



ÉCOLE POLYTECHNIQUE

CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE



**BALANCING TRAFFIC IN NETWORKS: REDUNDANCY, LEARNING
AND THE EFFECT OF STOCHASTIC FLUCTUATIONS**

Panayotis MERTIKOPOULOS
Aris L. MOUSTAKAS

March 2011

Cahier n° 2011-06

DEPARTEMENT D'ECONOMIE

Route de Saclay

91128 PALAISEAU CEDEX

(33) 1 69333033

<http://www.enseignement.polytechnique.fr/economie/>
<mailto:chantal.poujouly@polytechnique.edu>

BALANCING TRAFFIC IN NETWORKS: REDUNDANCY, LEARNING AND THE EFFECT OF STOCHASTIC FLUCTUATIONS

BY PANAYOTIS MERTIKOPOULOS¹ AND ARIS L. MOUSTAKAS²

¹*École Polytechnique*

²*National & Kapodistrian University of Athens*

We study the distribution of traffic in networks whose users try to minimise their delays by adhering to a simple learning scheme inspired by the replicator dynamics of evolutionary game theory. The stable steady states of these dynamics coincide with the network's *Wardrop equilibria* and form a convex polytope whose dimension is determined by the network's *redundancy* (an important concept which measures the “linear dependence” of the users' paths). Despite this abundance of stationary points, the long-term behaviour of the replicator dynamics turns out to be remarkably simple: every solution orbit converges to a Wardrop equilibrium.

On the other hand, a major challenge occurs when the users' delays fluctuate unpredictably due to random external factors. In that case, interior equilibria are no longer stationary, but *strict* equilibria remain stochastically stable irrespective of the fluctuations' magnitude. In fact, if the network has no redundancy and the users are patient enough, we show that the long-term average of the users' traffic flows converges to the vicinity of an equilibrium, and we also estimate the corresponding invariant measure.

1. Introduction. The underlying problem of managing the flow of traffic in a large-scale network is as simple to state as it is challenging to resolve: given the rates of traffic generated by the users of the network, one is asked to identify and realise the most “satisfactory” distribution of traffic among the network's routes.

Of course, given that this notion of “satisfaction” depends on the users' optimisation criteria, it would serve well to keep a concrete example in mind. Perhaps the most illustrative one is that of the Internet itself, where the primary concern of its users is to minimise the travel times of their data flows. However, since the time needed to traverse a link in the network increases (nonlinearly even) as the link becomes more congested, the users' concurrent minimisation efforts invariably lead to game-like interactions whose complexity precludes even the most rudimentary attempts at coordination. In this way, a traffic distribution will be considered “satisfactory” by a user when there is no unilateral move that he could make in order to further decrease the delays (or *latencies*) that he experiences.

This Nash-type condition is aptly captured by *Wardrop's principle* (Wardrop, 1952): given the level of congestion caused by other users, every user seeks to employ the minimum-latency path available to him. As might be expected, this principle has attracted a great deal of interest and it was shown early on that these *Wardrop equilibria* can be calculated by solving a convex optimisation problem (Beckmann, McGuire and Winsten, 1956; Dafermos and Sparrow, 1969). Among others, this characterisation enabled Roughgarden and Tardos (2002, 2004) to quantify the efficiency of these equilibrial states by estimating their “price of anarchy”, i.e. the ratio

between the aggregate delay of a flow at Wardrop equilibrium and the minimum achievable (aggregate) latency (Koutsoupias and Papadimitriou, 1999).

Still, the size of large-scale networks makes computing these equilibria a task of considerable 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). It is thus more reasonable to take a less centralised approach and instead ask: *is there a simple learning procedure which leads users to Wardrop equilibrium?*

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 that the work of Sandholm (2001) showed that a good candidate for such a learning scheme would be the *replicator dynamics* of evolutionary game theory, a dynamical system that was first introduced by Taylor and Jonker (1978) to model the evolution of (nonatomic) populations that interact with one another by means of random matchings in a Nash game. More precisely, these dynamics arise as the byproduct of an “imitation of the fittest” process which drives the per capita growth rate of a genotype (strategy) proportionately to the difference between the reproductive fitness (payoff) of the genotype itself and the population average. Thus, owing to this correlation between growth rates and payoffs, the game's Nash equilibria emerge as ω -limit points of the replicator trajectories – see also the excellent surveys by Weibull (1995) and by Hofbauer and Sigmund (1998, 2003).

In our congestion setting, these populations correspond to the users' traffic flows, so the convex optimisation formulation of Beckmann, McGuire and Winsten allows us to recast our problem in terms of a (nonatomic) *potential game* (Sandholm, 2001). Indeed, Wardrop equilibria can be located by looking at the minimum of the *Rosenthal potential* (Rosenthal, 1973) and, hence, Sandholm's analysis shows that they are Lyapunov stable rest points of the replicator dynamics. This fact was also recognized independently by Fischer and Vöcking (2004) who additionally showed that the (interior) solution orbits of the replicator dynamics converge to the set of Wardrop equilibria – actually, the authors suggest that these orbits converge to a point, but their analysis only holds when there is a unique equilibrium.

Rather surprisingly, when there is not a unique equilibrium, the structure of the Wardrop set itself seems to have been overlooked in the above considerations. Specifically, it has been widely assumed that if the network's delay functions are strictly increasing, then there exists a unique Wardrop equilibrium (for instance, see Sandholm, 2001, Corollary 5.6). As a matter of fact, this uniqueness property is only true in *irreducible* networks, i.e. networks whose paths are “independent” of one another (in a sense made precise by Definition 2.1). In general, the Wardrop set of a network is a convex polytope whose dimension is determined by the network's *redundancy*, a notion which quantifies precisely this “linear dependence”. Nonetheless, despite this added structure, we show that the expectations of Fischer and Vöcking are vindicated in that the long-term behaviour of the replicator dynamics remains disarmingly simple: (almost) every replicator orbit converges to a Wardrop flow and not merely to the *set* of such flows (Theorem 3.2).

Having said that, the imitation procedure inherent in the replicator dynamics implicitly presumes itself 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 gathered under the collective moniker “nature”). In population biology, these disturbances are usually modelled by introducing “aggregate shocks” to the replicator dynamics (Fudenberg and Harris, 1992) and, as one would expect, these shocks complicate the situation considerably. For instance, Cabrales (2000) proved that dominated strategies become extinct in the long run, but only if the variance of the shocks is mild enough compared to the payoffs of the game. More recently, Imhof (2005) showed that even equilibrial play arises over time but, again, conditionally on the noise processes not being too loud (see also Benaïm, Hofbauer and Sandholm, 2008; Hofbauer and Imhof, 2009). On the other hand, if one interprets the replicator dynamics as the derivative of an *exponential learning* procedure and perturbs them accordingly (i.e. not as an evolutionary birth-death process), it was shown that similar rationality properties continue to hold, no matter how loud the noise becomes (Mertikopoulos and Moustakas, 2009, 2010).

All the same, these approaches have chiefly focused on Nash-type games where payoffs are multilinear functions over a product of simplices; for example, payoffs in single-population evolutionary games are determined by the bilinear form which is associated to the matrix of the game. This linear structure simplifies things considerably but, unfortunately, congestion models rarely adhere to it; additionally, the notions of Nash and Wardrop equilibrium are at variance in many occasions, a disparity which also calls for a different approach; and, finally, the way that stochastic fluctuations propagate to the users’ choices in a network leads to a new stochastic version 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 cannot be understood by simply translating previous work on the stochastic replicator dynamics.

1.1. *Outline.* In this paper, we study the distribution of traffic in networks whose links are subject to constant stochastic perturbations that randomly affect the delays experienced by individual traffic elements. This model is presented in detail in Section 2, where we also develop our game-theoretic machinery: specifically, we introduce the notion of a network’s *redundancy* in Section 2.2, and we examine its connection to Wardrop equilibria in Section 2.3. We then derive the rationality properties of the deterministic replicator dynamics in Section 3, where we show that (almost) every solution trajectory converges to a Wardrop equilibrium.

Section 4 is devoted to the stochastic considerations which constitute the core of our paper. Our first result is that *strict* Wardrop equilibria remain stochastically asymptotically stable *irrespective* of the fluctuations’ magnitude (Theorem 4.3); in fact, if the users are “patient enough”, we are able to estimate the average time it takes them to hit a neighbourhood of the equilibrium in question (Theorem 4.4). In conjunction with stochastic stability, this allows us to conclude that when a strict equilibrium exists, users converge to it almost surely (Corollary 4.5). On the other

hand, given that such equilibria do not always exist, we also prove that the replicator dynamics in irreducible networks are recurrent (again under the assumption that the users are patient enough), and we use this fact to show that the long-term average of their traffic distributions concentrates mass in the neighbourhood of an interior Wardrop equilibrium (Theorem 4.6).

1.2. Notational Conventions. If $\mathcal{S} = \{s_\alpha\}_{\alpha=0}^n$ is a finite set, the vector space spanned by \mathcal{S} over \mathbb{R} is defined to be the set of all formal linear combinations of elements of \mathcal{S} with real coefficients, i.e. the set of all functions $x : \mathcal{S} \rightarrow \mathbb{R}$. In tune with standard set-theoretic notation, we will denote this space by $\mathbb{R}^{\mathcal{S}} \equiv \text{Maps}(\mathcal{S}, \mathbb{R})$. In this way, $\mathbb{R}^{\mathcal{S}}$ admits a canonical basis $\{e_\alpha\}_{\alpha=0}^n$ consisting of the indicator functions $e_\alpha : \mathcal{S} \rightarrow \mathbb{R}$ which take the value $e_\alpha(s_\alpha) = 1$ on s_α and vanish otherwise; in particular, if $x \in \mathbb{R}^{\mathcal{S}}$ has $x(s_\alpha) = x_\alpha$, we will have $x = \sum_\alpha x_\alpha e_\alpha$. Hence, under the natural identification $s_\alpha \mapsto e_\alpha$, we will make no distinction between the elements s_α of \mathcal{S} and the corresponding basis vectors e_α of $\mathbb{R}^{\mathcal{S}}$ – in fact, to avoid drowning in a morass of indices, we will routinely use α to refer interchangeably to either s_α or e_α , writing e.g. “ $\alpha \in \mathcal{S}$ ” instead of “ $s_\alpha \in \mathcal{S}$ ”. 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 follow the original convention of Nash (1951) and employ Latin indices (i, j, \dots) for players while reserving Greek ones (α, β, \dots) for their (pure) strategies; also, to differentiate between strategies, we will use α, β, \dots for indices that start at 0 and μ, ν, \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* $\mathcal{A} \equiv \coprod_i \mathcal{A}_i = \bigcup_i \{(\alpha, i) : \alpha \in \mathcal{A}_i\}$ by mapping $\alpha \in \mathcal{A}_i \mapsto (\alpha, i) \in \mathcal{A}$. Hence, 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 $x = \sum_\alpha x_\alpha e_\alpha \in \mathbb{R}^{\mathcal{A}}$ instead of $x = \sum_{i,\alpha} x_{i\alpha} e_{i\alpha} \in \prod_i \mathbb{R}^{\mathcal{A}_i}$. 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 $\sum_\alpha^i \equiv \sum_{\alpha \in \mathcal{A}_i}$.

Finally, if $X(t)$ is some stochastic process in \mathbb{R}^n starting at $X(0) = x$ and there is no doubt that we are referring to the process X , its law will be denoted by \mathbf{P}_x . In that case, we will also employ the term “almost surely” instead of the somewhat unwieldy “ \mathbf{P}_x -almost surely”.

2. Preliminaries.

2.1. Games in Normal Form. Our starting point for the definition of a game in normal form will be a set of *players* \mathcal{N} , together with a finite measure ν on \mathcal{N} which “accounts” for all players $i \in \mathcal{N}$ (in the sense that the singletons $\{i\} \subseteq \mathcal{N}$ are all ν -measurable).

The players’ possible actions in the game will then be represented by their *strategy sets* Δ_i , $i \in \mathcal{N}$. For our purposes, we will assume that these sets are locally compact Hausdorff spaces and that the relative topologies induced on $\Delta_i \cap \Delta_j$ agree for all $i, j \in \mathcal{N}$. Thanks to this compatibility condition, $\Delta_0 \equiv \bigcup_j \Delta_j$ inherits a natural Borel structure arising from the union topology (the finest topology in

which the inclusions $\Delta_i \hookrightarrow \Delta_0$ are continuous) and, in this way, an admissible *strategy profile* $x \in \prod_i \Delta_i$ will just be a measurable function $x : \mathcal{N} \rightarrow \Delta_0$ which maps $i \mapsto x_i \in \Delta_i$ for all players $i \in \mathcal{N}$. For technical reasons, we will also require that the push-forward measure $x_*\nu$ induced on Δ_0 by x (given by $x_*\nu(U) = \nu(x^{-1}(U))$) for any Borel $U \subseteq \Delta_0$) be inner regular, and, hence, Radon (since ν is finite).

As is customary, we will identify two profiles which agree ν -almost everywhere, except when we need to focus on the strategy of a particular player $i \in \mathcal{N}$ against that of his *opponents* $\mathcal{N}_{-i} \equiv \mathcal{N} \setminus \{i\}$; in that case, we will use the shorthand $(x_{-i}; q_i)$ to denote the profile which agrees with x on \mathcal{N}_{-i} (ν -a.e.) and maps $i \mapsto q_i \in \Delta_i$. The set Δ of all such profiles $x \in \prod_i \Delta_i$ will then be referred to as the *strategy space* of the game and is itself a Borel space because it inherits the subspace topology from the product $\prod_i \Delta_i$.

Bearing all this in mind, the fitness of the players' strategic choices will be determined by their *payoff functions* (or *utilities*) $u_i : \Delta \rightarrow \mathbb{R}$, $i \in \mathcal{N}$; in particular, $u_i(x) \equiv u_i(x_{-i}; x_i)$ will simply represent the reward that player $i \in \mathcal{N}$ receives in the strategy profile $x \equiv (x_{-i}; x_i) \in \Delta$, i.e. when he plays $x_i \in \Delta_i$ against his opponents' strategy $x_{-i} \in \prod_{j \neq i} \Delta_j$. The only further assumptions that we will make is that these payoff functions be (Borel) measurable and that $u_i(x_{-i}; x_i) = u_i(x'_{-i}; x_i)$ whenever x and x' agree ν -a.e. on \mathcal{N}_{-i} .

This collection of *players* $i \in \mathcal{N}$, their *strategy sets* Δ_i , and their *payoff functions* $u_i : \Delta \rightarrow \mathbb{R}$ will be our working definition for a *game in normal form*, usually denoted by $\mathfrak{G} \equiv \mathfrak{G}(\mathcal{N}, \Delta, u)$. Additionally, if the payoff functions $u_i : \Delta \rightarrow \mathbb{R}$ happen to be continuous, the game \mathfrak{G} will be called *continuous* as well.

Needless to say, this abstract definition might appear somewhat opaque, so we will immediately proceed with a few important examples to clarify the concept.

2.1.1. *N*-person Games. As the name suggests, the players here are indexed by the finite set $\mathcal{N} = \{1, 2, \dots, N\}$ (endowed with the usual counting measure) and the game's strategy space will be the finite product $\Delta \equiv \prod_i \Delta_i$ (thus doing away with some of the technical subtleties present in the more general definition).

This point is where we recover the original scenario of Nash (1951). To see how, assume that every player $i \in \mathcal{N}$ comes with a finite set \mathcal{A}_i of *actions* (or *pure strategies*) which can be "mixed" according to some probability distribution $x_i \in \Delta(\mathcal{A}_i)$. In this interpretation, the players' strategy sets are just the simplices $\Delta_i \equiv \Delta(\mathcal{A}_i)$ and their payoff functions $u_i : \Delta \equiv \prod_i \Delta_i \rightarrow \mathbb{R}$ are given by the multilinear expectations:

$$(2.1) \quad u_i(x) = u_i(x_1, \dots, x_N) = \sum_{\alpha_1 \in \mathcal{A}_1} \cdots \sum_{\alpha_N \in \mathcal{A}_N} x_{1, \alpha_1} \cdots x_{N, \alpha_N} u_{i, \alpha_1 \dots \alpha_N},$$

where $x_i = \sum_{\alpha} x_{i\alpha} e_{i\alpha}$ in the standard basis $\{e_{i\alpha}\}$ of $\mathbb{R}^{\mathcal{A}_i}$ and $u_{i, \alpha_1 \dots \alpha_N}$ is the reward that player i would obtain by choosing $\alpha_i \in \mathcal{A}_i$ against his opponents' action $\alpha_{-i} \in \mathcal{A}_{-i} \equiv \prod_{j \neq i} \mathcal{A}_j$. Because of this (multi)linear structure, we will commonly refer to Nash-type games as *linear games* to contrast them with more general *N*-person games where payoffs and strategy sets might fail to have any sort of linear structure – as is the case for example with *concave games* (Rosen, 1965).

2.1.2. *Population Games.* The cornerstone of evolutionary game theory concerns games played by an uncountable number of players – for instance, see Schmeidler (1973). As such, these nonatomic *population games* require the full breadth afforded by our more abstract definition.

The first piece of additional structure encountered in these games is a measurable partition $\mathcal{N} = \bigcup_{r=1}^N \mathcal{N}_r$ of the player set \mathcal{N} into N disjoint *populations* (or *classes*) $\mathcal{N}_r \subseteq \mathcal{N}$; accordingly, every player $i \in \mathcal{N}$ belongs to a unique class \mathcal{N}_r which we denote by $\text{class}(i)$. Each of these populations is then “measured” by the corresponding restriction ν_r of the measure ν on \mathcal{N}_r (i.e. $\nu_r(B) = \nu(B \cap \mathcal{N}_r)$ for any Borel $B \subseteq \mathcal{N}$), and the basic underlying assumption is that these measures are nonatomic.

The second fundamental assumption is that this classification of players also determines how they interact with their environment and with each other. More precisely, this means that the strategy sets of two players that belong to the same population coincide: $\Delta_i = \Delta_j$ whenever $\text{class}(i) = \text{class}(j)$. Because of this, we will write \mathcal{A}_r for the common strategy set of the r -th population and \mathcal{A}_0 for the corresponding union: $\mathcal{A}_0 = \bigcup_{r=1}^N \mathcal{A}_r = \bigcup_{i \in \mathcal{N}} \Delta_i$.

Now, every strategy profile $x : \mathcal{N} \rightarrow \mathcal{A}_0$ pushes forward a (Radon) measure \hat{x}_r on \mathcal{A}_r in the usual way:

$$(2.2) \quad \hat{x}_r(U) \equiv (x_* \nu_r)(U) = \nu_r(x^{-1}(U)) = \nu\{i \in \mathcal{N}_r : x_i \in U\}$$

for any Borel $U \subseteq \mathcal{A}_r$ – in other words, $\hat{x}_r(U)$ is just the measure of the players in the r -th population whose chosen strategy lies in $U \subseteq \mathcal{A}_r$. Then, the final (and perhaps most significant) requirement in population games is that the players’ payoffs depend only on the *strategy distribution* $\hat{x} = (\hat{x}_1, \dots, \hat{x}_N)$ and not on the players’ individual strategic choices. Specifically, if $P_0(\mathcal{A})$ denotes the space of all such strategy distributions equipped with the topology of vague convergence, we assume that there exist continuous functions $\hat{u}_r : P_0(\mathcal{A}) \times \mathcal{A}_r \rightarrow \mathbb{R}$, $r = 1, \dots, N$, such that:

$$(2.3) \quad u_i(x) = \hat{u}_r(\hat{x}; x_i) \text{ for all } i \in \mathcal{N}_r.$$

Consequently, as long as the overall strategy distribution \hat{x} stays the same, payoffs remain unaffected even by positive-mass migrations of players from one strategy to another (and not only by migrations of measure zero).

Again, it would serve well to illustrate this abstract definition by means of a more concrete example. To wit, in evolutionary game theory, populations are usually represented by the intervals $\mathcal{N}_r = [0, m_r]$ where $m_r > 0$ denotes the “mass” of the population under Lebesgue measure. The strategy spaces \mathcal{A}_r are typically assumed to be finite, so that a strategy distribution is simply a point in the (finite-dimensional) product of simplices $\prod_r m_r \Delta(\mathcal{A}_r)$. Hence, if player $i \in \mathcal{N}_r$ picks the strategy $\alpha \in \mathcal{A}_r$, his payoff will be given by

$$(2.4) \quad u_{r\alpha}(x) \equiv u_r(x; \alpha),$$

where, in a slight abuse of notation, we removed the hats from \hat{u}_r and \hat{x} in order to stress that they are the fundamental quantities that describe the game (it will

always be clear from the context whether we are referring to the distribution $\hat{x} \in P_0(\mathcal{A})$ or to the actual strategy profile $x : \mathcal{N} \rightarrow \mathcal{A}_0$.

This choice of notation is very suggestive for another reason as well: if we set $\Delta_r \equiv m_r \Delta(\mathcal{A}_r)$, then these simplices may be taken as the strategy sets of an associated N -person game whose players are indexed by $r = 1, 2, \dots, N$ (that is, they correspond to the populations themselves). The only thing needed to complete this description is to define the payoff functions $u_r : \Delta \equiv \prod_r \Delta_r \rightarrow \mathbb{R}$ in this picture, and a natural choice would be to take the *population averages*:

$$(2.5) \quad u_r(x) = \frac{1}{m_r} \sum_{\alpha} x_{r\alpha} u_{r\alpha}(x),$$

where $x_{r\alpha}$ are the coordinates of x in Δ . However, it is worth keeping in mind that, depending on the situation at hand, this need not be the only reasonable choice for a payoff function (we will explore this issue further in the next section).

Potential Games. An important subclass of population games arises when the payoffs $u_{r\alpha}$ satisfy the closedness condition:

$$(2.6) \quad \frac{\partial u_{r\alpha}}{\partial x_{s\beta}} = \frac{\partial u_{s\beta}}{\partial x_{r\alpha}} \text{ for all populations } r, s \text{ and for all strategies } \alpha \in \mathcal{A}_r, \beta \in \mathcal{A}_s.$$

This condition is commonly referred to as “externality symmetry” (Sandholm, 2001) and it describes games where a marginal increase in the population of players using strategy α has the same effect on the payoffs to players playing strategy β as the converse increase. Clearly, since the strategy distributions of these games live in the simply connected polytope $\Delta = \prod_r \Delta_r$, condition (2.6) amounts to the existence of a *potential function* $F : \Delta \rightarrow \mathbb{R}$ such that:

$$(2.7) \quad u_{r\alpha}(x) = -\frac{\partial F}{\partial x_{r\alpha}}.$$

Hence, if a player $i \in \mathcal{N}_r$ makes the switch $\alpha \rightarrow \beta$, his payoff will change by:

$$(2.8) \quad u_{r\beta}(x) - u_{r\alpha}(x) = -\left(\frac{\partial F}{\partial x_{r\beta}} - \frac{\partial F}{\partial x_{r\alpha}} \right) = -dF(e_{r\beta} - e_{r\alpha}),$$

where $\{e_{r\beta}\}$ denotes the standard basis of $\prod_r \mathbb{R}^{\mathcal{A}_r}$. In other words, the strategy migration $\alpha \rightarrow \beta$ is profitable to a player iff the direction $e_{r\beta} - e_{r\alpha}$ descends the potential F . This property of potential games will be extremely important for our purposes and its ramifications underlie a large part of our work.

2.1.3. Nash Equilibrium and Wardrop’s Principle. Under the umbrella of rationality, selfish players will seek to play those strategies which deliver the best rewards against the choices of their opponents. This leads to the celebrated notion of a *Nash equilibrium*, i.e. a strategy profile q which discourages unilateral deviations:

$$(NEQ) \quad u_i(q) \geq u_i(q_{-i}; q'_i) \text{ for almost every } i \in \mathcal{N} \text{ and all strategies } q'_i \in \Delta_i$$

(see also Schmeidler (1973) or Milchtaich (2000)).

The seminal result of Nash (1951) was that N -person linear games always possess equilibria of this kind. Rosen (1965) subsequently extended this result to the class of *concave* games (continuous concave payoffs over convex strategy sets), while Schmeidler (1973) essentially settled the issue for population games with finite strategy sets (see also Ali Khan, 1986).

In this last instance, Nash equilibria are aptly captured by *Wardrop's principle*:

$$(2.9) \quad u_{r\alpha}(q) \geq u_{r\beta}(q) \text{ for all } \alpha, \beta \in \mathcal{A}_r \text{ s.t. } q \text{ assigns positive mass to } \alpha.$$

To see this, note that if $\alpha \in \mathcal{A}_r$ has positive measure in the strategy distribution q , then there exists a player $i \in \mathcal{N}_r$ (actually a positive mass of such players) with $q_i = \alpha$ and such that (NEQ) holds. Hence, for every $\beta \in \mathcal{A}_r$, we immediately get:

$$(2.10) \quad u_{r\alpha}(q) = u_i(q_{-i}; \alpha) \geq u_i(q_{-i}; \beta) = u_{r\beta}(q).$$

If the game in question is also a potential one, we have seen that beneficial migrations descend the potential function, so the minima of the potential correspond to strategy distributions where no unilateral improvement is possible. In fact, the Kuhn-Tucker conditions for the game's potential coincide precisely with the Wardrop characterisation (2.9) and, hence, the game's equilibria will be the critical points of the potential (Sandholm, 2001, Proposition 3.1).

On account of the above, the equilibrium characterisation (2.9) will be central in our analysis, so we will examine it in depth in the sections that follow. En passant, we only note here that a similar condition can be laid down for population games with continuous strategy sets. This case has recently attracted quite a bit of interest, but since we will not need this added generality, we will not press the issue further – see instead Cressman (2005) or Hofbauer, Oechssler and Riedel (2009).

2.2. Networks and Flows. Stated somewhat informally, our chief interest lies in networks whose nodes produce traffic that seeks to reach its destination as quickly as possible. However, since the time taken to traverse a path in a network increases as the network becomes congested, it is hardly an easy task to pick the “path of least resistance” – especially given that users compete against each other in their endeavours. As a result, the game-theoretic setup of the previous section turns out to be remarkably appropriate for the analysis of these traffic flows.

Following Roughgarden and Tardos (2002, 2004), let $\mathcal{G} \equiv \mathcal{G}(\mathcal{V}, \mathcal{E})$ be a (finite) directed graph with node set \mathcal{V} and edge set \mathcal{E} , and let $\sigma = (v, w)$ be an *origin-destination* pair in \mathcal{G} (i.e. an ordered pair of nodes $v, w \in \mathcal{V}$ that can be joined by a path in \mathcal{G}). Suppose further that the origin v of σ outputs traffic towards the destination node w at some rate $\rho > 0$; then, 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 consist of a set of *users* \mathcal{N} (indexed by $i = 1, \dots, N$), together with an associated collection $\mathcal{A} \equiv \prod_i \mathcal{A}_i$ 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 origin-destination pair of user $i \in \mathcal{N}$).

Two remarks of a book-keeping nature are now in order: first, since we will only be interested in users with at least a modicum of choice on 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 traffic 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 2.3). This means that the sets \mathcal{A}_i can be assumed disjoint and, as a pleasant byproduct, the path index $\alpha \in \mathcal{A}_i$ fully characterizes the user i to whom it belongs – cf. the conventions of Section 1.2.

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 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 individual flows in a single profile, a *flow in the network* Ω 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:

$$(2.11) \quad y_r = \sum_i y_{ir} = \sum_i \sum_{\alpha \ni r}^i x_{i\alpha}$$

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 whether these two descriptions are equivalent; put differently, whether one can recover the flow distribution $x \in \Delta$ from the loads y_r on the edges of the network.

To answer this question, 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* $P^i : V_i \rightarrow W$ which sends 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 $P^i(e_{i\alpha}) = \sum_r P_{r\alpha}^i \varepsilon_r$, we see that the entries of P^i will be $P_{r\alpha}^i = 1$ if $r \in \alpha$ and 0 otherwise. We can then aggregate this construction over all $i \in \mathcal{N}$ by considering the product space $V \equiv \mathbb{R}^{\mathcal{A}} \cong \prod_i V_i$ and the corresponding indicator matrix $P = P^1 \oplus \dots \oplus P^N$ 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, (2.11) 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 $P : V \rightarrow W$ is injective.

This, however, 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 $p_i = \frac{\rho_i}{|\mathcal{A}_i|} \sum_{\alpha}^i e_{i\alpha}$ is the barycentre of Δ_i and $Z_i \equiv T_{p_i} \Delta_i = \{z_i \in V_i : \sum_{\alpha}^i z_{i\alpha} = 0\}$ is the tangent space to Δ_i at p_i – it is also worth keeping in mind that if we set $\mathcal{A}_i^* = \mathcal{A}_i \setminus \{\alpha_{i,0}\}$, then $Z_i \cong \mathbb{R}^{\mathcal{A}_i^*}$. 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 $Q \equiv P|_Z : Z \rightarrow W$ of P on the subspace $Z \equiv T_p \Delta \cong \prod_i Z_i$, where $p = (p_1, \dots, p_N)$ is the barycentre of Δ . 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:

$$(2.12) \quad y' - y = P(x') - P(x) = P(z) = Q(z),$$

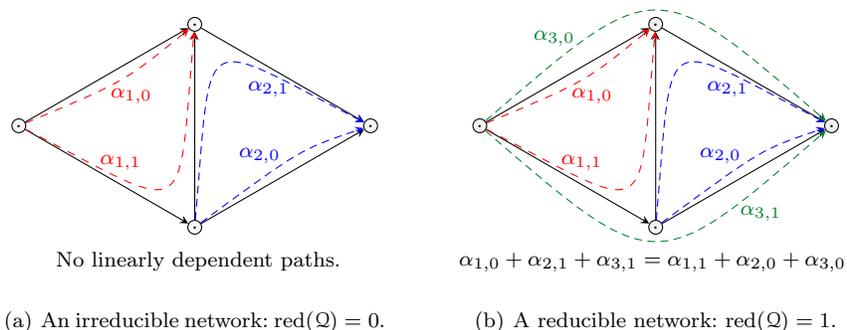


FIG 1. The addition of a user may increase the redundancy of a network.

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 $Q : Z \rightarrow W$ will be called the *redundancy matrix of the network* \mathcal{Q} , giving rise to:

DEFINITION 2.1. Let \mathcal{Q} be a network in a graph \mathcal{G} and let Q be the redundancy matrix of \mathcal{Q} . The *redundancy* $\text{red}(\mathcal{Q})$ of \mathcal{Q} is defined to be:

$$(2.13) \quad \text{red}(\mathcal{Q}) \equiv \dim(\ker Q).$$

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

The rationale behind this terminology should be clear enough: when a network \mathcal{Q} is reducible, some of its routes are “linearly dependent” and the respective directions in $\ker Q$ are “redundant” (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.

This dichotomy between reducible and irreducible networks will be quite significant for our purposes, so it is worth dwelling on Definition 2.1 for a bit more; specifically, it will be important to have a simple recipe with which to compute the redundancy matrix Q of a network \mathcal{Q} . To that end, let $Q^i \equiv P^i|_{Z_i}$ be the restriction of P^i on Z_i and, as before, let $\{e_{i,0}, e_{i,1}, \dots\}$ be the standard basis of $V_i = \mathbb{R}^{A_i}$. Then, the vectors $\tilde{e}_{i\mu} = e_{i\mu} - e_{i,0}$, $\mu \in \mathcal{A}_i^* \equiv \mathcal{A}_i \setminus \{0\}$, constitute a basis for Z_i and it is easy to see that the matrix elements of Q^i in this basis will be given by:

$$(2.14) \quad Q_{r\mu}^i = P_{r\mu}^i - P_{r,0}^i.$$

The above suggests that if there are too many users in a network, then it is highly unlikely that the network will be irreducible. Indeed, we have:

PROPOSITION 2.2. *Let $\mathcal{Q}(\mathcal{N}, \mathcal{A})$ be a network in the graph $\mathcal{G}(\mathcal{V}, \mathcal{E})$ and let $\mathcal{E}' \subseteq \mathcal{E}$ be the set of edges that are present in \mathcal{Q} . Then:*

$$(2.15) \quad \text{red}(\mathcal{Q}) \geq |\mathcal{N}| - |\mathcal{E}'|.$$

Hence, a network will always be reducible if the number of users exceeds the number of available links.

PROOF. From the definition of $Q : Z \rightarrow W$ we can easily see that $\text{im } Q$ is contained in the subspace of W that is spanned by \mathcal{E}' ; furthermore, since every user has $|\mathcal{A}_i| \geq 2$ routes to choose from, it follows that $\dim Z = \sum_i (|\mathcal{A}_i| - 1) \geq |\mathcal{N}|$. Therefore: $\text{red}(\mathcal{Q}) = \dim(\ker Q) = \dim Z - \dim(\text{im } Q) \geq |\mathcal{N}| - |\mathcal{E}'|$. \square

2.3. *Congestion Models and Equilibrium.* The time spent by a 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 also, to keep things simple, that they are at least C^1 with $\phi'_r > 0$.

On that account, the time needed to traverse an entire route $\alpha \in \mathcal{A}_i$ will be:

$$(2.16) \quad \omega_{i\alpha}(x) = \sum_{r \in \alpha} \phi_r(y_r) = \sum_r P_{r\alpha}^i \phi_r(y_r),$$

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

DEFINITION 2.3. 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 $\phi_r, r \in \mathcal{E}$.

The similarities between this definition and that of a game in normal form should be evident: all that is needed to turn Definition 2.3 into a N -person game is to specify its payoff functions. One way to go about this is to consider the *user averages*:

$$(2.17) \quad \omega_i(x) = \frac{1}{\rho_i} \sum_\alpha^i x_{i\alpha} \omega_{i\alpha}(x) = \frac{1}{\rho_i} \sum_r y_{ir} \phi_r(y_r),$$

where the last equality follows from (2.16) and the definition of $y_{ir} = \sum_\alpha^i x_{i\alpha}$. Thus, in keeping with the equilibrium condition (NEQ), a flow q will be at Nash equilibrium in the game $\mathfrak{G}_1 \equiv \mathfrak{G}_1(\mathcal{N}, \Delta, -\omega)$ when:

$$(NE1) \quad \omega_i(q) \leq \omega_i(q_{-i}; q'_i) \text{ for every user } i \in \mathcal{N} \text{ and all flows } q'_i \in \Delta_i.$$

For many classes of latency functions ϕ_r , the average delays ω_i turn out to be convex and the existence of equilibria is assured by the results of Rosen (1965). However, not only is this not always the case but, more importantly, the user averages (2.17) do not necessarily reflect the users' actual optimisation objectives either.

Indeed, another equally justified choice of payoffs is given by the *worst delays*:

$$(2.18) \quad \tilde{\omega}_i(x) = \max_{\alpha: x_{i\alpha} > 0} \{\omega_{i\alpha}(x)\},$$

i.e. the time at which a user's last traffic packet reaches its destination. In that case, a flow q will be at equilibrium for the game $\mathfrak{G}_2 \equiv \mathfrak{G}_2(\mathcal{N}, \Delta, -\tilde{\omega}_i)$ when:

$$(NE2) \quad \tilde{\omega}_i(q) \leq \tilde{\omega}_i(q_{-i}; q'_i) \text{ for every user } i \in \mathcal{N} \text{ and all flows } q'_i \in \Delta_i.$$

Unfortunately, the payoff functions $\tilde{\omega}_i$ may be discontinuous along any intersection of faces of Δ_i because the support $\text{supp}(x_i) = \{\alpha \in \mathcal{A}_i : x_{i\alpha} > 0\}$ of x_i changes there as well. Consequently, the existence of equilibrial flows cannot be inferred from the general theory in this instance either.

On the other hand, if we go back to our original motivation (the Internet), we see that our notion of a “user” more accurately portrays the network's *routers* and not its “real-life” users (humans, applications, etc.). However, since these routers are not selfish in themselves, conditions (NE1) and (NE2) do not necessarily point to the right direction either. Instead, the routers' selfless task is to ensure that the *nonatomic* traffic elements circulating in the network (the actual selfish entities) remain satisfied. It is thus more reasonable to go back to Wardrop's principle (2.9):

DEFINITION 2.4. A flow $q \in \Delta$ will be at *Wardrop equilibrium* when

$$(WEQ) \quad \omega_{i\alpha}(q) \leq \omega_{i\beta}(q) \text{ for all } i \in \mathcal{N} \text{ and for all routes } \alpha, \beta \in \mathcal{A}_i \text{ with } q_{i\alpha} > 0,$$

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

Condition (WEQ) 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 $\alpha \in \text{supp}(q_i)$ and leads to the following alternative characterisation of Wardrop flows:

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

Even more importantly however, Wardrop equilibria can also be harvested from the (global) minimum of the *Rosenthal potential* (Rosenthal, 1973):

$$(2.19) \quad \Phi(y) = \sum_r \Phi_r(y_r) = \sum_r \int_0^{y_r} \phi_r(w) dw.$$

The reason for calling this function a potential is twofold: firstly, it is the nonatomic generalisation of the potential function introduced by Monderer and Shapley (1996) to describe finite congestion games; secondly, the payoff functions $\omega_{i\alpha}$ can be obtained from Φ by a simple differentiation. To be sure, if we set $F(x) = \Phi(y)$ where $y = P(x)$, we readily obtain:

$$(2.20) \quad \frac{\partial F}{\partial x_{i\alpha}} = \sum_r \frac{\partial \Phi}{\partial y_r} \frac{\partial y_r}{\partial x_{i\alpha}} = \sum_r \phi_r(y_r) P_{r\alpha}^i = \sum_{r \in \alpha} \phi_r(y_r) = \omega_{i\alpha}(x),$$

which is exactly the definition of a potential function in the sense of (2.7) – note also that the “externality symmetry” condition (2.6) can be verified independently:

$$(2.21) \quad \frac{\partial \omega_{i\alpha}}{\partial x_{j\beta}} = \sum_r P_{r\alpha}^i \phi'_r(y_r) P_{r\beta}^j = \sum_{r \in \alpha \cap \beta} \phi'_r(y_r) = \frac{\partial \omega_{j\beta}}{\partial x_{i\alpha}}.$$

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 characterise this minimum coincide with the Wardrop condition (2.9) (Beckmann, McGuire and Winsten, 1956; Dafermos and Sparrow, 1969; Roughgarden and Tardos, 2002; Sandholm, 2001), so the Wardrop set of the congestion model \mathfrak{C} will be given by:

$$(2.22) \quad \Delta^* = \{x \in \Delta : P(x) = y^*\} = P^{-1}(y^*) \cap \Delta.$$

PROPOSITION 2.5. *Let $\mathfrak{C} \equiv \mathfrak{C}(\mathcal{Q}, \phi_r)$ be a congestion model with strictly increasing latencies ϕ_r and let Δ^* be its set of Wardrop equilibria. Then:*

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

Since $P^{-1}(y^*)$ is an affine subspace of \mathbb{R}^A and Δ is a product of simplices, there is really nothing left to prove (simply observe that if q is an interior Wardrop flow, then $P^{-1}(y^*)$ intersects the full-dimensional interior of Δ). The only surprise here is that this result seems to have been overlooked in much of the literature concerning congestion models: for instance, both Sandholm (2001, Corollary 5.6) and Fischer and Vöcking (2004, Propositions 2 and 3) 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 the simple network of Fig. 1(b)), then the potential function $F(x) \equiv \Phi(P(x))$ is no longer strictly convex: it is in fact constant along every null direction of the redundancy matrix $Q = P|_{T\Delta}$.

We thus see that a Wardrop equilibrium is unique iff *a) the network \mathcal{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 2.6. A Wardrop equilibrium q will be called *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 $\beta \in \mathcal{A}_i \setminus \{\alpha_i\}$.*

In Nash games, a pure equilibrium occasionally fails to be strict, but only by a hair: an arbitrarily small perturbation of a player's pure payoffs $u_{i, \alpha_1, \dots, \alpha_N}$ resolves a pure equilibrium into a strict one without affecting the payoffs of the other players. In congestion models however, there is no such guarantee because, whenever two users' paths overlap, one cannot perturb the delays of one user independently of the other's. As a matter of fact, the existence of a strict Wardrop equilibrium actually *precludes* the existence of any other equilibria:

PROPOSITION 2.7. *Let \mathfrak{C} be a congestion model. If q is a strict Wardrop equilibrium of \mathfrak{C} , then q is the unique Wardrop equilibrium of \mathfrak{C} .*

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. \square

In other words, even if q is a strict equilibrium of a reducible network, then the redundant directions which constitute the affine subspace $q + \ker Q$ will only intersect Δ at q . On the other hand, if q is merely a pure equilibrium, $q + \ker Q$ might well intersect the open interior of Δ ; in that case, there is no arbitrarily small perturbation of the delay functions that could make q into a strict equilibrium.

Equilibria and Objectives. On account of the above, we will focus our investigations on the concept of Wardrop equilibrium. However, we should mention here that this equilibril notion can also be reconciled (to some extent at least) with the optimisation objectives represented by the payoffs (2.17) and (2.18) as well.

First, with respect to the average delays $\omega_i(x) = \rho_i^{-1} \sum_\alpha x_{i\alpha} \omega_{i\alpha}(x)$, the *optimal* traffic distributions which minimise 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 by Beckmann, McGuire and Winsten (1956), who observed the inherent duality in Wardrop's principle: 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 Roughgarden and Tardos, 2004). Then, to study these "socially optimal" flows, we simply have to redress our analysis to fit these "marginal latencies" instead (see Section 5 for more details).

Secondly, Wardrop equilibria also have close ties with the Nash condition (NE2) which corresponds to the "worst-delays" (2.18). Specifically, one can easily see that the Nash condition (NE2) is equivalent to the Wardrop condition (2.9) when every user only has 2 possible paths to choose from (every amount of traffic diverted from one path increases the delay at the user's other path). However, if a user has 3 or more paths at his disposal, then the situation can change dramatically because of *Braess's paradox* (Braess, 1968).

The essence of this paradox is that there exist networks which perform better if one removes their fastest link. An example of such a network is given in Fig. 2, where it is assumed that a user seeks to route 6 units of traffic from A to D using the three paths $A \rightarrow B \rightarrow D$ (blue), $A \rightarrow C \rightarrow D$ (red) and $A \rightarrow B \rightarrow C \rightarrow D$ (green). In that case, the Wardrop condition (WEQ) calls for equidistribution: 2 units are routed via each path, leading to a delay of 92 time units along all paths. Alternatively, if the user sends 3 traffic units via the red and blue paths and ignores the green one, all traffic will experience a delay of 83. Paradoxically, even though

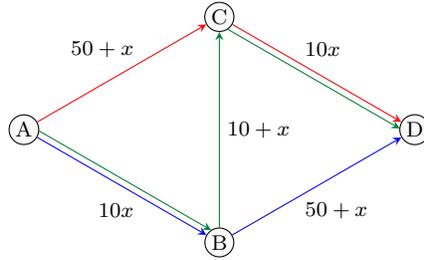


FIG 2. Braess's paradox and the disparity between Wardrop and Nash equilibria.

the green path has a latency of only 70, the Nash condition (NE2) is satisfied: if traffic is diverted from, say, the red path to the faster green one, then the latency of the blue path will also increase, thus increasing the worst delay $\tilde{\omega}_i$ as well.

This paradox is what led to the original investigations in the efficiency of selfish routing (Koutsoupias and Papadimitriou, 1999; Roughgarden and Tardos, 2002), and it seems that it is also what causes this disparity between Wardrop and Nash equilibria. A thorough investigation of this matter is a worthy project but, since it would take us too far afield, we will not pursue it here. Henceforward, we will focus almost exclusively on Wardrop flows, which represent the most relevant equilibrium concept for our purposes.

3. Learning, Evolution and Rational Behaviour. Unfortunately, locating the Wardrop equilibria of a network is a rather arduous process which entails a good deal of global calculations (namely the minimisation of a nonlinear convex functional with exponentially many variables over a convex polytope). Since such calculations clearly exceed the deductive capabilities of individual users (especially if they do not have access to global information), it is of great interest to see whether there are simple learning schemes which allow users to reach an equilibrium without having to rely on centralised computations.

3.1. Learning and the Replicator Dynamics. For our purposes, a learning scheme will be a rule which trains users to route their traffic in an efficient way by processing information that is readily available. On the other hand, since this information must be “local” in nature, the learning scheme should be similarly “distributed”: for example, the play of one’s opponents or the exact form of the network’s latency functions are not easily accessible pieces of information. Furthermore, we should also be looking for a learning scheme which is simple enough for users to apply in real-time, without having to perform a huge number of calculations at each instant.

In continuous time, such a learning scheme may be cast as a dynamical system:

$$(3.1) \quad \frac{dx}{dt} = v(x) \text{ or, in coordinates: } \frac{dx_{i\alpha}}{dt} = v_{i\alpha}(x),$$

where $x(t) \in \Delta$ denotes the flow at time t and the vector field $v : \Delta \rightarrow \mathbb{R}^A$ plays the part of the “learning rule” in question – for simplicity, we will also take v to be

smooth. Of course, since the flow $x(t)$ evolves in Δ , v itself must lie on the tangent space Z of Δ ; we thus require that $\sum_{\alpha}^i v_{i\alpha}(x) = 0$ for all $i \in \mathcal{N}$.

Furthermore, v should also leave the *faces* of Δ invariant in the sense that any individual trajectory $x_i(t)$ that begins at some face of Δ_i must always remain in said face. This is actually an essential consequence of our postulates: if a user does not employ a particular route $\alpha \in \mathcal{A}_i$, then he has *no* information on the route and, as such, there is no a priori reason that an adaptive learning rule would induce the user to sample it. In effect, such a learning rule would either fail to rely solely on readily observable information or would not necessarily be a very simple one.

This shows that $v_{i\alpha}(x)$ must vanish if $x_{i\alpha} = 0$, so if we set $v_{i\alpha}(x) = x_{i\alpha} \tilde{v}_{i\alpha}(x)$, we obtain the orthogonality condition $\sum_{\alpha}^i x_{i\alpha} \tilde{v}_{i\alpha}(x) = 0$. Accordingly, $\tilde{v}_{i\alpha}$ may be written in the form:

$$(3.2) \quad \tilde{v}_{i\alpha}(x) = u_{i\alpha}(x) - u_i(x)$$

where the $u_{i\alpha}$ satisfy no further constraints and, as can be shown by a simple summation, the function $u_i(x)$ is just the *user average*: $u_i(x) = \rho_i^{-1} \sum_{\beta}^i x_{i\beta} u_{i\beta}(x)$ (recall that $\sum_{\beta}^i x_{i\beta} = \rho_i$). This shows that *any* learning rule which leaves the faces of Δ invariant must necessarily be of the form:

$$(3.3) \quad \frac{dx_{i\alpha}}{dt} = x_{i\alpha} (u_{i\alpha}(x) - u_i(x)).$$

Dynamics of this type were first derived in the context of population biology by Taylor and Jonker (1978), initially for different genotypes within a species (single-population models), and then for different species altogether (multi-population models; Weibull (1995) provides an excellent survey). In these evolutionary games, the key objects of interest are large populations of different species, each of them subdivided into distinct *genotypes* that are “programmed” to a specific behaviour (e.g. “hawks” fight, while “doves” take flight). Then, at each instance of biological interaction, it is assumed that one representative from each species is selected at random, and they are all matched to play some Nash game \mathfrak{G} whose payoffs represent a proportionate increase in their reproductive fitness (measured by the number of offsprings in the unit of time). In this fashion, if $u_{i\alpha}(x)$ denotes the population average of the payoff to the α -th genotype, it turns out that the evolution of the species will be governed by the *replicator dynamics* (3.3).

In our case, the most natural choice for the payoffs $u_{i\alpha}$ of (3.3) is to use the delay functions $\omega_{i\alpha}(x)$ and set $u_{i\alpha} = -\omega_{i\alpha}$. In so doing, we obtain:

$$(3.4) \quad \frac{dx_{i\alpha}}{dt} = x_{i\alpha} (\omega_i(x) - \omega_{i\alpha}(x)).$$

In keeping with our “local information” mantra, we see that users do not need to know the delays along paths that they do not employ because the replicator vector field vanishes when $x_{i\alpha} = 0$. Thus, users that evolve according to (3.4) are oblivious to their surroundings, even to the existence of other users: they simply use (3.4) to respond to the stimuli $\omega_{i\alpha}(x)$ in the hope of minimising their delays.

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

$$(3.5) \quad \frac{dx_{i\alpha}}{dt} = \lambda_i x_{i\alpha} (\omega_i(x) - \omega_{i\alpha}(x))$$

(naturally, the uniform case (3.4) 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 $\lambda_i \rightarrow \infty$, equation (3.5) “freezes” to a rigid (and myopic) best-reply process (see also Börgers and Sarin, 1997).

3.2. Entropy and Rationality. An immediate observation concerning the replicator dynamics (3.5) is that Wardrop equilibria are rest points: if q is a Wardrop flow, the characterisation (WEQ') 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 Sandholm (2001), 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.

In that direction, if y^* is the minimum of the Rosenthal potential $\Phi(y)$, it is easy to see that the function $F_0(x) = \Phi(P(x)) - \Phi(y^*)$ is a *semi-definite Lyapunov function* for the dynamics (3.5). Indeed, F_0 vanishes on the Wardrop set Δ^* , is positive otherwise, and its evolution under (3.5) satisfies:

$$(3.6) \quad \frac{dF_0}{dt} = \sum_{i,\alpha} \frac{\partial F_{i\alpha}}{\partial x_{i\alpha}} \frac{dx_{i\alpha}}{dt} = \sum_i \lambda_i \rho_i \left[\omega_i^2(x) - \rho_i^{-1} \left(\sum_{\alpha} x_{i\alpha} \omega_{i\alpha}^2(x) \right) \right] \leq 0,$$

the last step following from Jensen’s inequality – equality only holds when $\omega_{i\alpha}(x) = \omega_i(x)$ for all $\alpha \in \text{supp}(x)$. Thus, by standard results in the theory of dynamical systems, it follows that the solution orbits of (3.5) descend the potential F_0 and eventually converge to a connected subset of rest points – see also Sandholm (2001), where the property (3.6) is referred to as “positive correlation”.

Nevertheless, since not all stationary points of (3.5) are Wardrop equilibria, this result tells us little about the rationality properties of the replicator dynamics in congestion models. A much more important role is played by the *relative entropy* (also known as the *Kullback-Leibler divergence*):

$$(3.7) \quad H_q(x) \equiv d_{\text{KL}}(q, x) = \sum_{\alpha \in \text{supp}(q)} q_{i\alpha} \log \frac{q_{i\alpha}}{x_{i\alpha}}$$

where the sum is taken over the support of q : $\text{supp}(q) = \{\alpha : q_{i\alpha} > 0\}$. Of course, this sum is finite only when x employs with positive probability all $\alpha \in \mathcal{A}$ that are present in q ; that is, the domain of definition of H_q is $\Delta_q \equiv \{x \in \Delta : q \ll x\}$ (here, “ \ll ” denotes absolute continuity of measures). Even so though, it will matter little

if we extend H_q continuously to all of Δ by setting $H_q = \infty$ outside Δ_q , so we will occasionally act as if H_q were defined over all of Δ .

Technicalities aside, the significance of the relative entropy lies in that it measures distance in probability space. Indeed, even though it is not a distance function per se (it fails to be symmetric and does not satisfy the triangle inequality), it is positive definite (see below) and strictly convex (Weibull, 1995). More importantly for our purposes, it is also a (semi-definite) Lyapunov function for the dynamics (3.4):

LEMMA 3.1. *Let $\mathfrak{C}(\Omega, \{\phi_r\})$ be a congestion model with increasing latencies, and let $q \in \Delta$ be a Wardrop flow of \mathfrak{C} . Then, the relative entropy $H_q(x)$ satisfies:*

1. $H_q(q) = 0$ and $H_q(x) > 0$ for all $x \neq q$;
2. \dot{H}_q vanishes on the Wardrop set $\Delta^*(\mathfrak{C})$ and is negative otherwise (where \dot{H}_q denotes the time derivative with respect to (3.4): $\dot{H}_q = \sum_{i,\alpha} \frac{\partial H_q}{\partial x_{i\alpha}} \dot{x}_{i\alpha}$).

In particular, if the network Ω is irreducible ($\text{red}(\Omega) = 0$), then H_q is Lyapunov for the replicator dynamics (3.4).

PROOF. The first part of the lemma (positive-definiteness) is an easy consequence of Jensen's inequality (Weibull, 1995, pp. 95–100). As for the second part:

$$(3.8) \quad \dot{H}_q(x) = \sum_{i,\alpha} \frac{\partial H_q}{\partial x_{i\alpha}} \dot{x}_{i\alpha} = - \sum_{i,\alpha} q_{i\alpha} (\omega_i(x) - \omega_{i\alpha}(x)) \equiv -L_q(x),$$

where we have set $L_q(x) \equiv \sum_{i,\alpha} q_{i\alpha} (\omega_i(x) - \omega_{i\alpha}(x))$. Then, the simple rearrangement $\sum_{i,\alpha} q_{i\alpha} \omega_i(x) = \rho_i \omega_i(x) = \sum_{i,\alpha} x_{i\alpha} \omega_{i\alpha}(x)$ and some trivial linear algebra yield:

$$(3.9) \quad L_q(x) \equiv \sum_{i,\alpha} (x_{i\alpha} - q_{i\alpha}) \omega_{i\alpha}(x) = \sum_r (y_r - y_r^*) \phi_r(y_r) \equiv \Lambda(y),$$

where $y = P(x)$ and $y^* = P(q)$ are the load profiles which correspond to the flows x and q respectively.

We will refer to the expression L_q (or, interchangeably, to Λ) as the *adjoint potential* of \mathfrak{C} because, similarly to the Rosenthal potential Φ , it measures distance from the Wardrop set Δ^* . The properties of L_q will be discussed at length in Appendix A where, among others, we establish the easy (but crucial!) inequality:

$$(3.10) \quad \Lambda(y) \geq \Phi(y) - \Phi(y^*).$$

Hence, with y^* being the global minimum of Φ , we conclude that Λ is positive definite and the lemma follows by noting that $P(x) = y^*$ iff x is Wardrop. \square

REMARK 1 (*The Rate-adjusted Case*). It is also reasonable to ask whether the relative entropy function enjoys the same properties in the rate-adjusted dynamics (3.5). Unfortunately, this is not true unless all players learn at the same rate; however, if we consider the *rate-adjusted* relative entropy:

$$(3.11) \quad H_q(x; \lambda) = \sum_{i \in \mathcal{N}} \lambda_i^{-1} \sum_{\alpha: q_{i\alpha} > 0} q_{i\alpha} \log \frac{q_{i\alpha}}{x_{i\alpha}},$$

the same calculations show that $\dot{H}_q(x; \lambda) = -L_q(x)$ and provide us with the analogue of Lemma 3.1 for the rate-adjusted dynamics (3.5).

In view of the above, it would be tempting to infer that the replicator dynamics converge to Wardrop equilibrium. Nevertheless, a semi-definite Lyapunov function is not enough to guarantee convergence by itself, even if we rule out the existence of limit cycles. For instance, if we consider the homogeneous system:

$$(3.12) \quad \dot{x} = yz, \quad \dot{y} = -xz, \quad \dot{z} = -z^2,$$

with $z \geq 0$, we see that it admits the semi-definite Lyapunov function $H(x, y, z) = x^2 + y^2 + z^2$ whose time derivative only vanishes on the x - y plane. However, the general solution of (3.12) in cylindrical coordinates (ρ, ϕ, z) is just:

$$(3.13) \quad \rho(t) = \rho_0, \quad \phi(t) = \phi_0 - \log(1 + z_0 t), \quad z(t) = \frac{z_0}{1 + z_0 t},$$

and this represents a helix of constant radius whose coils become topologically dense as the solution orbits approach the x - y plane. We thus see that the solutions of (3.12) approach a set of stationary points, but do not converge to a specific one.

That said, there is much more at work in the replicator dynamics (3.5) than a single semi-definite Lyapunov function: there exists a whole *family* of such functions, one for each Wardrop flow $q \in \Delta$. So, undeterred by potential pathologies, the replicator dynamics actually *do* converge to equilibrium:

THEOREM 3.2. *Let $\mathfrak{C}(\Omega, \{\phi_r\})$ be a congestion model in a network Ω . Then, every interior solution trajectory of the replicator dynamics (3.5) converges to a Wardrop equilibrium of \mathfrak{C} ; in particular, if the network Ω is irreducible, $x(t)$ converges to the unique Wardrop equilibrium of \mathfrak{C} .*

PROOF. It will be useful to shift our point of view to the *evolution function* $\theta(x, t)$ of the dynamics (3.5) which describes the solution trajectory that starts at x at time $t = 0$ and which satisfies the consistency condition:

$$(3.14) \quad \theta(x, t + s) = \theta(\theta(x, t), s) \text{ for all } t, s \geq 0 \text{ and for all } x \in \Delta.$$

Now, fix the initial condition $x \in \text{Int}(\Delta)$ and let $x(t) = \theta(x, t)$ be the corresponding solution orbit. If $q \in \Delta^*$ is a Wardrop equilibrium of \mathfrak{C} , then, in view of Lemma 3.1, the function $V_q(t) \equiv H_q(\theta(x, t))$ will be decreasing and will converge to some $m \geq 0$ as $t \rightarrow \infty$. It thus follows that $x(t)$ converges itself to the level set $H_q^{-1}(m)$.

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

$$(3.15) \quad H_q(\theta(x, t_n + t)) = H_q(\theta(\theta(x, t_n), t)) \rightarrow H_q(\theta(x^*, t)) < H_q(x^*) = m$$

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

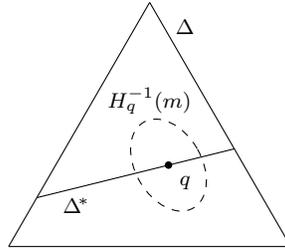


FIG 3. The various sets in the proof of Theorem 3.2.

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

REMARK 1 (*Previous Work*). In the context of potential games, Sandholm (2001) examined a class of learning dynamics $v_{i\alpha}$ which are “positively correlated” to the game’s payoff functions $u_{i\alpha} = -\omega_{i\alpha}$, in the sense that $\sum_{i,\alpha} v_{i\alpha}(x)\omega_{i\alpha}(x) \geq 0$. It was then shown that if the rest points of these dynamics coincide with the game’s Wardrop equilibria (the “non-complacency” condition), then all solution orbits converge to set of Wardrop equilibria. Unfortunately, as we have already pointed out, the replicator dynamics are “complacent” and, in that case, Sandholm’s results only ensure that Wardrop equilibria are Lyapunov stable.

To the best of our knowledge, the stronger convergence properties of Theorem 3.2 were first suggested by Fischer and Vöcking (2004) who identified the link between Wardrop equilibrium and evolutionary stability (Maynard Smith, 1974). In particular, the authors showed that Wardrop equilibria are robust against “mutations” that lead to greater delays but, in networks with more than one users (the “multi-commodity” case as they call it), their approach rests heavily on the (implicit) assumption of irreducibility. If this is not the case, the adjoint potential L_q is only positive *semi-definite* and the approach of Fischer and Vöcking breaks down because Wardrop equilibria are only *neutrally* stable – this is also the problem with the formulation of Corrolary 5.1 in Sandholm (2001).

REMARK 2 (*Non-interior Trajectories*). One might also ask what happens if the initial condition $x(0)$ is not an interior point of Δ . Clearly, if $x(0)$ does not employ the routes that are present in a Wardrop flow q , $x(t)$ cannot have q as a limit point – a simple consequence of the fact that the replicator dynamics leave the faces of Δ invariant. All the same, one can simply quotient out the routes that are not initially present until $x(0)$ *becomes* an interior point in the reduced strategy space Δ_{eff} that ensues. In that case, Theorem 3.2 can be applied to the (similarly reduced) congestion model $\mathfrak{C}_{\text{eff}}$ to show that $x(t)$ converges to Wardrop equilibrium in $\mathfrak{C}_{\text{eff}}$ (cf. the “restricted equilibria” of Fischer and Vöcking, 2004).

REMARK 3 (*Evolution and Friction*). Since the replicator trajectories converge to a Wardrop equilibrium, it follows that there can be no limit cycles. On the other hand, limit cycles are a common occurrence in evolutionary games: for example, Matching Pennies and Rock-Paper-Scissors both exhibit limit cycles in the standard replicator dynamics (Weibull, 1995). So, while the evolutionary energy of large populations may remain undiminished over time, Theorem 3.2 shows that congestion models are *dissipative* and traffic flows settle down to a steady state.

4. The Effect of Stochastic Fluctuations. Going back to our original discussion on learning schemes, 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 or imperfect readings of one's payoffs might perturb this information considerably; additionally, if the users' traffic flows are not continuous in time but consist of discrete segments instead (e.g. datagrams in communication networks), the queueing latencies $\omega_{i\alpha}$ only represent the users' expected delays. Hence, the delays that users actually observe might only be a randomly fluctuating estimate of the underlying payoffs, and this could negatively affect the rationality properties of the replicator dynamics.

4.1. *Stochastic Replicator Dynamics.* Our goal here will be to determine the behaviour 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, $\eta_{i\alpha} = \sum_{\alpha} P_{r\alpha}^i \eta_r$. In this way, the replicator dynamics (3.4) become:

$$(4.1) \quad \frac{dx_{i\alpha}}{dt} = x_{i\alpha} (\hat{\omega}_i - \hat{\omega}_{i\alpha}) = x_{i\alpha} (\omega_i - \omega_{i\alpha}) + x_{i\alpha} (\eta_i - \eta_{i\alpha})$$

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

The exact form of the perturbations η_r clearly depends on the particular situation at hand. Still, since we are chiefly interested in stochastic fluctuations around the underlying delays $\omega_{i\alpha}$, it is reasonable to take these perturbations to be some sort of white noise that does not bias users towards one direction or another. In that case, we should rewrite (4.1) as a *stochastic* differential equation:

$$(4.2) \quad dX_{i\alpha} = X_{i\alpha} [\omega_i(X) - \omega_{i\alpha}(X)] dt + X_{i\alpha} \left[dU_{i\alpha} - \rho_i^{-1} \sum_{\beta}^i X_{i\beta} dU_{i\beta} \right]$$

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

$$(4.3) \quad dU_{i\alpha} = \sum_{r \in \alpha} \sigma_r dW_r = \sum_r P_{r\alpha}^i \sigma_r dW_r$$

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:

$$(4.4) \quad \begin{aligned} dX_{i\alpha} &= \lambda_i X_{i\alpha} [\omega_i(X) - \omega_{i\alpha}(X)] + \lambda_i X_{i\alpha} \left[dU_{i\alpha} - \rho_i^{-1} \sum_{\beta}^i X_{i\beta} dU_{i\beta} \right] \\ &= \lambda_i b_{i\alpha}(X) dt + \lambda_i \sum_{\beta}^i c_{i,\alpha\beta}(X) dU_{i\beta} \end{aligned}$$

where b and c are the drift and diffusion coefficients that appear in (4.2).

The rate-adjusted equation (4.4) will constitute our stochastic version of the replicator dynamics and, as such, it warrants some discussion in and by itself. A first remark to be made concerns the noise coefficients σ_r : 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 $\sigma_r \leftrightarrow \text{ess sup}_y \sigma_r(y)$.

Secondly, it is also important to compare (4.4) to other stochastic incarnations of the replicator dynamics, namely the ‘‘aggregate shocks’’ version of Fudenberg and Harris (1992) and the authors’ own ‘‘exponential learning’’ approach (Mertikopoulos and Moustakas, 2009, 2010). In the case of the former, one perturbs the replicator equation (3.3) by accounting for the (stochastic) interference of nature with reproduction rates (Fudenberg and Harris, 1992; Imhof, 2005):

$$(4.5) \quad \begin{aligned} dX_{i\alpha} &= X_{i\alpha} (u_{i\alpha}(X) - u_i(X)) dt - \left(\sigma_{i\alpha}^2 X_{i\alpha} - \sum_{\beta}^i \sigma_{i\beta}^2 X_{i\beta} \right) dt \\ &\quad + X_{i\alpha} \left[\sigma_{i\alpha} dW_{i\alpha} - \sum_{\beta}^i \sigma_{i\beta} X_{i\beta} dW_{i\beta} \right], \end{aligned}$$

where $W = \sum_{i,\alpha} W_{i\alpha} e_{i\alpha}$ is a Wiener process in $\prod_i \mathbb{R}^{A_i}$. Then, if the ‘‘aggregate shocks’’ $\sigma_{i\alpha}$ are mild enough, Cabrales (2000) and Imhof (2005) showed that dominated strategies become extinct and that the game’s strict Nash equilibria are asymptotically stable with arbitrarily high probability.

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 (Fudenberg and Levine, 1998, pp. 118–129). However, if the information that players have is imperfect, the errors propagate to their learning curves and instead lead to the stochastic dynamics:

$$(4.6) \quad \begin{aligned} dX_{i\alpha} &= \lambda_i X_{i\alpha} [(u_{i\alpha}(X) - u_i(X))] dt + \lambda_i X_{i\alpha} \left[\sigma_{i\alpha} dW_{i\alpha} - \sum_{\beta}^i \sigma_{i\beta} X_{i\beta} dW_{i\beta} \right] \\ &\quad + \frac{\lambda_i^2}{2} X_{i\alpha} \left[\sigma_{i\alpha}^2 (1 - 2X_{i\alpha}) - \sum_{\beta}^i \sigma_{i\beta}^2 X_{i\beta} (1 - 2X_{i\beta}) \right] dt. \end{aligned}$$

The rationality properties of these learning dynamics are somewhat stronger than in the biological setting: irrespective of the perturbations’ magnitude, strategies

which are not rationally admissible die out at an exponential rate and the strict Nash equilibria of the game are always (stochastically) stable (Mertikopoulos and Moustakas, 2009, 2010).

In light of the above, there are two notable traits of (4.4) that set it apart from its other stochastic versions. First off, the drift of (4.4) coincides with the deterministic replicator dynamics (3.5) whereas the drift coefficients of (4.5) and (4.6) do not. On the other hand, the martingale processes U that appear in (4.4) are not uncorrelated components of some Wiener process (as is the case for both (4.5) and (4.6)): instead, 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 last observation 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:

$$(4.7) \quad dU_\alpha \cdot dU_\beta = \left(\sum_r P_{r\alpha} \sigma_r dW_r \right) \cdot \left(\sum_s P_{s\beta} \sigma_s dW_s \right) \\ = \sum_{r,s} P_{r\alpha} P_{s\beta} \sigma_r \sigma_s \delta_{rs} dt = \sum_{r \in \alpha \cap \beta} \sigma_r^2 dt = \sigma_{\alpha\beta}^2 dt,$$

where $\sigma_{\alpha\beta}^2 = \sum_r P_{r\alpha} P_{r\beta} \sigma_r^2$ gives the variance of the noise along the intersection $\alpha\beta \equiv \alpha \cap \beta$ of the paths $\alpha, \beta \in \mathcal{A}$ (note also that we used our notational conventions to avoid cumbersome expressions such as $\sigma_{i\alpha, j\beta}^2$). 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:

$$(4.8) \quad (dU_\alpha)^2 = \sum_{r \in \alpha} \sigma_r^2 dt = \sigma_\alpha^2 dt$$

where $\sigma_\alpha^2 \equiv \sigma_{\alpha\alpha}^2 = \sum_r P_{r\alpha} \sigma_r^2$ measures the intensity of the noise on the route $\alpha \in \mathcal{A}$.¹ 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 (4.4). To begin with, note that (4.4) admits a (unique) strong solution for any initial state $X(0) = x \in \Delta$, even though its coefficients do not necessarily grow linearly – a common requisite for existence and uniqueness of strong solutions to SDE's. Indeed, an addition over $\alpha \in \mathcal{A}_i$ reveals that every component simplex Δ_i of Δ remains invariant under these dynamics: if $X_i(0) = x_i \in \Delta_i$, then $d(\sum_\alpha X_{i\alpha}) = 0$ and, hence, $X_i(t)$ stays in Δ_i for all $t \geq 0$. So, if $U \supseteq \Delta$ is open and ϕ is a smooth bump function on U that vanishes outside some compact set $K \supseteq U$, the SDE

$$(4.9) \quad dX_{i\alpha} = \lambda_i \phi(x) \left(b_{i\alpha}(X) dt + \sum_\beta^i c_{i,\alpha\beta}(X) dU_{i\beta} \right)$$

has bounded coefficients and will thus admit a unique strong solution. But since this last equation agrees with (4.4) on Δ and all solutions of (4.4) always stay in Δ , our claim follows.

¹This notation is also consistent with our intersection notation: $\alpha\alpha \equiv \alpha \cap \alpha = \alpha$.

Now, as in the deterministic setting, our main tool will be the (rate-adjusted) relative entropy $H_q(x; \lambda) = \sum_i \lambda_i^{-1} \sum_{\alpha} q_{i\alpha} \log(q_{i\alpha}/x_{i\alpha})$ which we will study with the help of the *generator* \mathcal{L} of the diffusion (4.4). To that end, recall that the generator \mathcal{L} of the Itô diffusion:

$$(4.10) \quad dX_{\alpha}(t) = \mu_{\alpha}(X(t)) dt + \sum_{\beta} \sigma_{\alpha\beta}(X(t)) dW_{\beta}(t),$$

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

$$(4.11) \quad \mathcal{L} = \sum_{\alpha} \mu_{\alpha}(x) \frac{\partial}{\partial x_{\alpha}} + \frac{1}{2} \sum_{\alpha, \beta} (\sigma(x) \sigma^T(x))_{\alpha\beta} \frac{\partial^2}{\partial x_{\alpha} \partial x_{\beta}}$$

(for a comprehensive account, consult the excellent book by Øksendal (2007)). In this manner, if f is a sufficiently smooth function, $\mathcal{L}f$ captures the drift of the process $f(X(t))$:

$$(4.12) \quad df(X(t)) = \mathcal{L}f(X(t)) dt + \sum_{\alpha, \beta} \frac{\partial f}{\partial x_{\alpha}} \Big|_{X(t)} \sigma_{\alpha\beta}(X(t)) dW_{\beta}(t).$$

Of course, in the case of the diffusion (4.4), the martingales U are not the components of a Wiener process, so (4.12) cannot be applied right off the shelf. However, a straightforward application of Itô's lemma (see appendix B) yields:

LEMMA 4.1. *Let \mathcal{L} be the generator of (4.4). Then, for any $q \in \Delta$:*

$$(4.13) \quad \begin{aligned} \mathcal{L}H_q(x; \lambda) = & -L_q(x) + \frac{1}{2} \sum_i \frac{\lambda_i}{\rho_i} \sum_{\beta, \gamma}^i \sigma_{\beta\gamma}^2(x_{i\beta} - q_{i\beta})(x_{i\gamma} - q_{i\gamma}) \\ & + \frac{1}{2} \sum_i \frac{\lambda_i}{\rho_i} \sum_{\beta, \gamma}^i \sigma_{\beta\gamma}^2 q_{i\beta} (\rho_i \delta_{\beta\gamma} - q_{i\gamma}), \end{aligned}$$

where $L_q(x) = \sum_{i\alpha} (x_{i\alpha} - q_{i\alpha}) \omega_{i\alpha}(x)$ is the adjoint potential of (3.9).

In a certain sense, this lemma can be viewed as the stochastic analogue of Lemma 3.1 (which is recovered immediately if we set $\sigma = 0$). However, it also shows that the stochastic situation is much more intricate than the deterministic one. For example, if q is a Wardrop equilibrium, (4.13) gives:

$$(4.14) \quad \mathcal{L}H_q(q; \lambda) = \frac{1}{2} \sum_i \frac{\lambda_i}{\rho_i} \sum_{\beta, \gamma}^i q_{i\beta} (\rho_i \delta_{\beta\gamma} - q_{i\gamma}) \sigma_{\beta\gamma}^2,$$

and if we focus on user $i \in \mathcal{N}$, we readily obtain:

$$(4.15) \quad \begin{aligned} \sum_{\beta, \gamma}^i q_{i\beta} (\rho_i \delta_{\beta\gamma} - q_{i\gamma}) \sigma_{\beta\gamma}^2 &= \sum_{\beta, \gamma}^i q_{i\beta} (\rho_i \delta_{\beta\gamma} - q_{i\gamma}) \sum_r P_{r\beta}^i P_{r\gamma}^i \sigma_r^2 \\ &= \sum_r \sigma_r^2 \rho_i y_{ir} - \sum_r \sigma_r^2 y_{ir}^2 = \sum_r \sigma_r^2 y_{ir} (\rho_i - y_{ir}), \end{aligned}$$

where $y_{ir} = \sum_r P_{r\alpha}^i q_{i\alpha} \leq \rho_i$ is the load induced on edge $r \in \mathcal{E}$ by the i -th user. This shows that (4.14) is *positive* if at least one user mixes his routes, thus ruling out negative definiteness (even semi-definiteness) for $\mathcal{L}H_q$.

In view of the above, unconditional convergence to Wardrop equilibrium appears to be a “bridge too far” in our stochastic environment, especially when the equilibrium in question is not pure – after all, mixed equilibria are not even traps of (4.4). This leads us to the notion of *stochastic stability*:

DEFINITION 4.2 (Arnold, 1974; Gikhman and Skorokhod, 1971). We will say that $q \in \mathbb{R}^n$ is *stochastically asymptotically stable* with respect to the process $X(t)$ when, for every neighbourhood U of q and every $\varepsilon > 0$, there exists a neighbourhood V of q such that:

$$(4.16) \quad \mathbf{P}_x \left\{ X(t) \in U \text{ for all } t \geq 0 \text{ and } \lim_{t \rightarrow \infty} X(t) = q \right\} \geq 1 - \varepsilon.$$

for all initial conditions $X(0) = x \in V$.

As we mentioned before, the notion of stochastic stability features prominently in the analysis of stochastically perturbed evolutionary or Nash-type games because it is precisely the type of stability that the strict equilibria of these games exhibit (Imhof, 2005; Mertikopoulos and Moustakas, 2010). Motivated by these results, we are finally in a position to state our analogue of the folk theorem for stochastically perturbed congestion models:

THEOREM 4.3. *Strict Wardrop equilibria are stochastically asymptotically stable in the replicator dynamics (4.4).*

PROOF. By relabeling indices if necessary, assume that $q = \sum_i \rho_i e_{i,0}$ is the strict Wardrop equilibrium of \mathfrak{C} . Inspired by the deterministic setting and the original idea of Imhof (2005), we will show that H_q is a local stochastic Lyapunov function, i.e. that $\mathcal{L}H_q(x) \leq -kH_q(x)$ for some $k > 0$ and for all x sufficiently close to q . Our result will then follow from Theorem 4 in Gikhman and Skorokhod (1971, pp. 314–315).

To that end, consider a perturbed flow $x = \sum_{i,\alpha} x_{i\alpha} q_{i\alpha}$ close to q :

$$(4.17) \quad x_{i,0} = \rho_i(1 - \varepsilon_i), \quad x_{i\mu} = \varepsilon_i \xi_{i\mu} \text{ for } \mu = 1, 2 \dots \in \mathcal{A}_i \setminus \{0\}$$

where $\varepsilon_i > 0$ controls the L^1 distance between x_i and q_i , and ξ_i is a point in the face of Δ_i lying opposite to q (i.e. $\xi_{i\mu} \geq 0$ and $\sum_\mu \xi_{i\mu} = \rho_i$). Then, in view of Lemma A.1 in Appendix A, the adjoint potential L_q will be bounded below by:

$$(4.18) \quad L_q(x) \geq \sum_i \rho_i \varepsilon_i \Delta \omega_i,$$

where $\Delta \omega_i = \min_\mu \{\omega_{i\mu}(q) - \omega_{i,0}(q)\} > 0$ (recall that q is a strict Wardrop equilibrium). Therefore, since the second term of $\mathcal{L}H_q(x; \lambda)$ in (4.13) is clearly of order $\mathcal{O}(\varepsilon^2)$, we obtain:

$$(4.19) \quad \mathcal{L}H_q(x; \lambda) \leq - \sum_i \rho_i \varepsilon_i \Delta \omega_i + \mathcal{O}(\varepsilon^2)$$

where $\varepsilon^2 = \sum_i \varepsilon_i^2$. On the other hand, we also have:

$$(4.20) \quad H_q(x; \lambda) = \sum_i \frac{\rho_i}{\lambda_i} \log \frac{\rho_i}{x_{i,0}} = - \sum_i \frac{\rho_i}{\lambda_i} \log(1 - \varepsilon_i) = \sum_i \frac{\rho_i}{\lambda_i} \varepsilon_i + \mathcal{O}(\varepsilon^2).$$

Thus, if we pick some positive $k < \min_i \{\lambda_i \Delta \omega_i\}$, some elementary algebra gives:

$$(4.21) \quad \mathcal{L}H_q(x; \lambda) \leq -k \sum_i \frac{\rho_i}{\lambda_i} \varepsilon_i + \mathcal{O}(\varepsilon^2) = -kH_q(x; \lambda) + \mathcal{O}(\varepsilon^2),$$

thus showing that (4.13) holds whenever ε is small enough. \square

In other words, Theorem 4.3 implies that trajectories which start sufficiently close to a strict equilibrium will remain in its vicinity and will eventually converge to it with arbitrarily high probability. Nonetheless, this is a *local* result: if the users' initial traffic distribution is not close to a strict equilibrium itself, Theorem 4.3 does not apply; specifically, if $X(0)$ is an arbitrary initial condition in Δ , we cannot even tell if the trajectory $X(t)$ will *ever* approach q .

To put this in more precise terms, it will be convenient to measure distances in Δ with the L^1 -norm: $\|\sum_\alpha z_\alpha e_\alpha\|_1 = \sum_\alpha |z_\alpha|$. In this norm, it is not too hard to see that Δ has a diameter of $2 \sum_i \rho_i$, so pick some positive $\delta < 2 \sum_i \rho_i$ and let $K_\delta = \{x \in \Delta : \|x - q\|_1 \leq \delta\}$ be the corresponding compact neighbourhood of q . Then, to see if $X(t)$ ever hits K_δ , we will examine the hitting time τ_δ :

$$(4.22) \quad \tau_\delta \equiv \tau_{K_\delta} = \inf\{t > 0 : X(t) \in K_\delta\} = \inf\{t > 0 : \|X(t) - q\|_1 \leq \delta\}.$$

Thereby, our chief concern is this: is the hitting time τ_δ finite with high probability? And if it is, is its expected value also finite?

To make our lives easier, let us consider the collective expressions:

$$(4.23) \quad \begin{aligned} \rho &= \sum_i \rho_i, & \Delta\omega &= \rho^{-1} \sum_i \rho_i \Delta\omega_i \\ \sigma^2 &= \sum_r \sigma_r^2, & \lambda &= \rho^{-1} \sum_i \rho_i \lambda_i. \end{aligned}$$

where $\Delta\omega_i = \min_{\mu \neq \alpha_i} \{\omega_{i\mu}(q) - \omega_{i\alpha_i}(q)\} > 0$ is the minimum delay difference between a user's equilibrium path α_i and his other choices. We then have:

THEOREM 4.4. *Let $q = \sum_i \rho_i e_{i\alpha_i}$ be a strict Wardrop equilibrium of a congestion model \mathfrak{C} and assume that the users' learning rates satisfy the condition:*

$$(4.24) \quad \lambda\sigma^2 < \Delta\omega.$$

Then, for any $\delta < 2\rho$ and any initial condition $X(0) = x \in \Delta$ with finite relative entropy $H_q(x; \lambda) < \infty$, the hitting time τ_δ has finite mean:

$$(4.25) \quad \mathbf{E}_x[\tau_\delta] \leq \frac{2H_q(x; \lambda)}{\Delta\omega} \frac{2\rho}{\delta(2\rho - \delta)}.$$

PROOF. As in the case of Theorem 4.3, our proof hinges on the expression:

$$(4.26) \quad -\mathcal{L}H_q(x; \lambda) = L_q(x) - \frac{1}{2} \sum_i \frac{\lambda_i}{\rho_i} \sum_{\beta, \gamma}^i \sigma_{\beta\gamma}^2 (x_{i\beta} - q_{i\beta})(x_{i\gamma} - q_{i\gamma}),$$

where $q = \sum_i \rho_i e_{i,0}$ is the strict equilibrium in question. In particular, set $x = q + \theta z$, where $z = \sum_{i,\alpha} z_{i\alpha} e_{i\alpha} \in T_q \Delta$ is the ‘‘inward’’ direction:

$$(4.27) \quad z_{i,0} = -\rho_i, \quad z_{i\mu} \geq 0 \text{ for } \mu = 1, 2 \dots \in \mathcal{A}_i^* \equiv \mathcal{A}_i \setminus \{0\} \text{ and } \sum_\mu z_{i\mu} = \rho_i.$$

Then, regarding the first term of (4.26), Lemma A.1 in Appendix A readily yields:

$$(4.28) \quad L_q(q + \theta z) \geq \Phi(q + \theta z) - \Phi(q) \geq \theta \sum_i \rho_i \Delta\omega_i \text{ for all } \theta \in [0, 1].$$

In a similar vein, the second term of (4.26) becomes:

$$(4.29) \quad \begin{aligned} \sum_{\beta, \gamma}^i \sigma_{\beta\gamma}^2 (x_{i\beta} - q_{i\beta})(x_{i\gamma} - q_{i\gamma}) &= \theta^2 \sum_{\beta, \gamma}^i \sigma_{\beta\gamma}^2 z_{i\beta} z_{i\gamma} \\ &= \theta^2 \sum_r \sum_{\beta, \gamma}^i \sigma_r^2 P_{r\beta}^i P_{r\gamma}^i z_{i\beta} z_{i\gamma} = \theta^2 \sum_r \sigma_r^2 w_{ir}^2, \end{aligned}$$

where $w_i = P^i(z)$. Since $w_{ir} \leq \rho_i$ for all $r \in \mathcal{E}$, we then obtain the inequality:

$$(4.30) \quad -\mathcal{L}H_q(x; \lambda) \geq \theta \sum_i \rho_i \Delta\omega_i - \frac{\theta^2}{2} \sigma^2 \sum_i \lambda_i \rho_i = (\rho \Delta\omega) \theta - \frac{\theta^2}{2} \rho \lambda \sigma^2$$

where ρ, σ^2 and $\lambda, \Delta\omega$ are the respective aggregates and averages of (4.23).

Suppose now that the rates λ_i satisfy (4.24), i.e. $\lambda \sigma^2 < \Delta\omega$. In that case, the RHS of (4.30) will be increasing for all $\theta \in [0, 1]$ and we will have:

$$(4.31) \quad -\mathcal{L}H_q(x; \lambda) \geq \rho \left[\theta \Delta\omega - \frac{1}{2} \lambda \sigma^2 \theta^2 \right] > 0$$

for all x with $\|x - q\|_1 \geq \theta \|z\|_1 = 2\theta \sum_i \rho_i = 2\rho\theta$. So, if $\|x - q\|_1 \geq \delta$, we get:

$$(4.32) \quad -\mathcal{L}H_q(x; \lambda) \geq \frac{\delta}{2} \Delta\omega - \frac{1}{2} \frac{\delta^2}{4\rho^2} \sigma^2 \rho \lambda \geq \frac{\delta}{2} \Delta\omega \left(1 - \frac{\delta}{2\rho} \right) > 0.$$

Therefore, if K_δ is the compact neighbourhood $K_\delta = \{x \in \Delta : \|x - q\|_1 \leq \delta\}$, we will have $\mathcal{L}H_q(x) \leq -\frac{\delta}{2} \Delta\omega \left(1 - \frac{\delta}{2\rho} \right) < 0$ for all $x \notin K_\delta$. Then, by a simple (but very useful!) estimate of Durrett (1996, Theorem 5.3 in page 268) we get:

$$(4.33) \quad \mathbf{E}_x[\tau_\delta] \leq \frac{2H_q(x; \lambda)}{\Delta\omega(1 - \delta/2\rho)} \frac{1}{\delta}. \quad \square$$

Recall now that Theorem 4.3 ensures that a trajectory $X(t)$ which starts sufficiently close to a strict equilibrium q will converge to q with arbitrarily high probability. Therefore, since Theorem 4.4 shows that $X(t)$ will come arbitrarily close to q in finite time, a tandem application of these two theorems yields:

COROLLARY 4.5. *If q is a strict equilibrium of \mathfrak{C} and the players' learning rates λ_i satisfy (4.24), the trajectories $X(t)$ converge to q almost surely.*

Of course, if a strict Wardrop equilibrium exists, then it is the unique equilibrium of the game (Proposition 2.7). In that case, Corollary 4.5 is the stochastic counterpart of Theorem 3.2: if the players' learning rates are soft enough compared to the level of the noise, then the solution orbits of the stochastic replicator dynamics (4.2) converge to a stationary traffic distribution almost surely. A few remarks are thus in order:

REMARK 1 (Temperance and Temperature). Condition (4.24) shows that the replicator dynamics reward *patience*: players who take their time in learning the game manage to weed out the noise and eventually converge to equilibrium. This begs to be compared with the (inverse) temperature analogy for the learning rates: if the "learning temperature" $T = 1/\lambda$ is too low, the players' learning scheme becomes very rigid and this intemperance amplifies any random variations in the experienced delays. On the other hand, when the temperature rises above the threshold $T_c = \sigma^2/\Delta\omega$, the stochastic fluctuations are toned down and the deterministic drift draws users to equilibrium.

REMARK 2. Admittedly, the form of (4.25) is a bit opaque for practical purposes. To lighten it up, note that we are only interested in small δ , so the term $2\rho/(2\rho - \delta)$ may be ignored to leading order. Therefore, if we also assume for simplicity that all players learn at the same rate $\lambda_i = \lambda$, we get:

$$(4.34) \quad \mathbf{E}_x[\tau_\delta] \leq \frac{2h}{\lambda\Delta\omega} \frac{\rho}{\delta}$$

where $h = \rho^{-1} \sum_i \rho_i \log(\rho_i/x_i)$ is the "average" Kullback-Leibler distance between x and q .

This very rough estimate is pretty illuminating on its own. First and foremost, it shows that our bound for $\mathbf{E}_x[\tau_\delta]$ is inversely proportional to the learning rate λ , much the same as in the deterministic setting where λ essentially rescales time to λt . Moreover, because ρ and δ are both $\mathcal{O}(N)$, one might be tempted to think that our time estimates are *intensive*, i.e. independent of N . However, since delays increase (nonlinearly even) with the aggregate load ρ , the dependence on N is actually hidden in $\Delta\omega$ – this also shows that the learning rates λ_i do not have to be $\mathcal{O}(1/|\mathcal{E}|)$ -small in order to satisfy (4.24).

In any event, since strict equilibria do not always exist, we should return to the generic case of *interior* equilibria $q \in \text{Int}(\Delta)$. We have already seen that these equilibria are not very well-behaved in stochastic environments: they are not stationary in (4.4) and (4.15) shows that $\mathcal{L}H_q$ is actually *positive* in their vicinity. Despite all that, if the network \mathcal{Q} is irreducible and the users' learning rates λ_i are slow enough, we will see that *the replicator dynamics (4.4) admit a finite invariant measure which concentrates mass around the (unique) equilibrium of \mathfrak{C} .*

To state this result precisely, a little more groundwork is required. First off, akin to the case of strict equilibria, it will be convenient to measure distances from q with a scaled variant of the L^1 norm. In particular, let $S_q = \{z \in T_q\Delta : q + z \in \text{bd}(\Delta)\}$. 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:

$$(4.35) \quad \Theta_q(x) = \theta \Leftrightarrow x = q + \theta z \text{ for some } z \in S_q \text{ and } 0 \leq \theta \leq 1.$$

Θ_q is not a bona fide distance function by itself, but it closely resembles the L^1 norm: the ‘‘projective balls’’ $B_\theta = \{x : \Theta_q(x) \leq \theta\}$ are rescaled copies of Δ (S_q is the ‘‘unit sphere’’ in this picture), and the graph $\text{gr}(\Theta_q) \equiv \{(x, \theta) \in \Delta \times \mathbb{R} : \Theta_q(x) = \theta\}$ of Θ_q is simply a cone over the polytope Δ .

In a similar vein, we define the *essence* of a point $q \in \Delta$ to be:

$$(4.36) \quad \text{ess}(q) = \rho^{-1} \min \{\|P(z)\| : z \in S_q\},$$

where $\|\cdot\|$ denotes the ordinary Euclidean norm and the factor of ρ was included for scaling purposes. Comparably to $\text{red}(\mathcal{Q})$, $\text{ess}(q)$ measures redundancy (or rather, the lack thereof): $\text{ess}(q) = 0$ only if some direction $z \in S_q$ is null for P , i.e. only if \mathcal{Q} is reducible.²

We are finally in a position to state and prove:

THEOREM 4.6. *Let $q \in \text{Int}(\Delta)$ be an interior equilibrium of an irreducible congestion model \mathfrak{C} , and assume that the users’ learning rates satisfy the condition:*

$$(4.37) \quad \lambda < \frac{4m\rho\kappa^2}{5\sigma^2}, \text{ where } m = \inf\{\phi'_r(y_r) : r \in \mathcal{E}, y \in P(\Delta)\} \text{ and } \kappa = \text{ess}(q).$$

Then, for any interior initial condition $X(0) = x \in \text{Int}(\Delta)$, the trajectories $X(t)$ are recurrent (a.s.) and their time averages are concentrated in a neighbourhood of q . Specifically, if $\Theta_q(\cdot)$ denotes the projective distance (4.35) from q , then:

$$(4.38) \quad \mathbf{E}_x \left[\frac{1}{t} \int_0^t \Theta_q^2(X(s)) ds \right] \leq \theta_\lambda^2 + \mathcal{O}(1/t), \text{ where } \theta_\lambda^2 = \frac{1}{4} \left(\frac{m\rho\kappa^2}{\lambda\sigma^2} - 1 \right)^{-1}.$$

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 $B_\theta = \{x \in \Delta : \Theta_q(x) \leq \theta\}$ is a ‘‘projective ball’’ around q , we have:

$$(4.39) \quad \pi(B_\theta) \geq 1 - \theta_\lambda^2/\theta^2.$$

Following Bhattacharya (1978), recurrence here means that for every $\xi \in \text{Int}(\Delta)$ and every neighbourhood U_ξ of ξ , the diffusion $X(t)$ has the property:

$$(4.40) \quad \mathbf{P}_x\{X(t_k) \in U_\xi\} = 1,$$

²With some trivial modifications, these definitions can be extended to points in $\text{bd}(\Delta)$ as well. In particular, if q is strict we get $\Theta_{q_i}(x_i) = \|z_i\|_1/2\rho_i$ and, as a consequence of Proposition 2.7, we will have $\text{ess}(q) > 0$ irrespective of whether \mathcal{Q} is reducible or not.

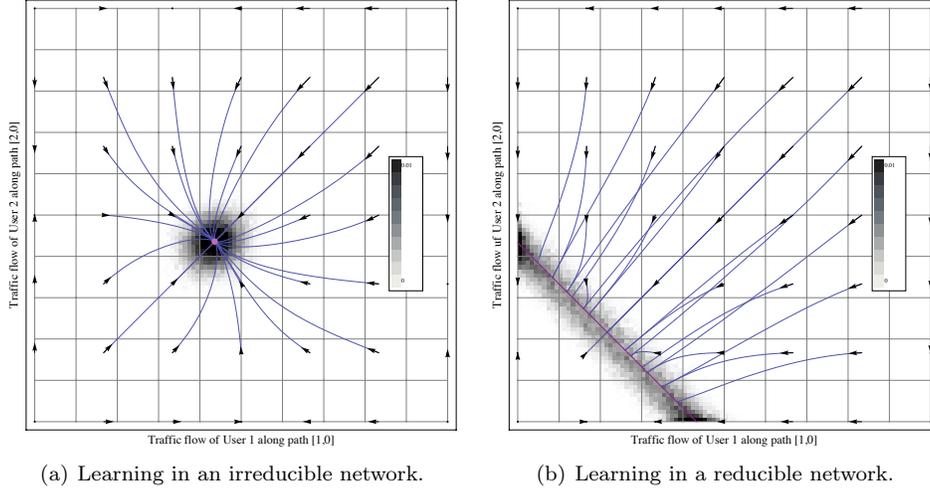


FIG 4. Learning in the networks of Figure 1 with $M/M/1$ latency functions $\phi_r(y_r) = (\mu_r - y_r)^{-1}$ and arbitrarily chosen capacities μ_r . The shades of gray represent the invariant distribution of (4.4) (obtained by numerical simulations) and the flow lines (blue) are the solution trajectories of (3.5) – in (b) they are actually projections because there is a third user as well.

for some sequence of (random) times t_k that increases to infinity. Hence, using the recurrence criteria of Bhattacharya (1978), we will prove our claim by showing that $X(t)$ hits a compact neighbourhood of q in finite time (this is the hard part), and that the generator of a suitably transformed process is elliptic.

PROOF OF THEOREM 4.6. 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 “sphere” $S_q = \{z' \in T_q\Delta : q + z' \in \text{bd}(\Delta)\}$. In this manner, (4.13) becomes:

$$(4.41) \quad -\mathcal{L}H_q(x; \lambda) = L_q(q + \theta z) - \frac{1}{2} \sum_i \frac{\lambda_i}{\rho_i} \theta^2 \sum_{\beta, \gamma}^i \sigma_{\beta\gamma}^2 z_{i\beta} z_{i\gamma} - \frac{1}{2} \sum_i \frac{\lambda_i}{\rho_i} \sum_{\beta, \gamma}^i \sigma_{\beta\gamma}^2 q_{i\beta} (\rho_i \delta_{\beta\gamma} - q_{i\gamma}).$$

With regards to the first term of (4.41), Lemma A.2 in Appendix A and the definition (4.36) 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, we have already seen in the proof of Theorem 4.4 that the second term of (4.41) is bounded above:

$$(4.42) \quad \frac{1}{2} \sum_i \frac{\lambda_i}{\rho_i} \theta^2 \sum_{\beta, \gamma}^i \sigma_{\beta\gamma}^2 z_{i\beta} z_{i\gamma} \leq \frac{1}{2} \rho \lambda \sigma^2 \theta^2.$$

We are thus left to estimate the last term of (4.41). To that end, (4.15) gives:

$$(4.43) \quad \sum_{\beta, \gamma}^i q_{i\beta} (\rho_i \delta_{\beta\gamma} - q_{i\gamma}) \sigma_{\beta\gamma}^2 = \sum_r \sigma_r^2 y_{ir} (\rho_i - y_{ir}) \leq \frac{1}{4} \rho_i^2 \sigma^2$$

where the last inequality stems from the bound $y_{ir}(\rho_i - y_{ir}) \leq \frac{1}{4} \rho_i^2$ (recall that $0 \leq y_{ir} \leq \rho_i$). Combining all of the above, we then get:

$$(4.44) \quad -\mathcal{L}H_q(x; \lambda) \geq \frac{1}{2} m \kappa^2 \rho^2 \theta^2 - \frac{1}{2} \rho \lambda \sigma^2 \theta^2 - \frac{1}{8} \rho \lambda \sigma^2 \equiv g(\theta), \quad 0 \leq \theta \leq 1.$$

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 (4.44) 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 : $\theta_\lambda = \frac{1}{2} (\lambda_0 / \lambda - 1)^{-1/2}$.

So, pick some positive $a < g(1) = \frac{1}{2} \rho \sigma^2 (\lambda_0 - \frac{5}{4} \lambda)$ and consider the set $K_a = \{q + \theta z : z \in S_q, g(\theta) \leq a\}$. By construction, K_a is a compact neighbourhood of q which does not intersect $\text{bd}(\Delta)$ and, by (4.44), we have $\mathcal{L}H_q(x; \lambda) \leq -a$ outside K_a . Therefore, if $\tau_a \equiv \tau_{K_a}$ denotes the hitting time $\tau_a = \inf\{t : X(t) \in K_a\}$, Theorem 5.3 in Durrett (1996, p. 268) yields:

$$(4.45) \quad \mathbf{E}_x[\tau_a] \leq \frac{H_q(x; \lambda)}{a} < \infty$$

for every interior initial condition $X(0) = x \in \text{Int}(\Delta)$.

Inspired by a trick of Imhof (2005), let us now consider the transformed process $Y(t) = \Psi(X(t))$ given by $\Psi_{i\mu}(x) = \log x_{i\mu} / x_{i,0}$, $\mu \in \mathcal{A}_i^* \equiv \mathcal{A}_i \setminus \{\alpha_{i,0}\}$. 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:

$$(4.46) \quad dY_{i\mu} = \mathcal{L}\Psi_{i\mu}(X) dt + dU_{i\mu} - dU_{i,0} = \mathcal{L}\Psi_{i\mu}(X) dt + \sum_r Q_{r\mu}^i \sigma_r dW_r,$$

where $Q_{r\mu}^i = P_{r\mu}^i - P_{r,0}^i$ are the components of the redundancy matrix Q of \mathcal{Q} in the basis $\tilde{e}_{i\mu} = e_{i\mu} - e_{i,0}$ of $T_q\Delta$ – recall also (2.14) and the relevant discussion in Section 2.2.

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:

$$(4.47) \quad \langle Az, Az \rangle = \sum_{\mu, \nu} (AA^T)_{\mu\nu} z_\mu z_\nu = \sum_{\mu, \nu} \sum_r Q_{r\mu} Q_{r\nu} \sigma_r^2 z_\mu z_\nu = \sum_r \sigma_r^2 w_r^2,$$

where $w = Q(z)$. Since \mathcal{Q} is irreducible, we will have $w \neq 0$, and in view of (4.47) above, this proves our assertion.

We have thus shown that the process $Y(t)$ hits a compact neighbourhood of $\Psi(q)$ in finite time (on average), and also that the generator of Y is elliptic. From the criteria of Bhattacharya (1978, Lemma 3.4) it follows that Y is recurrent, and since Ψ is invertible in $\text{Int}(\Delta)$, 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 estimate (4.38), note that Dynkin's formula (see e.g. Øksendal, 2007, Theorem 7.4.1) applied to (4.44) yields:

$$(4.48) \quad \begin{aligned} \mathbf{E}_x [H_q(X(t); \lambda)] &= H_q(x; \lambda) + \mathbf{E}_x \left[\int_0^t \mathcal{L} H_q(X(s); \lambda) ds \right] \\ &\leq H_q(x; \lambda) - \frac{1}{2} \rho \sigma^2 (\lambda_0 - \lambda) \mathbf{E}_x \left[\int_0^t \Theta_q^2(X(s)) ds \right] + \frac{1}{8} \rho \lambda \sigma^2 t, \end{aligned}$$

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

$$(4.49) \quad \mathbf{E}_x \left[\frac{1}{t} \int_0^t \Theta_q^2(X(s)) ds \right] \leq \theta_\lambda^2 + \frac{C}{t}, \text{ where } C = \frac{2}{\rho \sigma^2} \frac{H_q(x; \lambda)}{\lambda_0 - \lambda}.$$

We are thus left to establish the bound $\pi(B_\theta) \geq 1 - \theta_\lambda^2/\theta^2$ which shows that the invariant measure π concentrates its mass around the ‘‘projective balls’’ B_θ . For that, we will use the ergodic property of $X(t)$, namely that:

$$(4.50) \quad \pi(B_\theta) = \lim_{t \rightarrow \infty} \mathbf{E}_x \left[\frac{1}{t} \int_0^t \chi_{B_\theta}(X(s)) ds \right],$$

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:

$$(4.51) \quad \mathbf{E}_x \left[\frac{1}{t} \int_0^t \chi_{B_\theta}(X(s)) ds \right] \geq \mathbf{E}_x \left[\frac{1}{t} \int_0^t \left(1 - \Theta_q^2(X(s))/\theta^2 \right) ds \right]$$

and the bound (4.39) follows by letting $t \rightarrow \infty$ in (4.49). \square

We conclude this section with a few remarks on our results so far:

REMARK 1 (*Learning vs. Noise*). The nature of our bounds reveals a most interesting feature of the replicator equation (4.4). On the one hand, as $\lambda \rightarrow 0$, we also get $\theta_\lambda \rightarrow 0$ and the invariant measure π converges vaguely to a point mass at q . Hence, if the learning rate λ is slow enough (or if the noise σ is low enough), we recover Theorem 3.2 (as we should!). On the other hand, there is a clear downside to using very slow learning rates: the expected time to hit a neighbourhood of an equilibrium is inversely proportional to λ . As a result, choosing learning rates is a delicate process and users will have to balance the rate versus the desired sharpness of their convergence.

REMARK 2 (*The Effect of Redundancy*). The irreducibility assumption is actually quite important: it appears both in the ‘‘slow-learning’’ condition (4.37) (recall that $\text{ess}(q) = 0$ if q is an interior point of a reducible network) and also in the proof

that the generator of $Y(t) = \Psi(X(t))$ is elliptic. This shows that the stochastic dynamics (4.4) are not oblivious to redundant degrees of freedom, in stark contrast with the deterministic case (Theorem 3.2).

Regardless, we expect that an analogue for Theorem 4.6 still holds for reducible networks if we replace q with the entire (affine) set Δ^* . 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^* (see Fig. 4(b)). One way to prove this claim would be to find a suitable way to “quotient out” $\ker Q$ but, since the replicator equation (4.4) is not invariant over the redundant 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 (4.37) can be tightened and the assumption that $\phi'_r > 0$ can actually be dropped. In that case however, the corresponding expressions would become significantly more complicated without adding much essence, so we have opted to keep our analysis focused on the simpler estimates.

5. Discussion. In this last section, we will discuss some issues that have not been thoroughly addressed in the rest of the paper and provide some directions for future work.

Learning and Optimality. We have already noted that the traffic flows which minimise the aggregate latency $\omega(x) = \sum_i \rho_i \omega_i(x)$ in a network correspond precisely to the Wardrop equilibria of a congestion model which is defined over the same network and whose delay functions are given by the “marginal latencies” $\phi_r^*(y_r) = \phi_r(y_r) + y_r \phi'_r(y_r)$ (see e.g. Roughgarden and Tardos, 2002). Hence, if we set $\omega_{i\alpha}^*(x) = \sum_{\alpha}^i P_{r\alpha}^i \phi_r^*(y_r)$ and substitute $\omega_{i\alpha}^*$ instead of $\omega_{i\alpha}$ in the replicator dynamics (3.5) and (4.4), our analysis yields:

THEOREM 5.1. *Let $\mathfrak{C} \equiv \mathfrak{C}(Q, \phi)$ be a congestion model with strictly convex latency functions ϕ_r , $r \in \mathcal{E}$, and assume that users follow a replicator learning scheme with cost functions $\omega_{i\alpha}^*$. Then:*

1. *In the deterministic case (3.5), players converge to a traffic flow which minimises the aggregate delay $\omega(x) = \sum \rho_i \omega_i(x)$.*
2. *In the stochastic case (4.4), if the network is irreducible and the players’ learning rates are slow enough, their time-averaged flows will be concentrated near the (necessarily unique) optimal distribution q which minimises ω .*

Of course, for a sharper statement one need only reformulate Theorems 3.2, 4.3, and 4.6 accordingly (the convexity of ϕ_r^* replaces the monotonicity requirement for ϕ_r). The only thing worthy of note here 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 publicised to all users that employ the edge

$r \in \mathcal{E}$. Consequently, if the administrators 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.

Equilibrium Classes. In a certain sense, interior and strict equilibria represent the extreme ends of the Wardrop spectrum, so it was a reasonable choice to focus our analysis on them. Nevertheless, there are equilibrium classes that we did not consider: for instance, there are pure Wardrop equilibria which are not strict, or there could be “quasi-strict” equilibria q in the boundary of Δ with the property that $\omega_{i\alpha}(q) > \omega_i(q)$ for all α which are not present in q .

Strictly speaking, such equilibria are not covered by either Theorem 4.3 or Theorem 4.6. Still, by a suitable modification of our stochastic calculations, we may obtain similar convergence and stability results for these types of equilibria as well. For example, modulo a “slow-learning” condition similar to (4.37), it is easy to see that pure equilibria that are not strict are still stochastically stable. The reason we have opted not to consider all these special cases is that it would be too much trouble for little gain: the assortment of similar-looking results that we would obtain in this way would confuse things more than it would clarify them.

Exponential Learning. In the context of N -person Nash games, we have already mentioned that the replicator dynamics also arise as the result of an “exponential learning” process, itself a variant of logistic fictitious play (Fudenberg and Levine, 1998; Mertikopoulos and Moustakas, 2010). The way this scheme works is that players keep cumulative scores of their strategies' performance and they employ each strategy with a probability which is exponentially proportional to these scores. As such, it is not too hard to adapt this method directly to our congestion setting.

In more detail, assume that all users $i \in \mathcal{N}$ keep performance scores $V_{i\alpha}$ of the paths at their disposal as specified by the differential equation:

$$(5.1) \quad dV_{i\alpha}(t) = -\omega_{i\alpha}(x(t)) dt,$$

where $x(t)$ is the traffic profile at time t . Based on these scores, the users then update their traffic flows according to the Boltzmann distribution:

$$(5.2) \quad x_{i\alpha}(t) = \frac{e^{\lambda_i V_{i\alpha}(t)}}{\sum_{\beta}^i e^{\lambda_i V_{i\beta}(t)}},$$

where λ_i denotes the learning rate of player $i \in \mathcal{N}$ (this expression also explains why these rates can be seen as inverse temperatures). In this way, by decoupling these expressions, one obtains the deterministic replicator equation (3.5).

We thus see that exponential learning tells us nothing new in deterministic environments. In the presence of noise however, the scores $V_{i\alpha}$ also reflect any fluctuations in the observed delays, so we obtain instead:

$$(5.3) \quad dV_{i\alpha}(t) = -\omega_{i\alpha}(X) dt + \sigma_{i\alpha} dU_{i\alpha},$$

where, as in (4.3), $dU_{i\alpha}$ describes the total noise along the path $\alpha \in \mathcal{A}_i$. Therefore,

if the users' flow profile $X(t)$ is updated according to (5.2), Itô's lemma now gives:

$$(5.4) \quad dX_{i\alpha} = \lambda_i X_{i\alpha} [\omega_i(X) - \omega_{i\alpha}(X)] dt + \lambda_i X_{i\alpha} \left[dU_{i\alpha} - \rho_i^{-1} \sum_{\beta}^i X_{i\beta} dU_{i\beta} \right] \\ + \frac{\lambda_i^2}{2} X_{i\alpha} \left[\sum_{\beta}^i (\rho_i \delta_{\alpha\beta} - 2X_{i\beta}) \sigma_{\alpha\beta}^2 - \rho_i^{-1} \sum_{\beta, \gamma}^i \sigma_{\beta\gamma}^2 X_{i\gamma} (\rho_i \delta_{\beta\gamma} - 2X_{i\beta}) \right] dt.$$

As far as the rationality properties of these new dynamics are concerned, a simple modification in the proof of Theorem 4.3 suffices to show that strict Wardrop equilibria are stochastically stable in (5.4). Just the same, the extra drift term in (5.4) complicates things considerably, so results containing explicit estimates of hitting times are significantly harder to obtain. Of course, this approach might well lead to improved convergence rates, but since the calculations would take us too far afield, we prefer to postpone this analysis for the future.

The Brown-von Neumann-Nash Dynamics. Another powerful learning scheme is given by the Brown-von Neumann-Nash (BNN) dynamics (see e.g. Fudenberg and Levine, 1998) where users look at the “excess delays”

$$(5.5) \quad \psi_{i\alpha}(x) = [\omega_i(x) - \omega_{i\alpha}(x)]^+ = \max \{ \omega_i(x) - \omega_{i\alpha}(x), 0 \}$$

and update their traffic flows according to the differential equation:

$$(5.6) \quad \frac{dx_{i\alpha}}{dt} = \psi_{i\alpha}(x(t)) - \psi_i(x(t)),$$

where $\psi_i(x) = \rho_i^{-1} \sum_{\alpha}^i x_{i\alpha} \psi_{i\alpha}(x)$. On the negative side, these dynamics require users to monitor delays even along paths that they do not employ. On the other hand, they satisfy the pleasant property of “non-complacency” (Sandholm, 2001): the stationary states of (5.6) coincide with the game's Wardrop equilibria and every solution trajectory converges to a connected set of such equilibria.

In terms of convergence to a Wardrop equilibrium, Theorem 3.2 shows that the replicator dynamics behave at least as well as the BNN dynamics (except perhaps on the boundary of Δ), so there is no real reason to pick the more complicated expressions (5.5), (5.6). However, this might not be true in the presence of stochastic fluctuations: in fact, virtually nothing is known about the behaviour of the BNN dynamics in stochastic environments so this question alone makes pursuing this direction a worthwhile project.

Acknowledgements. We would like to extend our gratitude to M. Scarsini for a series of fruitful discussions on Braess's paradox which helped us clarify the differences between the equilibril conditions that arise in congestion models.

APPENDIX A: PROPERTIES OF THE ADJOINT POTENTIAL

We collect here some of the most useful properties of the *adjoint potential*:

$$(3.9) \quad L_q(x) \equiv \sum_{i, \alpha} (x_{i\alpha} - q_{i\alpha}) \omega_{i\alpha}(x) = \sum_r (y_r - y_r^*) \phi_r(y_r) \equiv \Lambda(y).$$

To begin with, the equality in (3.9) stems from the invariance identity:

$$(A.1) \quad \sum_{\alpha} z_{\alpha} \omega_{\alpha} = \sum_{\alpha} \sum_r z_{\alpha} P_{r\alpha} \phi_r = \sum_r w_r \phi_r, \quad z \in V$$

where $P : V \rightarrow W$ is the indicator matrix of the network \mathcal{Q} and $w = P(z)$. It is then easy to verify that $L_q(x) = L_q(x')$ whenever $x' - x \in \ker Q$ (thus justifying (3.9)), and also that $L_q = L_{q'}$ iff $q' - q \in \ker Q$. As a result, the notation $L_q(x) \equiv \Lambda(y)$ is consistent with any choice of $q \in \Delta^*$.

This ‘‘adjoint’’ potential owes its name to the formula for integration by parts:

$$(A.2) \quad \sum_r \int_{y_r^*}^{y_r} \phi_r(w) dw = \sum_r (y_r - y_r^*) \phi_r(y_r) - \sum_r \int_{y_r^*}^{y_r} w \phi_r'(w) dw.$$

Since the latencies ϕ_r are increasing, this expression immediately yields the estimate (3.9): $\Phi(y) - \Phi(y^*) \leq \Lambda(y)$. However, if we also assume that q is strict (say $q = \sum_i \rho_i e_{i,0}$ for convenience), we can get a more direct bound:

LEMMA A.1. *Let $q = \sum_i \rho_i e_{i,0}$ be a strict Wardrop equilibrium and let $z \in T_q \Delta$. Then, for all $t \geq 0$ such that $x(t) = q + tz \in \Delta$, we have:*

$$(A.3) \quad L_q(q + tz) \geq \frac{1}{2} \sum_i \Delta \omega_i \|z_i\|_1 t, \quad \text{where } \Delta \omega_i = \min_{\mu \neq 0} \{\omega_{i\mu}(q) - \omega_{i,0}(q)\}.$$

PROOF. Clearly, to have $q + tz \in \Delta$ for some $t > 0$, z must be of the form:

$$(A.4) \quad z = \sum_i z_i \text{ with } z_i = \sum_{\mu} z_{i\mu} (e_{i\mu} - e_{i,0}) \text{ and } z_{i\mu} \geq 0 \text{ for all } \mu \in \mathcal{A}_i^* \equiv \mathcal{A}_i \setminus \{0\}.$$

So, let $f(t) = \Phi(y(t))$ where $y(t) = P(x(t)) = y^* + tw$ and $w = P(z)$. With Φ convex, we get $f(t) \geq f(0) + f'(0)t$ and a simple differentiation yields: $f'(0) = \left. \frac{d}{dt} \right|_{t=0} \sum_r \Phi_r(y_r^* + tw_r) = \sum_r w_r \phi_r(y_r^*)$. However, thanks to (A.1) and (A.4), we may rewrite this sum as:

$$(A.5) \quad \begin{aligned} \sum_r w_r \phi_r(y_r^*) &= \sum_i \sum_{\alpha} z_{i\alpha} \omega_{i\alpha}(q) = \sum_i \sum_{\mu} z_{i\mu} [\omega_{i\mu}(q) - \omega_{i,0}(q)] \\ &\geq \sum_i \sum_{\mu} z_{i\mu} \Delta \omega_i = \frac{1}{2} \sum_i \Delta \omega_i \|z_i\|_1, \end{aligned}$$

because $\|z_i\|_1 = \sum_{\alpha} |z_{i\alpha}| = \left| -\sum_{\mu} z_{i\mu} \right| + \sum_{\mu} |z_{i\mu}| = 2 \sum_{\mu} z_{i\mu}$. We thus obtain:

$$(A.6) \quad L_q(q + tz) = \Lambda(y^* + tw) \geq f(t) - f(0) \geq \frac{1}{2} \sum_i \Delta \omega_i \|z_i\|_1 t. \quad \square$$

This lemma shows that L_q increases at least linearly along all ‘‘inward’’ rays $q + tz$.³ This is not so if q is an *interior* equilibrium:

³It is interesting to note here the relation with Proposition 2.7: if the ray $q + tz$ is inward-pointing, then z cannot be ‘‘redundant’’, i.e. we cannot have $z \in \ker P$.

LEMMA A.2. *Let $q \in \text{Int}(\Delta)$ be an interior Wardrop equilibrium and let $z \in T_q\Delta$. Then, for all $t \geq 0$ such that $x(t) = q + tz \in \Delta$, we have:*

$$(A.7) \quad L_q(q + tz) \geq \frac{1}{2}m\|P(z)\|^2t^2, \quad \text{where } m = \inf\{\phi'_r(y_r) : r \in \mathcal{E}, y \in P(\Delta)\}.$$

PROOF. Following the proof of Lemma A.1 above, we obtain:

$$(A.8) \quad f'(0) = \sum_i \sum_\alpha^i z_{i\alpha} \omega_{i\alpha}(q) = \sum_i \sum_\alpha^i z_{i\alpha} \omega_i(q) = 0$$

where the second equality follows from the fact that q is an interior equilibrium (that is, $\omega_{i\alpha}(q) = \omega_i(q)$ for all paths $\alpha \in \mathcal{A}_i$), and the last one is a consequence of z being tangent to Δ (meaning that $\sum_\alpha^i z_{i\alpha} = 0$). On the other hand, we also get:

$$(A.9) \quad f''(t) = \frac{d^2}{dt^2} \sum_r \Phi_r(y_r^* + tw_r) = \sum_r w_r^2 \phi'_r(y_r^* + tw_r).$$

Clearly, since the set $P(\Delta)$ of load profiles y is compact and the (continuous) functions ϕ'_r are positive, we will also have $m = \inf\{\phi'_r(y_r) : y \in P(\Delta), r \in \mathcal{E}\} > 0$. We will thus have $f(t) \geq \frac{1}{2}mt^2$, and a first order Taylor expansion with Lagrange remainder easily yields:

$$(A.10) \quad L_q(q + tz) = \Lambda(y^* + tw) \geq f(t) - f(0) \geq \frac{1}{2}m\|P(z)\|^2t^2. \quad \square$$

APPENDIX B: STOCHASTIC CALCULATIONS

This appendix is devoted to the calculations that are hidden under the hood of (4.13), the equation that describes the evolution of the relative entropy $H_q(x; \lambda)$.

PROOF OF LEMMA 4.1. Let $V_q(t) = H_q(X(t); \lambda)$. We then have:

$$(B.1) \quad \begin{aligned} dV_q &= \sum_{i,\alpha} \frac{\partial H_q}{\partial x_{i\alpha}} dX_{i\alpha} + \frac{1}{2} \sum_{i,\alpha} \sum_{j,\beta} \frac{\partial^2 H_q}{\partial x_{i\alpha} \partial x_{j\beta}} (dX_{i\alpha}) \cdot (dX_{j\beta}) \\ &= - \sum_{i,\alpha} \frac{1}{\lambda_i} \frac{q_{i\alpha}}{x_{i\alpha}} dX_{i\alpha} + \frac{1}{2} \sum_{i,\alpha} \frac{1}{\lambda_i} \frac{q_{i\alpha}}{x_{i\alpha}^2} (dX_{i\alpha})^2. \end{aligned}$$

However, with $X(t)$ being as in (4.4), we readily obtain:

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

As a result, we may combine the two equations (B.1) and (B.2) to obtain:

$$(B.3) \quad \begin{aligned} dV_q &= - \sum_{i,\alpha} q_{i\alpha} [\omega_i(X) - \omega_{i\alpha}(X)] dt - \sum_{i,\alpha} q_{i\alpha} \left[dU_{i\alpha} - \rho_i^{-1} \sum_\beta^i X_{i\beta} dU_{i\beta} \right] \\ &\quad + \frac{1}{2} \sum_{i,\alpha} \lambda_i q_{i\alpha} \left[\sigma_{i\alpha}^2 - \frac{2}{\rho_i} \sum_\beta^i \sigma_{\alpha\beta}^2 X_{i\beta} + \frac{1}{\rho_i^2} \sum_{\beta,\gamma}^i \sigma_{\beta\gamma}^2 X_{i\beta} X_{i\gamma} \right] dt. \end{aligned}$$

Therefore, if we focus at a particular user $i \in \mathcal{N}$, the last term of (B.3) gives:

$$\begin{aligned}
& \sum_{\alpha}^i q_{i\alpha} \left[\sigma_{i\alpha}^2 - \frac{2}{\rho_i} \sum_{\beta}^i \sigma_{\alpha\beta}^2 X_{i\beta} + \frac{1}{\rho_i^2} \sum_{\beta,\gamma}^i \sigma_{\beta\gamma}^2 X_{i\beta} X_{i\gamma} \right] \\
&= \sum_{\alpha}^i q_{i\alpha} \sigma_{i\alpha}^2 - \frac{2}{\rho_i} \sum_{\alpha,\beta}^i \sigma_{\alpha\beta}^2 q_{i\alpha} X_{i\beta} + \frac{1}{\rho_i} \sum_{\beta,\gamma}^i \sigma_{\beta\gamma}^2 X_{i\beta} X_{i\gamma} \\
&= \sum_{\alpha}^i q_{i\alpha} \sigma_{i\alpha}^2 - \frac{1}{\rho_i} \sum_{\beta,\gamma}^i q_{i\beta} q_{i\gamma} \sigma_{\beta\gamma}^2 \\
&+ \frac{1}{\rho_i} \left[\sum_{\beta,\gamma}^i q_{i\beta} q_{i\gamma} \sigma_{\beta\gamma}^2 - 2 \sum_{\beta,\gamma}^i X_{i\beta} q_{i\gamma} \sigma_{\beta\gamma}^2 + \sum_{\beta,\gamma}^i X_{i\beta} X_{i\gamma} \sigma_{\beta\gamma}^2 \right] \\
\text{(B.4)} \quad &= \frac{1}{\rho_i} \left[\sum_{\beta,\gamma}^i q_{i\beta} (\rho_i \delta_{\beta\gamma} - q_{i\gamma}) \sigma_{\beta\gamma}^2 + \sum_{\beta,\gamma}^i \sigma_{\beta\gamma}^2 (X_{i\beta} - q_{i\beta})(X_{i\gamma} - q_{i\gamma}) \right]
\end{aligned}$$

and the lemma follows by substituting (B.4) into (B.3) and keeping only the resulting drift, that is, the first and third terms of (B.3). \square

REFERENCES

- ALI KHAN, M. (1986). Equilibrium points of nonatomic games over a Banach space. *Transactions of the American Mathematical Society* **293** 737–749. 8
- ARNOLD, L. (1974). *Stochastic Differential Equations: Theory and Applications*. Wiley. 25
- BECKMANN, M., MCGUIRE, C. B. and WINSTEN, C. (1956). *Studies in the Economics of Transportation*. Yale University Press. 1, 2, 13, 14
- BENAÏM, M., HOFBAUER, J. and SANDHOLM, W. H. (2008). Robust Permanence and Impermanence for Stochastic Replicator Dynamics. *Journal of Biological Dynamics* **2** 180–195. 3
- BHATTACHARYA, R. N. (1978). Criteria for recurrence and existence of invariant measures for multidimensional diffusions. *The Annals of Probability* **6** 541–553. 29, 30, 31
- BÖRGER, T. and SARIN, R. (1997). Learning through reinforcement and replicator dynamics. *Journal of Economic Theory* **77** 1–14. 17
- BRAESS, D. (1968). Über ein paradoxon der verkehrsplanung. *Unternehmensforschung* **12** 258–268. 14
- CABRALES, A. (2000). Stochastic Replicator Dynamics. *International Economic Review* **41** 451–81. 3, 22
- CRESSMAN, R. (2005). Stability of the Replicator Equation with Continuous Strategy Space. *Mathematical Social Sciences* **50** 127–147. 8
- DAFERMOS, S. C. and SPARROW, F. T. (1969). The traffic assignment problem for a general network. *Journal of Research of the National Bureau of Standards* **73B** 91–118. 1, 13
- DURRETT, R. (1996). *Stochastic Calculus*. CRC Press, Boca Raton, FL. 27, 31
- FISCHER, S. and VÖCKING, B. (2004). On the Evolution of Selfish Routing. In *Proceedings of the 12th European Symposium on Algorithms*. 2, 13, 20
- FUDENBERG, D. and HARRIS, C. (1992). Evolutionary dynamics with aggregate shocks. *Journal of Economic Theory* **57** 420–441. 3, 22
- FUDENBERG, D. and LEVINE, D. K. (1998). *The Theory of Learning in Games*. MIT Press Series on Economic Learning and Social Evolution **2**. The MIT Press. 22, 34, 35
- GIKHMAN, I. I. and SKOROKHOD, A. V. (1971). *Stochastische Differentialgleichungen*. Akademie-Verlag. 25
- HOFBAUER, J. and IMHOF, L. A. (2009). Time averages, recurrence and transience in the stochastic replicator dynamics. *The Annals of Applied Probability* **19** 1347–1368. 3
- HOFBAUER, J., OECHSSLER, J. and RIEDEL, F. (2009). Brown-von Neumann-Nash Dynamics: the Continuous Strategy Case. *Games and Economic Behavior* **65** 406–429. 8
- HOFBAUER, J. and SIGMUND, K. (1998). *Evolutionary Games and Population Dynamics*. Cambridge University Press. 2
- HOFBAUER, J. and SIGMUND, K. (2003). Evolutionary Game Dynamics. *Bulletin of the American Mathematical Society* **40** 479–519. 2

- IMHOF, L. A. (2005). The long-run behavior of the stochastic replicator dynamics. *The Annals of Applied Probability* **15** 1019–1045. 3, 22, 25, 31
- KOUTSOPIAS, E. and PAPADIMITRIOU, C. (1999). Worst-case equilibria. In *Proceedings of the 16th Annual Symposium on Theoretical Aspects of Computer Science* 404–413. 2, 15
- MAYNARD SMITH, J. (1974). The theory of games and the evolution of animal conflicts. *Journal of Theoretical Biology* **47** 209–221. 20
- MERTIKOPOULOS, P. and MOUSTAKAS, A. L. (2009). Learning in the Presence of Noise. In *GameNets '09: Proceedings of the 1st International Conference on Game Theory for Networks*. 3, 22, 23
- MERTIKOPOULOS, P. and MOUSTAKAS, A. L. (2010). The Emergence of Rational Behavior in the Presence of Stochastic Perturbations. *The Annals of Applied Probability* **20** 1359–1388. 3, 22, 23, 25, 34
- MILCHTAICH, I. (2000). Generic Uniqueness of Equilibrium in Large Crowding Games. *Mathematics of Operations Research* **25** 349–364. 7
- MONDERER, D. and SHAPLEY, L. S. (1996). Potential Games. *Games and Economic Behavior* **14** 124–143. 12
- NASH, J. F. (1951). Non-Cooperative Games. *The Annals of Mathematics* **54** 286–295. 4, 5, 8
- ØKSENDAL, B. (2007). *Stochastic Differential Equations*, 6 ed. Springer-Verlag, Berlin. 24, 32
- ROSEN, J. B. (1965). Existence and Uniqueness of Equilibrium Points for Concave N-Person Games. *Econometrica* **33** 520–534. 5, 8, 11
- ROSENTHAL, R. W. (1973). A class of games possessing pure-strategy Nash equilibria. *International Journal of Game Theory* **2** 65–67. 2, 12
- ROUGHGARDEN, T. and TARDOS, E. (2002). How bad is selfish routing? *Journal of the ACM* **49** 236–259. 1, 8, 13, 15, 33
- ROUGHGARDEN, T. and TARDOS, E. (2004). Bounding the Inefficiency of Equilibria in Nonatomic Congestion Games. *Games and Economic Behavior* **47** 389–403. 1, 8, 14
- SANDHOLM, W. H. (2001). Potential Games with Continuous Player Sets. *Journal of Economic Theory* **97** 81–108. 2, 7, 8, 13, 17, 20, 35
- SCHMEIDLER, D. (1973). Equilibrium points of nonatomic games. *Journal of Statistical Physics* **7** 295–300. 6, 7, 8
- TAYLOR, P. D. and JONKER, L. B. (1978). Evolutionary stable strategies and game dynamics. *Mathematical Biosciences* **40** 145–156. 2, 16
- WARDROP, J. G. (1952). Some theoretical aspects of road traffic research. In *Proceedings of the Institute of Civil Engineers, Part II* **1** 325–78. 1
- WEIBULL, J. W. (1995). *Evolutionary Game Theory*. MIT Press, Cambridge, MA. 2, 16, 18, 21