Selfish Routing Revisited: Degeneracy, Evolution and Stochastic Fluctuations

Panayotis Mertikopoulos*
Department of Economics
École Polytechnique
91128 Palaiseau Cedex, France

Aris L. Moustakas†
Department of Physics
National & Kapodistrian University of Athens
15784 Athens, Greece

ABSTRACT

We study the traffic routing problem in networks whose users try to minimize their latencies by employing a distributed learning rule inspired by the replicator dynamics of evolutionary game theory. The stable states of these dynamics coincide with the network's (Wardrop) equilibrium points and we find that they form a convex polytope whose dimension is determined by the network's *degeneracy index* (an important concept which measures the overlap of the users' paths). Still, despite this abundance of stable states, we find that (almost) every solution trajectory converges to an equilibrium point at an exponential rate.

On the other hand, a major challenge occurs when network latencies fluctuate unpredictably due to random exogenous factors. In that case, we show that the time-average of the traffic flows of sufficiently patient users is still concentrated in a neighborhood of evolutionarily stable equilibria and we estimate analytically the corresponding stationary distribution and convergence times.

1. INTRODUCTION

The problem of managing the flow of traffic in large-scale networks (the Internet being a prime example) is as simple to state as it is challenging to resolve: given the traffic rates generated by the network's users, one is asked to identify and realize the most "satisfactory" distribution of traffic among the network's routes. However, since the time needed to traverse a link in the network increases as the link becomes more congested, the users' concurrent optimization efforts invariably lead to competitive interactions whose complexity precludes even the most rudimentary attempts at coordination. Thus, a traffic distribution is considered "satisfactory" by a user when there is no unilateral move that could further decrease the delays (or *latencies*) that he experiences.

In this way, a most natural way to approach this problem is by means of game theory, an idea which has given renewed impetus to the subject – see [1] for a panoramic survey. In particular, in the seminal papers [2–5], it was identified that the relevant solution concept is described by *Wardrop's principle* [6]: given the level of congestion caused by other users, every user seeks to employ the minimum-latency path available to him.

On the other hand, this raises the issue of whether traffic flows which are at Wardrop equilibrium are also "socially optimal" in the sense that they minimize the aggregate latency in the network. The answer to this question is a negative one, but Roughgarden and Tardos showed that even in paradoxical, Braess-type networks, the *price of anarchy* (an equilibrium efficiency ratio that was introduced in [3]) is actually bounded by the celebrated factor 4/3 [4, 5]. More importantly, it is also well known that the network's optimal state can be interpreted as a Wardrop equilibrium for a slightly modified network [7, 8], so, in effect, equilibrium and optimality can be seen as different sides of the same coin.

Still, the size of large-scale networks makes computing these equilibria a task of formidable difficulty, clearly beyond the users' individual deductive capabilities. Moreover, a user has no incentive to actually play out his component of an equilibrial traffic allocation unless he is convinced that his opponents will also employ theirs (an argument which gains additional momentum if there are multiple equilibria). Finally, in real-world applications, the information that users have at their disposal is not only local in nature, but might also be inaccurate as well (a consequence of the stochastic interference of nature with the network). Hence, our goal in this paper is to take a less centralized approach and ask: *is there a distributed learning scheme which leads users to an equilibrium and which remains robust in the presence of uncertainties and noise?*

Even though the static properties of Wardrop equilibria have been studied quite extensively, this question has been left relatively unexplored. In fact, it was only recently shown in [9–11] that a good candidate for such a learning scheme would be the *replicator dynamics* of evolutionary game theory, a dynamical system which was first introduced in [12] to model the evolution of (nonatomic) populations that interact with one another by means of random matchings in a Nash game. In our congestion setting, these populations correspond to the users' traffic flows, so the convex optimization formulation of [7] allows us to recast our problem in terms of a (nonatomic) *potential game* [9]. As a result, the game's Wardrop equilibria emerge as the game's only Lyapunov stable states in deterministic environments [9], and when there is a unique equilibrium, it was shown that interior trajectories converge to it [10, 11].

Rather surprisingly, the structure of the Wardrop set itself seems to have been overlooked in the above considerations when it does not consist of a single equilibrium. Specifically, a subtle mistake that is present in [9] (and in the multi-commodity considerations of [10] as well) is that if the network's delay functions are strictly increasing, then there exists a unique Wardrop equilibrium – see e.g. [9, Corollary 5.6]. As a matter of fact, this is only true in

*panayotis.mertikopoulos@polytechnique.edu
†arislm@phys.uoa.gr

VALUETOOLS '11 Paris, France

irreducible networks, i.e. networks whose paths are "independent" of one another (in a sense made precise in Definition 1). In general, the Wardrop set of a network is a convex polytope whose dimension is determined by the network's *degeneracy index*, a notion which precisely quantifies this "linear dependence".

Nonetheless, despite this added structure, we show that the expectations of [10] are vindicated even when then are multiple Wardrop equilibria: (almost) every replicator orbit converges to a Wardrop equilibrium (Theorem 7) and not merely to the Wardrop set. More importantly, we find that the rate of this convergence is actually *exponential*, in the sense that users hit an ε-neighborhood of an equilibrium in time of order $\mathcal{O}(\log 1/\varepsilon)$ – a significant improvement over the $\mathcal{O}(\varepsilon^{-3})$ polynomial bound that was predicted by [10].

That said, an assumption which is central (albeit implicit) in the replicator dynamics is that users have perfectly accurate information at their disposal. Unfortunately however, this assumption is not very realistic in networks which exhibit wild delay fluctuations as the result of interference by random exogenous factors (commonly referred to as "nature"). In population biology, these disturbances are modelled by introducing "aggregate shocks" to the replicator dynamics [13] and, as one would expect, these shocks complicate the situation by quite a bit. For instance, if the variance of the shocks is mild enough, it was shown in [14] that dominated strategies become extinct in the long run, and, under similar hypotheses, even equilibrial play arises over time [15, 16]. On the other hand, if one interprets the replicator dynamics as the derivative of an *exponential learning* process and perturbs them accordingly, it was only recently shown that similar rationality properties continue to hold, *regardless* of the noise level [17, 18].

Nevertheless, all of these approaches have been focused on Nash-type finite games with multilinear payoffs. This linear structure simplifies things considerably but, unfortunately, congestion models rarely adhere to it; even more to the point, the way that stochastic fluctuations propagate to the users' choices in a network leads to a new stochastic incarnation of the replicator dynamics where the noise processes are no longer independent across users (different paths might share a common subset of links over which disturbances are strongly correlated). On that account, the effect of stochastic fluctuations in congestion models is radically different than in previous work on the stochastic replicator dynamics and, as such, our aim in this paper will be to chart out and identify the rationality properties that remain true in the presence of noise.

1.1 Outline

Our network model is presented in detail in Section 2 where we develop the "static" aspect of our work: specifically, we introduce the *degeneracy index* of a network in Section 2.1, and we examine its ties to Wardrop equilibria in Section 2.2. We then derive the rationality properties of the deterministic replicator dynamics in Section 3 and show that replicator solution orbits converge to Wardrop equilibrium at an exponential rate which we are able to estimate (Theorems 7 and 8).

Section 4 is devoted to the stochastic considerations which constitute the core of our paper. We find that if the users are "patient enough" (in the sense that their learning rates are sufficiently slow), then the long-term time average of the stochastic replicator dynamics in irreducible networks concentrates mass in the vicinity of an evolutionarily stable Wardrop equilibrium (Theorem 10). This not only provides a significant extension of the results of [15, 16] to nonlinear multi-population environments, but it also highlights the interplay between learning and uncertainty: no matter how loud the noise becomes, players who take their time will always be able to weed out the effect of stochastic fluctuations.

Notational Conventions

If $\mathcal{S} = \{s_\alpha\}_{\alpha=0}^n$ is a finite set, we will denote the vector space spanned by $\mathcal{S}$ over $\mathbb{R}$ by $\mathbb{R}^{\mathcal{S}} \equiv \mathrm{Hom}(\mathcal{S}, \mathbb{R})$. The canonical basis $\{e_\alpha\}_{\alpha=0}^n$ of $\mathbb{R}^{\mathcal{S}}$ then consists of the indicator functions $e_\alpha : \mathcal{S} \to \mathbb{R}$ which take the value $e_\alpha(s_\alpha) = 1$ on s_α and vanish otherwise; hence, under the identification $s_\alpha \mapsto e_\alpha$, we will use α to refer interchangeably to either s_α or e_α. In the same vein, we will also identify the set $\Delta(\mathcal{S})$ of probability measures on $\mathcal{S}$ with the standard n-dimensional simplex of $\mathbb{R}^{\mathcal{S}}$: $\Delta(\mathcal{S}) = \{x \in \mathbb{R}^{\mathcal{S}} : \sum_\alpha x_\alpha = 1 \text{ and } x_\alpha \geq 0\}$.

Concerning players and their strategies, we will employ Latin indices $(i, j, \dots)$ for players while reserving Greek ones $(\alpha, \beta \dots)$ for their (pure) strategies; also, to differentiate between strategies, we will use $\alpha, \beta, \dots$ for indices that start at 0 and $\mu, \nu, \dots$ for those that start at 1. Moreover, if the players' action sets $\mathcal{A}_i$ are disjoint (as is typically the case), we will identify their union $\bigcup_i \mathcal{A}_i$ with their *disjoint union* $\mathcal{A} \equiv \coprod_i \mathcal{A}_i = \bigcup_i \{(\alpha, i) : \alpha \in \mathcal{A}_i\}$. Thus, if $\{e_{i\alpha}\}$ is the natural basis of $\mathbb{R}^{\mathcal{A}_i}$ and $\{e_\alpha\}$ is the corresponding basis of $\mathbb{R}^{\mathcal{A}} \cong \prod_i \mathbb{R}^{\mathcal{A}_i}$, we will occasionally drop the index i altogether and write $x = \sum_\alpha x_\alpha e_\alpha$ instead of $x = \sum_{i,\alpha} x_{i\alpha} e_{i\alpha}$. Similarly, when it is clear from the context that we are summing over the strategy set $\mathcal{A}_i$ of player i, we will use the shorthand $\sum_\alpha^i \equiv \sum_{\alpha \in \mathcal{A}_i}$.

2. PRELIMINARIES

2.1 Networks and Flows

Following the widely used flow model of [4, 5], we will model our network on a finite directed graph $\mathcal{G} \equiv \mathcal{G}(\mathcal{V}, \mathcal{E})$ with node set $\mathcal{V}$ and edge set $\mathcal{E}$. Then, an ordered pair $\sigma = (v, w)$ will be called an *origin-destination* pair of $\mathcal{G}$ if w can be joined to v by a path in $\mathcal{G}$. Furthermore, if we assume that the origin v of σ outputs traffic towards the destination node w at some rate $\rho > 0$, the pair σ together with the rate ρ will be referred to as a *user* of $\mathcal{G}$. In this way, a *network* $\mathcal{Q} \equiv \mathcal{Q}(\mathcal{N}, \mathcal{A})$ in $\mathcal{G}$ will comprise a set of *users* $\mathcal{N}$ (indexed by $i = 1, \dots N$), together with an associated collection $\mathcal{A} \equiv \coprod_i \mathcal{A}_i$ of sets of paths (or *routes*) $\mathcal{A}_i = \{\alpha_{i,0}, \alpha_{i,1} \dots\}$ joining v_i to w_i (where $\sigma_i = (v_i, w_i)$ is the i-th origin-destination pair).

Two remarks of a book-keeping nature are now in order: first, since we will only be interested in users who can choose how to route their traffic, we will take $|\mathcal{A}_i| \geq 2$ for all i. Secondly, we will be assuming that the origin-destination pairs of distinct users are themselves distinct. Fortunately, neither assumption is crucial: if there is only one route available to user i, the rate ρ_i can be considered as a constant load on the route; and if two users $i, j \in \mathcal{N}$ with rates ρ_i, ρ_j share the same origin-destination pair, we will replace them by a single user with rate $\rho_i + \rho_j$ (see also Section 2.2).

So, if $x_{i\alpha} \equiv x_\alpha$ denotes the amount of traffic that user i routes via the path $\alpha \in \mathcal{A}_i$, the corresponding traffic flow may be represented as $x_i = \sum_\alpha^i x_{i\alpha} e_{i\alpha}$, where $\{e_{i\alpha}\}$ is the standard basis of the vector space $V_i \equiv \mathbb{R}^{\mathcal{A}_i}$. However, for such a flow to be admissible, we must also have $x_{i\alpha} \geq 0$ and $\sum_\alpha^i x_{i\alpha} = \rho_i$; hence, the set of admissible flows for user i will be the simplex $\Delta_i \equiv \rho_i \Delta(\mathcal{A}_i) = \{x_i \in V_i : x_{i\alpha} \geq 0 \text{ and } \sum_\alpha^i x_{i\alpha} = \rho_i\}$. Then, by collecting all these individuals flows in a single profile, a *flow in the network* $\mathcal{Q}$ will simply be a point $x = \sum_i x_i \in \Delta \equiv \prod_i \Delta_i$.

An alternative (and very useful!) description of a flow $x \in \Delta$ can be obtained by looking at the traffic load that the flow induces on the edges of the network, i.e. at the amount of traffic y_r that circulates in each edge $r \in \mathcal{E}$ of $\mathcal{G}$. In particular, we set:

$$y_r = \sum_i y_{ir} = \sum_i \sum_{\alpha \ni r}^i x_{i\alpha} \tag{1}$$

where $y_{ir} = \sum_{\alpha \ni r}^i x_{i\alpha}$ is the load induced on $r \in \mathcal{E}$ by the individual flow $x_i \in \Delta$. In this manner, a very important question that arises is the following: *can one recover the flow distribution $x \in \Delta$ from the loads y_r on the edges of the network?*

To answer this, let $\{\varepsilon_r\}$ be the standard basis of the space $W \equiv \mathbb{R}^{\mathcal{E}}$ spanned by the edges $\mathcal{E}$ of $\mathcal{G}$ and consider the *indicator map* $P^i : V_i \to W$ which maps a path $\alpha \in \mathcal{A}_i$ to the sum of its constituent edges: $P^i(e_{i\alpha}) = \sum_{r\in\alpha} \varepsilon_r$; obviously, if we set $P^i(e_{i\alpha}) = \sum_r P^i_{r\alpha}\varepsilon_r$, we see that the entries of P^i will be $P^i_{r\alpha} = 1$ if $r \in \alpha$ and 0 otherwise. We can then aggregate this construction over all $i \in \mathcal{N}$ by considering the product $V \equiv \mathbb{R}^{\mathcal{A}} \cong \prod_i V_i$ and the corresponding indicator matrix $P = P^1 \oplus \cdots \oplus P^N$ whose entries take the value $P_{r\alpha} = 1$ if the path $\alpha \in \mathcal{A}$ employs the edge r and vanish otherwise. By doing just that, (1) takes the simpler form $y_r = \sum_\alpha P_{r\alpha} x_\alpha$ or, even more succinctly, $y = P(x)$. Therefore, the question of whether a flow can be recovered from a load profile can be answered in the positive if the indicator map $P : V \to W$ is injective.

Still, this is not the end of the matter because the individual flows $x_i \in \Delta_i$ actually live in the affine subspaces $p_i + Z_i$ where p_i is the (bary)center of Δ_i and $Z_i \equiv T_{p_i}\Delta_i = \{z_i \in V_i : \sum^i_\alpha z_{i\alpha} = 0\}$ is the tangent space to Δ_i at p_i. As a result, what is actually of essence here is the action of P on the subspaces $Z_i \leq V_i$, i.e. the restriction $Q \equiv P|_Z : Z \to W$ of P on the subspace $Z \equiv T_p\Delta \cong \prod_i Z_i$, where $p = (p_1, \ldots p_N)$. In this way, any two flows $x, x' \in \Delta$ will have $z = x' - x \in Z$ and the respective loads $y, y' \in W$ will satisfy:

$$y' - y = P(x') - P(x) = P(z) = Q(z), \tag{2}$$

so that $y' = y$ iff $x' - x \in \ker Q$. Under this light, it becomes clear that a flow $x \in \Delta$ can be recovered from the corresponding load profile $y \in W$ if and only if Q is injective.

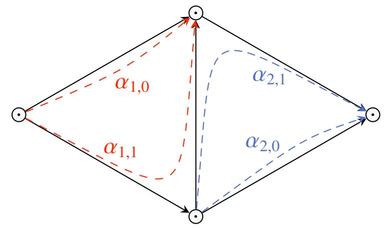
For this reason, the map $Q : Z \to W$ will be called the *degeneracy matrix of the network* $\mathcal{Q}$. In its turn, this leads to:

Definition 1. Let $\mathcal{Q}$ be a network in a graph $\mathcal{G}$ and let Q be its degeneracy matrix. Then, the *degeneracy index* $\operatorname{ind}(\mathcal{Q})$ of $\mathcal{Q}$ is:

$$\operatorname{ind}(\mathcal{Q}) \equiv \dim(\ker Q). \tag{3}$$

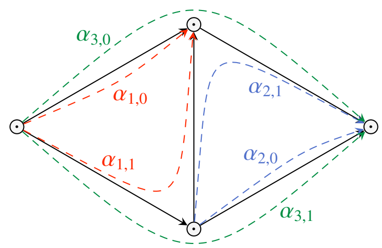
If $\operatorname{ind}(\mathcal{Q}) = 0$, the network $\mathcal{Q}$ will be called *irreducible*; otherwise, $\mathcal{Q}$ will be called *reducible*.

The rationale behind this terminology should be clear: when a network $\mathcal{Q}$ is reducible, some of its routes are "linearly dependent" and the respective directions in $\ker Q$ are "degenerate" (in the sense that they are not reflected on the edge loads). By comparison, the degrees of freedom of irreducible networks are all active and any statement concerning the network's edges may be translated to one concerning its routes.



No linearly dependent paths.

(a) An irreducible network: $\operatorname{ind}(\mathcal{Q}) = 0$.



$\alpha_{1,0} + \alpha_{2,1} + \alpha_{3,1} = \alpha_{1,1} + \alpha_{2,0} + \alpha_{3,0}$

(b) A reducible network: $\operatorname{ind}(\mathcal{Q}) = 1$.

Figure 1: The degeneracy index of a network.

2.2 Congestion Models and Equilibrium

The time spent by an infinitesimal traffic element on an edge $r \in \mathcal{E}$ of the graph $\mathcal{G}$ will be a function $\phi_r(y_r)$ of the traffic load y_r on the edge in question – for example, if the edge represents an M/M/1 queue with capacity μ_r, then $\phi_r(y_r) = 1/(\mu_r - y_r)$. In tune with tradition, we will assume that these *latency* (or *delay*) functions are strictly increasing, and, to keep things simple, that they are C^1 with $\phi'_r(y_r) > 0$.

In this way, the time needed to traverse the route $\alpha \in \mathcal{A}_i$ will be:

$$\omega_{i\alpha}(x) = \sum\nolimits_{r\in\alpha} \phi_r(y_r) = \sum\nolimits_r P^i_{r\alpha}\phi_r(y_r), \tag{4}$$

where, as before, $y_r = \sum_\beta P_{r\beta} x_\beta$. In summary, we then have:

Definition 2. A *congestion model* $\mathfrak{C} \equiv \mathfrak{C}(\mathcal{Q}, \phi)$ in a graph $\mathcal{G}(\mathcal{V}, \mathcal{E})$ is a network $\mathcal{Q}(\mathcal{N}, \mathcal{A})$ of $\mathcal{G}$ equipped with a family of increasing latency functions $\phi_r, r \in \mathcal{E}$.

Definition 2 will constitute our network model in terms of congestion and delay characteristics – see also the seminal papers [4, 5] for a more in-depth discussion on this flow model. On the other hand, we see that our notion of a "user" more accurately portrays a network's *routers* and not its "real-life" users (humans, applications, etc.), so care must be taken in how to phrase the equilibrial conditions for the network. Therefore, given that the routers' selfless task is to satisfy the *nonatomic* traffic elements circulating in the network (the actual selfish entities that correspond to the requests of the network's actual users), the relevant equilibrium concept is given by Wardrop's principle [6]:

Definition 3. A flow profile $q \in \Delta$ is at *Wardrop equilibrium* when, for all users $i \in \mathcal{N}$ and for all routes $\alpha, \beta \in \mathcal{A}_i$ with $q_{i\alpha} > 0$, we have:

$$\omega_{i\alpha}(q) \leq \omega_{i\beta}(q), \tag{5}$$

i.e. when every nonatomic traffic element employs the fastest path available to it.

Condition (5) holds as an equality for all routes $\alpha, \beta \in \mathcal{A}_i$ that are employed in a Wardrop profile q. This gives $\omega_i(q) = \omega_{i\alpha}(q)$ for all $\alpha \in \operatorname{supp}(q_i)$ and leads to the following alternative characterization of Wardrop flows:

$$\omega_i(q) \leq \omega_{i\beta}(q) \text{ for all } i \in \mathcal{N} \text{ and for all } \beta \in \mathcal{A}_i. \tag{5'}$$

Even more importantly, Wardrop equilibria can also be harvested from the (global) minimum of the *Rosenthal potential* [19]:

$$\Phi(y) = \sum\nolimits_r \Phi_r(y_r) = \sum\nolimits_r \int_0^{y_r} \phi_r(w)\, dw. \tag{6}$$

The reason for calling this function a potential is twofold: firstly, it is the nonatomic generalization of the potential function introduced in [20] to describe finite congestion games; secondly, the latencies $\omega_{i\alpha}$ can be obtained from Φ by a simple differentiation.

To be sure, if we set $F(x) = \Phi(y)$ where $y = P(x)$ is the load profile which corresponds to the traffic distribution $x \in \Delta$, we obtain:

$$\frac{\partial F}{\partial x_{i\alpha}} = \sum\nolimits_r \frac{\partial \Phi}{\partial y_r}\frac{\partial y_r}{\partial x_{i\alpha}} = \sum_{r\in\alpha} \phi_r(y_r) = \omega_{i\alpha}(x), \tag{7}$$

which connects to the continuous population setting of [9].

To describe the exact relation between Wardrop flows and the minima of Φ, consider the (convex) set $P(\Delta)$ of all load profiles y that result from admissible flows $x \in \Delta$. Since the latency functions ϕ_r are increasing, Φ will be strictly convex over $P(\Delta)$ and it will thus have a unique (global) minimum $y^* \in P(\Delta)$. Amazingly enough, the Kuhn-Tucker conditions that characterize this minimum coincide with the Wardrop condition (5) [4, 7–9], so the Wardrop set of the congestion model $\mathfrak{C}$ will be:

$$\Delta^* = \{x \in \Delta : P(x) = y^*\} = P^{-1}(y^*) \cap \Delta. \tag{8}$$

This leads us to the following proposition, whose first part is well-known, but whose second part seems to have been overlooked by the existing literature, despite its simplicity:

Proposition 4. *Let $\mathfrak{C} \equiv \mathfrak{C}(\mathfrak{Q}, \phi)$ be a congestion model with strictly increasing latencies ϕ_r and let Δ^* be its Wardrop set. Then:*

1. *Any two Wardrop flows exhibit equal loads and delays.*
2. *Δ^* is a nonempty convex polytope with* $\dim(\Delta^*) \leq \operatorname{ind}(\mathfrak{Q})$*; moreover, if there exists an interior equilibrium $q \in \operatorname{Int}(\Delta)$, then* $\dim(\Delta^*) = \operatorname{ind}(\mathfrak{Q})$.

Since $P^{-1}(y^*)$ is an affine subspace of $\mathbb{R}^{\mathcal{A}} \equiv \prod_i \mathbb{R}^{\mathcal{A}_i}$ and Δ is a product of simplices, there is not much to prove (simply observe that if q is an interior Wardrop flow, then $P^{-1}(y^*)$ intersects the full-dimensional interior of Δ). As we noted above, the only surprise here is that this result seems to have been overlooked in most of the literature concerning congestion models: for instance, in both [9, Corollary 5.6] and [10, Propositions 2, 3], the authors presume that Wardrop equilibria are unique in networks with increasing latencies. However, if there are two distinct flows x, x' leading to the same load profile y (e.g. as in Fig. 1(b)), then the potential function $F(x) \equiv \Phi(P(x))$ is no longer strictly convex: in fact, it is constant along every null direction of the degeneracy matrix $Q \equiv P|_{T\Delta}$.

We thus see that a Wardrop equilibrium is unique iff *a*) the network $\mathfrak{Q}$ is *irreducible*, or *b*) $P^{-1}(y^*)$ only intersects Δ at a vertex. This last condition suggests that the vertices of Δ play a special role so, in analogy with Nash games, we define:

Definition 5. A Wardrop equilibrium q is *strict* if *a*) q is *pure*: $q = \sum_i \rho_i e_{i,\alpha_i}$, $\alpha_i \in \mathcal{A}_i$; and *b*) $\omega_{i\alpha_i}(q) < \omega_{i\beta}(q)$ for all paths $\beta \in \mathcal{A}_i \setminus \{\alpha_i\}$.

As a matter of fact, the existence of a strict Wardrop equilibrium actually *precludes* the existence of any other equilibria:

Proposition 6. *Let $\mathfrak{C}$ be a congestion model. If q is a strict Wardrop equilibrium of $\mathfrak{C}$, then $\mathfrak{C}$ has no other equilibria.*

Proof. Without loss of generality, let $q = \sum_i \rho_i e_{i,0}$ be a strict Wardrop equilibrium of $\mathfrak{C}$ and suppose ad absurdum that $q' \neq q$ is another Wardrop flow. If we set $z = q' - q \in \ker Q$, it follows that the convex combinations $q + \theta z$ will also be Wardrop for all $\theta \in [0, 1]$; moreover, for small enough $\theta > 0$, $q + \theta z$ employs at least one path $\mu \in \mathcal{A}_i \setminus \{0\}$ that is not present in q (recall that q is pure). As a result, we get $\omega_{i\mu}(q + \theta z) = \omega_{i,0}(q + \theta z)$ for all sufficiently small $\theta > 0$, and because the latency functions $\omega_{i\alpha}$ are continuous, this yields $\omega_{i,0}(q) = \omega_{i\mu}(q)$. However, since q is a *strict* Wardrop equilibrium which does not employ μ, we must also have $\omega_{i,0}(q) < \omega_{i\mu}(q)$, a contradiction. □

Equilibria and Optimality

It is important to note here that, in terms of the average latencies

$$\omega_i(x) = \rho_i^{-1} \sum\nolimits_\alpha^i x_{i\alpha}\omega_{i\alpha}(x), \tag{9}$$

the *optimal* traffic distributions which minimize the aggregate delay $\omega(x) = \sum_i \rho_i \omega_i(x)$ coincide with the Wardrop equilibria of a suitably modified game. This was first noted in [7]: just as Wardrop equilibria occur at the minimum of the Rosenthal potential, so can one obtain the minimum of the aggregate latency ω by looking at the Wardrop equilibria of an associated congestion model. More precisely, the only change that needs to be made is to consider the "marginal" latency functions $\phi_r^*(y_r) = \phi_r(y_r) + y_r\phi_r'(y_r)$ – see also [4, 5]. Therefore, to study these "socially optimal" flows, we simply have to redress our analysis to fit these "marginal latencies" instead (see also Section 5 for more details).

3. DETERMINISTIC EVOLUTION

Unfortunately, locating the Wardrop equilibria of a network is a relatively arduous process which entails a good deal of global calculations (namely, the minimization of a nonlinear convex functional with exponentially many variables over a convex polytope). Since such calculations exceed the deductive capabilities of individual users (especially if they do not have access to global information), we will examine whether there are simple learning schemes which allow users to reach an equilibrium in an efficient manner, without having to rely on centralized computations.

3.1 The Replicator Dynamics

A most natural choice for such a learning scheme is the following: the i-th user contrasts the delay $\omega_{i\alpha}(x)$ that he observes along the path $\alpha \in \mathcal{A}_i$ to his average delay $\omega_i(x) = \rho_i^{-1}\sum_\alpha^i x_{i\alpha}\omega_{i\alpha}(x)$, and then increases or decreases the amount of traffic $x_{i\alpha}$ routed via $\alpha \in \mathcal{A}_i$ proportionately to this difference. In continuous time, this leads to the *replicator equation*:

$$\frac{dx_{i\alpha}}{dt} = x_{i\alpha}\left(\omega_i(x) - \omega_{i\alpha}(x)\right). \tag{10}$$

Alternatively, if players learn at different rates $\lambda_i > 0$ as a result of varied stimulus-response characteristics, we obtain the *rate-adjusted* dynamics:

$$\frac{dx_{i\alpha}}{dt} = \lambda_i x_{i\alpha}\left(\omega_i(x) - \omega_{i\alpha}(x)\right) \tag{11}$$

(naturally, the uniform case (10) is recovered when all players learn at the "standard" rate $\lambda_i = 1$). Interestingly enough, these learning rates can also be viewed as (player-specific) inverse temperatures: in high temperatures (small λ_i), the differences between routes are toned down and players evolve along the slow time-scales $\lambda_i t$; at the other end of the spectrum, if $\lambda_i \to \infty$, equation (11) "freezes" instead to a rigid (and myopic) best-reply process [18, 21].

At any rate, we see that this equation is in perfect harmony with our "local information" mantra because users do not need to know the delays along paths that they do not employ – the replicator vector field vanishes when $x_{i\alpha} = 0$. Thus, users that evolve according to (10) are oblivious to their surroundings, even to the existence of other users: they simply use (10) to respond to the stimuli $\omega_{i\alpha}(x)$ in the hope of minimizing their delays.

3.2 Entropy and Rationality

As is well-known, Wardrop equilibria are rest points of (11): if q is Wardrop, the characterization (5′) gives $\omega_{i\alpha}(q) = \omega_i(q)$ whenever $x_{i\alpha} > 0$. However, the same holds for *all* flows q' which exhibit equal latencies along the paths in their support, and these flows are not necessarily Wardrop (in the terminology of [9], this means that the replicator dynamics are "complacent"). Consequently, the issue at hand is whether or not the replicator dynamics manage to single out Wardrop equilibria among other stationary states.

A key role in this question is played by the *relative entropy* (also known as the *Kullback-Leibler divergence*):

$$H_q(x) \equiv d_{\mathrm{KL}}(q,x) = \sum_{\alpha\in\mathrm{supp}(q)} q_{i\alpha} \log \frac{q_{i\alpha}}{x_{i\alpha}} \tag{12}$$

where the sum is over the support of q: $\mathrm{supp}(q) \equiv \{\alpha : q_{i\alpha} > 0\}$. The significance of the relative entropy function lies in that it measures distance in probability space. Even more importantly, an easy calculation shows that its time-derivative $\dot{H}_q(x)$ is just:

$$\dot{H}_q(x) = -\sum\nolimits_{i,\alpha} q_{i\alpha}\left(\omega_i(x) - \omega_{i\alpha}(x)\right) \equiv -L_q(x), \tag{13}$$

where we have set $L_q(x) \equiv \sum_\alpha^i q_{i\alpha}(\omega_i(x) - \omega_{i\alpha}(x))$.

This last quantity will be very important to our analysis, so $L_q(x)$ will be called *the evolutionary index of x w.r.t. q*, on account of the fact that q is *evolutionarily stable* (in the sense of [22]) if and only if $L_q(x) > 0$ near q.[1] The properties of L_q will be discussed at length in Appendix A where, among others, we establish the crucial link with the game's potential $F(x) = \Phi(P(x))$:

$$L_q(x) \geq F(x) - F(q) \geq 0. \tag{14}$$

First off, this inequality reveals that q is weakly evolutionarily stable and, in fact, *strictly* so if $\mathcal{Q}$ is irreducible (otherwise the network's degenerate directions are evolutionarily neutral).[2] Secondly (and more importantly from a dynamical standpoint), it shows that the relative entropy is a (semi-definite) Lyapunov function for the replicator dynamics (10). In view of this, it would be tempting to infer that the replicator dynamics converge to Wardrop equilibrium, but, unfortunately, counterexamples abound where a semi-definite Lyapunov function is not enough to guarantee convergence to a *point* by itself. Be that as it may, there is much more at work in (10) than a single semi-definite Lyapunov function: there exists a whole *family* of such functions, one for each Wardrop flow $q \in \Delta$. So, undettered by potential degeneracies, we show in appendix B that the dynamics (10) really *do* converge to an equilibrium point:

Theorem 7. *Every interior solution trajectory $x(t)$ of the replicator dynamics (11) converges to a Wardrop equilibrium; in particular, if the network $\mathcal{Q}$ is irreducible, $x(t)$ converges to the network's (necessarily) unique equilibrium point.*

Remark. In continuous potential games, the replicator dynamics (10) are "complacent" because their rest points include but do not coincide with the network's Wardrop set [9]. As a result, Sandholm's work only ensures that Wardrop equilibria are Lyapunov stable. To the best of our knowledge, the stronger asymptotic properties of Theorem 7 were first suggested in [10], but, in networks with more than one user (the "multi-commodity" case), the authors implicitly rely on the network's irreducibility.[3] If this is not the case, $\dot{H}_q(x)$ is only positive *semi-definite* and their approach breaks down because Wardrop equilibria are only *neutrally* stable – this is also the problem with Corrolary 5.1 in [9].

3.3 Rate of Convergence

Of course, in order for the replicator dynamics to constitute a relevant learning scheme from an applicational perspective, the convergence to equilibrium must occur at a timely fashion. Theorem 7 is moot on this point, but as we prove in Appendix B, the convergence is, in fact, remarkably fast:

Theorem 8. *Let $x(t)$ be an interior solution orbit of the replicator dynamics (10), and let $q = \lim_{t\to\infty} x(t)$ be its (well-defined by Theorem 7) ω-limit. Then:*

$$H_q(x(t)) \leq h_0 e^{-ct}, \tag{15}$$

where $h_0 = H_q(x(0))$ and c is a positive constant. In other words, the replicator dynamics hit an ε-neighborhood of a Wardrop equilibrium in time which is at most of order $\mathcal{O}(\log(1/\varepsilon))$.

In particular:

1. *If q is strict, say $q = \sum_i \rho_i e_{i,0}$, and we set $\Delta\omega_i = \min_{\mu\neq 0}\{\omega_{i\mu}(q) - \omega_{i,0}(q)\}$, then $c = h_0^{-1} \min_i \left\{\rho_i\left(1 - e^{-h_0/\rho_i}\right)\Delta\omega_i\right\}$.*
2. *If $q \in \mathrm{Int}(\Delta)$ and the network $\mathcal{Q}$ is irreducible, then $c = \frac{1}{2}\frac{r\zeta_c^2}{h_c}$, where r is the minimum eigenvalue of the Hessian of F over Δ, $h_c \geq h_0$ is a constant and $\zeta_c^2 = \min_x\{\|x-q\|^2 : H_q(x) = h_c\}$.*

Remark 1. Even though we estimate the asymptotic convergence rate c for two special cases (strict equilibria and interior equilibria of irreducible networks), it should be stressed that the exponential bound (15) is valid for *any* congestion model with strictly increasing latency functions, whether the network is irreducible or not. The only difference in that case is that the expressions for c become more complicated (see Appendix B for details), so we opted to present only the two most representative cases.

Remark 2. It is important to note that the dependence of c on the various parameters of the network also allows us to express the speed of convergence in terms of whichever parameter we wish to study. For example, in the strict equilibrium case, we see that the convergence time is of order $\mathcal{O}(1/\Delta\omega)$ in the payoff differences $\Delta\omega_i$ or, in the interior case, of order $\mathcal{O}(1/r)$ in the minimum eigenvalue r of the game's Hessian (which reflects the network's latencies). In this way, (15) can be used to obtain a number of very useful estimates on the equilibration time of the dynamics (10).

Remark 3. In comparing Theorem 8 with the $\mathcal{O}(1/\varepsilon^3)$ bound of [10], one should bear in mind that hitting an ε-neighborhood of an equilibrium is not the same as hitting an ε-approximate equilibrium. Of course, the two notions are closely related, but the precise relation between them depends on the character of the equilibrium in question. At any rate, since payoffs are approximately linear near strict equilibria while they grow quadratically near (isolated) interior equilibria, we see that the logarithmic bound of (15) represents a significant improvement over the polynomial bounds of [10].

Finally, it should be mentioned here that the lower bound $\mathcal{O}(1/\varepsilon)$ for the convergence time that can be found in [10] does not apply in our case because it concerns constant latency functions (and ε-dependent to boot).

[1] In multi-population settings such our own, there is no consensus among game theorists on how to define evolutionary stability. Taylor's definition [22] is the middle (and most useful in terms of applications) road between the stronger definition of [23] and the weaker one of [24].

[2] Evolutionary stability was also noted in [10] but the important distinction between weak and strong stability was missed, again because of potential degeneracies.

[3] Specifically, their claim that the proof of Proposition 2 in [10] "covers the multi-commodity case" is false, because their assertion that different flow profiles carry edges with different loads only holds in irreducible networks.

4. THE ADVENT OF STOCHASTICITY

Going back to our original discussion on learning, we see that the users' evolution hinges on the feedback that they receive about their choices, namely the delays $\omega_{i\alpha}(x)$ that they record. We have already noted that this information is based on actual observations, but this does not necessarily mean that it is also *accurate* as well. For instance, the interference of nature with the game (manifesting e.g. as packet drops which lead to retransmissions that perturb the load on a link), or imperfect estimates of the packets' round-trip times might perturb this information considerably. Additionally, it is important to remember that the latency functions $\omega_{i\alpha}$ only represent the users' *expected* delays in queueing theory, so the delays that users actually observe will fluctuate randomly around their average latencies, and this could negatively affect the rationality properties of our replicator learning scheme.

4.1 Stochastic Replicator Dynamics

Our goal here will be to determine the behavior of the replicator dynamics under stochastic perturbations of the kind outlined above. To that end, write the delay that users experience along the edge $r \in \mathcal{E}$ as $\hat{\phi}_r = \phi_r + \eta_r$ where η_r denotes the perturbation process. Then, the latency $\hat{\omega}_{i\alpha}$ along $\alpha \in \mathcal{A}_i$ will just be $\hat{\omega}_{i\alpha} = \omega_{i\alpha} + \eta_{i\alpha}$, where, in obvious notation, $\eta_{i\alpha} = \sum_\alpha P^i_{r\alpha}\eta_r$. In this way, the replicator dynamics (10) become:

$$\frac{dx_{i\alpha}}{dt} = x_{i\alpha}\left(\hat{\omega}_i - \hat{\omega}_{i\alpha}\right) = x_{i\alpha}\left(\omega_i - \omega_{i\alpha}\right) + x_{i\alpha}(\eta_i - \eta_{i\alpha}) \tag{16}$$

where $\hat{\omega}_i = \rho_i^{-1}\sum^i_\beta x_{i\beta}\hat{\omega}_{i\beta}$ and $\eta_i = \rho_i^{-1}\sum^i_\beta x_{i\beta}\eta_{i\beta}$.

The exact form of the perturbations η_r clearly depends on the situation at hand, but since we are interested in stochastic fluctuations around the average latencies $\omega_{i\alpha}$, it is reasonable to take these perturbations to be a driftless ergodic process which does not bias users towards one direction or another. We will thus rewrite (16) as a *stochastic* differential equation:

$$dX_{i\alpha} = X_{i\alpha}\left[\omega_i(X) - \omega_{i\alpha}(X)\right]dt + X_{i\alpha}\left[dU_{i\alpha} - \tfrac{1}{\rho_i}\sum\nolimits^i_\beta X_{i\beta}\,dU_{i\beta}\right] \tag{17}$$

where $dU_{i\alpha}$ is the total noise along the path $\alpha \in \mathcal{A}_i$:

$$dU_{i\alpha} = \sum\nolimits_{r\in\alpha}\sigma_r\,dW_r = \sum\nolimits_r P^i_{r\alpha}\sigma_r\,dW_r \tag{18}$$

and $W(t) = \sum_r W_r(t)\varepsilon_r$ is a Wiener process in $\mathbb{R}^{\mathcal{E}}$, the space spanned by the edges $\mathcal{E}$ of the network. Similarly, if players learn at different rates λ_i, we get:

$$\begin{aligned} dX_{i\alpha} &= \lambda_i X_{i\alpha}\left[\omega_i(X) - \omega_{i\alpha}(X)\right] + \lambda_i X_{i\alpha}\left[dU_{i\alpha} - \rho_i^{-1}\sum\nolimits^i_\beta X_{i\beta}\,dU_{i\beta}\right] \\ &= \lambda_i b_{i\alpha}(X)dt + \lambda_i\sum\nolimits^i_\beta c_{i,\alpha\beta}(X)\,dU_{i\beta} \end{aligned} \tag{19}$$

where b and c are the drift and diffusion coefficients of (17).

Remark 1. The rate-adjusted equation (19) will be our stochastic version of the replicator dynamics and, as such, the noise coefficients σ_r warrant some discussion. Indeed, even though we have written them in a form that suggests they are constant, they need not be so: after all, the intensity of the noise on an edge might well depend on the edge loads $Y_r = \sum_\alpha P_{r\alpha}X_\alpha$. On that account, we will only assume that these coefficients are essentially bounded functions of the loads y. Nonetheless, in an effort to reduce notational clutter, we will not indicate this dependence explicitly; instead, we simply remark here that our results continue to hold if we replace σ_r with the worst-case scenario $\sigma_r \leftrightarrow \operatorname{ess\,sup}_y \sigma_r(y)$.

Remark 2. This last remark also highlights the generality of the stochastic replicator equation (17). Indeed, if the noise coefficients $\sigma_{i\alpha}$ depend on the state of the process $X(t)$, it is not hard to see that the the diffusion term of (17) represents the most general diffusion term that respects the simplicial structure of the strategy space Δ. In this sense, we lose little generality by restricting ourselves to Wiener-type perturbation processes.[4]

It is also important to compare (19) to other stochastic incarnations of the replicator dynamics, namely the "aggregate shocks" version of [13–16] and the "exponential learning" approach of [17, 18]. In the case of the former, one perturbs the replicator equation (10) by accounting for the (stochastic) interference of nature with a species' growth rate and obtains [13]:

$$\begin{aligned} dX_{i\alpha} &= X_{i\alpha}\left(u_{i\alpha}(X) - u_i(X)\right)dt - \left(\sigma^2_{i\alpha}X_{i\alpha} - \sum\nolimits^i_\beta \sigma^2_{i\beta}X_{i\beta}\right)dt \\ &\quad + X_{i\alpha}\left[\sigma_{i\alpha}\,dW_{i\alpha} - \sum\nolimits^i_\beta \sigma_{i\beta}X_{i\beta}\,dW_{i\beta}\right], \end{aligned} \tag{20}$$

where $W = \sum_{i,\alpha} W_{i\alpha}e_{i\alpha}$ is a Wiener process in $\prod_i \mathbb{R}^{\mathcal{A}_i}$.

By comparison, in the "exponential learning" case it is assumed that the players of a Nash game employ a learning scheme akin to logistic fictitious play [25]. However, if the information that players have is imperfect, the errors propagate to their learning curves and instead lead to the stochastic dynamics:

$$\begin{aligned} dX_{i\alpha} &= X_{i\alpha}\left[\left(u_{i\alpha}(X) - u_i(X)\right)\right]dt + X_{i\alpha}\left[\sigma_{i\alpha}dW_{i\alpha} - \sum\nolimits^i_\beta \sigma_{i\beta}X_{i\beta}dW_{i\beta}\right] \\ &\quad + \tfrac{1}{2}X_{i\alpha}\left[\sigma^2_{i\alpha}(1 - 2X_{i\alpha}) - \sum\nolimits^i_\beta \sigma^2_{i\beta}X_{i\beta}(1 - 2X_{i\beta})\right]dt. \end{aligned} \tag{21}$$

In light of the above, there are two notable traits of (19) that set it apart from its other stochastic versions. First, the drift of (19) coincides with the deterministic replicator dynamics (11) whereas the drifts of (20) and (21) do not. Secondly, the martingales U that appear in (19) are not uncorrelated components of some Wiener process (as is the case for both (20) and (21)), but, depending on whether the paths $\alpha, \beta \in \mathcal{A}$ have edges in common or not, the processes U_α, U_β might be highly correlated or not at all.

To make this more precise, recall that the Wiener differentials dW_r are orthogonal: $dW_r \cdot dW_s = d[W_r, W_s] = \delta_{rs}\,dt$. In its turn, this implies that the stochastic differentials dU_α, dU_β satisfy:

$$\begin{aligned} dU_\alpha \cdot dU_\beta &= \left(\sum\nolimits_r P_{r\alpha}\sigma_r\,dW_r\right)\cdot\left(\sum\nolimits_s P_{s\beta}\sigma_s\,dW_s\right) \\ &= \sum\nolimits_{r,s} P_{r\alpha}P_{s\beta}\sigma_r\sigma_s\delta_{rs}\,dt = \sum\nolimits_{r\in\alpha\beta}\sigma_r^2\,dt = \sigma^2_{\alpha\beta}\,dt, \end{aligned} \tag{22}$$

where $\sigma^2_{\alpha\beta} = \sum_r P_{r\alpha}P_{r\beta}\sigma_r^2$ is the variance of the noise along the intersection $\alpha\beta \equiv \alpha \cap \beta$ of the paths $\alpha, \beta \in \mathcal{A}$. We thus see that the processes U_α and U_β are uncorrelated iff the paths $\alpha, \beta \in \mathcal{A}$ have no common edges. At the other extreme, we have:

$$(dU_\alpha)^2 = \sum\nolimits_{r\in\alpha}\sigma_r^2\,dt = \sigma_\alpha^2\,dt \tag{23}$$

where $\sigma^2_\alpha \equiv \sigma^2_{\alpha\alpha} = \sum_r P_{r\alpha}\sigma_r^2$ measures the intensity of the noise on the route $\alpha \in \mathcal{A}$. These expressions will be key to our analysis and we will make liberal use of them in the rest of our paper.

4.2 Stochastic Fluctuations and Rationality

Our goal in this section will be to explore the rationality properties of the stochastic replicator dynamics (19). Similarly to the

[4] Recall also that a Brownian motion with drift is the most general Lévy process with no jumps, so the real limitation of the dynamics (17) is that they cannot account for catastrophic events (such as a discontinuous drop in the capacity of a link).

deterministic setting, our main tool will be the (rate-adjusted) relative entropy:

$$H_q(x;\lambda) = \sum_i \lambda_i^{-1} \sum_\alpha^i q_{i\alpha} \log(q_{i\alpha}/x_{i\alpha}) \tag{24}$$

which we will study with the help of the *generator* $\mathcal{L}$ of the diffusion (19). To that end, recall that the generator $\mathcal{L}$ of the Itô diffusion:

$$dX_\alpha(t) = \mu_\alpha(X(t))\,dt + \sum\nolimits_\beta \sigma_{\alpha\beta}(X(t))\,dW_\beta(t), \tag{25}$$

where W is a Wiener process, is just the differential operator:

$$\mathcal{L} = \sum\nolimits_\alpha \mu_\alpha(x)\frac{\partial}{\partial x_\alpha} + \frac{1}{2}\sum\nolimits_{\alpha\beta}\left(\sigma(x)\sigma^T(x)\right)_{\alpha\beta}\frac{\partial^2}{\partial x_\alpha \partial x_\beta} \tag{26}$$

(for a comprehensive account, consult the excellent book [26]). In this manner, if f is sufficiently smooth (C^2 suffices), $\mathcal{L}f$ captures the drift of the process $f(X(t))$:

$$df(X(t)) = \mathcal{L}f(X(t))\,dt + \sum\nolimits_{\alpha\beta} \left.\frac{\partial f}{\partial x_\alpha}\right|_{X(t)} \sigma_{\alpha\beta}(X(t))\,dW_\beta(t). \tag{27}$$

Of course, in the case of the diffusion (19), the martingales U are not the components of a Wiener process, so (27) cannot be applied right off the shelf. However, a straightforward application of Itô's lemma (see appendix C for this section's proofs) yields:

Lemma 9. *Let $\mathcal{L}$ be the generator of (19). Then, for any $q \in \Delta$:*

$$\begin{aligned}\mathcal{L}H_q(x;\lambda) = -L_q(x) &+ \frac{1}{2}\sum\nolimits_i \frac{\lambda_i}{\rho_i}\sum\nolimits_{\beta,\gamma}^i \sigma_{\beta\gamma}^2 (x_{i\beta} - q_{i\beta})(x_{i\gamma} - q_{i\gamma}) \\ &+ \frac{1}{2}\sum\nolimits_i \frac{\lambda_i}{\rho_i}\sum\nolimits_{\beta,\gamma}^i \sigma_{\beta\gamma}^2 q_{i\beta}(\rho_i\delta_{\beta\gamma} - q_{i\gamma}),\end{aligned} \tag{28}$$

where, as before, $L_q(x) = \sum_{i\alpha}(x_{i\alpha} - q_{i\alpha})\omega_{i\alpha}(x)$.

Based on this lemma, we see that $\mathcal{L}H_q$ (that is, the "average" evolution of $H_q(X(t))$) is actually *positive* in a neighborhood of an interior equilibrium q. As a result, unconditional convergence to Wardrop equilibrium appears to be a "bridge too far" in stochastic environments, especially when the equilibrium in question is not pure – after all, mixed equilibria are not even traps (stationary states with probability 1) of (19). So, instead of asking for almost-sure convergence or the like, we will concentrate on a "stochasticized" version of "convergence in the mean" by looking at the long-time averages of the replicator dynamics (10).

Before embarking on this analysis, a little more groundwork is required. First, we introduce for convenience the aggregates:

$$\rho = \sum\nolimits_i \rho_i, \tag{29a}$$

$$\sigma^2 = \sum\nolimits_r \sigma_r^2, \tag{29b}$$

$$\lambda = \sum\nolimits_i \frac{\rho_i}{\rho}\lambda_i. \tag{29c}$$

Secondly, it will be more practical to measure distances from q with a variant of the L^1 norm. Indeed, let $S_q \equiv \{z \in T_q\Delta : q + z \in \mathrm{bd}(\Delta)\}$ be the set of tangent vectors $z \in Z$ which connect $q \in \mathrm{Int}(\Delta)$ to the boundary $\mathrm{bd}(\Delta)$ of Δ. Since Δ is convex, any $x \in \Delta$ can be uniquely expressed as $x = q + \theta z$ for some $z \in S_q$ and some $\theta \in [0, 1]$, so we define the *projective distance* $\Theta_q(x)$ of x from q to be:

$$\Theta_q(x) = \theta \Leftrightarrow x = q + \theta z \text{ for some } z \in S_q \text{ and } 0 \le \theta \le 1. \tag{30}$$

Of course, Θ_q is not a bona fide distance function by itself, but it closely resembles the L^1 norm: the "projective balls" $B_\theta = \{x : \Theta_q(x) \le \theta\}$ are rescaled copies of Δ (S_q is the "unit sphere" in this picture). In a similar vein, we define the *essence* of $q \in \Delta$ to be:

$$\mathrm{ess}(q) = \rho^{-1}\min\{\|P(z)\| : z \in S_q\}, \tag{31}$$

where $\|\cdot\|$ denotes the ordinary Euclidean norm and the factor of ρ was included for scaling purposes. Comparably to $\mathrm{ind}(\mathfrak{Q})$, $\mathrm{ess}(q)$ measures degeneracy (or rather, the lack thereof): $\mathrm{ess}(q) = 0$ only if some direction $z \in S_q$ is null for P, i.e. only if $\mathfrak{Q}$ is reducible.

We are finally in a position to state and prove:

Theorem 10. *Let $q \in \mathrm{Int}(\Delta)$ be an interior equilibrium of an irreducible congestion model $\mathfrak{C}$, and assume that the users' learning rates satisfy the condition:*

$$\lambda < \frac{4}{5}\frac{m\rho\kappa^2}{\sigma^2}, \text{where } m = \inf\{\phi_r'(y_r) :\} \text{ and } \kappa = \mathrm{ess}(q). \tag{32}$$

Then, for any interior initial condition $X(0) = x \in \mathrm{Int}(\Delta)$, the time averages of $X(t)$ are concentrated in a neighborhood of q; specifically, if $\Theta_q(\cdot)$ denotes the projective distance (30) from q, then:

$$\mathbf{E}_x\left[\frac{1}{t}\int_0^t \Theta_q^2(X(s))\,ds\right] \le \theta_\lambda^2 + \mathcal{O}(1/t), \text{ where } \theta_\lambda^2 = \frac{1}{4}\left(\frac{m\rho\kappa^2}{\lambda\sigma^2} - 1\right)^{-1}.$$

Accordingly, the transition probabilities of $X(t)$ converge in total variation to an invariant probability measure π on Δ which concentrates mass around q. In particular, if $B_\theta = \{x \in \Delta : \Theta_q(x) \le \theta\}$ is a "projective ball" around q, we have:

$$\pi(B_\theta) \ge 1 - \theta_\lambda^2/\theta^2. \tag{33}$$

Since these last results constitute the stochastic rationality properties of the replicator dynamics, a few remarks are in order:

Remark 1 (*Convergence Speed*). Since the invariant measure which corresponds to the stationary distribution of the stochastic replicator dynamics is concentrated around a "projective ball" of radius θ_λ around q, the best we can do in order to estimate the convergence speed of the stochastic replicator dynamics is to calculate the average time it takes to hit a neighborhood of this ball. In Appendix C we show that if K_ε is an ε-neighborhood of this "projective ball", then the corresponding expected hitting time $\mathbf{E}_x[\tau_\varepsilon]$ is bounded by:

$$\mathbf{E}_x[\tau_\varepsilon] \le \frac{H_q(x;\lambda)}{\varepsilon}. \tag{34}$$

To all intents and purposes, this is the analogue of the convergence time estimate of Theorem 8: we see that the replicator trajectories hit an ε-neighborhood of the projective ball where the long-term average of the stochastic dynamics is concentrated in time which is at most of order $\mathcal{O}(1/\varepsilon)$. In other words, the effects of noise can be summarized as follows:

1. They "blur" the dynamics' convergence: only the time-average of the stochastic dynamics is concentrated in a projective ball around an evolutionarily stable flow profile – see Figure 2.

2. The equilibration time scale of the dynamics jumps up to $\mathcal{O}(1/\varepsilon)$ from $\mathcal{O}(\log(1/\varepsilon))$ in the deterministic setting.

Remark 2 (*Degeneracy*). The irreducibility assumption is actually quite important: it appears both in the "slow-learning" condition (32) (recall that $\mathrm{ess}(q) = 0$ if q is an interior point of a reducible network) and also in the proof of Theorem 10 (Appendix C). This shows that the stochastic dynamics (19) are not oblivious to degenerate degrees of freedom, in stark contrast with the deterministic case (Theorems 7 and 8).

Regardless, an analogue of Theorem 10 should still hold for reducible networks if we replace q with the entire (affine) set Δ^*.

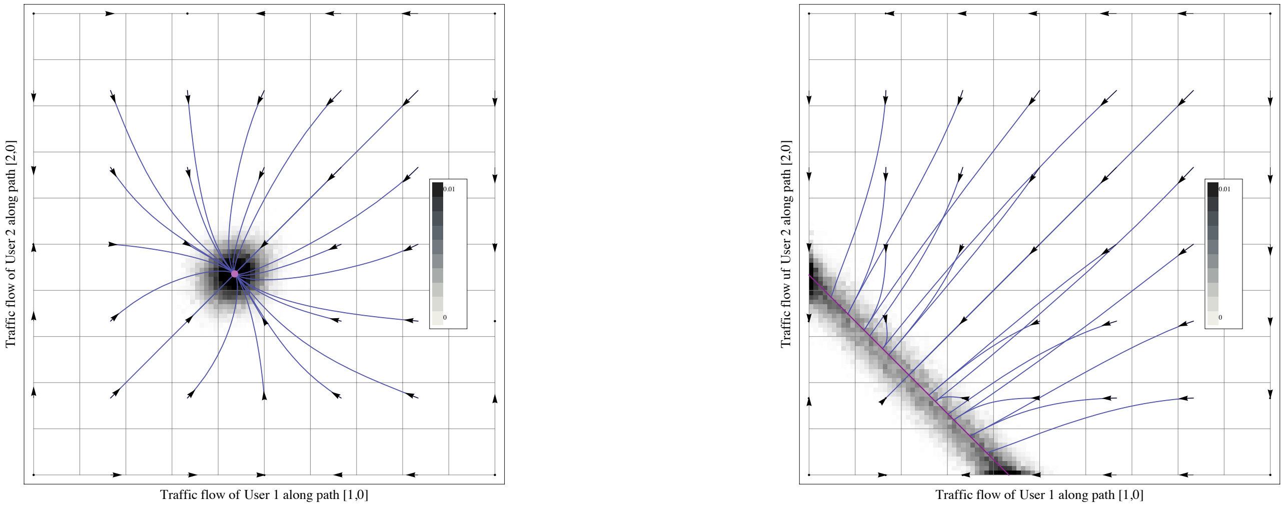


Figure 2: Learning in the networks of Figure 1 with M/M/1 latency functions $\phi_r(y_r) = (\mu_r - y_r)^{-1}$ and arbitrarily chosen capacities μ_r. The shades of gray represent the stationary distribution of (19) (obtained by numerical simulations) and the flow lines (blue) represent the solution trajectories of (11) – in (b) they are actually projections because there are three users.

More precisely, we conjecture that under a suitably modified learning condition, the transition probabilities of $X(t)$ converge to an invariant distribution which concentrates mass around Δ^* (Fig. 2(b)). One way to prove this would be to find a suitable way to "quotient out" $\ker Q$ but, since (19) is not invariant over the degenerate fibres $x + \ker Q$, $x \in \Delta$, we have not yet been able to do so.

Remark 3 (*Sharpness*). We should also note here that the bounds we obtained are not the sharpest possible ones. For example, the learning condition (32) can be tightened (by quite a bit actually!) and the assumption that $\phi_r' > 0$ can be dropped. In that case however, the corresponding expressions are significantly more complicated, so we have opted to focus on the simpler estimates.

5. DISCUSSION

In this paper, we studied the evolution of traffic flows in networks whose users try to minimize their latencies by employing a learning scheme based on the replicator dynamics of evolutionary game theory. The stable states of these dynamics coincide with the network's Wardrop equilibria and we showed that these points constitute an affine polytope whose dimension is determined by the network's *degeneracy index*, an important new notion which measures the linear dependence of the users' paths.

When the information that they observe is perfect (deterministic), we showed that the replicator dynamics converge to a Wardrop equilibrium at an exponential rate: users get within ε of an equilibrium in time of order $\mathcal{O}(\log(1/\varepsilon))$. However, if the network latencies are subject to random perturbations, this convergence is "blurred": in networks with no degeneracy, it is the time average of the replicator dynamics which is concentrated in a ball around the network's (evolutionarily stable) equilibrium point; furthermore, the equilibration time scale of the dynamics is also affected by the noise and jumps up to order $\mathcal{O}(1/\varepsilon)$.

These results also carry significant applicational potential from an optimization point of view as well. Indeed, we have already noted that the traffic flows which minimize the *aggregate* latency function $\omega(x) = \sum_i \rho_i \omega_i(x)$ in a network are precisely the Wardrop equilibria of a modified congestion model which is defined over the same network and whose delay functions are given by the "marginal latencies" $\phi_r^*(y_r) = \phi_r(y_r) + y_r\phi_r'(y_r)$. Hence, if we set $\omega_{i\alpha}^*(x) = \sum_\alpha^i P_{r\alpha}^i \phi_r^*(y_r)$ and use $\omega_{i\alpha}^*$ in place of $\omega_{i\alpha}$ in the replicator dynamics (11) and (19), our analysis instead yields convergence to the *optimal* traffic allocation in a network.

The only limiting factor in this optimization approach is that the marginal costs $\phi_r^*(y_r)$ do not really constitute "local information" that users can acquire simply by routing their traffic and recording the delays that they experience. However, the missing components $y_r\phi_r'(y_r)$ can easily be measured by observers monitoring the edges of the network and could be subsequently publicized to all users that employ the edge $r \in \mathcal{E}$. Consequently, if the adminstrators of a network wish users to figure out the optimal traffic allocation on their own, they simply have to go the (small) extra distance of providing such monitors on the network's links.

Important extensions of this work include the asymptotic convergence properties of the stochastic replicator dynamics with respect to strict equilibria; in fact, it can be shown that strict equilibria are *always* stochastically stable, irrespective of the level of the noise. However, of more immediate interest is the extension of our stochastic results to *reducible* networks: as evidenced by Figure 2(b), it is plausible to expect our results to extend to the reducible case as well, but due to the inherent complications of this analysis, we prefer to leave it as a future project.

APPENDIX

A. THE EVOLUTIONARY INDEX

This appendix is devoted to the *evolutionary index*:

$$L_q(x) \equiv \sum\nolimits_{i,\alpha} (x_{i\alpha} - q_{i\alpha})\omega_{i\alpha}(x) = \sum\nolimits_r (y_r - y_r^*)\phi_r(y_r) \equiv \Lambda(y). \quad (35)$$

The first important property of the evolutionary index is obtained by a simple integration by parts:

$$\sum\nolimits_r \int_{y_r^*}^{y_r} \phi_r(y)\,dy = \sum\nolimits_r (y_r - y_r^*)\phi_r(y_r) - \sum\nolimits_r \int_{y_r^*}^{y_r} y\phi_r'(y)\,dy.$$

Since the latencies ϕ_r are increasing, this expression immediately yields the estimate (14). In the special case $q \in \mathrm{Int}(\Delta)$ we can actually refine this bound to a *quadratic* one:

Lemma 11. *Let $q \in \mathrm{Int}(\Delta)$ be an interior Wardrop equilibrium and let $z \in T_q\Delta$. Then, for all $\theta \geq 0$ such that $q + \theta z \in \Delta$, we have:*

$L_q(q+\theta z) \geq \frac{1}{2} m \|P(z)\|^2 \theta^2$, *where* $m = \inf\{\phi'_r(y_r) : r \in \mathcal{E}, y \in P(\Delta)\}$.

Proof. Let $f(\theta) = F(q + \theta z)$. A simple differentiation then yields $f'(0) = \sum_i \sum_\alpha^i z_{i\alpha}\omega_{i\alpha}(q) = \sum_i \sum_\alpha^i z_{i\alpha}\omega_i(q) = 0$, the second equality following from the fact that q is an interior equilibrium, and the last one being a consequence of $z \in T_q\Delta$ (so that $\sum_\alpha^i z_{i\alpha} = 0$). On the other hand, we also have:

$$f''(\theta) = \frac{d^2}{d\theta^2} \sum\nolimits_r \Phi_r(y_r^* + \theta w_r) = \sum\nolimits_r w_r^2 \phi'_r(y_r^* + \theta w_r). \qquad (36)$$

Clearly, since the set $P(\Delta)$ of load profiles y is compact and the $\phi'_r(y_r) > 0$, we will also have $m = \inf\{\phi'_r(y_r) : y \in P(\Delta), r \in \mathcal{E}\} > 0$. This gives $f(\theta) \geq \frac{1}{2} m\theta^2$, and a first order Taylor expansion with Lagrange remainder easily completes the proof. □

B. DETERMINISTIC CONVERGENCE

In this appendix, our goal is to prove the deterministic convergence results of Section 3. We begin with:

Proof of Theorem 7. Let $\psi(x, t)$ be the *evolution function* of the dynamics (11) which describes the solution trajectory that starts at x at time $t = 0$. Clearly, ψ satisfies the consistency condition:

$$\psi(x, t + s) = \psi(\psi(x, t), s) \text{ for all } t, s \geq 0 \text{ and for all } x \in \Delta. \qquad (37)$$

So, fix the initial condition $x \in \mathrm{Int}(\Delta)$ and let $x(t) = \psi(x, t)$ be the corresponding solution orbit. If $q \in \Delta^*$ is Wardop, then we have seen that the function $V_q(t) \equiv H_q(\psi(x, t))$ is decreasing and will converge to some $m \geq 0$ as $t \to \infty$. It thus follows that $x(t)$ converges to the level set $H_q^{-1}(m)$.

Suppose now that there exists some increasing sequence of times $t_n \to \infty$ such that $x_n \equiv x(t_n)$ does not converge to Δ^*. By compactness of Δ, we may then assume that $x_n = \psi(x, t_n)$ converges to some $x^* \notin \Delta^*$ (but necessarily in $H_q^{-1}(m)$). Hence, for any $t > 0$:

$$H_q(\psi(x, t_n + t)) = H_q(\psi(\psi(x, t_n), t)) \to H_q(\psi(x^*, t)) < H_q(x^*) = m,$$

where the (strict) inequality stems from the fact that $\dot{H}_q(x) < 0$ outside Δ^*. On the other hand, $H_q(\psi(x, t_n + t)) = V_q(t_n + t) \to m$, a contradiction.

Since t_n was arbitrary, this shows that $x(t)$ converges to the set Δ^*. So, let q' be a limit point of $x(t)$ with $x(t'_n) \to q'$ for some sequence of times $t'_n \to \infty$. Then, $V_{q'}(t'_n) = H_{q'}(x(t'_n))$ will converge to zero and, with $V_{q'}$ decreasing, we will have $\lim_{t\to\infty} V_{q'}(t) = 0$ as well. Seeing as $H_{q'}$ only vanishes at q', we conclude that $x(t) \to q'$. □

Proof of Theorem 8. The basic idea of the proof is to establish the inequality $L_q(x(t)) \geq cH_q(x(t))$ for some $c > 0$; with $\dot{H}_q(x) = -L_q(x)$, our claim will then follow from Grönwall's lemma. Also, we will only present the case where $x(t)$ converges to an interior equilibrium $q \in \mathrm{Int}(\Delta)$ – the strict case is analogous (but easier) and thus omitted for space considerations.

We first show that $H_q(x(t)) \leq \frac{h_c}{\zeta_c^2}\|x(t)-q\|^2$ for some constant $h_c \geq h_0$ (to be determined), and for $\zeta_c^2 = \min_x\{\|x - q\|^2 : H_q(x) = h_c\}$. Though a bit tedious, it is not too hard to show that for any $a > 1$ and $z \in S_q$ (the "projective" unit sphere around q), the equation $H_q(q + \theta z) = \frac{a}{2}\left(\sum_{k,\alpha} z_{k\alpha}^2/q_{k\alpha}\right)\theta^2$ has a unique positive root $\theta_a \equiv \theta_a(z)$ such that $H_q(q + \theta z) \leq h_a(z)\theta^2/\theta_a(z)^2$ iff $\theta \leq \theta_a(z)$ (where, in obvious notation, $h_a(z) = H_q(q + \theta_a z)$. So, if $h_a = \max_{z\in S_q}\{h_a(z)\}$ and $h_c = \max\{h_0, h_a\}$, the equation $H_q(q + \theta z) = h_c$ will also have a unique positive root $\theta_c \equiv \theta_c(z)$ so that $H_q(q + \theta z) \leq h_c\theta^2/\theta_c^2$ iff $\theta \leq \theta_c$.

Now, write $x(t)$ in the projective form $x(t) = q + \theta(t)z(t)$; we then claim that $\theta(t) \leq \theta_c(z(t))$ for all $t \geq 0$. Indeed, if $\theta(t)$ ever exceeded $\theta_c(t)$, we would also have $H_q(x(t)) > h_c \geq h_0 = H_q(x(0))$, a contradiction since $\dot{H}_q \leq 0$. Thus, for all $t \geq 0$, we will have:

$$H_q(x(t)) \leq h_c \frac{\theta^2(t)}{\theta_c^2(z(t))} = h_c \frac{\|\theta(t)z(t)\|^2}{\|\theta_c(t)z(t)\|^2} \leq \frac{h_c}{\zeta_c^2}\|x(t) - q\|^2. \qquad (38)$$

Now, let $K = \ker Q$ be the space of degenerate directions of the network, and let $K^\perp$ be its orthogonal complement in the tangent space $Z \equiv T_q\Delta$. Then, if we decompose $z \in Z$ as $z = z_{/\!/} + z_\perp$, it can be shown that:

$$F(q + \theta z) - F(q) \geq \tfrac{1}{2} r\theta^2\|z\|_\perp^2 \qquad (39)$$

where r is the minimum of the (restricted) Rayleigh quotient $R_x(z) = \langle z, M_x z\rangle/\|z\|^2$ over $x \in \Delta$ and M_x is the Hessian of $F(x)$ restricted over $z \in K^\perp$. This shows that $L_q(x(t)) \geq \frac{1}{2} r\|x(t) - q\|_\perp^2$, and, noting that the replicator vector field is (a.s.) transversal to the equilibrial set Δ^*, we will have:

$$L_q(x(t)) \geq \frac{k}{2} r^2\|x(t) - q\|^2 \qquad (40)$$

for some $k \leq 1$ (which is equal to 1 if the network is irreducible). The theorem then follows by combining (38) and (40). □

C. STOCHASTIC CONSIDERATIONS

This appendix is devoted to the calculations that are hidden under the hood of Section 4. We begin with:

Proof of Lemma 9. Let $V_q(t) = H_q(X(t); \lambda)$. We then have:

$$\begin{aligned} dV_q &= \sum\nolimits_{i,\alpha} \frac{\partial H_q}{\partial x_{i\alpha}} dX_{i\alpha} + \frac{1}{2}\sum\nolimits_{i,\alpha}\sum\nolimits_{j,\beta} \frac{\partial^2 H_q}{\partial x_{i\alpha}\partial x_{j\beta}} (dX_{i\alpha})\cdot\left(dX_{j\beta}\right) \\ &= -\sum\nolimits_{i,\alpha} \frac{1}{\lambda_i}\frac{q_{i\alpha}}{x_{i\alpha}} dX_{i\alpha} + \frac{1}{2}\sum\nolimits_{i,\alpha} \frac{1}{\lambda_i}\frac{q_{i\alpha}}{x_{i\alpha}^2} (dX_{i\alpha})^2 . \end{aligned} \qquad (41)$$

However, with $X(t)$ being as in (19), we readily obtain:

$$\begin{aligned} (dX_{i\alpha})^2 &= \lambda_i^2 X_{i\alpha}^2 \left(dU_{i\alpha} - \rho_i^{-1}\sum\nolimits_\beta^i X_{i\beta}\, dU_{i\beta}\right)^2 \\ &= \lambda_i^2 X_{i\alpha}^2 \left[\sigma_{i\alpha}^2 - \frac{2}{\rho_i}\sum\nolimits_\beta^i \sigma_{\alpha\beta}^2 X_{i\beta} + \frac{1}{\rho_i^2}\sum\nolimits_{\beta,\gamma}^i \sigma_{\beta\gamma}^2 X_{i\beta}X_{i\gamma}\right] dt. \end{aligned} \qquad (42)$$

We may thus combine the two equations (41) and (42) into:

$$\begin{aligned} dV_q = &-\sum\nolimits_{i,\alpha} q_{i\alpha}\left[\omega_i(X) - \omega_{i\alpha}(X)\right] dt \\ &-\sum\nolimits_{i,\alpha} q_{i\alpha}\left[dU_{i\alpha} - \rho_i^{-1}\sum\nolimits_\beta^i X_{i\beta}\, dU_{i\beta}\right] \\ &+\frac{1}{2}\sum\nolimits_{i,\alpha} \lambda_i q_{i\alpha}\left[\sigma_{i\alpha}^2 - \frac{2}{\rho_i}\sum\nolimits_\beta^i \sigma_{\alpha\beta}^2 X_{i\beta} + \frac{1}{\rho_i^2}\sum\nolimits_{\beta,\gamma}^i \sigma_{\beta\gamma}^2 X_{i\beta}X_{i\gamma}\right] dt. \end{aligned} \qquad (43)$$

So, if we focus at a particular $i \in \mathcal{N}$, the last term of (43) gives:

$$\begin{aligned} &\sum\nolimits_\alpha^i q_{i\alpha}\left[\sigma_{i\alpha}^2 - \frac{2}{\rho_i}\sum\nolimits_\beta^i \sigma_{\alpha\beta}^2 X_{i\beta} + \frac{1}{\rho_i^2}\sum\nolimits_{\beta,\gamma}^i \sigma_{\beta\gamma}^2 X_{i\beta}X_{i\gamma}\right] \\ &= \sum\nolimits_\alpha^i q_{i\alpha}\sigma_{i\alpha}^2 - \frac{1}{\rho_i}\sum\nolimits_{\beta,\gamma}^i q_{i\beta}q_{i\gamma}\sigma_{\beta\gamma}^2 \\ &+ \frac{1}{\rho_i}\left[\sum\nolimits_{\beta,\gamma}^i q_{i\beta}q_{i\gamma}\sigma_{\beta\gamma}^2 - 2\sum\nolimits_{\beta,\gamma}^i X_{i\beta}q_{i\gamma}\sigma_{\beta\gamma}^2 + \sum\nolimits_{\beta,\gamma}^i X_{i\beta}X_{i\gamma}\sigma_{\beta\gamma}^2\right] \\ &= \frac{1}{\rho_i}\left[\sum\nolimits_{\beta,\gamma}^i q_{i\beta}(\rho_i\delta_{\beta\gamma} - q_{i\gamma})\sigma_{\beta\gamma}^2 + \sum\nolimits_{\beta,\gamma}^i \sigma_{\beta\gamma}^2(X_{i\beta} - q_{i\beta})(X_{i\gamma} - q_{i\gamma})\right] \end{aligned}$$

and the lemma follows by substituting the last equation into (43) and keeping only the resulting drift. □

Proof of Theorem 10. As we mentioned before, any $x \in \Delta$ may be (uniquely) written in the "projective" form $x = q + \theta z$, where $\theta = \Theta_q(x) \in [0, 1]$ is the projective distance of x from q, and z is a point in the "projective sphere" $S_q = \{z' \in T_q\Delta : q + z' \in \mathrm{bd}(\Delta)\}$. In this manner, (28) becomes:

$$-\mathcal{L}H_q(x;\lambda) = L_q(q+\theta z) - \frac{1}{2}\sum\nolimits_i \frac{\lambda_i}{\rho_i}\theta^2 \sum\nolimits_{\beta,\gamma}^i \sigma_{\beta\gamma}^2 z_{i\beta} z_{i\gamma} - \frac{1}{2}\sum\nolimits_i \frac{\lambda_i}{\rho_i}\sum\nolimits_{\beta,\gamma}^i \sigma_{\beta\gamma}^2 q_{i\beta}(\rho_i\delta_{\beta\gamma} - q_{i\gamma}). \tag{44}$$

With regards to the first term of (44), Lemma 11 and the definition (31) of κ yield $L_q(q+\theta z) \geq \frac{1}{2}m\|P(z)\|^2\theta^2 \geq \frac{1}{2}m\kappa^2\rho^2\theta^2$. Moreover, the second term of (44) is bounded above:

$$\frac{1}{2}\sum\nolimits_i \frac{\lambda_i}{\rho_i}\theta^2 \sum\nolimits_{\beta,\gamma}^i \sigma_{\beta\gamma}^2 z_{i\beta} z_{i\gamma} \leq \frac{1}{2}\rho\lambda\sigma^2\theta^2. \tag{45}$$

We are thus left to estimate the last term of (44). To that end:

$$\sum\nolimits_{\beta,\gamma}^i q_{i\beta}\left(\rho_i\delta_{\beta\gamma} - q_{i\gamma}\right)\sigma_{\beta\gamma}^2 = \sum\nolimits_r \sigma_r^2 y_{ir}(\rho_i - y_{ir}) \leq \frac{1}{4}\rho_i^2\sigma^2, \tag{46}$$

where y_{ir} is given by (1) and the last inequality stems from the bound $y_{ir}(\rho_i - y_{ir}) \leq \frac{1}{4}\rho_i^2$. Combining all of the above, we then get:

$$-\mathcal{L}H_q(x;\lambda) \geq \tfrac{1}{2}m\kappa^2\rho^2\theta^2 - \tfrac{1}{2}\rho\lambda\sigma^2\theta^2 - \tfrac{1}{8}\rho\lambda\sigma^2 \equiv g(\theta). \tag{47}$$

As a result, if $\lambda < \frac{4}{5}\lambda_0$ where $\lambda_0 = \frac{m\rho\kappa^2}{\sigma^2}$, it is easy to see that the RHS of (47) will be increasing for $0 \leq \theta \leq 1$. Moreover, it will also be positive for $\theta_\lambda < \theta \leq 1$, where θ_λ is the positive root of g (that is, $\theta_\lambda = \frac{1}{2}(\lambda_0/\lambda - 1)^{-1/2}$). So, pick some positive $a < g(1) = \frac{1}{2}\rho\sigma^2\left(\lambda_0 - \frac{5}{4}\lambda\right)$ and consider the set $K_a = \{q+\theta z : z \in S_q, g(\theta) \leq a\}$. By construction, K_a is a compact neighborhood of q which does not intersect $\mathrm{bd}(\Delta)$ and, by (47), we have $\mathcal{L}H_q(x;\lambda) \leq -a$ outside K_a. Therefore, if $\tau_a \equiv \inf\{t : X(t) \in K_a\}$, Theorem 5.3 in [27] yields:

$$\mathbf{E}_x[\tau_a] \leq \frac{H_q(x;\lambda)}{a} < \infty \tag{48}$$

for every interior initial condition $X(0) = x \in \mathrm{Int}(\Delta)$ – this is also the source of the estimate (34).

Inspired by a trick of [15], let us consider the transformed process $Y(t) = \Psi(X(t))$ where $\Psi_{i\mu}(x) = \log x_{i\mu}/x_{i,0}, \mu \in \mathcal{A}_i^* \equiv \mathcal{A}_i \setminus\{\alpha_{i,0}\}$. With $\frac{\partial\Psi_{i\mu}}{\partial x_{i\mu}} = 1/x_{i\mu}$ and $\frac{\partial\Psi_{i\mu}}{\partial x_{i,0}} = -1/x_{i,0}$, Itô's formula gives:

$$dY_{i\mu} = \mathcal{L}\Psi_{i\mu}(X(t))\,dt + dU_{i\mu} - dU_{i,0} = \mathcal{L}\Psi_{i\mu}(X(t))\,dt + \sum\nolimits_r Q_{r\mu}^i\sigma_r\,dW_r$$

where $Q_{r\mu}^i = P_{r\mu}^i - P_{r,0}^i$ are the components of the degeneracy matrix of $\mathcal{Q}$ in the basis $\tilde{e}_{i\mu} = e_{i\mu} - e_{i,0}$ of $T_q\Delta$ – see also Section 2.1.

We now claim that the generator of Y is elliptic. Indeed, if we drop the user index i for convenience and set $A_{\mu r} = Q_{r\mu}\sigma_r$, $\mu \in \coprod_i \mathcal{A}_i^*$, it suffices to show that the matrix AA^T is positive-definite. Sure enough, for any tangent vector $z = \sum_\mu z_\mu\tilde{e}_\mu \in T_q\Delta$, we get:

$$\langle Az, Az\rangle = \sum\nolimits_{\mu,\nu}\sum\nolimits_r Q_{r\mu}Q_{r\nu}\sigma_r^2 z_\mu z_\nu = \sum\nolimits_r \sigma_r^2 w_r^2, \tag{49}$$

where $w = Q(z)$. Since $\mathcal{Q}$ is irreducible, we will have $w \neq 0$, and in view of (49) above, this proves our assertion.

We have thus shown that the process $Y(t)$ hits a compact neighborhood of $\Psi(q)$ in finite time (on average), and also that the generator of Y is elliptic. From Lemma 3.4 in [28], it then follows that Y is recurrent, and since Ψ is invertible in $\mathrm{Int}(\Delta)$, the same must hold for $X(t)$ as well. In a similar fashion, these criteria also ensure that the transition probabilities of the diffusion $X(t)$ converge in total variation to an invariant probability measure π on Δ, thus proving the first part of our theorem.

To obtain the time-average estimate of the theorem, note that Dynkin's formula [26, Theorem 7.4.1] applied to (47) yields:

$$\begin{aligned}\mathbf{E}_x\left[H_q(X(t);\lambda)\right] &= H_q(x;\lambda) + \mathbf{E}_x\left[\int_0^t \mathcal{L}H_q(X(s);\lambda)\,ds\right]\\ &\leq H_q(x;\lambda) - \tfrac{1}{2}\rho\sigma^2(\lambda_0 - \lambda)\,\mathbf{E}_x\left[\int_0^t \Theta_q^2(X(s))\,ds\right] + \tfrac{1}{8}\rho\lambda\sigma^2 t,\end{aligned}$$

and with $\mathbf{E}_x[H_q(X(t);\lambda)] \geq 0$, we easily get:

$$\mathbf{E}_x\left[\frac{1}{t}\int_0^t \Theta_q^2(X(s))\,ds\right] \leq \theta_\lambda^2 + \frac{C}{t}, \text{ where } C = \frac{2}{\rho\sigma^2}\frac{H_q(x;\lambda)}{\lambda_0 - \lambda}. \tag{50}$$

We are thus left to establish the bound $\pi(B_\theta) \geq 1 - \theta_\lambda^2/\theta^2$ which shows that the mass of the invariant measure π is concentrated in the "projective balls" B_θ. For that, we will use the ergodic property:

$$\pi(B_\theta) = \lim_{t\to\infty}\mathbf{E}_x\left[\frac{1}{t}\int_0^t \chi_{B_\theta}(X(s))\,ds\right], \tag{51}$$

where χ_{B_θ} is the indicator function of B_θ. However, with $\Theta_q^2(x)/\theta^2 \geq 1$ outside B_θ by definition, it easily follows that:

$$\mathbf{E}_x\left[\frac{1}{t}\int_0^t \chi_{B_\theta}(X(s))\,ds\right] \geq \mathbf{E}_x\left[\frac{1}{t}\int_0^t \left(1 - \Theta_q^2(X(s))\big/\theta^2\right)ds\right] \tag{52}$$

and the bound (33) follows by letting $t \to \infty$ in (50). □

References

[1] E. Altman, T. Boulogne, R. el Azouzi, T. Jiménez, and L. Wynter, "A survey on networking games in telecommunications," *Computers and Operations Research*, vol. 33, no. 2, pp. 286–311, 2006.

[2] A. Orda, N. Rom, and N. Shimkin, "Competitive routing in multi-user communication networks," *IEEE/ACM Trans. Netw.*, vol. 1, pp. 614–27, 1993.

[3] E. Koutsoupias and C. Papadimitriou, "Worst-case equilibria," in *Proceedings of the 16th Annual Symposium on Theoretical Aspects of Computer Science*, 1999, pp. 404–413.

[4] T. Roughgarden and E. Tardos, "How bad is selfish routing?" *Journal of the ACM*, vol. 49, no. 2, pp. 236–259, 2002.

[5] ——, "Bounding the inefficiency of equilibria in nonatomic congestion games," *Games and Economic Behavior*, vol. 47, no. 2, pp. 389–403, May 2004.

[6] J. G. Wardrop, "Some theoretical aspects of road traffic research," in *Proceedings of the Institute of Civil Engineers, Part II*, vol. 1, 1952, pp. 325–78.

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

[8] S. C. Dafermos and F. T. Sparrow, "The traffic assignment problem for a general network," *Journal of Research of the National Bureau of Standards*, vol. 73B, no. 2, pp. 91–118, 1969.

[9] W. H. Sandholm, "Potential games with continuous player sets," *Journal of Economic Theory*, vol. 97, pp. 81–108, 2001.

[10] S. Fischer and B. Vöcking, "On the evolution of selfish routing," in *Proceedings of the 12th European Symposium on Algorithms*, 2004.

[11] E. Altman, V. Kamble, and V. Borkar, "Convergence of population dynamics in symmetric routing games with a finite number of players," in *GameNets '09: Proceedings of the 1st International Conference on Game Theory for Networks*, 2009.

[12] P. D. Taylor and L. B. Jonker, "Evolutionary stable strategies and game dynamics," *Mathematical Biosciences*, vol. 40, no. 1-2, pp. 145–156, 1978.

[13] D. Fudenberg and C. Harris, "Evolutionary dynamics with aggregate shocks," *Journal of Economic Theory*, vol. 57, no. 2, pp. 420–441, August 1992.

[14] A. Cabrales, "Stochastic replicator dynamics," *International Economic Review*, vol. 41, no. 2, pp. 451–81, May 2000.

[15] L. A. Imhof, "The long-run behavior of the stochastic replicator dynamics," *The Annals of Applied Probability*, vol. 15, no. 1B, pp. 1019–1045, 2005.

[16] J. Hofbauer and L. A. Imhof, "Time averages, recurrence and transience in the stochastic replicator dynamics," *The Annals of Applied Probability*, vol. 19, no. 4, pp. 1347–1368, 2009.

[17] P. Mertikopoulos and A. L. Moustakas, "Learning in the presence of noise," in *GameNets '09: Proceedings of the 1st International Conference on Game Theory for Networks*, 2009.

[18] ——, "The emergence of rational behavior in the presence of stochastic perturbations," *The Annals of Applied Probability*, vol. 20, no. 4, pp. 1359–1388, 2010.

[19] R. W. Rosenthal, "A class of games possessing pure-strategy nash equilibria," *International Journal of Game Theory*, vol. 2, pp. 65–67, 1973.

[20] D. Monderer and L. S. Shapley, "Potential games," *Games and Economic Behavior*, vol. 14, no. 1, pp. 124 – 143, 1996.

[21] T. Börgers and R. Sarin, "Learning through reinforcement and replicator dynamics," *Journal of Economic Theory*, vol. 77, pp. 1–14, 1997.

[22] P. D. Taylor, "Evolutionarily stable strategies with two types of player," *Journal of Applied Probability*, vol. 16, pp. 76–83, 1979.

[23] J. M. Swinkels, "Evolutionary stability with equilibrium entrants," *Journal of Economic Theory*, vol. 57, no. 2, pp. 306–332, August 1992.

[24] R. Cressman, *The stability concept of evolutionary game theory*. Berlin: Springer-Verlag, 1992.

[25] D. Fudenberg and D. K. Levine, *The Theory of Learning in Games*, ser. MIT Press Series on Economic Learning and Social Evolution. The MIT Press, 1998, vol. 2.

[26] B. Øksendal, *Stochastic Differential Equations*, 6th ed. Berlin: Springer-Verlag, 2007.

[27] R. Durrett, *Stochastic Calculus*. Boca Raton, FL: CRC Press, 1996.

[28] R. N. Bhattacharya, "Criteria for recurrence and existence of invariant measures for multidimensional diffusions," *The Annals of Probability*, vol. 6, pp. 541–553, 1978.