Fast Convergence to Wardrop Equilibria by Adaptive Sampling Methods

Simon Fischer[*]
Dept. of Computer Science
RWTH Aachen, Germany
fischer@cs.rwth-aachen.de

Harald Räcke[†]
Toyota Technological Institute
Chicago, Illinois, USA
harry@cs.cmu.edu

Berthold Vöcking[‡]
Dept. of Computer Science
RWTH Aachen, Germany
voecking@cs.rwth-aachen.de

ABSTRACT

We study rerouting policies in a dynamic round-based variant of a well known game theoretic traffic model due to Wardrop. Previous analyses (mostly in the context of selfish routing) based on Wardrop's model focus mostly on the static analysis of equilibria. In this paper, we ask the question whether the population of agents responsible for routing the traffic can jointly *compute* or better *learn* a Wardrop equilibrium efficiently. The rerouting policies that we study are of the following kind. In each round, each agent samples an alternative routing path and compares the latency on this path with its current latency. If the agent observes that it can improve its latency then it switches with some probability depending on the possible improvement to the better path.

We can show various positive results based on a rerouting policy using an adaptive sampling rule that implicitly amplifies paths that carry a large amount of traffic in the Wardrop equilibrium. For general asymmetric games, we show that a simple *replication protocol* in which agents adopt strategies of more successful agents reaches a certain kind of bicriteria equilibrium within a time bound that is independent of the size and the structure of the network but only depends on a parameter of the latency functions, that we call the *relative slope*. For symmetric games, this result has an intuitive interpretation: *Replication approximately satisfies almost everyone very quickly.*

In order to achieve convergence to a Wardrop equilibrium besides replication one also needs an exploration component discovering possibly unused strategies. We present a sampling based *replication-exploration protocol* and analyze its convergence time for symmetric games. For example, if the latency functions are defined by positive polynomials in coefficient representation, the convergence time is polynomial in the representation length of the latency functions. To the best of our knowledge, all previous results on the speed of convergence towards Wardrop equilibria, even when restricted to linear latency functions, were pseudopolynomial.

In addition to the upper bounds on the speed of convergence, we can also present a lower bound demonstrating the necessity of adaptive sampling by showing that static sampling methods result in a slowdown that is exponential in the size of the network. A further lower bound illustrates that the relative slope is, in fact, the relevant parameter that determines the speed of convergence.

Categories and Subject Descriptors

F.2.2 [**Analysis of algorithms and problem complexity**]: Nonnumerical Algorithms and Problems—*Routing and layout*; C.2.2 [**Computer-communication networks**]: Network Protocols—*Routing protocols*

General Terms

Theory, Algorithms

Keywords

Adaptive routing, Wardrop equilibria, convergence time

1. INTRODUCTION

Recent contributions in the field of algorithmic game theory have provided much insight into the structure of Nash equilibria for routing in networks that lack central coordination. Prominent results include bounds on the *price of anarchy* measuring the performance loss due to selfishness in relation to the centrally optimized solution, see, e. g., [21, 22], and questions regarding how to design networks such that equilibria induced by selfish agents coincide with the globally optimal solution, e. g., by imposing taxes [8, 13] or by introducing a global instance that controls a small fraction of the traffic [14, 20]. These static analyses of Nash equilibria disregard the question of how an equilibrium is actually reached. Classical game theory does not give an answer to this question either. The motivation of Nash equilibria is based on idealistic assumptions like unbounded rationality

[*]Supported by DFG grant Vo889/1-3.

[†]Part of this work was conducted while the author stayed at the RWTH Aachen under support of the EU within the 6th Framework Programme under contract 001907 (DELIS).

[‡]Supported in part by the EU within the 6th Framework Programme under contract 001907 (DELIS) and by DFG grant Vo889/1-3.

STOC'06, May 21–23, 2006, Seattle, Washington, USA.

and global knowledge that, however, are rarely fulfilled in real-world networks like the Internet.

In this paper, we study the question of how a large population of agents can *compute* or *learn* an equilibrium efficiently based on simple sampling and adaption policies. Our motivation is twofold. On the one hand, we want to support the previous results about Nash equilibria by showing that a population of agents following simple, myopic, and reasonable rules quickly converges to a Nash equilibrium. On the other hand, we think that our analysis may contribute to the design of distributed adaptive re-routing protocols that quickly converge to stable routing allocations. Our study is based on the well known traffic model of Wardrop [23] (see also [22]) in which each of an infinite number of agents is responsible for an infinitesimal amount of traffic. We imagine that the agents play a repeated game in rounds. In each round, each agent may compare the latency of his current route with the latency of another route and switch to the other route if it promises a better latency. The problem with this natural approach is that other agents might switch simultaneously to the same route so that the latency of an agent may not improve or even get worse. This way, the game may get stuck in oscillations. This phenomenon is also well known in the *networks* community as the instabilities due to oscillations observed within the ARPANET project are one of the major reasons why the Internet does not support adaptive routing, see e. g. [15, 16, 19].

In [12], it was shown that such oscillation effects can be avoided by letting the agents sample alternative routes at random and migrate with a probability depending on the observed latency difference. The weakness of the routing protocols presented in [12] is that the migration policy depends heavily on the first derivatives of the latency functions: In order to avoid oscillation the probability to switch to another path is scaled down by a factor that is linear in the maximum first derivative over all latency functions. While this is effective in avoiding oscillation effects it also slows down the routing process in a dramatic way. For example, when assuming linear latency functions the obtained bounds on the convergence time depend in a pseudopolynomial way on the ratio between the largest and the smallest coefficient over all latency functions. Several other approaches (see, e. g. [2, 3, 6, 7]) tackling similar problems are discussed in the literature. All of them depend on some kind of network parameter like, e. g., the maximum first derivative or the maximum latency, in a pseudopolynomial fashion. For a more detailed discussion on related work see Section 1.3.

In this work, we show that the first derivative of the latency functions is not the limiting factor in the speed of convergence towards Nash equilibria. We will provide upper and lower bounds that identify the "relative slope" instead of the first derivative as the relevant parameter that determines the convergence time of adaptive routing policies. This parameter is a generalization of the polynomial degree of a function. Our approach enables us to obtain the first polynomial bounds on the convergence time of adaptive rerouting policies for classes of latency functions with bounded relative slope, especially for latency functions defined by positive polynomials. Remarkably, some of our upper bounds are completely independent of any parameter reflecting the size or the structure of the network but depend only on the latency functions. More specifically, they depend in a linear fashion on the maximum relative slope over all latency functions. Before describing our results in more detail, let us give the necessary formal definitions.

1.1 The Model

1.1.1 *Wardrop's traffic model.*

We consider a model for selfish routing where an infinite population of agents carries an infinitesimal amount of load each [22, 23]. Let E denote a set of *resources (edges)* with continuous, non-decreasing *latency functions* $\ell_e : [0, 1] \mapsto \mathbb{R}_{\geq 0}$. Furthermore, let $[k] = \{1, \ldots, k\}$ denote a set of *commodities* with *flow demands* or *rates* r_i, $i \in [k]$ such that $\sum_{i=1}^{k} r_i = r$. We normalize $r = 1$. For every commodity $i \in [k]$ let $\mathcal{P}_i \subseteq 2^E$ denote a set of strategies *(paths)* available for commodity i. Let $\mathcal{P} = \cup_{i\in[k]}\mathcal{P}_i$ and let $L = \max_{P\in\mathcal{P}} |P|$. By $\Gamma = (E, (\ell_e)_{e\in E}, (\mathcal{P}_i)_{i\in[k]}, (r_i)_{i\in[k]})$ we denote an *instance* of the routing game. The instance is *symmetric* if $k = 1$ and *asymmetric* otherwise. An instance is *single-resource* if for all $P \in \mathcal{P}$, $|P| = 1$.

For $P \in \mathcal{P}$, let f_P denote the volume of agents utilizing strategy P. A population or flow vector $(f_P)_{P\in\mathcal{P}}$ is *feasible* if for all $i \in [k]$, $\sum_{P\in\mathcal{P}_i} f_P = r_i$. Let $f_e = \sum_{P\ni e} f_P$ denote the load of resource $e \in E$. Then, the latency of a resource $e \in E$ is $\ell_e(f) = \ell_e(f_e)$ and the latency of a strategy is $\ell_P(f) = \sum_{e\in P} \ell_e(f)$. By $\bar{\ell}(f) = \sum_{P\in\mathcal{P}} f_P\, \ell_P(f) = \sum_{e\in E} f_e\, \ell_e(f)$ denote the overall average latency, and for $i \in [k]$ let $\bar{\ell}_i = \sum_{P\in\mathcal{P}_i} (f_P/r_i) \cdot \ell_P(f)$ denote the average latency of commodity i.

We are interested in flow assignments that are stable in the sense that no agent can improve their situation by changing their strategy unilaterally.

Definition 1 (Wardrop equilibrium [23]). *A feasible flow vector $(f_P)_{P\in\mathcal{P}}$ is at a* Wardrop equilibrium *for the instance Γ if for every commodity $i \in [k]$ and every $P, P' \in \mathcal{P}_i$ with $f_P > 0$ it holds that $\ell_P(f) \leq \ell_{P'}(f)$.*

1.1.2 *Potential and α-shifted potential.*

A natural and nice potential function by Beckmann *et al.* [4] allows to formulate the problem of computing a Wardrop equilibrium in form of a convex optimization problem, see also [22]. The set of allocations in equilibrium coincides with the set of allocations minimizing the potential function

$$\Phi(f) = \sum_{e\in E} \int_0^{f_e} \ell(x)\, dx \ .$$

The allocations in equilibrium do not only all have the same (optimal) potential but they also impose the same latencies on all edges. In this sense, the Wardrop equilibrium is essentially unique. Our goal is the design of distributed rerouting policies that approximate the Wardrop equilibrium. As a measure for the quality of approximation, we upper-bound the factor between the potential achieved after a certain amount of time divided by the minimal potential Φ^*. Observe, however, for certain instances of the routing game, Φ^* might be zero. In this case, we suggest to shift the potential by some positive additive term. In general, we consider an *α-shifted potential* of the form $\Phi + \alpha$ where $\alpha \geq 0$ can be chosen arbitrarily in such a way that, for the given instance, $\Phi^* + \alpha$ is strictly positive. Let us remark that shifting the potential can be interpreted as adding a virtual amount of α to the latency observed on every path.

In fact, this is the way how our algorithms make use of this parameter.

1.1.3 Relative slope

A rerouting policy cannot guarantee convergence to Wardrop equilibria if the latency functions make arbitrarily large leaps due to minor shifts of the flow. To restrict the number of agents migrating simultaneously, it must have some information about the behavior of the latency functions. Our analysis shows that the following parameter is relevant.

DEFINITION 2 (RELATIVE SLOPE). *A differentiable latency function ℓ has* relative slope d *at x if $\ell'(x) \le d \cdot \ell(x)/x$. A latency function has relative slope d if it has relative slope d over the entire range $[0,1]$ and a class of latency functions $\mathcal{L}$ has relative slope d if every $\ell \in \mathcal{L}$ has relative slope d.*

The polynomial function $\ell(x) = a\,x^d$ has relative slope d over the entire range. The exponential function $\ell(x) = a \cdot \exp(\lambda\,x)$ has relative slope at most λ for $x \in [0,1]$ reaching its maximum λ for $x = 1$.

We will use the following two facts frequently.

FACT 1. *If the function ℓ has relative slope d and if $0 \le \delta \le 1/(2\,d)$, then $\ell(x\,(1+\delta)) \le (1 + 2\,d\,\delta)\,\ell(x)$.*

PROOF. Let $\delta \le 1/(2\,d)$. The derivative $\ell'(y)$ of the latency function ℓ at a point $y \in [x,(1+\delta)x]$ is at most $d \cdot \ell(y)/y \le d \cdot \ell((1+\delta)x)/x$. This gives

$$\begin{aligned} \ell((1+\delta)\,x) &\le \ell(x) + \int_x^{(1+\delta)x} \ell'(u)\,du \\ &\le \ell(x) + \delta\,x \frac{d\,\ell((1+\delta)\,x)}{x} \ . \end{aligned}$$

Hence, $\ell((1+\delta)x) \le \frac{1}{1-\delta\,d}\ell(x)$, and for $\delta \le 1/(2\,d)$, $\ell((1+\delta)\,x) \le (1 + 2\,\delta\,d)\,\ell(x)$, as desired. □

We can generalize the definition of relative slope to latency functions that are not differentiable by requiring $\ell(x(1+\delta)) \le (1 + \mathcal{O}(d\,\delta))\,\ell(x)$ for $x \in [0,1]$ and $\delta = \mathcal{O}(1/d)$.

FACT 2. *For every flow f, $\bar{\ell}(f)/(d+1) \le \Phi(f) \le \bar{\ell}(f)$.*

PROOF. We compare $\bar{\ell}$ and Φ termwise. For the upper bound consider a resource e and note that the contribution to the average is $\ell_e(f_e) \cdot f_e$ whereas the contribution to the potential is $\int_0^{f_e} \ell_e(x)dx \le \ell_e(f_e) \cdot f_e$ by monotonicity of ℓ_e. For the lower bound, observe that the ratio $\int_0^x \ell_e(u)\,du\,/\,(x\,\ell_e(x))$ is minimized if ℓ has relative slope d over the entire range $[0,1]$, i.e. $\ell(\cdot)$ is a solution of the functional equation $\ell'(x) = d \cdot \ell(x)/x$ for all $x \in [0,1]$. Solutions of this equation are of the form $a\,x^d$. For these functions the contribution to the average is $a\,f_e^{d+1}$ whereas its contribution to the potential is $a\,\frac{1}{d+1} f_e^{d+1}$ which yields the desired ratio. □

1.1.4 A dynamic extension to Wardrop's traffic model.

We analyze rerouting policies on the basis of a dynamic, round-based variant of Wardrop's traffic model. Each round starts and ends with a feasible traffic allocation in the static model. The rerouting policies describe the behavior of the agents from a local point of view. They consist of simple probabilistic rules specifying whether an agent stays with his current routing strategy or switches to another apparently better strategy. Our rerouting policies are *Markovian*, that is, the behavior of an agent only depends on the traffic allocation at the beginning of the current round and not on observations made in previous rounds. Following Wardrop's traffic model, we assume that traffic is controlled by an infinite population of agents and we investigate the dynamical behavior of the system in the *fluid limit*, that is, the rerouting policies of the agents are translated into mean field equations that specify how the population shares allocated to the paths change from round to round.

1.2 Summary of our results

Let us first present our algorithms and results for symmetric games. Some of the results can be generalized towards asymmetric games as noted after their presentation.

Our algorithms use an adaptive sampling rule in which paths are sampled with a probability that increases with the fraction of agents using this path. In its simplest form our approach works as follows. Consider an agent currently using a path P. The agent chooses an alternative path Q with a probability that is proportional to the fraction of agents using this path. Then the agent compares the latencies on the paths P and Q. If ℓ_Q is smaller than ℓ_P the agent switches from P to Q with a probability proportional to $(\ell_P - \ell_Q)/(d \cdot (\ell_P + \alpha))$, where $d \ge 1$ is an upper bound on the relative slope of the latency functions and $\alpha \ge 0$. We can illustrate the power of this simple protocol in terms of a bicriteria result. We say that a traffic allocation is in a δ-ϵ-equilibrium if almost all agents, i.e. at least a $1-\epsilon$ fraction of the agents, have a latency close to the average latency, i.e., their latencies are within a factor of $(1 \pm \delta)$ of the average latency. Applying the simple policy described above, the total number of rounds in which the traffic allocation is not in a δ-ϵ-equilibrium is upper-bounded by

$$\mathcal{O}\left(\frac{d}{\epsilon\,\delta^2} \cdot \log\left(\frac{\Phi_{\text{init}} + \alpha}{\Phi^* + \alpha}\right)\right) \ ,$$

where Φ_{init} and Φ^* refer to the initial and the optimal potential, respectively. Remarkably, this bicriteria bound does not depend on any parameters describing the size or the structure of the network. It only depends on a single parameter of the latency functions, namely the maximum relative slope. We also show how this bicriteria result generalizes to asymmetric routing games. The major disadvantage of the bicriteria result is that δ-ϵ-equilibria are transient in that they can be left again once they are reached. For this reason we extend our analysis to approximations of the optimal potential.

The adaptive sampling rule is inspired by the so-called *replicator dynamics* from evolutionary game theory where players compare their payoffs with the payoffs of other players that are picked uniformly at random from the set of all players [24, 11]. This way, the probability to choose a strategy is proportional to the fraction of players using this strategy. The strength of this approach is that it amplifies good strategies in an exponential fashion. In our analysis this is reflected by a geometric improvement of the potential from round to round until the allocation reaches a state in which almost all agents have almost the same latency. The major weakness of this approach, however, is that it only replicates strategies used by other agents but does not explore new, unused strategies. Exploring unused strategies, however, is obviously necessary to ensure convergence to a Wardrop equilibrium.

We combine exploration based on static, uniform sampling with replication based on adaptive sampling into a distributed algorithm that we call *exploration-replication policy*. The agents perform uniform sampling with sufficiently small probability. This way, we can guarantee *monotonicity* with respect to the potential, i. e., the value of the potential function decreases from round to round. Unfortunately, this requires that the probability for uniform sampling is bounded from above in terms of the reciprocal of the first derivative of the latency functions. Fortunately, because of the amplification effects of the replication, this parameter appears only logarithmically in our result on the speed of convergence. We need a few more parameters to describe this result. Let m denote the number of edges, and L the maximum path length. Let $\ell_{\min}$ denote a lower bound on the latency on any edge and $\ell'_{\max}$ an upper bound on the first derivative of the latencies. (In fact, it is sufficient to upper-bound the first derivate of the latency functions in a small region around 0.) W. l. o. g., let $\ell'_{\max} > \ell_{\min} + \alpha$. In our analysis, we need to assume $\ell_{\min} + \alpha > 0$, that is, the latency functions are strictly positive or, alternatively, we use a positive shift of the potential function. We show that the exploration-replication policy if parameterized in the right way reaches a $(1 + \epsilon)$-approximation of the (α-shifted) optimal potential in a number of rounds of order at most

$$\operatorname{poly}\left(\frac{d \cdot L}{\epsilon}\right) \cdot \operatorname{polylog}\left(m \cdot \frac{\ell'_{\max}}{\ell_{\min} + \alpha}\right) \cdot \log\left(\frac{\Phi_{\text{init}} + \alpha}{\Phi^* + \alpha}\right) .$$

Let us remark that this bound is polynomial in $1/\epsilon$ and the description length of the instance if the latency functions are, e. g., defined in terms of positive polynomials of arbitrary degree in coefficient representation. This result about the speed of convergence with respect to the potential holds only for symmetric games.

We conclude our analysis with two lower bounds that substantiate the quality of our results. An important and unusual characteristic of our policies is that they make use of the parameter *maximum relative slope* in order to ensure monotonicity with respect to the potential. Our upper bounds depend in a polynomial fashion on this parameter. We present a lower bound showing that this dependence cannot be avoided by the class of policies under consideration. In particular, we show a lower bound of $\Omega(d/\sqrt{\epsilon})$ rounds to approximate the optimal potential within a factor of $1 + \epsilon$, for all Markovian policies that ensure monotonicity with respect to the potential over a basic class of latency functions with relative slope at most d. The network underlying this analysis consists only of two parallel links.

Furthermore, we study the necessity of adaptive sampling. As explained above, we use this technique to enable the population of agents to efficiently find those paths that can support a large amount of traffic. Observe that we achieve this goal without providing the agents with pre-knowledge about the latency functions on particular edges. Our lower bound documents that adaptive sampling is, in fact, necessary to quickly converge to a Wardrop equilibrium under these conditions, that is, we present an example where the restriction to static sampling results into a slowdown of the speed of convergence that is exponential in the size of the network.

1.3 Related Work

A policy similar to the one described above has been considered in [11] in a naive way, implicitly assuming that the agents act sequentially. In particular, all effects of simultaneous migrations that potentially cause oscillation effects and harm network performance are ignored. These problems are addressed in [12] applying a bulletin board model inspired by an analysis of load balancing with stale information by Mitzenmacher [17]. However, there the bounds on the time of convergence towards approximate equilibria are only pseudopolynomial. In particular, the bounds depend polynomially on the maximum slope of the latency functions and the maximum path length.

Well established heuristics for convex optimization use similar techniques to ensure convergence. For example, Bertsekas and Tsitsiklis [6] describe a distributed algorithm for non-linear multi-commodity flow in which the amount of flow that is moved in one step from one path to another depends in a linear way on the reciprocal of the second derivative of the latency functions. This algorithm can also be applied to compute Nash equilibria in the Wardrop model in a distributed way in which case the slowdown is again linear in the first derivative.

The convergence rate of adaptive rerouting policies has also been studied from the perspective of online learning, where one aims at minimizing the *regret* which is defined as the difference between a user's average latency over time and the latency of the best path in hindsight (see, e. g., [2, 3, 7]). The bounds obtained here also depend pseudopolynomially on network parameters.

Even-Dar and Mansour [9] study distributed and concurrent rerouting policies in a discrete model. Their study is restricted to networks with parallel links with speeds. Upper bounds are presented for the case of agents with identical weights. Their algorithms use static sampling rules that explicitly take into account the speeds of the individual links. Berenbrink *et al.* [5] present an efficient distributed protocol for balancing identical jobs on identical machines.

Beckmann *et al.* [4] show that Wardrop equilibria can be computed in polynomial time. Fabrikant *et al.* [10] consider the complexity of computing Nash equilibria in a discrete model. They show that computing Nash equilibria is PLS-complete in general whereas there exists a polynomial time algorithm for the case of symmetric network congestion games. In contrast to our work, these are centralized algorithms whereas we analyze how agents can compute or learn an equilibrium in a distributed fashion.

Finally, let us remark that our dynamic systems are similar to quadratic dynamic systems [1, 18] in that there is an infinite number of individuals that are *mated* at random to produce two individuals as *offspring*. In general, it is known that such systems with an infinite number of agents can solve PSPACE-complete problems in a polynomial number of rounds and can hence also compute Wardrop equilibria. However, this approach again only yields centralized algorithms since here, individuals do not have a natural interpretation as participants in a network routing game.

2. THE EXPLORATION-REPLICATION POLICY

We now formally introduce our rerouting policy for a class of latency functions with relative slope d. The policy takes two parameters α and β. In every round, an agent is activated with constant probability $\lambda = 1/32$. It then performs the following two steps. Consider an agent in commodity $i \in [k]$ currently utilizing path P.

1. *Sampling:* With probability $(1-\beta)$ perform step 1(a) and with probability β perform step 1(b).
   (a) *Proportional sampling:* Sample path $Q \in \mathcal{P}_i$ with probability f_Q/r_i.
   (b) *Uniform sampling:* Sample path $Q \in \mathcal{P}_i$ with probability $1/|\mathcal{P}_i|$.
2. *Migration:* If $\ell_Q < \ell_P$, migrate to path Q with probability $\frac{\ell_P - \ell_Q}{d\,(\ell_P + \alpha)}$.

Whereas the parameter $\alpha \geq 0$ can be chosen arbitrarily, the parameter β must be chosen subject to the constraint

$$\beta \leq \frac{\min_{P \in \mathcal{P}} \ell_P(0) + \alpha}{L \cdot \max_{e \in E} \max_{x \in [0,\beta]} \ell'_e(x)} \quad . \tag{1}$$

In the following, we always assume that this constraint is satisfied. We define our policy formally by specifying the amount of flow that is shifted between any pair of paths within one round.

Definition 3 (Exploration-Replication Policy). *For an instance Γ let $d \geq 1$ be an upper bound on the relative slope of the latency functions and let β be chosen as in Equation (1). For every commodity $i \in [k]$ and every path $P, Q \in \mathcal{P}_i$ with $\ell_Q \leq \ell_P$, the* (α, β)-exploration-replication policy *migrates a fraction of*

$$\mu_{PQ} = \lambda \cdot \frac{1}{d} \left((1-\beta) \cdot \frac{f_Q}{r_i} + \beta \cdot \frac{1}{|\mathcal{P}_i|} \right) \frac{\ell_P - \ell_Q}{\ell_P + \alpha}$$

with $\lambda = \frac{1}{32}$ agents from path P to path Q.

In our proofs we will simulate the (α, β)-exploration-replication policy by applying the $(0, \beta)$-exploration-replication policy to a modified instance with additional offsets α added to the path latencies.

Fact 3. *Let Γ be an instance of the congestion game and let $\Gamma^{+\alpha}$ be an instance that we obtain from Γ by inserting a new resource e_P for every $P \in \mathcal{P}$ with constant latency function $\ell_{e_P}(x) = \alpha$. Let Φ and $\Phi^{+\alpha}$ denote the respective potential functions.*

1. *The (α, β)-exploration-replication policy behaves on Γ precisely as the $(0, \beta)$-exploration-replication policy does on $\Gamma^{+\alpha}$.*
2. *If $\Phi^{+\alpha}(f) \leq (1+\epsilon)\,(\Phi^{+\alpha})^*$, then $\Phi(f) \leq (1+\epsilon)\,\Phi^* + \epsilon\,\alpha$.* □

2.1 Convergence

In this section we show that our rerouting policy decreases the potential in every round and that it therefore converges to a Wardrop equilibrium. Intuitively, the potential decreases since agents shift flow from high latency strategies to low latency strategies. Ideally the agents migrating from a strategy P to a strategy Q change the potential by $\mu_{PQ}(\ell_Q - \ell_P)$ where μ_{PQ} denotes the fraction of these agents. However, this is not true since the latencies ℓ_Q and ℓ_P are not constant.

The following lemma establishes that the potential gain in one round of our strategy is at least half of this ideal value. A similar result has been shown in [12]. However, here we improve this result for the (α, β)-exploration-replicaiton policy, which does not satisfy the requirements of [12], and for latency functions with bounded relative slope (instead of bounded first derivative).

For two flow vectors f and f' of consecutive rounds, the *virtual potential gain* is the potential gain that would occur if the latencies were fixed at the beginning of the round, i. e.

$$\mathcal{V}(f, f') = \sum_{e \in E} \ell_e(f) \cdot (f'_e - f_e) \enspace .$$

By our policy, this value is always negative. We show that the true potential gain $\Delta\Phi = \Phi(f') - \Phi(f)$ is at least half of $\mathcal{V}(f, f')$.

Lemma 4. *Consider an instance Γ and the (α, β)-exploration-replication policy changing the flow vector from f to f' in one step. Then we have*

$$\Delta\Phi = \Phi(f') - \Phi(f) \geq \frac{1}{2} \sum_{P,Q \in \mathcal{P}} \mu_{PQ}\,(\ell_Q - \ell_P) = \frac{\mathcal{V}(f, f')}{2}.$$

The proof extends the arguments given in [12]. We give bounds on the error terms by which the true potential gain differs from the virtual potential gain. The proof makes use of the fact that latency functions have relative slope d to bound the error terms caused by proportional sampling. The contribution to the error terms caused by uniform sampling can be bounded since β is chosen according to Equation (1). Due to space limitations we defer the proof to the full version.

Corollary 5. *The (α, β)-exploration-replication policy converges towards a Wardrop equilibrium.*

3. SYMMETRIC GAMES

In this section, we consider the case of symmetric games where the number of commodities is $k = 1$. We will first derive upper bounds for the time of convergence towards approximate equilibria and proceed by giving upper bounds for the number of rounds until the potential is close to the optimum.

3.1 Bicriteria Approximation

In this section we will use the following bicriterial definition of approximate equilibria.

Definition 4 (δ-ϵ-equilibrium). *For a flow vector f let $\mathcal{P}^+(\delta) = \{P \in \mathcal{P} \mid \ell_P(f) \geq (1+\delta)\,\bar{\ell}(f)\}$ denote the set of* δ-expensive strategies *and let $\mathcal{P}^-(\delta) = \{P \in \mathcal{P} \mid \ell_P(f) \leq (1-\delta)\,\bar{\ell}(f)\}$ denote the set of* δ-cheap strategies. *The population f is in a δ-ϵ-equilibrium iff at most ϵ agents utilize δ-expensive and δ-cheap strategies. We write $\mathcal{P}^+$ and $\mathcal{P}^-$ if δ is clear from the context.*

Note that our policy may leave such approximate equilibria again once they are reached. Hence, we bound the total number of rounds that are not at an approximate equilibrium.

Theorem 6. *Consider a symmetric congestion game Γ and an initial flow vector f_{init}. For the (α, β)-exploration-replication policy, the number of rounds in which the population vector is not at a δ-ϵ-equilibrium w. r. t. $\Gamma^{+\alpha}$ (as defined in Fact 3) is bounded from above by*

$$\mathcal{O}\left(\frac{d}{\epsilon\,\delta^2} \log\left(\frac{\Phi(f_{\text{init}}) + \alpha}{\Phi^* + \alpha}\right)\right) \enspace .$$

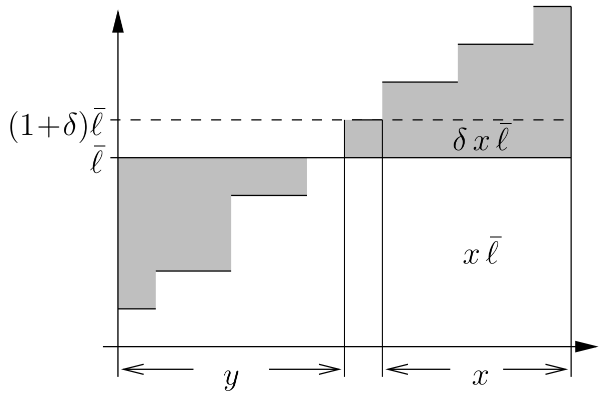

Figure 1: The figure depicts the distribution of the agents' latencies. The shaded areas are of the same size. Since the left area represents the conditional expectation of the difference between $\bar{\ell}$ and the latency of an agent with latency in the range $[0, \bar{\ell}]$, this expectation has a value of at least $\delta\, x\, \bar{\ell}$.

In particular, this bound holds for $\alpha = \beta = 0$ implying $\Gamma^{+\alpha} = \Gamma$.

Proof. It is sufficient to consider the case $\alpha = 0$ (and $\Phi^* > 0$). To see this, assume that the lemma is valid for $\alpha = 0$ and consider an instance $\hat{\Gamma}$ and some $\hat{\alpha} > 0$. By Fact 3, applying the lemma to $\Gamma = \hat{\Gamma}^{+\hat{\alpha}}$ and $\alpha = 0$ yields the assertion of the lemma applied to $\hat{\Gamma}$ and $\hat{\alpha}$.

We estimate the virtual potential gain of a round that starts which a population that is not at a δ-ϵ-equilibrium. Recall that the virtual potential gain is the difference between the potential of two consecutive rounds assuming that the latency functions were fixed at the beginning of the round. The virtual potential gain is actually negative. For simplicity, here we denote by $\mathcal{V}$ the absolute value of the virtual potential gain. By Lemma 4, the true potential gain is at least half of the virtual potential gain.

Consider a strategy $P \in \mathcal{P}$. As long as we are not at a δ-ϵ-equilibrium, at least one of the following cases holds.

Case 1. At least $\epsilon/2$ agents utilize δ-expensive paths.

Case 1a. At least a fraction of $x \geq 1/2$ agents utilizes δ-expensive paths. Let $y > 0$ denote the fraction of agents utilizing paths with latency at most $\bar{\ell}$. Consider a random agent that is sampled by proportional sampling and let Y denote the random variable that represents its current latency. Let $D = \mathrm{E}[\bar{\ell} - Y \mid Y \leq \bar{\ell}]$ denote the conditional expectation of the difference between Y and $\bar{\ell}$. We have

$$D = \bar{\ell} - \frac{1}{y}\left(\sum_{P:\ell_P \leq \bar{\ell}} f_P \ell_P\right) .$$

Substituting this into the definition of $\bar{\ell}$,

$$\begin{aligned}\bar{\ell} &= \sum_{P:\ell_P \leq \bar{\ell}} f_P \ell_P + \sum_{P:\ell_P > \bar{\ell}} f_P \ell_P \\ &= y \cdot (\bar{\ell} - D) + \sum_{P:\bar{\ell} < \ell_P \leq (1+\delta)\bar{\ell}} f_P \ell_P + \sum_{P:\ell_P > (1+\delta)\bar{\ell}} f_P \ell_P \ .\end{aligned}$$

By definition of x (for an illustration see Fig. 1),

$$\sum_{P:\ell_P > (1+\delta)\bar{\ell}} f_P \ell_P \geq \delta\, x\, \bar{\ell} + x\, \bar{\ell}$$

Substituting this for the last sum in the previous equation and solving for $\delta\, x\, \bar{\ell}$ yields

$$\begin{aligned}\delta\, x\, \bar{\ell} &\leq (1-x)\,\bar{\ell} - y \cdot (\bar{\ell} - D) - \sum_{P:\bar{\ell} < \ell_P \leq (1+\delta)\bar{\ell}} f_P \ell_P \\ &\leq (1-x)\,\bar{\ell} - y \cdot (\bar{\ell} - D) - (1 - x - y)\,\bar{\ell}\end{aligned}$$

implying $D \geq (\delta\, x/y)\,\bar{\ell}$. All agents with latency at least $(1+\delta)\,\bar{\ell}$ that sample a path with latency at most $\bar{\ell}$ migrate to the new path with probability at least δ/d. The probability to sample such a path is at least $(1-\beta)\, y$. Their expected latency gain which is equivalent to their infinitesimal contribution to the virtual potential gain is at least D. In total there are $x \geq 1/2$ such agents. Altogether, the expected virtual potential gain is

$$\mathcal{V} \geq \frac{\lambda\, y\, (1-\beta)\, \delta\, D}{d} \geq \frac{\lambda\, x\, \delta^2}{2\, d}\bar{\ell} \geq \frac{\lambda\, \delta^2}{4\, d}\bar{\ell} \ .$$

Case 1b. At least $\epsilon/2$ but at most $1/2$ agents utilize δ-expensive strategies. Then, the virtual potential gain of the agents leaving δ-expensive strategies is at least

$$\begin{aligned}\mathcal{V} &\geq \sum_{P \in \mathcal{P}^+} f_P \sum_{Q \notin \mathcal{P}^+} \mu_{PQ}(\ell_P - \ell_Q) \\ &\geq \lambda \sum_{P \in \mathcal{P}^+} f_P \sum_{Q \notin \mathcal{P}^+} \left((1-\beta) f_Q + \frac{\beta}{|\mathcal{P}|}\right) \frac{(\ell_P - \ell_Q)^2}{d\, \ell_P}.\end{aligned}$$

Omitting the term $\beta/|\mathcal{P}|$, substituting $\ell_P \geq \bar{\ell} + \delta\bar{\ell}$ and $(1-\beta) \geq 1/2$ and applying Jensen's inequality to the last sum yields

$$\begin{aligned}\mathcal{V} &\geq \frac{\lambda}{2\, d\, \bar{\ell}} \sum_{P \in \mathcal{P}^+} f_P \left(\sum_{Q \notin \mathcal{P}^+} f_Q(\bar{\ell} + \delta\bar{\ell} - \ell_Q)\right)^2 \\ &\geq \frac{\lambda}{2\, d\, \bar{\ell}} \sum_{P \in \mathcal{P}^+} f_P \left((\bar{\ell} + \delta\bar{\ell}) \sum_{Q \notin \mathcal{P}^+} f_Q - \sum_{Q \notin \mathcal{P}^+} f_Q \ell_Q\right)^2\end{aligned}$$

Note that $\sum_{Q \notin \mathcal{P}^+} f_Q \ell_Q \leq \bar{\ell} \sum_{Q \notin \mathcal{P}^+} f_Q$ since the sum omits the terms of the expensive strategies $Q \in \mathcal{P}^+$. Hence,

$$\begin{aligned}\mathcal{V} &\geq \frac{\lambda\, \delta^2\, \bar{\ell}^2}{2\, d\, \bar{\ell}} \sum_{P \in \mathcal{P}^+} f_P \left(\sum_{Q \notin \mathcal{P}^+} f_Q\right)^2 \\ &\geq \frac{\lambda\, (\epsilon/2)(1 - 1/2)^2\, \delta^2}{2\, d} \bar{\ell} \ .\end{aligned}$$

In the last inequality we used the above assumption that $\sum_{Q \notin \mathcal{P}^+} f_Q \leq 1/2$.

Case 2. At least $\epsilon/2$ agents utilize δ-cheap paths.

Case 2a. At least a fraction of $x \geq 1/2$ agents utilizes δ-cheap paths. This case is symmetric to case 1a. Let $y > 0$ denote the fraction of agents utilizing paths with latency at least $\bar{\ell}$. Consider a random agent that is activated in a round and let Y denote the random variable that represents its current latency. Let $D = \mathrm{E}[Y - \bar{\ell} \mid Y \geq \bar{\ell}]$ denote the conditional expectation of the difference by which Y exceeds $\bar{\ell}$. The proof is now identical to case 1*a* and again yields

$$\mathcal{V} \geq \frac{\lambda\, \delta^2}{4\, d}\bar{\ell} \ .$$

Case 2b. At least $\epsilon/2$ but at most $1/2$ agents utilize δ-cheap strategies. We can treat the virtual potential gain of agents moving towards δ-cheap strategies in a way similar to case 1b.

$$\begin{aligned}\mathcal{V} &\geq \sum_{Q\notin\mathcal{P}^-} f_Q \sum_{P\in\mathcal{P}^-} \mu_{QP}(\ell_Q-\ell_P) \\ &\geq \lambda \sum_{P\in\mathcal{P}^-}\left((1-\beta)f_P+\frac{\beta}{|\mathcal{P}|}\right)\sum_{Q\notin\mathcal{P}^-} f_Q\frac{(\ell_Q-\ell_P)^2}{d\,\ell_Q}.\end{aligned}$$

Now, we use that $\ell_P \leq \bar\ell - \delta\bar\ell$ and apply Jensen's inequality.

$$\begin{aligned}\mathcal{V} &\geq \frac{\lambda}{2\,d\,\bar\ell}\sum_{P\in\mathcal{P}^-} f_P\left(\sum_{Q\notin\mathcal{P}^-} f_Q(\ell_Q-\bar\ell+\delta\bar\ell)\right)^2 \\ &\geq \frac{\lambda}{2\,d\,\bar\ell}\sum_{P\in\mathcal{P}^-} f_P\left(\sum_{Q\notin\mathcal{P}^-} f_Q\ell_Q+(\delta\bar\ell-\bar\ell)\sum_{Q\notin\mathcal{P}_i^-} f_Q\right)^2 \\ &\geq \frac{\lambda\,\delta^2\,\bar\ell^2}{2\,d\,\bar\ell}\sum_{P\in\mathcal{P}^-} f_P\left(\sum_{Q\notin\mathcal{P}^-} f_Q\right)^2 \\ &\geq \frac{\lambda\,(\epsilon/2)(1-1/2)^2\,\delta^2}{2\,d}\bar\ell\ .\end{aligned}$$

In all four cases we have $\mathcal{V} \geq \frac{\lambda\,\epsilon\,\delta^2\bar\ell}{16\,d}$. Due to Lemma 4 and Fact 2, the true potential gain is

$$\Delta\Phi \leq -\mathcal{V} \leq -\frac{\lambda\,\epsilon\,\delta^2}{32\,d}\Phi.$$

Let $\Phi(t)$ denote the potential in the t-th round. Then,

$$\Phi(t) = \Phi(f_{\text{init}})\cdot\left(1-\frac{\lambda\,\epsilon\,\delta^2}{32\,d}\right)^t.$$

Since Φ is lower bounded by Φ^* we obtain the desired upper bound on the number of unbalanced phases. □

3.2 Approximation of the Potential

Since δ-ϵ-equilibria are transient we are interested in approximations of the potential. If $\Phi^*=0$, the potential cannot be approximated up to a relative factor of $(1+\epsilon)$ unless exactly optimized. We therefore allow an additional deviation by an additive term $\epsilon\,\alpha$, i. e., we want to reach a population with potential at most $\Phi(f) \leq (1+\epsilon)\,\Phi^*+\epsilon\,\alpha$. As discussed above, this is equivalent to adding a virtual constant latency α to every path. We will start with several lemmas that can be applied to symmetric games in general and proceed by analyzing the single-resource case and the general symmetric case separately.

First we show, that the value of the average $\bar\ell = \sum_{e\in E}\ell_e\,f_e$ does not change much within one round unless the potential does also decrease significantly.

Lemma 7. *Consider a symmetric routing game and a flow at δ-ϵ-equilibrium. If the (α,β)-exploration-replication policy changes the average latency $\bar\ell$ in one round by $\Delta > 10\,\lambda\cdot(2\,\epsilon+2\,\delta+\beta)\,\bar\ell$, it reduces the potential Φ by at least $\Delta/(10\,(d+1))$.*

Due to space limitations we defer the proof to the full version.

Our analysis of the time of convergence will proceed by constructing a flow vector in which the latencies of the cheapest path and the most expensive used path are close to the average latency. The following theorem shows that such a configuration in which all strategies deviate by no more than a fixed fraction of $\bar\ell$ from the average latency of their commodity, are also approximations of the optimal potential.

Definition 5 (δ-equilibrium). *A population vector f is at a δ-equilibrium if for every commodity $i\in[k]$ and for every $P\in\mathcal{P}_i$ it holds that $\ell_P(f)\geq\bar\ell_i-\delta\,\bar\ell$ and, in addition, if $f_P>0$, $\ell_P(f)\leq\bar\ell_i+\delta\,\bar\ell$.*

Lemma 8. *For every instance and feasible flow vector f at a δ-equilibrium, $\Phi(f)/\Phi^*\leq 1+\mathcal{O}(\delta\,d)$.*

Proof. Consider the δ-equilibrium flow f for the instance Γ. For every $i\in[k]$ let $\ell_{\max,i}=\max_{P\in\mathcal{P}_i,f_P>0}\ell_P(f)$. We extend every strategy $P\in\mathcal{P}_i$ by a new resource with constant latency $\delta_P=\ell_{\max,i}(f)-\ell_P(f)$. Since f is at a δ-equilibrium, we can be sure that $\delta_P\leq 2\,\delta\,\bar\ell$. Let Γ' denote the thus obtained network. Since in Γ' all latencies of used strategies are equal, f is a Wardrop equilibrium in Γ' with $\Phi_{\Gamma'}(f)\geq\Phi_\Gamma(f)$.

Now, consider the minimal potential in Γ, denoted by Φ^*_Γ. We have to show that $\Phi_{\Gamma'}(f)\leq\Phi^*_\Gamma+2\,\delta\,\bar\ell$ since this and Fact 2 imply that $\Phi_\Gamma(f)\leq\Phi^*_\Gamma+2\,\delta\,(d+1)\,\Phi_\Gamma(f)$ which is the assertion of our theorem. To see the claim, note that for every $\tilde f$ it holds that $\Phi_{\Gamma'}(\tilde f)\leq\Phi_\Gamma(\tilde f)+2\,\delta\,\bar\ell$. Also, the constraints for $\tilde f$ under which $\Phi_{\Gamma'}$ and Φ_Γ are to be minimized are identical, i. e., every $\tilde f$ feasible for Γ is also feasible for Γ' and vice versa. By optimality of f in Γ', we have $\Phi_{\Gamma'}(f)\leq\Phi_{\Gamma'}(\tilde f)\leq\Phi_\Gamma(\tilde f)+2\,\delta\,\bar\ell$ for any $\tilde f$ and specifically for an $\tilde f$ that satisfies $\Phi_\Gamma(\tilde f)=\Phi^*_\Gamma$. □

There exist instances showing that the parameters δ and d in Lemma 8 are actually required.

3.2.1 Symmetric Single-Resource Games

For the case in which every strategy utilizes only one resource, i. e., for all $P\in\mathcal{P}$, $|P|=1$, we can now show convergence towards potential approximations.

Lemma 9. *Consider a symmetric single-resource instance Γ and the $(0,\beta)$-exploration-replication-policy. For every $\epsilon>0$ define the following constants.*

$$\begin{aligned}\delta &= \max\left\{\frac{c\,\epsilon}{d}, c\,\beta\right\}, \\ \delta' &= \epsilon' = c'\frac{\delta^2}{d\log(|\mathcal{P}|/\beta)},\ \textit{and} \\ T &= c''\frac{d}{\delta}\log\left(\frac{\epsilon'\,d\,|\mathcal{P}|}{\delta\,\beta}\right),\end{aligned}$$

where c, c' and c'' are positive constants independent of ϵ, β, and d. Then, in every phase consisting of T rounds starting with a flow vector f^0, there exists a round $t\in\{1,\ldots,T\}$ with flow vector $f(t)$ that satisfies one of the following three properties.

1. *The population $f(t)$ is not at a δ'-ϵ'-equilibrium or*
2. *the potential decreases by at least $\frac{2\,\delta'}{d+1}\bar\ell(f^0)$, i. e., $\Phi(f^0)-\Phi(f)\geq\frac{2\,\delta'}{d+1}\bar\ell(f^0)$ or*

(3) *the population is a $(1+\epsilon)$-approximation of the optimal potential, i. e., $\Phi(f(t)) \le (1+\epsilon)\Phi^*$.*

Let us give an intuition of the proof. Unless the phase contains a round t satisfying property (2), we can use Lemma 7 to fix the value of the average within a small interval around its initial value throughout the phase. Then we can partition the paths into categories according to their latency. The first category contains all paths with latency at most $(1-\delta)\,\bar\ell(f^0)$. We show that within one phase, paths do not change to the first category from any other category. This is, because the load of paths in the other categories is either increasing or their latency is too large to drop to at most $(1-\delta)\,\bar\ell(f^0)$ within one round. Therefore, the total load on δ-cheap paths must grow exponentially implying that we quickly reach a configuration $f(t)$ which is either not a δ'-ϵ'-equilibrium any more (i. e. $f(t)$ satisfies property (1)) or $f(t)$ is a δ-equilibrium, which, together with Lemma 8, implies that property (3) is satisfied. Due to space limitations we leave the proof for the full version.

Theorem 10. *Consider a symmetric single-resource instance Γ and an initial flow vector f_{init}. If $\beta \le \epsilon/d$, the (α,β)-exploration-replication-policy generates a configuration with potential $\Phi \le (1+\epsilon)\Phi^* + \epsilon\,\alpha$ in at most*

$$\mathcal{O}\left(\frac{d^{12}}{\epsilon^7}\log^4\left(\frac{|E|}{\beta}\right)\log\left(\frac{\Phi(f_{\text{init}})+\alpha}{\Phi^*+\alpha}\right)\right)$$

rounds.

Proof. Again, by Fact 3, it is sufficient to prove the theorem for the case $\alpha = 0$ (see the first paragraph of the proof of Theorem 6). Consider a phase of length T as defined in Lemma 9. By Lemma 9, the phase terminates with one of the following events after at most T rounds:

(1) The population is no longer at a δ'-ϵ'-equilibrium. By Theorem 6, this can happen at most

$$T_1 = \mathcal{O}\left(\frac{d}{\epsilon'^2\delta'^2}\log\left(\frac{\Phi(f^0)}{\Phi^*}\right)\right)$$

times.

(2) The potential decreases by at least $2\delta'/(d+1)\,\bar\ell(f^0)$. This decreases the potential by a factor of at least $(1-2\,\delta'/(d+1))$. Therefore, this can happen at most

$$T_2 = \mathcal{O}\left(\frac{d}{\delta'}\log\left(\frac{\Phi(f^0)}{\Phi^*}\right)\right)$$

times.

(3) The policy has reached a $(1+\epsilon)$-approximation of the potential.

Hence, after at most $T \cdot \max\{T_1, T_2\}$ rounds, event (3) must occur. □

3.2.2 General Symmetric Games

In the general symmetric case, a lemma similar to Lemma 9 holds with modified definitions of δ, ϵ, δ', and T. We leave the proof, which is more involved than in the symmetric case, for the full version. The modified values imply the following theorem:

Theorem 11. *Consider a symmetric instance Γ and an initial flow vector f_{init}. If $\beta \le \epsilon^2/(L^3\,d^2)$, then the (α,β)-exploration-replication-policy generates a configuration with potential $\Phi \le (1+\epsilon)\Phi^* + \epsilon\,\alpha$ in at most*

$$\text{poly}\left(d, \frac{1}{\epsilon}, L\right)\ln^4\left(\frac{|E|}{\beta}\right)\ln\left(\frac{\Phi(f_{\text{init}})+\alpha}{\Phi^*+\alpha}\right)$$

rounds.

Substituting Equation (1) for β, this yields the bound as presented in Section 1.2. The proof is identical with the proof of Theorem 10 with the modified definitions of δ, ϵ', δ', and T. In addition, we use that $|\mathcal{P}| = \mathcal{O}\left(|E|^L\right)$.

4. ASYMMETRIC GAMES

For the asymmetric, or multi-commodity, case, we generalize the bicriteria definition of approximate equilibria in the following manner.

Definition 6 (δ-ϵ-equilibrium). *For a flow vector f, for every commodity $i \in [k]$, let $\mathcal{P}_i^+(\delta) = \{P \in \mathcal{P}_i \mid \ell_P(f) \ge \bar\ell_i(f) + \delta\bar\ell\}$ denote the set of* δ-expensive strategies *and let $\mathcal{P}_i^-(\delta) = \{P \in \mathcal{P}_i \mid \ell_P(f) \le \bar\ell_i(f) - \delta\bar\ell\}$ denote the set of* δ-cheap strategies*. The population f is called a δ-ϵ-equilibrium iff at most ϵ agents utilize δ-expensive and δ-cheap strategies.*

The analysis of the time of convergence is slightly more involved than in the symmetric case.

Theorem 12. *Consider an asymmetric congestion game Γ and an initial flow vector f_{init}. For the (α,β)-exploration-replication policy, the number of rounds in which the population vector is not at a δ-ϵ-equilibrium w. r. t. $\Gamma^{+\alpha}$ (as defined in Fact 3) is bounded from above by*

$$\mathcal{O}\left(\frac{d}{\epsilon^2\,\delta^2}\log\left(\frac{\Phi(f_{\text{init}})+\alpha}{\Phi^*+\alpha}\right)\right) .$$

In particular, this bound holds for $\alpha = \beta = 0$ (and hence $\Gamma^{+\alpha} = \Gamma$).

Proof (Sketch). As in the proof of Theorem 6, it is sufficient to consider the case $\alpha = 0$. Again, we estimate the virtual potential gain $\mathcal{V}$. Consider a path $P \in \mathcal{P}$. There are two cases.

1. $\ell_P > \bar\ell \cdot 2/\epsilon$. By Markov's inequality, at most a fraction of $\epsilon/2$ of the agents utilizes such paths. We ignore the potential gain of these agents.

2. $\ell_P \le \bar\ell \cdot 2/\epsilon$. The remaining $1-\epsilon/2$ agents utilize these paths. As long as we are not at a δ-ϵ-equilibrium, there must be at least ϵ agents utilizing δ-expensive or δ-cheap paths. Since at most $\epsilon/2$ of them are in case 1, at least $\epsilon/2$ of them are in this case.

Throughout the proof we only consider agents of the second case. For these agents, the proof is an extension of the proof of Theorem 6. Consider commodity i.
Case 1. At least an $\epsilon/4$-fraction of the agents utilizes δ-expensive paths. First, fix one commodity i with total rate r_i.
Case 1a. In commodity i, at least $r_i/2$ agents utilize δ-expensive paths. As in case 1a of the proof of Theorem 6,

this yields an expected contribution to the virtual potential gain of

$$\mathcal{V}_i \geq \frac{r_i\,\lambda\,\delta^2\,\epsilon}{4\,d}\bar{\ell}\ .$$

Case 1b. The number of agents in commodity i utilizing δ-expensive paths in commodity i is $x_i \leq r_i/2$. With an argument similar as in the proof of Theorem 6 we see that

$$\mathcal{V}_i \geq \frac{\lambda\,\epsilon\,\delta^2\,\bar{\ell}}{3\,d}\cdot\frac{x_i(r_i - x_i)^2}{r_i^2} \geq x_i\frac{\lambda\,\epsilon\,\delta^2\,\bar{\ell}}{12\,d}$$

Now, summing up over all commodities, the virtual potential gain is

$$\begin{aligned}\mathcal{V} &\geq \sum_{i\in[k]}\mathcal{V}_i\\ &\geq \sum_{i\in[k]}\min\left\{r_i\frac{\lambda\,\delta^2\,\epsilon\,\bar{\ell}}{4\,d}, x_i\frac{\lambda\,\epsilon\,\delta^2\,\bar{\ell}}{12\,d}\right\}\\ &\geq \frac{\lambda\,\epsilon\,\delta^2\,\bar{\ell}}{12\,d}\sum_{i\in[k]}x_i\\ &\geq \frac{\lambda\,\epsilon^2\,\delta^2\,\bar{\ell}}{48\,d}\ .\end{aligned}$$

Here, the minimum is over the two cases 1a and 1b.
Case 2. Again, cases 2a and 2b are symmetric. □

5. LOWER BOUNDS

5.1 Relative Slope is Necessary

In this section we show that the relative slope is the relevant parameter in our analysis. We provide a lower bound that shows that for a class $\mathcal{L}$ of latency functions the relative slope d of $\mathcal{L}$ is a lower bound for the time of convergence towards approximate equilibria if the rerouting policy is monotone in the following sense. A policy is *monotone* for a class of latency functions $\mathcal{L}$ if for every instance with latency functions taken from $\mathcal{L}$ and for every feasible flow vector f, the policy does not increase the value of the potential in one round. It is *Markovian* if the policy maintains no state and the migration rates only depend on the current flow vector.

THEOREM 13. *For every d, there exists a class $\mathcal{L}$ of latency functions with relative slope d together with an initial flow vector f, such that any Markovian rerouting policy monotone for $\mathcal{L}$ requires $\Omega(d/\sqrt{\epsilon})$ rounds in order to obtain a $(1+\epsilon)$ approximation to the optimum potential.*

PROOF. We choose a latency function $\ell : [0,1] \mapsto \mathbb{R}_{\geq 0}$ and numbers $x, y \in [0,1]$, $\sqrt{\epsilon} \leq x < y$, such that the relative slope of ℓ in the interval (x, y) is at least d. We define the class $\mathcal{L}_\ell := \{\ell + a \mid a \in [0, \ell(1)]\} \cup \{a \mid a \in [0, \ell(1)]\}$ to contain all constant latency functions in the interval $[0, \ell(1)]$ and all latency functions that can be obtained by adding a constant from this interval to ℓ.

Now, consider an instance with two strategies each consisting of one resource with a constant latency function. Define the latency function of the first resource as $\ell_1(f_1) = L_1 := \ell(y)/(1+\sqrt{\epsilon})$, and the latency function of the second resource as $\ell_2(f_2) = L_2 := \ell(y)$. Clearly, the global optimum puts all flow on resource one which results in a potential $\Phi^* = L_1$.

Assume a starting configuration where the flows f_1 and f_2 for the two strategies are $f_1 = 1 - y$ and $f_2 = y$. The potential of a flow vector f' is $\Phi(f') = L_1 \cdot f'_1 + L_2 \cdot f'_2$. In order to obtain a $(1+\epsilon)$-approximation to Φ^* the flow f_2 over resource 2 must drop from its initial value y to a value less or equal to $\sqrt{\epsilon}$.

Consider a configuration f' with $f'_2 \in [x, y]$. The rerouting process must guarantee that it does not increase the potential for all possible latency functions in $\mathcal{L}_\ell$. Since the process is Markovian, it has no knowledge about ℓ_1 and ℓ_2 but can only observe the values $\ell_1(f'_1)$ and $\ell_2(f'_2)$.

In particular, the process must assume that the latency function of resource 2 is $\tilde{\ell}(f_2) = \ell(f_2) + c$ where $c \geq 0$ is chosen such that $\tilde{\ell}(f'_2) = L_2$. Since $c \leq L_2 - L_1$ and since the relative slope of ℓ at f'_2 is at least d, we know that the relative slope of $\tilde{\ell}$ at f'_2 is at least $d/2$ in the interval $[x, y]$. Hence, the rerouting policy may move at most

$$\Delta = \mathcal{O}\left(\frac{\ell_2 - \ell_1}{\ell_1}\frac{1}{d}f_1\right) = \mathcal{O}\left(\frac{\sqrt{\epsilon}}{d}f_1\right)$$

agents from strategy 2 to strategy 1. Since we started with a population of y utilizing strategy 2, it takes at least $(y-x)/\Delta$ rounds until the population utilizing this strategy decreases to below x. □

The proof shows that the theorem actually holds for any $\mathcal{L}$ containing at least one function ℓ that has relative slope d on an interval of constant width plus all functions $\ell + c$ for constants $c > 0$ and the constant functions.

5.2 Sampling with Static Probabilities is Slow

The following theorem shows that every rerouting policy that samples with static probabilities that are independent of the latency functions needs at least $\Omega(|P|)$ rounds to approach a Wardrop equilibrium. We will formalize the notion of static sampling probabilities in the following way. If strategy P is sampled with static probability σ_P, at most $(1-f_P)\,\sigma_P$ agents may migrate towards P in one round since $(1 - f_P)$ agents utilize other paths. We say that a rerouting policy has *static sampling probabilities* (for a set of strategies $\mathcal{P}$) denoted by $(\sigma_P)_{P\in|\mathcal{P}|}$ with $\sum_{P\in\mathcal{P}_i}\sigma_P = 1$ for all $i \in [k]$, if for every feasible flow vector f, every commodity $i \in [k]$, and every strategy $P \in \mathcal{P}_i$ it holds that the total volume of flow that the policy shifts to strategy P in one round is bounded from above by $\sigma_P(r_i - f_P)$.

THEOREM 14. *For every m, there exist a set of resources E with $|E| = m$ and strategy set $\mathcal{P}$ with $|\mathcal{P}| = 2^{m/4}$ such that for every rerouting policy with static sampling probabilities for $\mathcal{P}$ there exist a set of latency functions $(\ell_e)_{e\in E}$ and an initial population such that the rerouting policy needs at least $\Omega(|\mathcal{P}|\log(1/\epsilon))$ rounds to reach a $(1+\epsilon)$-approximation of the optimal potential for the symmetric instance $\Gamma = (E, \mathcal{P}, (\ell_e))$.*

PROOF. Consider a directed network $G = (V, E)$ with node set $V = \{s, t, v_1, \ldots, v_n, w_1, \ldots, w_n\}$. The edge set consists of the edges (s, v_1), (s, w_1), (w_1, t), (w_n, t) and, for $1 \leq i < n$, (v_i, v_{i+1}), (v_i, w_{i+1}), (w_i, v_{i+1}), and (w_i, w_{i+1}). This network has $m = 4\,n$ resources and $|\mathcal{P}| = 2^n = 2^{m/4}$ paths. Edges in this network correspond to resources and paths correspond to strategies.

Let $(\sigma_P)_{P\in\mathcal{P}}$ be the vector of sampling probabilities chosen by the rerouting policy and let $\tilde{P}$ be the strategy that minimizes σ_P. Then, $\sigma_P \leq 1/|\mathcal{P}|$.

For $e \in \{(s, v_1), (s, w_1), (w_1, t), (w_n, t)\}$, let $\ell_e = 0$. For every resource $e \in P$ with $s, t \notin e$, let $\ell_e = 1$. For all other resources e', let $\ell_{e'} = n$. Hence, strategy $\tilde{P}$ is the unique optimal strategy with constant latency $n-1$ and $\Phi^* = n-1$. All other strategies have latency at least $2\,n-2$. In order to reach a $(1+\epsilon)$ approximation of Φ^*, the population on $\tilde{P}$ must be at least $1-\epsilon$ and hence the population on the remaining strategies must be at most ϵ. By our choice of $\tilde{P}$, we have

$$f_{\tilde{P}}(t+1) \leq f_{\tilde{P}}(t) + (1 - f_{\tilde{P}}(t))\sigma_{\tilde{P}} \leq f_{\tilde{P}}(t) + \frac{1 - f_{\tilde{P}}(t)}{|\mathcal{P}|}$$

or, writing $\bar{f}_{\tilde{P}} = 1 - f_{\tilde{P}}$,

$$\bar{f}_{\tilde{P}}(t+1) \geq \bar{f}_{\tilde{P}}(t) - \frac{\bar{f}_{\tilde{P}}(t)}{|\mathcal{P}|} = \bar{f}_{\tilde{P}}(t)\left(1 - \frac{1}{|\mathcal{P}|}\right) \ .$$

Therefore, $\bar{f}_{\tilde{P}}(t) \geq f_{\tilde{P}}(0) \cdot (1 - 1/|\mathcal{P}|)^t$ implying that it takes $\Omega(|\mathcal{P}|\log(1/\epsilon))$ rounds until $\bar{f}_{\tilde{P}}(t) < \epsilon$ and hence $f_{\tilde{P}} \geq 1-\epsilon$ if we choose $f_{\tilde{P}}(0) > 0$ constant. $\square$

6. OPEN PROBLEMS

Let us say that an agent is *almost satisfied* if her latency is within a factor of $1+\delta$ of the average latency in the same commodity. In the symmetric case, our bicriteria result for the replication protocol has an intuitive interpretation: Replication approximately satisfies almost all agents very quickly. Unfortunately, our definition of δ-ϵ-equilibria from Section 4 does not allow to extend this intuition to asymmetric games. Can one obtain a similar result for the multi-commodity case?

Is it possible to achieve a polynomial upper bound on the time of convergence of the replication-exploration protocol or a protocol of similar flavor w. r. t. the potential in the multi-commodity case?

In a discrete setting, our replication-exploration protocol would require an exponential number of agents as the strategy space can be exponentially large in the number of resources. Is it possible to explore the large strategy space in a distributed fashion with a smaller (polynomial) number of agents?

7. REFERENCES

[1] S. Arora, Y. Rabani, and U. Vazirani. Simulating quadratic dynamical systems is PSPACE-complete. In *Proceedings of the 26th ACM Symposium on Theory of Computing (STOC)*, pages 459–467, 1994.

[2] B. Awerbuch and R. D. Kleinberg. Adaptive routing with end-to-end feedback: Distributed learning and geometric approaches. In *Proceedings of the 36th ACM Symposium on Theory of Computing (STOC)*, pages 45–53, 2004.

[3] B. Awerbuch and R. D. Kleinberg. Competitive collaborative learning. In *Proceedings of the 18th Annual Conference on Learning Theory (COLT)*, pages 233–248, 2005.

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

[5] P. Berenbrink, T. Friedetzky, L. A. Goldberg, P. Goldberg, Z. Hu, and R. Martin. Distributed selfish load balancing. In *Proceedings of the 17th ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 354–363, 2006.

[6] D. P. Bertsekas and J. N. Tistiklis. *Parallel and Distributed Computing: Numerical Methods*. Athena Scientific, 1989.

[7] A. Blum, E. Even-Dar, and K. Ligett. Routing without regret, 2005. Manuscript.

[8] R. Cole, Y. Dodis, and T. Roughgarden. Pricing network edges for heterogeneous selfish users. In *Proceedings of the 35th ACM Symposium on Theory of Computing (STOC)*, pages 521–530, 2003.

[9] E. Even-Dar and Y. Mansour. Fast convergence of selfish rerouting. In *Proceedings of the 16th ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 772–781, 2005.

[10] A. Fabrikant, C. Papadimitriou, and K. Talwar. The complexity of pure Nash equilibria. In *Proceedings of the 36th ACM Symposium on Theory of Computing (STOC)*, pages 604–612, 2004.

[11] S. Fischer and B. Vöcking. On the evolution of selfish routing. In *Proceedings of the 12th European Symposium on Algorithms (ESA)*, pages 323–334, 2004.

[12] S. Fischer and B. Vöcking. Adaptive routing with stale information. In *Proceedings of the 24th ACM Symposium on Principles of Distributed Computing (PODC)*, pages 276–283, 2005.

[13] L. Fleischer. Linear tolls suffice: New bounds and algorithms for tolls in single source networks. In *Proceedings of the 31st International Colloquium on Automata, Languages and Programming (ICALP)*, pages 544–554, 2004.

[14] A. C. Kaporis, E. I. Politopoulou, and P. G. Spirakis. The price of optimum in Stackelberg games. Technical report, Electronic Colloquium on Computational Complexity (ECCC), 2005.

[15] A. Khanna and J. A. Zinky. The revised ARPANET routing metric. In *Proceedings of the ACM Symposium on Communications Architectures & Protocols (SIGCOMM)*, pages 45–56, 1998.

[16] J. F. Kurose and K. W. Ross. *Computer Networking: A Top Down Approach Featuring the Internet, 3rd ed.* Addison–Wesley Longman Publishing Co., Boston, MA, USA, 2004.

[17] M. Mitzenmacher. How useful is old information? In *Proceedings of the 16th ACM Symposium on Principles of Distributed Computing (PODC)*, pages 83–91, 1997.

[18] Y. Rabinovich, A. Sinclair, and A. Wigderson. Quadratic dynamical systems. In *Proceedings of the 33rd IEEE Symposium on Foundations of Computer Science (FOCS)*, pages 304–313, 1992.

[19] J. Rexford. *Handbook of Optimization in Telecommunications*, chapter Route optimization in IP networks. Kluwer Academic Publishers, 2005.

[20] T. Roughgarden. Stackelberg scheduling strategies. In *Proceedings of the 33rd ACM Symposium on Theory of Computing (STOC)*, pages 104–113, 2001.

[21] T. Roughgarden. How unfair is optimal routing? In *Proceedings of the 13th ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 203–204, 2002.

[22] T. Roughgarden and É. Tardos. How bad is selfish routing? *Journal of the ACM*, 49(2):236–259, 2002.

[23] 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.
[24] J. W. Weibull. *Evolutionary Game Theory*. MIT press, 1995.