\left(\sum_{k\in[n],k\neq i} p_k^j w_k\right) = \sum_{j\in[m]}\left(\sum_{k\in[n],k\neq i} f_k^j w_k\right),$$

and therefore there exists some link $j_0 \in [m]$ such that

$$\sum_{k\in[n],k\neq i} p_k^{j_0} w_k \leq \sum_{k\in[n],k\neq i} f_k^{j_0}\, w_k.$$

Then,

$$\begin{aligned}
&\lambda_i(\mathbf{w}, \mathbf{P}) \\
\leq\;& \lambda_i^{j_0}(\mathbf{w}, \mathbf{P}) && (\text{since } \lambda_i \text{ is the minimum of all } \lambda_i^j, j \in [n]) \\
=\;& \frac{w_i + \sum_{k\in[n],k\neq i} p_k^{j_0} w_k}{c^{j_0}} \\
\leq\;& \frac{w_i + \sum_{k\in[n],k\neq i} f_k^{j_0} w_k}{c^{j_0}} \\
=\;& \lambda_i^{j_0}(\mathbf{w}, \mathbf{F}) \\
=\;& \lambda_i(\mathbf{w}, \mathbf{F}) && (\text{since } f_i^{j_0} > 0 \text{ and } \mathbf{F} \text{ is a Nash equilibrium}).
\end{aligned}$$

■

We now prove:

Theorem 4.2 *Fix any traffic vector* $\mathbf{w}$ *and pure Nash equilibrium* $\mathbf{L}$. *Then,* $\mathsf{SC}(\mathbf{w}, \mathbf{L}) \leq \mathsf{SC}(\mathbf{w}, \mathbf{F})$.

Proof: For each user $i \in [n]$, $\lambda_i(\mathbf{w}, \mathbf{P})$ is the minimum, over all links $j \in [m]$, of the expected latency cost for user i on link j, and $\mathsf{SC}(\mathbf{w}, \mathbf{P})$ is the expectation of the maximum (over all links) latency of traffic through a link. This implies that $\lambda_i(\mathbf{w}, \mathbf{P}) \leq \mathsf{SC}(\mathbf{w}, \mathbf{P})$ for every mixed Nash equilibrium $\mathbf{P}$. Hence,

$$\begin{aligned} & \lambda_i(\mathbf{w}, \mathbf{P}) \\ \leq \quad & \lambda_i(\mathbf{w}, \mathbf{F}) \quad \text{(by Lemma 4.1)} \\ \leq \quad & \mathsf{SC}(\mathbf{w}, \mathbf{F}) \quad \text{(as shown above)}. \end{aligned}$$

The claim follows now since $\mathsf{SC}(\mathbf{w}, \mathbf{L}) = \max_{i \in [n]} \lambda_i(\mathbf{w}, \mathbf{L})$ holds for every pure Nash equilibrium $\mathbf{L}$. ■

5 Worst Mixed Nash Equilibria

In this section we show that if $n = m$ and m is suitable large then the social cost of any Nash equilibrium is at most $6 + \varepsilon$ times the social cost of the fully mixed Nash equilibrium.

Theorem 5.1 *Consider the model of identical capacities. Let* $n = m$, m *suitable large. Then, for any traffic vector* $\mathbf{w}$ *and Nash equilibrium* $\mathbf{P}$, $\mathsf{SC}(\mathbf{w}, \mathbf{P}) < (6 + \varepsilon)\,\mathsf{SC}(\mathbf{w}, \mathbf{F})$, *for any* $\varepsilon > 0$.

Proof: Fix any traffic vector $\mathbf{w}$ and Nash equilibrium $\mathbf{P}$. We start by showing a simple technical fact.

Claim 5.2 *Fix any pair of a link* $\ell \in [m]$ *and a user* $i \in view(\ell)$. *Then,* $p_i^\ell w_i \geq \theta^\ell - \frac{W}{m}$.

Proof: Clearly,

$$\begin{aligned} \sum_{j \in [m]} \theta^j &= \sum_{j \in [m]} \left(\sum_{i \in [n]} p_i^j w_i \right) \\ &= \sum_{i \in [n]} \left(\sum_{j \in [m]} p_i^j w_i \right) \end{aligned}$$

$$
\begin{aligned}
&= \sum_{i\in[n]} \left(w_i \sum_{j\in[m]} p_i^j \right) \\
&= \sum_{i\in[n]} w_i \\
&= W\,.
\end{aligned}
$$

This implies that there exists some link $\ell' \in [m]$ such that $\theta^{\ell'} \leq \frac{W}{m}$. Note that by definition of social cost, $\lambda_i^{\ell'} = (1-p_i)w_i + \theta^{\ell'}$. It follows that $\lambda_i^{\ell'} \leq w_i + \frac{W}{m}$. On the other hand, $\lambda_i^{\ell} = (1-p_i^{\ell})w_i + \theta^{\ell}$.

Since $i \in view(\ell)$, we have, by definition of Nash equilibria, that $\lambda_i^{\ell} \leq \lambda_i^{\ell'}$ (with equality holding when $i \in view(\ell')$). It follows that $(1-p_i^{\ell})w_i + \theta^{\ell} \leq w_i + \frac{W}{m}$, or that $p_i^{\ell} w_i \geq \theta^{\ell} - \frac{W}{m}$, as needed. ∎

As an immediate consequence of Claim 5.2, we obtain:

Corollary 5.3 *Fix any link $\ell \in [m]$. Then, $\theta^{\ell} \leq \frac{V^{\ell}}{V^{\ell}-1}\frac{W}{m}$.*

Proof: Clearly, by Claim 5.2, $\theta^{\ell} = \sum_{i\in[n]} p_i^{\ell} w_i = \sum_{i\in view(\ell)} p_i^{\ell} w_i \geq \sum_{i\in view(\ell)} \left(\theta^{\ell} - \frac{W}{m}\right) = V^{\ell}\left(\theta^{\ell} - \frac{W}{m}\right)$, or, by rearrangement of terms, $\theta^{\ell} \leq \frac{V^{\ell}}{V^{\ell}-1}\frac{W}{m}$, as needed. ∎

Since $V^{\ell} \geq 2$, $\frac{V^{\ell}}{V^{\ell}-1} \leq 2$. Thus, by Corollary 5.3:

Lemma 5.4 *Fix any link $\ell \in [m]$ with $V^{\ell} \geq 2$. Then, $\theta^{\ell} \leq 2\,\frac{W}{m}$.*

We now prove a complementary lemma. Fix any link $\ell \in [m]$ with $V^{\ell} = 1$. Let $view(l) = \{i\}$. Then $\theta^{l} = w_i \leq \max_i w_i \leq \mathsf{OPT}(\mathbf{w}) \leq \mathsf{SC}(\mathbf{w}, \mathbf{F})$. Thus:

Lemma 5.5 *Fix any link $\ell \in [m]$ with $V^{\ell} = 1$. Then, $\theta^{\ell} \leq \mathsf{SC}(\mathbf{w}, \mathbf{F})$.*

Use $\mathbf{w}$ to define the vector $\widetilde{\mathbf{w}}$ with all entries equal to $\frac{W}{n}$. By definition of social cost, $\mathsf{SC}(\widetilde{\mathbf{w}}, \mathbf{F})$ is the load $\frac{W}{m}$ of each ball times the expected maximum number of balls thrown uniformly at random into m bins. Since $n = m$, we can state $\mathsf{SC}(\widetilde{\mathbf{w}}, \mathbf{F}) = R(m) \cdot \frac{W}{m}$, or $\frac{W}{m} = \frac{\mathsf{SC}(\widetilde{\mathbf{w}}, \mathbf{F})}{R(m)}$. Fix now any link $j \in [n]$ with $V^{j} \geq 2$. Then,

$$
\begin{aligned}
&\theta^{j} \\
&\leq 2\,\frac{W}{m} \qquad \text{(by Lemma 5.4)} \\
&= \frac{2}{R(m)}\,\mathsf{SC}(\widetilde{\mathbf{w}}, \mathbf{F}) \\
&\leq \frac{2}{R(m)}\,\mathsf{SC}(\mathbf{w}, \mathbf{F}) \quad \text{(by Corollary 3.4)}\,.
\end{aligned}
$$

Thus, for any constant $\varepsilon > 0$, arbitrarily close to 0,

$$\begin{array}{ll} & \mathbf{Pr}\left(\delta^j > 4\,(1+\varepsilon)\,\mathsf{SC}\,(\mathbf{w},\mathbf{F})\right) \\ \leq & \mathbf{Pr}\left(\delta^j > 4\,(1+\varepsilon)\,\frac{R(m)}{2}\,\theta^j\right) \qquad (\text{since } \theta^j \leq \frac{2}{R(m)}\,\mathsf{SC}\,(\mathbf{w},\mathbf{F})) \\ = & \mathbf{Pr}\left(\delta^j > 2\,(1+\varepsilon)\,R(m)\,\theta^j\right) \\ = & \mathbf{Pr}\left(\delta^j > 2\,(1+\varepsilon)\,R(m)\,\mathcal{E}\left(\delta^j\right)\right) \quad (\text{by definition of } \delta^j \text{ and } \theta^j) \end{array}$$

From Theorem 2.1 it follows that for any $\beta > 0$,

$$\mathbf{Pr}\left(\delta^j \geq (1+\beta)\mathcal{E}(\delta^j)\right) \leq e^{-((1+\beta)ln(1+\beta)-\beta)\mathcal{E}(\delta^j)} = \frac{e^{\beta\mathcal{E}(\delta^j)}}{(1+\beta)^{(1+\beta)\mathcal{E}(\delta^j)}} < \left(\frac{e}{1+\beta}\right)^{(1+\beta)\mathcal{E}(\delta^j)}.$$

With $(1+\beta) = 2(1+\varepsilon)R(m)$ we get:

$$\mathbf{Pr}\left(\delta^j > 4\,(1+\varepsilon)\,\mathsf{SC}\,(\mathbf{w},\mathbf{F})\right) < \left(\frac{e}{2(1+\varepsilon)R(m)}\right)^{2(1+\varepsilon)R(m)\,\mathcal{E}\left(\delta^j\right)}.$$

Note that

$$\begin{array}{ll} & \frac{e}{2(1+\varepsilon)R(m)} \\ < & \frac{e}{2R(m)} \qquad (\text{since } \varepsilon > 0) \\ = & \frac{e}{2\left(\Gamma^{-1}(m)-\frac{3}{2}+o(1)\right)} \quad (\text{by definition of } R(m)) \\ < & \frac{e}{2\Gamma^{-1}(m)-3}\,. \end{array}$$

Thus, $\frac{e}{2\Gamma^{-1}(m)-3} < 1$ if and only if $\Gamma^{-1}(m) > \frac{e+3}{2}$, which holds for all integers $m \geq 3$. Thus, for all such integers $\frac{e}{2(1+\varepsilon)R(m)} < 1$ and $\left(\frac{e}{2(1+\varepsilon)R(m)}\right)^{2(1+\varepsilon)R(m)} < 1$ as well. Hence, $\left(\frac{e}{2(1+\varepsilon)R(m)}\right)^{2(1+\varepsilon)R(m)\mathcal{E}(\delta^j)} < \left(\frac{e}{2(1+\varepsilon)R(m)}\right)^{2(1+\varepsilon)R(m)}$. It follows that

$$\mathbf{Pr}\left(\delta^j > 4\,(1+\varepsilon)\,\mathsf{SC}\,(\mathbf{w},\mathbf{F})\right) \quad < \quad \left(\frac{e}{2(1+\varepsilon)R(m)}\right)^{2(1+\varepsilon)R(m)}.$$

Note, however, that

$$\begin{array}{ll} & \left(\frac{e}{2(1+\varepsilon)R(m)}\right)^{2(1+\varepsilon)R(m)} \\ = & \left(\frac{1}{2}\right)^{2(1+\varepsilon)R(m)} \cdot \left(\left(\frac{e}{(1+\varepsilon)R(m)}\right)^{(1+\varepsilon)R(m)}\right)^2 \\ < & \left(\left(\frac{e}{(1+\varepsilon)R(m)}\right)^{(1+\varepsilon)R(m)}\right)^2 \qquad (\text{since } \left(\frac{1}{2}\right)^{2(1+\varepsilon)R(m)} < 1)\,. \end{array}$$

Define now $\alpha > 0$ so that $\left(\frac{\alpha}{e}\right)^{\alpha} = m$. Then, clearly, $\alpha = \Gamma^{-1}(m) + \Theta(1)$. Note that

$$\begin{aligned}
& (1+\varepsilon)R(m) \\
= \quad & (1+\varepsilon)\Gamma^{-1}(m) - (1+\varepsilon)\tfrac{3}{2} + o(1) \quad \text{(by definition of } R(m)\text{)} \\
= \quad & (1+\varepsilon)\Gamma^{-1}(m) + \Theta(1) \\
> \quad & \alpha \quad \text{(for suitable large } m\text{, since } \varepsilon > 0\text{)}.
\end{aligned}$$

Since $\left(\frac{x}{e}\right)^{x}$ is an increasing function of x, this implies that

$$\left(\frac{(1+\varepsilon)R(m)}{e}\right)^{(1+\varepsilon)R(m)} \quad > \quad \left(\frac{\alpha}{e}\right)^{\alpha} \quad = \quad m\,.$$

This implies that

$$\left(\left(\frac{e}{(1+\varepsilon)R(m)}\right)^{(1+\varepsilon)R(m)}\right)^{2} \quad < \quad \frac{1}{m^2}\,.$$

It follows that

$$\mathbf{Pr}\left(\delta^{j} > 4\,(1+\varepsilon)\,\mathsf{SC}\,(\mathbf{w},\mathbf{F})\right) \quad < \quad \frac{1}{m^2}\,.$$

Hence

$$\begin{aligned}
\mathbf{Pr}\left(\max_{\ell\in[m] \ \mid\ |V^{\ell}|\geq 2} \delta^{\ell} > 4(1+\varepsilon)\mathsf{SC}\,(\mathbf{w},\mathbf{F})\right) \quad &= \quad \mathbf{Pr}\left(\bigvee_{\ell\in[m] \ \mid\ |V^{\ell}|\geq 2} \delta^{\ell} > 4(1+\varepsilon)\mathsf{SC}\,(\mathbf{w},\mathbf{F})\right) \\
&\leq \quad \sum_{\ell\in[m] \ \mid\ |V^{\ell}|\geq 2} \mathbf{Pr}\left(\delta^{\ell} > 4(1+\varepsilon)\mathsf{SC}\,(\mathbf{w},\mathbf{F})\right) \\
&< \quad \sum_{\ell\in[m] \ \mid\ |V^{\ell}|\geq 2} \frac{1}{m^2} \\
&\leq \quad m \cdot \frac{1}{m^2} \\
&= \quad \frac{1}{m}\,.
\end{aligned}$$

Now, clearly,

$$\begin{aligned}
\max_{\ell\in[m]} \delta^{\ell} \quad &= \quad \max\left\{\max_{\ell\in[m] \ \mid\ |V^{\ell}|\geq 2} \delta^{\ell},\ \max_{\ell\in[m] \ \mid\ |V^{\ell}|=1} \delta^{\ell}\right\} \\
&\leq \quad \max_{\ell\in[m] \ \mid\ |V^{\ell}|\geq 2} \delta^{\ell} + \max_{\ell\in[m] \ \mid\ |V^{\ell}|=1} \delta^{\ell} \\
&\leq \quad \max_{\ell\in[m] \ \mid\ |V^{\ell}|\geq 2} \delta^{\ell} + \max_{i\in[n]} w_i \\
&\leq \quad \max_{\ell\in[m] \ \mid\ |V^{\ell}|\geq 2} \delta^{j} + \mathsf{OPT}(\mathbf{w})\,,
\end{aligned}$$

so that

$$\begin{aligned}
\mathcal{E}\left(\max_{\ell\in[m]}\delta^{\ell}\right) &\leq \mathcal{E}\left(\max_{\ell\in[m]\ \mid\ |V^{\ell}|\geq 2}\delta^{j}+\mathsf{OPT}(\mathbf{w})\right)\\
&= \mathcal{E}\left(\max_{\ell\in[m]\ \mid\ |V^{\ell}|\geq 2}\delta^{j}\right)+\mathsf{OPT}(\mathbf{w})\,.
\end{aligned}$$

Note, however, that

$$\begin{aligned}
\mathcal{E}\left(\max_{\ell\in[m]\ \mid\ |V^{\ell}|=1}\delta^{j}\right) &= \sum_{0\leq\delta\leq W}\delta\,\mathbf{Pr}\left(\max_{\ell\in[m]\ ||V^{\ell}|\geq 2}\delta^{\ell}=\delta\right)\\
&\quad \text{(by definition of expectation)}\\
&= \sum_{0\leq\delta\leq 4(1+\varepsilon)\mathsf{SC}(\mathbf{w},\mathbf{F})}\delta\,\mathbf{Pr}\left(\max_{\ell\in[m]\ ||V^{\ell}|\geq 2}\delta^{\ell}=\delta\right)\\
&\quad +\sum_{4(1+\varepsilon)\mathsf{SC}(\mathbf{w},\mathbf{F})<\delta\leq W}\delta\,\mathbf{Pr}\left(\max_{\ell\in[m]\ ||V^{\ell}|\geq 2}\delta^{\ell}=\delta\right)\\
&\leq \sum_{0\leq\delta\leq 4(1+\varepsilon)\mathsf{SC}(\mathbf{w},\mathbf{F})}4(1+\varepsilon)\mathsf{SC}\left(\mathbf{w},\mathbf{F}\right)\ \mathbf{Pr}\left(\max_{\ell\in[m]\ ||V^{\ell}|\geq 2}\delta^{\ell}=\delta\right)\\
&\quad +\sum_{4(1+\varepsilon)\mathsf{SC}(\mathbf{w},\mathbf{F})<\delta\leq W}W\,\mathbf{Pr}\left(\max_{\ell\in[m]\ ||V^{\ell}|\geq 2}\delta^{\ell}=\delta\right)\\
&= 4(1+\varepsilon)\mathsf{SC}\left(\mathbf{w},\mathbf{F}\right)\sum_{0\leq\delta\leq 4(1+\varepsilon)\mathsf{SC}(\mathbf{w},\mathbf{F})}\mathbf{Pr}\left(\max_{\ell\in[m]\ ||V^{\ell}|\geq 2}\delta^{\ell}=\delta\right)\\
&\quad +W\sum_{4(1+\varepsilon)\mathsf{SC}(\mathbf{w},\mathbf{F})<\delta\leq W}\mathbf{Pr}\left(\max_{\ell\in[m]\ ||V^{\ell}|\geq 2}\delta^{\ell}=\delta\right)\\
&= 4(1+\varepsilon)\mathsf{SC}\left(\mathbf{w},\mathbf{F}\right)\ \mathbf{Pr}\left(\max_{\ell\in[m]\ ||V^{\ell}|\geq 2}\delta^{\ell}\leq 4(1+\varepsilon)\mathsf{SC}\left(\mathbf{w},\mathbf{F}\right)\right)\\
&\quad +W\,\mathbf{Pr}\left(\max_{\ell\in[m]\ ||V^{\ell}|\geq 2}\delta^{\ell}>4(1+\varepsilon)\mathsf{SC}\left(\mathbf{w},\mathbf{F}\right)\right)\\
&< 4(1+\varepsilon)\mathsf{SC}\left(\mathbf{w},\mathbf{F}\right)\ \cdot 1+W\ \cdot\frac{1}{m}\\
&\quad \text{(since }\mathbf{Pr}\left(\max_{\ell\in[m]\ \mid\ |V^{\ell}|\geq 2}\delta^{\ell}>4(1+\varepsilon)\mathsf{SC}\left(\mathbf{w},\mathbf{F}\right)\right)<\tfrac{1}{m}\text{)}.
\end{aligned}$$

Hence,

$$\begin{aligned}
&\mathsf{SC}\left(\mathbf{w},\mathbf{P}\right)\\
=&\qquad \mathcal{E}\left(\max_{\ell\in[m]}\delta^{\ell}\right)\\
\leq&\ \mathcal{E}\left(\max_{\ell\in[m]\ \mid\ |V^{\ell}|=1}\delta^{j}\right)+\mathsf{OPT}(\mathbf{w})
\end{aligned}$$

$$\begin{aligned}
&\leq \quad 4(1+\varepsilon)\mathsf{SC}(\mathbf{w},\mathbf{F}) + \tfrac{W}{m} + \mathsf{OPT}(\mathbf{w}) \\
&\leq \quad 4(1+\varepsilon)\mathsf{SC}(\mathbf{w},\mathbf{F}) + 2\,\mathsf{OPT}(\mathbf{w}) \qquad (\text{since } \tfrac{W}{m} \leq \mathsf{OPT}) \\
&\leq \quad 4(1+\varepsilon)\mathsf{SC}(\mathbf{w},\mathbf{F}) + 2\mathsf{SC}(\mathbf{w},\mathbf{F}) \qquad (\text{since } \mathsf{OPT}(\mathbf{w}) \leq \mathsf{SC}(\mathbf{w},\mathbf{F})) \\
&= \quad (6+\epsilon)\mathsf{SC}(\mathbf{w},\mathbf{F}) \ ,
\end{aligned}$$

for any ϵ, where $0 < \epsilon < 1$, as needed. ■

Recall that there is a randomized, polynomial-time approximation scheme (RPTAS) to approximate the social cost of any Nash equilibrium (in particular, the fully mixed) within any arbitrary $\varepsilon > 0$ [7, Theorem 9]. Thus, since, by Theorem 5.1, the worst social cost is bounded by $6 + \varepsilon$ times the social cost of the fully mixed Nash equilibrium, this yields:

Theorem 5.6 *Consider the model of identical capacities. Let $n = m$, m suitable large. Then, there exists a randomized, polynomial-time algorithm with approximation factor $6 + \varepsilon$, for any $\varepsilon > 0$, for* WORST NASH EQUILIBRIUM SOCIAL COST.

We significantly improve Theorem 5.1 under a certain assumption on the traffics.

Theorem 5.7 *Consider any traffic vector $\mathbf{w}$ such that $w_1 \geq w_2 + \ldots + w_n$. Then, for any Nash equilibrium $\mathbf{P}$, $\mathsf{SC}(\mathbf{w},\mathbf{P}) \leq \mathsf{SC}(\mathbf{w},\mathbf{F})$.*

Proof: Since $w_1 \geq w_2 + \ldots + w_n$, it follows that the link with maximum latency has user 1 assigned to it in any pure strategy profile. Thus, in particular, $\mathsf{SC}(\mathbf{w},\mathbf{P}) = \lambda_1(\mathbf{w},\mathbf{P})$ and $\mathsf{SC}(\mathbf{w},\mathbf{F}) = \lambda_1(\mathbf{w},\mathbf{F})$. By Lemma 4.1, $\lambda_1(\mathbf{w},\mathbf{P}) \leq \lambda_1(\mathbf{w},\mathbf{F})$. It follows that $\mathsf{SC}(\mathbf{w},\mathbf{P}) \leq \mathsf{SC}(\mathbf{w},\mathbf{F})$, as needed. ■

6 Best Pure Nash Equilibria and Nashification

We start by establishing $\mathcal{NP}$-hardness for NASHIFY:

Theorem 6.1 NASHIFY *is $\mathcal{NP}$-hard, even if $m = 2$.*

Proof: By reduction from PARTITION. Consider any arbitrary instance of PARTITION consisting of of a set A of k items $a_1, a_2, \ldots, a_k$ with sizes $s(a_1), s(a_2), \ldots, s(a_k) \in \mathbf{N}$, for any integer k. Construct from it an instance of NASHIFY as follows: Set $n = 3k$ and $m = 2$. Set

$w_i = s(a_i)$ for $1 \leq i \leq k$, and $w_i = \frac{1}{2k}$ for $k+1 \leq i \leq 3k$. Take the pure strategy profile that assigns users $1, 2, \ldots, 2k$ to link 1 and users $2k+1, \ldots, 3k$ to link 2.

We establish that this yields a reduction from PARTITION to NASHIFY. Assume first that the instance of PARTITION is positive; that is, there exists a subset $A' \subseteq A$ such that $\sum_{a \in A'} s(a) = \sum_{a \in A \setminus A'} s(a)$. Since either $|A'| \leq \frac{k}{2}$ or $|A \setminus A'| \leq \frac{k}{2}$, assume, without loss of generality, that $|A'| \leq \frac{k}{2}$. Note that each user assigned to link 1 is unsatisfied in the constructed pure strategy profile since its latency cost on link 1 is $\sum_{a \in A} s(a) + k \cdot \frac{1}{2k} = \sum_{a \in A} s(a) + \frac{1}{2}$, while its latency cost on link 2 is $k \cdot \frac{1}{2k} = \frac{1}{2}$, which is less. Thus, each step that transfers an unsatisfied user that corresponds to an element $a \in A'$ from link 1 to link 2 is a selfish step, and the sequence of steps that transfer all users that correspond to elements of A' from link 1 to link 2 is a sequence of at most $\frac{k}{2} < k$ steps. As a result of this sequence of selfish steps, the latency of link 1 will be $\sum_{a \in A \setminus A'} s(a) + \frac{1}{2}$, while the latency of link 2 will be $\sum_{a \in A'} s(a) + \frac{1}{2}$. Since $\sum_{a \in A'} s(a) = \sum_{a \in A \setminus A'} s(a)$, these two latencies are equal and the resulting pure strategy profile is therefore a Nash equilibrium which implies that NASHIFY is positive.

Assume now that the instance of NASHIFY is positive; that is, there exists a sequence of at most k selfish steps that transforms the pure strategy profile in the constructed instance of NASHIFY to a Nash equilibrium. Assume that in the resulting pure strategy profile users corresponding to a subset $A' \subseteq A$ remain in link 1, users corresponding to the subset $A \setminus A' \subseteq A$ are transfered to link 2, while the sums of traffics of users with traffic $\frac{1}{2k}$ that reside in link 1 and link 2 are x and $1 - x$, respectively; thus, the latencies of links 1 and 2 are $\sum_{a \in A'} s(a) + x$ and $\sum_{a \in A \setminus A'} s(a) + 1 - x$, respectively. We consider two cases:

Assume first that $A' = A$. Then after at most k selfish steps the latency on link 2 is at most 1 whereas the latency on link 1 is at least $\sum_{a \in A} s(a) \geq k$. So there exists an unsatisfied user $a \in A$, a contradiction to the fact that NASHIFY is positive. So let $A' \neq A$. We show that this implies $\sum_{a \in A'} s(a) - \sum_{a \in A \setminus A'} s(a) = 0$. Assume $|\sum_{a \in A'} s(a) - \sum_{a \in A \setminus A'} s(a)| \neq 0$. Since the traffics of users in A are integer, this implies $|\sum_{a \in A'} s(a) - \sum_{a \in A \setminus A'} s(a)| \geq 1$. The fact that $A' \neq A$ shows that at least one user with large traffic was transformed to link 2. So we can make at most $k - 1$ selfish steps with the small traffics. However, transforming $k - 1$ small traffics to the link with smaller latency leaves one user with small traffic unsatisfied, a contradiction to the fact that NASHIFY is positive. So $|\sum_{a \in A'} s(a) - \sum_{a \in A \setminus A'} s(a)| = 0$ which implies that PARTITION is positive. ■

We remark that NASHIFY is $\mathcal{NP}$-complete in the strong sense (cf. [10, Section 4.2]) if m is part of the input. Thus, there is no pseudopolynomial-time algorithm for NASHIFY (unless $\mathcal{P} = \mathcal{NP}$). In contrast, there is a natural pseudopolynomial-time algorithm $\mathsf{A}_{k-nashify}$ for

Algorithm $\mathsf{A}_{nashify}$:
INPUT: A pure strategy profile $\mathbf{L}$ of n users with traffics $w_1, \ldots, w_n$.
OUTPUT: A pure strategy profile $\mathbf{L}'$ that is a Nash equilibrium.

- Sort the user traffics in non-increasing order so that $w_1 \geq \ldots \geq w_n$.
- For each user $i := 1$ to n, **do**
  - remove user i from the link it is currently assigned;
  - find the link ℓ with the minimum latency;
  - reassign user i to the link ℓ.

  od
- **Return** the resulting pure strategy profile $\mathbf{L}'$.

Figure 1: The algorithm $\mathsf{A}_{nashify}$

k-NASHIFY, which exhaustively searches all sequences of k selfish steps; since a selfish step involves a (unsatisfied) user and a link for a total of mn choices, the running time of $\mathsf{A}_{k-nashify}$ is $\Theta((mn)^k)$. We continue to present an algorithm $\mathsf{A}_{nashify}$ that solves NASHIFY when n selfish steps are allowed.

The algorithm $\mathsf{A}_{nashify}$ sorts the user traffics in non-increasing order so that $w_1 \geq \ldots \geq w_n$. Then for each user $i := 1$ to n, it removes user i from the link it is currently assigned, it finds the link ℓ with the minimum latency, and it reassigns user i to the link ℓ. We prove:

Lemma 6.2 *A greedy selfish step of an unsatisfied user i with traffic w_i makes no user k with traffic $w_k \geq w_i$ unsatisfied.*

Proof: Let $\mathbf{L} = \langle l_1, \ldots, l_n \rangle$ be a pure strategy profile. Furthermore, let $p = l_i$, and let q be the link with minimum latency. Denote λ^j and $\widehat{\lambda}^j$ the latency of link $j \in [m]$ before and after user i changed its strategy, respectively. Assume that user k becomes unsatisfied due to the move of user i. Since only the latency on link p and q changed, we have to distinguish between two cases. Either $l_k \neq q$ and user k wants to change its strategy to p, or $l_k = q$ and user k becomes unsatisfied due to the additional traffic w_i on link q.

First, assume that $l_k \neq q$, and that user k wants to change its strategy to p. Since user i changed its strategy from p to q we know that $\lambda^q < \widehat{\lambda}^p$ and therefore $w_k + \lambda^q < w_k + \widehat{\lambda}^p$. So if user k wants to change its strategy to p, then user k was already unsatisfied before user i changed its strategy, a contradiction.

For the case that the strategy of user k is q we define $\widetilde{\lambda}_q = \lambda^q - w_k$. We have $\forall j \in [m]$: $\lambda^j + w_k \geq \lambda^j + w_i \geq \lambda^q + w_i = \widetilde{\lambda}_q + w_k + w_i$. Therefore k stays satisfied. ■

Theorem 6.3 *Let* $\mathbf{L} = \langle l_1, \ldots, l_n \rangle$ *be a pure strategy profile for* n *users with traffics* $w_1, ..., w_n$ *on* m *links with social cost* $\mathsf{SC}(\mathbf{w}, \mathbf{L})$. *Then algorithm* $\mathsf{A}_{nashify}$ *computes a Nash equilibrium from* $\mathbf{L}$ *with social cost* $\leq \mathsf{SC}(\mathbf{w}, \mathbf{L})$ *using* $O(n \lg n)$ *time.*

Proof: In order to complete the proof of Theorem 6.3, we have to show that algorithm $\mathsf{A}_{nashify}$ returns a pure strategy profile $\mathbf{L}'$ that is a Nash equilibrium and has social cost $\mathsf{SC}(\mathbf{w}, \mathbf{L}') \leq \mathsf{SC}(\mathbf{w}, \mathbf{L})$. It is easy to see that $\mathsf{SC}(\mathbf{w}, \mathbf{L}') \leq \mathsf{SC}(\mathbf{w}, \mathbf{L})$, since for user j we always choose the link with lowest latency as its strategy. After every iteration the user that changed its strategy is satisfied. Since we go through the list of users in descending order of their traffic and because of Lemma 6.2, all users that changed their strategy in earlier iterations stay satisfied. Therefore after we went through the complete list of users, all users are satisfied and thus $\mathbf{L}'$ is a Nash equilibrium.

The running time of algorithm $\mathsf{A}_{nashify}$ is $O(n \lg n)$ for sorting the n user traffics, $O(m \lg m)$ for constructing a heap with all latencies in the input pure strategy profile $\mathbf{L}$, and $O(n \lg m)$ for finding the minimum element of the heap in each of the n iterations of the algorithm. Thus, the total running time is $O(n \lg n + m \lg m + n \lg m)$. The interesting case is when $m \leq n$ (since otherwise, a single user can be assigned to each link, achieving an optimal Nash equilibrium). Thus, in the interesting case, the total running time of $\mathsf{A}_{nashify}$ is $O(n \lg n)$. ■

Running the PTAS of Hochbaum and Shmoys [12] for scheduling n *jobs* on m identical *machines* yields a pure strategy profile $\mathbf{L}$ such that $\mathsf{SC}\,(\mathbf{w}, \mathbf{L}) \leq (1 + \varepsilon)\,\mathsf{OPT}(\mathbf{w})$. On the other hand, applying the algorithm $\mathsf{A}_{nashify}$ on $\mathbf{L}$ yields a Nash equilibrium $\mathbf{L}'$ such that $\mathsf{SC}\,(\mathbf{w}, \mathbf{L}') \leq \mathsf{SC}\,(\mathbf{w}, \mathbf{L})$. Thus, $\mathsf{SC}\,(\mathbf{w}, \mathbf{L}') \leq (1 + \varepsilon)\mathsf{OPT}\,(\mathbf{w})$. Since also $\mathsf{OPT}\,(\mathbf{w}) \leq \mathsf{SC}\,(\mathbf{w}, \mathbf{L}')$, it follows that:

Theorem 6.4 *There exists a* PTAS *for* BEST PURE NASH EQUILIBRIUM, *for the model of identical capacities.*

7 Worst Pure Nash Equilibria

Denote with m-WCpNE the decision problem corresponding to the problem to compute the worst-case pure Nash equilibrium for n users with traffics $w_1, \ldots, w_n$ on m links. If m is part of the input, then we call the problem WCpNE. We first show:

Theorem 7.1 *Fix any traffic vector* $\mathbf{w}$ *and pure Nash equilibrium* $\mathbf{L}$*. Then,* $\frac{\mathsf{SC}(\mathbf{w},\mathbf{L})}{\mathsf{OPT}(\mathbf{w})} \leq 2 - \frac{2}{m+1}$*. Furthermore, this upper bound is tight.*

Proof: Schuurman and Vredeveld [29] showed the tightness of the upper bound for jump optimal schedules proved by Finn and Horowitz [6]. Since every pure Nash equilibrium is also jump optimal, the upper bound follows directly. Greedy selfish steps on identical links can only increase the minimum load over all links. Thus, we can transform every jump optimal schedule into a Nash equilibrium without altering the makespan, proving tightness. ■

Theorem 7.2 *It is* $\mathcal{NP}$*-hard to find a pure Nash equilibrium* L *with* $\frac{\mathsf{WC}(\mathbf{w})}{\mathsf{SC}(\mathbf{w},\mathbf{L})} < 2 - \frac{2}{m+1} - \varepsilon$*, for any* $\varepsilon > 0$*. It is* $\mathcal{NP}$*-hard in the strong sense if the number of links* m *is part of the input.*

Proof: We show that for a certain class of instances we have to solve $\mathsf{BIN\ PACKING}$ in order to find a Nash equilibrium with desired property. $\mathsf{BIN\ PACKING}$ is $\mathcal{NP}$-complete in the strong sense [10]. Consider an arbitrary instance of $\mathsf{BIN\ PACKING}$ consisting of a set of items $\mathcal{U} = \{u_1, \ldots, u_{|\mathcal{U}|}\}$ with sizes $s(u_j) \leq \delta$, $\sum_{u_j \in \mathcal{U}} = m-1$, and $K = m-1$ bins of capacity $B = 1$. From this instance we construct an instance for the stated problem as follows: Set $\varepsilon = 2\delta$. There are $n - 2 = |\mathcal{U}|$ users with traffic $w_i = s(u_i)$ and two users with traffic $w_{n-1} = w_n = 1$. Note that the social cost of a Nash Equilibrium is either 2 when the users with traffic 1 are on the same link, or at most $\frac{m+1}{m} + \delta$ otherwise.

If $\mathsf{BIN\ PACKING}$ is negative, then there exists no Nash equilibrium with both users with traffic 1 on the same link. Thus every Nash equilibrium has the desired property. If $\mathsf{BIN\ PACKING}$ is positive, then there exists a Nash equilibrium with both users with traffic 1 on the same link. The social cost of this Nash equilibrium is $\mathsf{WC}(\mathbf{w}) = 2$. For any other Nash Equilibrium $\mathbf{L}$ where the users with traffic 1 use different links, $\mathsf{SC}(\mathbf{w},\mathbf{L}) \leq \frac{m+1}{m} + \delta$. This yields

$$\begin{aligned} \frac{\mathsf{WC}(\mathbf{w})}{\mathsf{SC}(\mathbf{w},\mathbf{L})} &\geq \frac{2}{\frac{m+1}{m} + \delta} \\ &= \frac{2}{\frac{m+1}{m} + \frac{\varepsilon}{2}} \end{aligned}$$

$$
\begin{aligned}
&= \frac{2m}{m+1+\frac{\varepsilon m}{2}} \\
&= 2 - \frac{2}{m+1+\frac{\varepsilon m}{2}} - \frac{\varepsilon m}{m+1+\frac{\varepsilon m}{2}} \\
&> 2 - \frac{2}{m+1} - \varepsilon\,.
\end{aligned}
$$

So, to find a Nash equilibrium with desired property, we have to find a distribution of the small traffics $w_1, \ldots, w_{n-2}$ to $m-1$ links which solves BIN PACKING.

Since BIN PACKING is $\mathcal{NP}$-hard in the strong sense, if the number of bins is part of the input, it follows that computing a pure Nash equilibrium L with $\frac{\mathsf{WC}(\mathbf{w})}{\mathsf{SC}(\mathbf{w},\mathbf{L})} < 2 - \frac{2}{m+1} - \varepsilon$ is also $\mathcal{NP}$-hard in the strong sense, if m is part of the input. ■

Since WCpNE is $\mathcal{NP}$-hard in the strong sense [7], there exists no pseudopolynomial algorithm to solve WCpNE. However, we can give such an algorithm for m-WCpNE.

Theorem 7.3 *There exists a pseudopolynomial-time algorithm for* m-WCpNE.

Proof: We start with the state set S_0 in which all links are empty. After inserting the first i traffics the state set S_i consists of all $(2m)$-tuples $(\lambda_1, \tilde{w}_1, \ldots, \lambda_m, \tilde{w}_m)$ describing a possible placement of the largest i traffics with λ_j being the latency on link j and $\tilde{w}_j$ the smallest traffic placed on link j. We need at most $m \cdot |S_i|$ steps to create S_{i+1} from S_i, and $|S_i| \leq (W_i)^m \cdot (w_1)^m$, where $W_i = \sum_{j=1}^{i} w_j$. Therefore the overall computation time is bounded by $O(n \cdot m \cdot W^m \cdot (w_1)^m)$. The best-case Nash equilibrium and the worst-case Nash equilibrium can be found by exhaustive search over the state set S_n using $O(n \cdot m \cdot W^m \cdot (w_1)^m)$ time. ■

Remark. Theorem 7.3 also holds for the case of arbitrary link capacities.

8 Conclusions and Discussion

In this work, we have studied the combinatorial structure and the computational complexity of the extreme (either *worst* or *best*) Nash equilibria for the selfish routing game introduced in the pioneering work of Koutsoupias and Papadimitriou [17].

Our study of the combinatorial structure has revealed an interesting, highly non-trivial, combinatorial conjecture about the worst such Nash equilibrium, namely the *Fully Mixed Nash Equilibrium Conjecture,* abbreviated as FMNE Conjecture; the conjecture states that the fully mixed Nash equilibrium [19] is the worst Nash equilibrium in the setting we consider. We have established that the FMNE Conjecture is valid when restricted to pure Nash equilibria. Furthermore, we have come close to establishing the FMNE Conjecture in its full generality by proving that the social cost of any (pure or mixed) Nash equilibrium is within a factor of $6+\varepsilon$, for any $\varepsilon > 0$, of that of the fully mixed Nash equilibrium, under the assumptions that all link capacities are identical, the number of users is equal to the number of links and the number of links is suitable large. The proof of this result has relied very heavily on applying and extending techniques from the theory of *stochastic orders* and *majorization* [18, 30]; such techniques are imported for the *first* time into the context of selfish routing, and their application and extension are both of independent interest. We hope that the application and extension of techniques from the theory of stochastic orders and majorization will be valuable to further studies of the selfish routing game considered in this paper and for the analysis and evaluation of mixed Nash equilibria for other games as well.

Our study of the computational complexity of extreme Nash equilibria has resulted in both positive and negative results. On the positive side, we have devised, for the case of identical link capacities, equal number of users and links and a suitable large number of links, a randomized, polynomial-time algorithm to approximate the worst social cost within a factor arbitrarily close to $6+\varepsilon$, for any $\varepsilon > 0$. The approximation factor $6+\varepsilon$ of this randomized algorithm will immediately improve upon reducing 6 further down in our combinatorial result described above, relating the social cost of any Nash equilibrium to that of the fully mixed. We have also introduced the technique of *Nashification* as a tool for converging to a Nash equilibrium starting with any assignment of users to links in a way that does not increase the social cost; coupling this technique with a polynomial-time approximation scheme for the optimal assignment of users to links [12] has yielded a polynomial-time approximation scheme for the social cost of the *best* Nash equilibrium. In sharp contrast, we have established a *tight* limit on the approximation factor of any polynomial-time algorithm that approximates the social cost of the *worst* Nash equilibrium (assuming $\mathcal{P} \neq \mathcal{NP}$). Our approximability and inapproximability results for the best and worst Nash equilibria, respectively, establish an essential difference between the approximation properties of the two types of extreme Nash equilibria.

The most obvious problem left open by our work is to establish the FMNE Conjecture. Some progress on this problem has been already reported by Lücking *et al.* [8], where the conjecture is proved in various special cases of the model of selfish routing introduced by Koutsoupias and Papadimitriou [17] and considered in this work; furthermore, Lücking *et al.* disprove the FMNE Conjecture in a different model for selfish routing that borrows from the model of *unrelated machines* [13] studied in the scheduling literature.

The technique of *Nashification,* as an algorithmic tool for the computation of Nash equilibria, deserves also further study. Some steps in this direction have been taken already by Feldmann *et al.* [5].

Establishment of the Fully Mixed Nash Equilibrium Conjecture will reveal an interesting complexity-theoretic contrast between the worst pure and mixed Nash equilibria. On one hand, computing the (supports of) the *worst* pure Nash equilibrium is an $\mathcal{NP}$-hard problem [7, Theorem 4]; however, computing the social cost of a worst pure Nash equilibrium is trivially in $\mathcal{P}$ (since it amounts to computing the maximum). On the other hand, if the Fully Mixed Nash Equilibrium Conjecture is true, computing the supports of a worst mixed Nash equilibrium is a trivial problem and, moreover, the polynomial characterization of the fully mixed Nash equilibrium shown in [19, Theorem 14] implies that a worst mixed Nash equilibrium can be computed in polynomial time; however, computing the social of a worst mixed Nash equilibrium remains $\#\mathcal{P}$-complete. This result follows from an inspection of the proof of [7, Theorem 8], which establishes that computing the social cost of a Nash equilibrium is a $\#\mathcal{P}$-complete problem. We consider this different behavior of pure and mixed Nash equilibria to be an interesting complexity-theoretic consequence of the Fully Mixed Nash Equilibrium Conjecture.

Acknowledgments. We would like to thank Rainer Feldmann and Manuel Rode for many fruitful discussions. We are also very grateful to Petra Berenbrink and Tasos Christophides for many helpful discussions on stochastic orders.

References

[1] P. Brucker, J. Hurink and F. Werner, "Improving Local Search Heuristics for Some Scheduling Problems. Part II," *Discrete Applied Mathematics,* Vol. 72, No.1-2, pp. 47–69, 1997.

[2] A. Czumaj and B. Vöcking, "Tight Bounds for Worst-Case Equilibria," *Proceedings of the 13th Annual ACM Symposium on Discrete Algorithms,* pp. 413–420, January 2002.

[3] R. Datta, "Universality of Nash Equilibria," unpublished manuscript, August 2002.

[4] X. Deng, C. Papadimitriou and S. Safra, "On the Complexity of Equilibria," *Proceedings of the 34th Annual ACM Symposium on Theory of Computing,* pp. 67–71, May 2002.

[5] R. Feldmann, M. Gairing, T. Lücking, B. Monien and M. Rode, "Nashification and the Coordination Ratio for a Selfish Routing Game," June/July 2003, to appear.

[6] G. Finn and E. Horowitz, "A linear time approximation algorithm for multiprocessor scheduling," *BIT,* Vol. 19, pp. 312–320, 1979.

[7] D. Fotakis, S. Kontogiannis, E. Koutsoupias, M. Mavronicolas and P. Spirakis, "The Structure and Complexity of Nash Equilibria for a Selfish Routing Game,' *Proceedings of the 29th International Colloquium on Automata, Languages and Programming,* P. Widmayer, F. Triguero, R. Morales, M. Hennessy, S. Eidenbenz and R. Conejo eds., pp. 123–134, Vol. 2380, Lecture Notes in Computer Science, Springer-Verlag, Málaga, Spain, July 2002.

[8] T. Lücking, M. Mavronicolas, B. Monien, M. Rode, P. Spirakis and I. Vrto, "Which is the Worst-case Nash Equilibrium?" *26th International Symposium on Mathematical Foundations of Computer Science,* August 2003, to appear.

[9] M. R. Garey and D. S. Johnson, "Complexity Results for Multiprocessor Scheduling Under Resoiurce Constraints," *SIAM Journal on Computing,* Vol. 4, pp. 397–411, 1975.

[10] M. R. Garey and D. S. Johnson, *Computers and intractability: A Guide to the Theory of NP-Completeness,* W. H. Freeman and Company, 1979.

[11] G. H. Gonnet, "Expected Length of the Longest Probe Sequence in Hash Code Searching," *Journal of the ACM,* Vol. 28, No. 2, pp. 289–304, April 1981.

[12] D. S. Hochbaum and D. Shmoys, "Using Dual Approximation Algorithms for Scheduling Problems: Theoretical and Practical Results," *Journal of the ACM,* Vol. 34, No. 1, pp. 144–162, 1987.

[13] E. Horowitz and S. Sahni, "Exact and Approximate Algorithms for Scheduling Non-Identical Processors," *Journal of the ACM,* Vol. 23, No. 2, pp. 317–327, 1976.

[14] R. M. Karp, "Reducibility among Combinatorial Problems," in R. E. Miller and J. W. Thatcher eds., *Complexity of Computer Computations,* pp. 85–103, Plenum Press, New York, 1972.

[15] V. F. Kolchin, V. P. Chistiakov and B. A. Sevastianov, *Random Allocations,* V. H. Winston, New York, 1978.

[16] E. Koutsoupias, M. Mavronicolas and P. Spirakis, "Approximate Equilibria and Ball Fusion," *Proceedings of the 9th International Colloquium on Structural Information and Communication Complexity,* Andros, Greece, June 2002. Accepted to *Theory of Computing Systems.*

Earlier version appeared as "A Tight Bound on Coordination Ratio," Technical Report 0100229, Department of Computer Science, University of California at Los Angeles, April 2001.

[17] E. Koutsoupias and C. H. Papadimitriou, "Worst-case Equilibria," *Proceedings of the 16th Annual Symposium on Theoretical Aspects of Computer Science,* G. Meinel and S. Tison eds., pp. 404–413, Vol. 1563, Lecture Notes in Computer Science, Springer-Verlag, Trier, Germany, March 1999.

[18] A. Marshall and I. Olkin, *Theory of Majorization and Its Applications,* Academic Press, Orlando, FL, 1979.

[19] M. Mavronicolas and P. Spirakis, "The Price of Selfish Routing," *Proceedings of the 33rd Annual ACM Symposium on Theory of Computing,* pp. 510–519, July 2001.

[20] C. McDiarmid, "Concentration," Chapter 9 in *Probabilistic Methods for Algorithmic Discrete Mathematics,* M. Habib, C. McDiarmidt, J. Ramires-Alfonsin and B. Reed eds., Springer, 1998.

[21] H. Moulin and L. Vial, "Strategically Zero-Sum Games: The Class of Games whose Completely Mixed Equilibria Cannot be Improved Upon," *International Journal of Game Theory,* Vol. 7, Nos. 3/4, pp. 201–221, 1978.

[22] J. F. Nash, "Equilibrium Points in N-Person Games," *Proceedings of the National Academy of Sciences,* Vol. 36, pp. 48–49, 1950.

[23] J. F. Nash, “Non-cooperative Games,” *Annals of Mathematics,* Vol. 54, No. 2, pp. 286–295, 1951.

[24] M. J. Osborne and A. Rubinstein, *A Course in Game Theory,* The MIT Press, 1994.

[25] C. H. Papadimitriou, *Computational Complexity,* Addison Wesley, 1994.

[26] C. H. Papadimitriou, "Algorithms, Games and the Internet," *Proceedings of the 33rd Annual ACM Symposium on Theory of Computing,* pp. 749–753, July 2001.

[27] T. E. S. Raghavan, "Completely Mixed Strategies in Bimatrix Games," *Journal of London Mathematical Society,* Vol. 2, No. 2, pp. 709–712, 1970.

[28] S. M. Ross, *Stochastic Processes,* Second Edition, John Wiley & Sons, Inc., 1996.

[29] P. Schuurman and T. Vredeveld, "Performance Guarantees of Load Search for Multiprocessor Scheduling," *Proceedings of the 8th Conference on Integer Programming and Combinatorial Optimization,* pp. 370–382, June 2001.

[30] M. Shaked and J. G. Shanthikumar, *Stochastic Orders and Their Applications,* Academic Press, San Diego, CA, 1994.

[31] A. Vetta, "Nash Equilibria in Competitive Societies, with Applications to Facility Location, Traffic Routing and Auctions," *Proceedings of the 43rd Annual IEEE Symposium on Foundations of Computer Science,* October 2002, pp. 416–425.