# Dynamic Cesaro-Wardrop Equilibration in Networks*


V. S. Borkar†

School of Technology and Computer Science, Tata Institute of Fundamental Research

Homi Bhabha Road, Mumbai 400005, India

P. R. Kumar‡

Department of Electrical and Computer Engineering, and Coordinated Science Laboratory

University of Illinois at Urbana-Champaign

1308 West Main Street, Urbana, IL 61801, USA



November 1, 2002



**Abstract**

We analyze a routing scheme for a broad class of networks which converges (in the Cesaro sense) with probability one to the set of approximate Cesaro-Wardrop equilibria, an extension of the notion of a Wardrop equilibrium. The network model allows for wireline networks where delays are caused by flows on links, as well as wireless networks, a primary motivation for us, where delays are caused by other flows in the vicinity of nodes.

The routing algorithm is distributed, using only the local information about observed delays by the nodes, and is moreover impervious to clock offsets at nodes. The scheme is also fully asynchronous, since different iterates have their own counters and the orders of packets and their acknowledgments may be scrambled. Finally, the scheme is adaptive to the traffic patterns in the network.

The demonstration of convergence in a fully dynamic context involves the treatment of two-time scale distributed asynchronous stochastic iterations. Using an ODE approach, the invariant measures are identified. Due to a randomization feature in the algorithm, a direct stochastic analysis shows that the algorithm avoids non-Wardrop equilibria.

Finally, some comments on the existence, uniqueness, stability, and other properties of Wardrop equilibria are made.



*Please address all correspondence to the second author at Univ. of Illinois, CSL, 1308 West Main St, Urbana, IL 61801, USA. Email: prkumar@uiuc.edu. Fax: 217-244-1653.



†Work done while visiting the Coordinated Science Laboratory, University of Illinois at Urbana-Champaign. Supported in part by grant no. DST/MS/104/99 from Dept. of Science and Technology, Govt. of India.

‡This material is based upon work partially supported by DARPA/AFOSR unfer Contract No. F49620-02-1-0325, AFOSR under Contract No. F49620-02-1-0217, DARPA under Contract No. N00014-01-1-0576, the USARO under Contract Nos. DAAD19-00-1-0466 and DAAD19-01010-465, DARPA under Contract No. F33615-01-C-1905, ONR under Contract No. N00014-99-1-0696, and NSF under Contract No. NSF ANI 02-21357. Any opinions, findings, and conclusions or recommendations expressed in this publication are those of the authors and do not necessarily reflect the views of the above agencies.

**Keywords:** Communication networks, wireless networks, routing, delay adaptive routing, Wardrop equilibrium, distributed asynchronous two-time scale stochastic approximation.

# 1 Introduction

We analyze an adaptive, distributed, asynchronous routing scheme for a general class of network models. The formulation covers both wireline networks where delays are caused by flows over links, as well as wireless networks where delays are better associated with nodes in the geographic vicinity of transmitters or receivers contending for the common shared channel. The main result is to show that the proposed routing scheme converges to the set of approximate "Cesaro-Wardrop equilibria," an extension of the notion of "Wardrop equilibria" enunciated in [26] in the context of transportation networks as follows: "The journey times on all the routes actually used are equal, and less than those which would be experienced by a single vehicle on any unused route."

The initial motivation for this problem for us was the challenge of designing a routing protocol which can adapt to the traffic requirements in wireless networks. Most routing protocols in this area have been designed to converge to the path with the minimum number of hops from source to destination, see [11], and are consequently completely insensitive to the actual delay incurred. In wireless networks, unlike wireline networks with optical links, the traffic carrying capacity can be scarce, and so it is valuable to utilize several paths simultaneously in carrying traffic from a source to a destination. However, standard "minimum delay routing" (see [3]), which attempts to minimize the delay averaged over all paths, routes traffic for a single source-destination pair over paths with very different delays (see [3]), since at optimality it equalizes only the derivatives of the delays with respect to the path flows. Once one takes resequencing delay into account (since packets are useful only when preceding packets have also been received), the actual delay can be quite large. In contrast, the Wardrop equilibrium based routing strives to maintain equal delay over all utilized paths, and less than that over unutilized paths, thus taking resequencing delay into consideration.

The adaptive algorithm we propose introduces some critical modifications to an earlier adaptive algorithm, the System and Traffic Adaptive Routing Algorithm (STARA), proposed for use in wireless networks [11]. It retains the important feature that the algorithm is immune to clock offsets at nodes which bias the delay estimates made by nodes. However, it introduces three critical modifications. The first is a modification of the update law which preserves the sum of the components so that they remain probabilities. This is achieved by considering a normalised version of STARA. The second is a probing term which ensures a positive probability of obtaining delay information on otherwise unutilized routes. Due to this $\epsilon$-probing feature, we seek Cesaro convergence of the scheme to the set of $\epsilon$-approximate Cesaro-Wardrop equilibria. The third is a decreasing randomization which forces the algorithm to escape from undesired equilibria / invariant sets where an unutilized route may offer a smaller delay. The adaptation scheme also has the following features desirable in any routing algorithm:

**i)** It is fully *distributed*: Each node adapts its own flow variables, using only the information available to it.

**ii)** It is *asynchronous*: The iterations are event-driven and therefore each iteration has its own counter, different for different iterations.

**iii)** It is *"stateless"*: It neither requires knowledge of the current state of the underlying system, nor even the origin of the flow.

Since its introduction the Wardrop equilibrium has been extensively studied in transportation research [7], [9], [13], [14], [19], [20], and, more recently, in telecommunications and distributed computing [1], [2], [15], [16]. These works, however, consider only the corresponding static optimization problem and/or deterministic dynamics. The Wardrop equilibrium can be regarded as the solution of an optimization problem where a link cost is the integral of the delay as a function of flow [20]. In our case the functional dependence of delay on flows is not known, and so neither is the integral; rather only noisy measurements of individual random delays of packets are obtained. For the general problem of routing we refer the reader to [3]. To our knowledge, the present work is the first to give a fully dynamic stochastic adaptive scheme with proven convergence properties. For communication networks, our results provide a routing algorithm. For transportation networks our results provide a rough behavioral explanation for a Cesaro-Wardrop equilibrium as resulting from distributed, asynchronous, stateless users diverting flow from routes with high delay to routes with low delay, though features in our algorithm such as time outs have no validity in the traffic context.

It is worth noting that this work has some kinship with the extensive literature on learning in evolutionary games [10], [23], [27], [25], and [28]. These works, however, study repeated games, a much simpler situation than a full-fledged stochastic dynamic game. Nevertheless, good convergence results are available only for some very simple cases. Our results provide a scheme for an apparently much more complex situation. It may also be noted that the "ordinary differential equation (ODE) limit" (15) of our scheme is in fact the "replicator dynamics" of [27]. See also [21], [22] for related developments. For earlier routing algorithms which do not, however, have a learning component, see [3], [24].

Our algorithm can also be regarded as a game theoretic learning problem with several interesting features. The most important is that each agent (node) uses only the information about its own observed delays; thus no additional communication or signaling between the agents is needed. Furthermore, despite the extremely simple nature of the adaptive scheme, it shows very interesting asymptotic behavior, viz., Cesaro convergence to the set of approximate Cesaro-Wardrop equilibria. As we shall see later, Wardrop equilibria in our point of view (distinct from its usage in transportation literature - more on this later) need not coincide with Nash equilibria, which have been the most dominant equilibrium concept in game theory.

Our algorithm can also be regarded as an extension of stochastic approximation, which however combines all the four "evils:" multiple time scales, asynchronism, projections, and presence of spurious equilibria or complex dynamics for the ODE. Their combined complexity is more than the sum of the parts. Asynchronism, in particular, arises because various updates are driven by various event completions (reception of acknowledgments or time-outs), which, among other things, may not occur in the same order as their initiations, leading to added complexity. Multiple time scales arise because the routing updates evolve on a slower time scale than the delay estimates.

It should be remarked, however, that an adaptive scheme such as this will typically be used with a small but constant stepsize (also known as the learning parameter) in practice so that it can track slowly varying environments. We have opted for the classical "diminishing stepsize" framework because of its analytic simplicity, as the constant stepsize framework is a lot messier to analyze. Loosely speaking, the latter leads to results similar to those of the diminishing stepsize algorithms with "almost surely" being replaced by "with high probability," thus the diminishing stepsize case does give an indication of what to expect for constant stepsize.

Finally, while natural in some sense in network flow situations for which they are defined, the usage of Wardrop equilibria in the electrical engineering literature seems to be limited, barring the notable exceptions already mentioned earlier. The Wardrop equilibrium is a curious animal, certainly deserving more attention from the telecommunications community than what it seems to have hitherto mustered.

A word on our proof technique: Our adaptive scheme has two components: an iterative scheme for averaging (truncated) delays and an iterative scheme for updating routing probabilities that incrementally increases the probability of routes that yield lower than average delays. This, however, requires a running estimate of average delays. This is achieved by having the stepsizes of the routing updates component smaller, in fact "asymptotically negligible" compared to those of the delay estimates. Thus the delay averaging occurs on a much faster time scale and sees routing probabilities as quasi-static. On the other hand, the routing probabilities see the delay averaging as quasi-equilibrated, giving out "running averages" as desired.

The discrete iterations asymptotically track appropriate ODEs. For delay averaging, it is a simple linear ODE, but for routing probabilities, one gets the replicator dynamics of [27]. One then shows that the "empirical measures'" of the ODE converge to a set of invariant probability measures contained in the set of Cesaro-Wardrop equilibria. The last major step exploits the decreasing randomization to show that the non-Wardrop equilibria are unstable for the stochastic algorithm, and hence avoided with probability one. Under an additional "monotonicity" condition, the results can be strengthened to a.s. convergence to a unique Wardrop equilibrium. Parenthetically, the conclusions about the replicator dynamics itself, implicit in our Theorems 7.1 and 9.1, are new and of independent interest for evolutionary game theory as well.

The claim, however, is that of Cesaro convergence to approximate Cesaro-Wardrop equilibria. The qualification "approximate" comes by because the actual routing probabilities used are a small perturbation of those suggested by the iterative scheme so that a minimum positive probability is guaranteed along all routes. This modification is essential to ensure that all options are probed sufficiently often so that the algorithm may learn their delay characteristics.

Section 2 provides background, Sections 3 and 4 describe the network model and the distributed asynchronous adaptive algorithm, Section 5 presents preliminaries for convergence analysis, and Section 6 analyzes the delay estimation scheme, while Section 7 the flow adaptation scheme. Section 8 shows avoidance of delay equalizing solutions which do not correspond to minimum delay paths. Section 9 establishes convergence to an unique Wardrop equilibrium in a special case.

## 2 Wardrop equilibria

The original notion of Wardrop equilibrium, in the context of transportation networks, is stated in terms of flow variables. At each node $i$ in a strongly connected directed graph, there originates traffic of rate $\lambda_{ij}$ destined for some node $j$. This traffic flow from $i$ to $j$ is routed over the set of paths $\Pi_{ij}$ from $i$ to $j$, with some paths carrying positive flow, and the others zero flow. Let $f_p$ denote the flow carried over path $p$, and $f$ the vector denoting the flows over all the paths. Typically, in transportation networks, delays are associated with links. (We explicitly do *not* make any such assumption in the more general model described below, since we wish to consider wireless networks too). If $f_{ij}$ denotes the total flow on link $(i, j)$, let $D_{ij}(f_{ij})$ be the delay on that link, which is assumed to be continuous and defined as a right limit for $f_{ij} = 0$, whence it represents the potential delay. The path delay $D_p(f)$ is the sum of the delays over the links in the path $p$. Then, $f$ is a Wardrop equilibrium if whenever $f_p, f_{p'} > 0$ for two paths $p, p'$ with a common destination and origin, $D_p(f) = D_{p'}(f)$ and if there is a third path with $f_{p''} = 0$, then $D_p(f) \leq D_{p''}(f)$. That is, all paths carrying nonzero traffic from a node $i$ to a node $j$ experience the same delay, and there is no unused path from $i$ to $j$ which offers a lower delay.

In this paper we work with routing probabilities over outgoing links rather than with flows over paths. This allows us to deal with a more general scenario in two aspects. First, we do away with the requirement that delays are associated with links, which allows application to wireless networks. Second, a node treats all flows passing through it with the same destination in the same way, using the same routing probabilities to choose an outgoing link over which to send packets, without regard to where the flow originated or what path the flow has followed until reaching that node. This renders the algorithm "stateless" rather than "stateful," in the language of networking, and has a significant implementation advantage.

Specifically, suppose that node $i$ has neighbors $j_1^i, j_2^i, \ldots, j_{m(i)}^i$, to which it has outgoing arcs. Let $q_\ell^{ik}$ be the probability that an incoming packet at node $i$ from wherever, is sent out on the outgoing link $(i, j_\ell)$ to node $k$. Thus in particular $\sum_\ell q_\ell^{ik} = 1$. These probabilities are defined only if there is a positive flow from $i$ to $k$. Let $q$ denote the vector $[q_\ell^{ik}]$ suitably ordered, and denote by $D_\ell^{ik}(q)$ the mean delay under $q$ experienced by packets sent from $i$ to $k$ via the immediate neighbor $j_\ell^i$. Then $q$ is a Wardrop equilibrium if $q_\ell^{ik} > 0$ implies $D_\ell^{ik}(q) = \sum_r q_r^{ik} D_r^{ik}(q) = \min_r D_r^{ik}(q)$. That is, positive routing probabilities for different $j_\ell^i$'s imply that these $j_\ell^i$'s lead to the same delay, and the $j_\ell^i$'s that have zero probability of routing from $i$ could not lead to a lower delay if they were to be tried. The flows resulting from these routing probabilities are determined once exogenous input flows are given. Our notion of Cesaro-Wardrop equilibrium (introduced later) is a characterization of probability measures w.r.t. which the above is required to hold in an average sense and our goal will be to show that suitably defined "empirical measures" converge to the set of such measures.

In adaptive control it is well known that control has a dual role, that of probing the system in order to learn about it, as well as to control it satisfactorily. Thus the routing scheme will need to sample other alternatives often enough to learn their delay characteristics, thus in particular ensuring that the potential delay along unutilized paths is higher than the minimum delay. Thus, we explicitly distinguish between the probabilities $q$ generated by the algorithm, and the actual routing probabilities $p$ employed to route packets. Specifically, we let the latter be a convex combination of the former and the uniform distribution on the possible directions for routing, with a very small positive weight for the latter. This ensures that all options are probed often enough and therefore "learnt." Correspondingly, we seek asymptotic Cesaro-Wardrop behavior only for the probabilities $q$ generated by the algorithm and not for the actual routing probabilities $p$. The latter then are $\epsilon$-approximate Cesaro-Wardrop equilibria.

We allow a very general functional dependence of $D_j^{ik}(q)$ on $q$, requiring only continuity. A standard argument based on the Kakutani fixed point theorem, as follows, shows that at least one Wardrop equilibrium exists: The set-valued map that maps $q$ to the set $A(q) = \{\tilde{q} = [\tilde{q}^{ik}] : \tilde{q}^{ik}$ minimizes $\sum_{\ell=1}^{m(i)} x_\ell D_\ell^{ik}(q)$ over all probability vectors $x = [x_1, \ldots, x_\ell]$, for all $i, k\}$ is easily seen to be compact convex-valued, and upper semicontinuous. The Kakutani fixed point theorem then guarantees the existence of a $q^*$ such that $q^* \in A(q^*)$, i.e., a Wardrop equilibrium.

It should be noted that Wardrop equilibria can exhibit a diversity of behavior vis-a-vis the kind of learning scheme we propose. We note that in the following examples, which are all covered by the theory, the "delays" need not even be monotone increasing in the flows. Consider the example shown in Figure 1. There are two links from a source to a destination, which have to carry a total of one unit of flow. Let $D_1(p)$ and $D_2(1-p)$ denote the delays incurred along the two links 1 and 2 respectively, when link 1 carries a proportion $p$ of the total flow. Then $p$ is a Wardrop equilibrium if (i) $0 < p < 1$ and $D_1(p) = D_2(1-p)$, or (ii) $p = 0$ and $D_1(0) \geq D_2(1)$, or (iii) $p = 1$ and $D_1(1) \leq D_2(0)$.

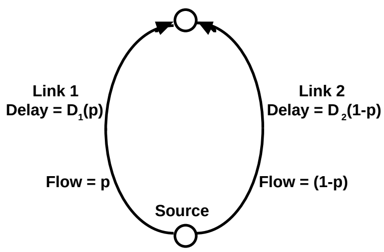

Figure 1

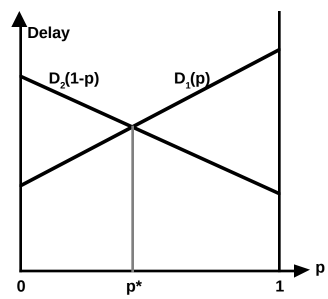

Figure 2

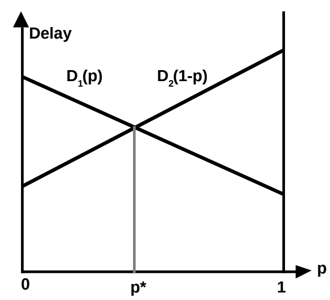

Figure 3

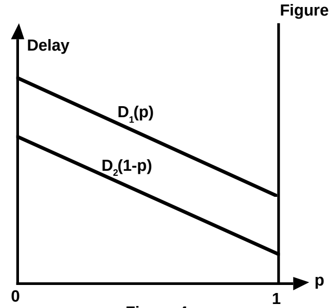

Figure 4

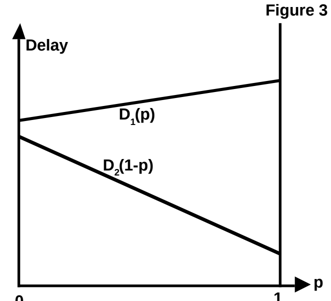

Figure 5

**(i)** For the delay functions in Figure 2, both $D_1(\cdot)$ and $D_2(\cdot)$ are increasing functions of the flow carried by them, which is the "normal" situation. Clearly $p^*$ is the unique Wardrop equilibrium. **(ii)** In Figure 3, both $D_1(\cdot)$ and $D_2(\cdot)$ are *decreasing* functions of the flow carried by them; this is not the "normal" situation, though what we have called the "delay" could represent some other quantity (for e.g., price per item) which decreases with the flow. Now $p = 0$, $p = 1$, and $p = p^*$ are all Wardrop equilibria. However, $p^*$ is an "unstable" Wardrop equilibrium for our Flow Adaptation Scheme (FAS) described in the sequel, which simply reduces the flow over a link whenever its delay is greater than the average delay over the two links. Hence, if we start with a small perturbation, say $p < p^*$, then it will decrease $p$ since $D_1(p) > D_2(1-p)$ and then converge to $p = 0$. Similarly, starting at $p > p^*$, it will converge to $p = 1$.

**(iii)** The Wardrop equilibrium does not correspond to a minimum delay solution. In Figure 4, $p = 0$ is a Wardrop equilibria, with a resulting delay larger than that of the assignment $p = 1$.

**(iv)** The FAS algorithm can move the system in the direction of increasing delay. For the delay functions shown in Figure 5, $p = 0$ is the Wardrop equilibria. No matter what $p > 0$ we start with, the Flow Adaptation Scheme converges to $p = 0$, all the while increasing the delay.

Finally, we note that the Wardrop equilibrium differs from the Nash equilibrium. For $q$ as above and a probability vector $x = [x_1, \ldots, x_{m(i)}]$, let $\hat{q}(i,k,x)$ denote the vector obtained from $q$ by replacing $q_\ell^{ik}$ by $x_\ell$ for $1 \leq \ell \leq m(i)$. For the Wardrop equilibrium the vector $[q_1^{ik}, \ldots, q_{m(i)}^{ik}]$ minimizes the function $\sum_{\ell=1}^{m(i)} x_\ell D_\ell^{ik}(q)$ over probability vectors $x$. In contrast, it would be a a Nash equilibrium if each $[q_1^{ik}, \ldots, q_{m(i)}^{ik}]$ minimized $\sum_{\ell=1}^{m(i)} x_\ell D_\ell^{ik}(\hat{q}(i,k,x))$ over all probability vectors $x$ as above. Thus the two need not coincide in general. The Wardrop equilibrium could however be regarded as a Nash equilibrium if every one of the potentially infinite number of packets is considered as a separate player whose individual decision does not affect delay, and is analogous to $\hat{q}$ resulting in the same delays as $q$. This is the traditional viewpoint in the transportation literature. In our formulation, not messages, but the nodes, are the finitely many agents making routing decisions.

# 3 The model of the system

Consider $M \geq 1$ nodes denoted by $i = 1, 2, \ldots, M$. Nodes may receive packets from the outside world, destined for particular destination nodes, which are then relayed through a sequence of nodes until they reach their final destination. Let $J(i)$ denote the set of destinations of packets arriving at $i$, whether from inside or outside the network. Call a packet destined for $k \in J(i)$ a "$k$-packet." Let $\mathcal{N}(i,k) = \{\ell_1, \ell_2, \ldots, \ell_{d(i,k)}\}$ denote the set of "neighbors of $i$ *en route* to $k$." Node $i$ directs the $n$-th $k$-packet it receives to some $j \in \mathcal{N}(i,k)$ with probability $p_j^{ik}(n)$, after noting its time of arrival. (Thus $p_j^{ik}(n) \geq 0$ and $\sum_{\ell \in \mathcal{N}(i,k)} p_\ell^{ik}(n) = 1$ for all $i$, all $k \in J(i)$, all $j \in \mathcal{N}(i,k)$.) Define $p^{ik}(n) = \left[p_1^{ik}(n), \ldots, p_{d(i,k)}^{ik}(n)\right]^T \in \mathcal{P}^{ik} \triangleq$ the simplex of probability vectors in $\mathcal{R}^{d(i,k)}$, and denote by $p(n) \in \mathcal{P} := \Pi_{i,k}\mathcal{P}^{ik}$ the vector obtained by concatenating the $p^{ik}(n)$'s in some order. Denote by $\{\sigma_j^{ik}(m)\}$ the successive times at which $i$ directs a $k$-packet to $j \in \mathcal{N}(i,k)$.

When a destination $k \in J(i)$ receives the $n$-th $k$-packet from $i$, it immediately sends a time-stamped acknowledgment to $i$. (Thus, acknowledgments are sent to all nodes on the path of a packet). Node $i$ waits for this acknowledgment till a "time-out" of $N > 0$ time units after it sent the packet. If it receives the acknowledgment by then, it notes the delay incurred by the message to reach its destination based on the difference between the time stamps, calling it $\Delta_j^{ik}(n)$, where $j \in \mathcal{N}(i,k)$ is the node to which it had originally directed the packet. If no acknowledgment is received before the time-out, it sets $\Delta_j^{ik}(n) = N$. We call $\Delta_j^{ik}(n)$'s the "truncated delays". (This is a slight abuse of terminology since it is possible that either the packet or the acknowledgement may have been lost and thus the delay is infinite in principle.) The above truncation does indeed make our estimation of average delays only approximate, but for large $N$, the approximation error should be small, assuming that the occurence of large delays or lost packets/acks is sufficiently rare. In case of a systemic fault where it is not so, there should be a different mechanism to override the proposed one to detect and amend it. The truncation not only aids our analysis, but is also a realistic feature of any implementation.

Let $\tau_j^{ik}(n)$ = the time when $i$ receives an acknowledgment for the $n$-th $k$-packet it sent via $j$, or the time $\sigma_j^{ik}(n) + N$, whichever is less.

We shall assume the following for all $i, k \in J(i), j \in \mathcal{N}(i,k)$.

**(A1)** There exists an $\overline{\epsilon} \in \left(0, \frac{1}{\max_{i,k} d(i,k)}\right)$ such that $p_j^{ik}(n) \geq \overline{\epsilon}$ for all $i$, $j$, $k$, and $n$. This will be ensured by the probing feature in the adaptive algorithm described below. The purpose is to ensure that all actions are tried often enough for the learning scheme to be able to learn their worth.

**(A2)** $\sigma_j^{ik}(m) \uparrow \infty$ a.s. for all $k \in J(i)$ and $j \in \mathcal{N}(i,k)$. Note that this implies that there are an infinite number of packets passing through $i$ to each destination in $J(i)$ and the transmission times thereof do not have any finite accumulation point.

**(A3)** The probability of $i$ receiving more than one acknowledgment at any time instant is zero. This can be relaxed to: The number of acknowledgments received by $i$ at any time instant are a.s. bounded by a finite deterministic constant. This is a standard technical assumption which is true, e.g., when the arrival times of acknowledgements have a continuous probability density.

**(A4)** If $W_j^{ik}(n)$ = the number of $k$-packets directed by $i$ to $j$ in the time interval $\left[\sigma_\ell^{ik}(n) - N, \sigma_\ell^{ik}(n) + N\right]$, for $\ell \in \mathcal{N}(i,k)$, then $\sup_n E\left[(W_j^{ik}(n))^m\right] < \infty$, for some $m \geq max(3, s)$, where $s$ is as in (5) below. Note that this is effectively a restriction on the service rate of packets at $i$.

Let $\tilde{p}^{ik}(t)$ denote the routing probability in effect at the time $t$, and define $\tilde{p}(t)$, $t \geq 0$, correspondingly. To allow for a general model, we shall assume that the "state" of the underlying network, denoted by $Z(t)$ at time $t$, is a continuous-time controlled Markov chain with state space $S = \Pi_{\ell=1}^L S_\ell$ for some $L \geq 1$, where each $S_i$ is a locally compact topological space, with the randomized control given by $\tilde{p}(t)$, which is the vector of the routing probabilities in effect at time $t$. (As an example, consider a queueing network. Then, with i.i.d. service times and i.i.d. interarrival times, the components of the state will be the various queue lengths, the elapsed service time already spent in service by the packets currently being served, and the elapsed times since the last arrivals at the nodes, taking values in $\{0, 1, 2, \ldots\}$, $[0, \infty)$, and $[0, \infty)$, respectively. $L$ will then typically be of the order of a polynomial in the number of edges plus the number of nodes of the graph of the network). Later on, we shall have reason to use the augmented state space $\overline{S} = \Pi_{\ell=1}^L \overline{S}_\ell$, where $\overline{S}_\ell$ is a one point compactification of $S_\ell$.

We assume that the trajectories of $Z(\cdot)$ are right continuous with left limits (r.c.l.l.). The only control being exercised is in the routing decisions $\tilde{p}(t)$. We define an r.c.l.l. process $\zeta(t)$, $t \geq 0$, taking values in $\mathcal{N} \triangleq \Pi_{i,k}\mathcal{N}(i,k)$, such that $\zeta^{ik}(t) = j$ if at time $t$, the most recent $k$-packet sent from $i$ was routed to $j \in \mathcal{N}(i,k)$. In other words, $\zeta(\cdot)$ is the actual realization of the randomized control $\tilde{p}(\cdot)$. If $\tilde{p}(\cdot) \equiv p$, i.e., it is held fixed at a specific $p \in \mathcal{P}$, $Z(\cdot)$, then $Z(\cdot)$ will be a time-homogeneous Markov chain on $S$. We assume it to be Feller (therefore strong Markov) with its infinitesimal generator depending continuously on $p$.

Let $D([0,\infty); S), D([0,\infty); \overline{S}), D([0,\infty); \mathcal{N})$ denote, respectively, the spaces of r.c.l.l. functions from $[0,\infty)$ to $S$ (respectively, $\overline{S}, \mathcal{N}$), endowed with Skorohod topology ([8], Chapter III). $(Z(\cdot), \zeta(\cdot))$ is a $D([0,\infty); S) \times D([0,\infty); \mathcal{N})$-valued random variable. Let $\mathcal{F}_t$ denote the right-continuous completion of $\sigma(Z(s), \zeta(s), s \leq t)$ for $t \geq 0$. We shall make the following assumptions:

**(A5)** For $\tilde{p}(\cdot) \equiv p$ and $Z(0) = z \in S$, the law of $Z(t)$ as $t \to \infty$ converges to a unique probability measure on $\overline{S}$ independent of $z$, which will then be an invariant probability measure for

$Z(\cdot)$ when viewed as an $\overline{S}$-valued process. Due to the one point compactification of $\overline{S}$, this allows for possible instability for some routing probability vectors. Furthermore, we assume that the corresponding limiting stationary law of $(Z(\cdot), \zeta(\cdot))$, viewed as a probability measure on $D([0,\infty);\overline{S}) \times D([0,\infty),\mathcal{N})$, depends continuously on $p$.

**(A6)** If $\tau$ is any $\{\mathcal{F}_t\}$-stopping time and $t_0 > 0$, then the probability of $i$ receiving an acknowledgment from $k \in J(i)$ precisely at time $\tau + t_0$ is zero. Intuitively, this captures the notion that the arrival times of acknowledgments cannot be predicted.

**(A7)** The actual (*not* truncated) delay encountered by the $k$-packet routed by $i$ to $j \in \mathcal{N}(i,k)$ at time $\sigma_j^{ik}(n)$ is a continuous function of the trajectories $(Z(\sigma_j^{ik}(n) + \cdot), \zeta(\sigma_j^{ik}(n) + \cdot))$. That is, there exists a continuous map $F : D([0,\infty);S) \times D([0,\infty);\mathcal{N}) \to R$ such that the delay in question is given by $F(Z(\sigma_j^{ik}(n) + \cdot), \zeta(\sigma_j^{ik}(n) + \cdot))$. Note that this implies in particular that $\Delta_j^{ik}(n) = \widetilde{F}(Z(\sigma_j^{ik}(n) + \cdot), \zeta(\sigma_j^{ik}(n) + \cdot))$, for a continuous $\widetilde{F} : D([0,\infty);S) \times D([0,\infty);\mathcal{N}) \to \mathcal{R}$.

We conclude this section with a useful consequence of (A6)-(A7). Let

$$\varphi_p(z) \quad = \quad E[\Delta_j^{ik}(n)/Z(\sigma_j^{ik}(n)) = z],$$

where the expectation is under $\tilde{p}(\cdot) \equiv p$.

**Lemma 3.1.** *The map $(p,z) \to \varphi_p(z)$ has a continuous version.*

**Proof (Sketch):** Let $(p^m, z^m) \to (p,z)$. By standard tightness arguments, the corresponding laws of $(Z(\sigma_j^{ik}(n) + \cdot), \zeta(\sigma_j^{ik}(n) + \cdot))$ converge along a subsequence. By Skorohod's theorem ([4], Chapter II), we may consider the convergence to be a.s. for our purposes. The corresponding $\Delta_j^{ik}(n)$'s then converge a.s. by virtue of (A7). Since these are bounded by $N$, use dominated convergence theorem to conclude. □

# 4 The distributed asynchronous adaptive algorithm

Our adaptation scheme is in two parts. The first consists of a Delay Estimation Scheme (DES) where estimates are made of the delays to destinations via differing nodes using the acknowledgments received. Second is the Flow Adaptation Scheme (FAS) where the routing probabilities are updated at the times that the acknowledgments are received.

Due to the distributed asynchronous nature of the algorithm, we need to keep track of the order in which events occur. Hence we define the following reorderings of the event times in order to define the sequence in which updates are made at the nodes:

$\{\tilde{\tau}_j^{ik}(n)\} := \{\tau_j^{ik}(m)\}$ rearranged in increasing order.

$\{\tilde{\sigma}_j^{ik}(n)\}, \{\widetilde{\Delta}_j^{ik}(n)\} := \{\sigma_j^{ik}(n)\}, \{\Delta_j^{ik}(n)\}$, respectively, rearranged in the same order as $\{\tilde{\tau}_j^{ik}(m)\}$.

$\{\widehat{\tau}^{ik}(n)\} := \cup_{j\in\mathcal{N}(i,k)}\{\tilde{\tau}_j^{ik}(m)\}$, arranged in increasing order.

$\{\theta^{ik}(n)\} := \cup_{j\in\mathcal{N}(i,k)}(\{\tilde{\tau}_j^{ik}(m)\} \cup \{\sigma_j^{ik}(m)\})$, arranged in increasing order.

We shall be updating the routing probabilities $\{p^{ik}(n)\}$ at the times $\{\widehat{\tau}^{ik}(n)\}$, and the delay estimates $\{X_j^{ik}(n)\}$ (of the average values of the truncated delays $\Delta_j^{ik}(m)$'s) at the times $\{\tilde{\tau}_j^{ik}(n)\}$.

We first specify the Delay Estimation Scheme (DES). Each node $i$ makes estimates $\{X_j^{ik}(n)\}$ of the average of the truncated delays $\Delta_j^{ik}(m)$ at the times $\{\tilde{\tau}_j^{ik}(n)\}$ by performing at time $\tilde{\tau}_j^{ik}(n+1), n \geq 0$, the iteration

$$X_j^{ik}(n+1) \;=\; (1-a(n))X_j^{ik}(n) + a(n)\widetilde{\Delta}_j^{ik}(n+1), \tag{1}$$

with $X_j^{ik}(0) = 0$. Here $\{a(n)\}$ is a sequence in $(0,\infty)$ satisfying:

**(i)** $a(n) \geq a(n+1)$. This condition simplifies our analysis, but as will become apparent later, can be easily replaced by the much weaker requirement $\sup_{m,n} \frac{a(n+m)}{a(n)} < \infty$.

**(ii)** $\sum_n a(n) = \infty$, $\sum_n a(n)^2 < \infty$,

**(iii)** $\sum_n \left(\frac{a(n)-a(n+1)}{a(n)}\right)^r < \infty$ for some $r \geq 1$. Note that this implies

$$\frac{a(n+1)}{a(n)} \to 1 \tag{2}$$

as well as

$$\sum_n \left(\frac{a(n)-a(n+k)}{a(n)}\right)^r \;<\; \infty, k \geq 1. \tag{3}$$

We also define the interpolated versions $\tilde{x}^{ik}(t)$, $t \geq 0$, of the iterates (1) as:

$$\begin{aligned}
\tilde{x}_j^{ik}(t) &= X_j^{ik}(n), \tilde{\tau}_j^{ik}(n) \leq t < \tilde{\tau}_j^{ik}(n+1),\\
\tilde{x}^{ik}(t) &= \left[\tilde{x}_1^{ik}(t), \ldots, \tilde{x}_{d(i,k)}^{ik}(t)\right]^T,\\
\hat{x}^{ik}(t) &= [\tilde{x}_1^{ik}(t), \ldots, \tilde{x}_{d(i,k)-1}^{ik}(t)]^T.
\end{aligned}$$

Now we turn to the Flow Adaptation Scheme. First we need some notation. Let

$$Q^{ik} = \{x = [x_1, \ldots, x_{d(i,k)-1}] : 0 \le x_\ell \le 1, 1 \le \ell < d(i,k), \sum_\ell x_\ell \le 1\}.$$

Note that each vector has one less dimension than $p^{ik}(n)$. Let $\Gamma^{ik} : \mathcal{R}^{d(i,k)-1} \to Q^{ik}$ denote the projection map. Also, let Diag $(q)$ for $q = [q_1, \ldots, q_{d(i,k)-1}] \in Q^{ik}$ denote the diagonal matrix whose $\ell$-th diagonal entry is $q_\ell$. Finally, for $q = [q_1, \ldots, q_{d(i,k)}] \in \mathcal{P}^{ik}$, let $\tilde{q} = [q_1, \ldots, q_{d(i,k)-1}] \in Q^{ik}$ be obtained by simply eliminating the last component, and let $\mathbf{1} = [1, 1, \ldots, 1]^T$ have the same dimension.

Each node $i$ updates $\{p^{ik}(n)\}$ at times $\{\tilde{\tau}^{ik}(n)\}$ in two steps:

**Step 1:** Iteratively generate $\{q^{ik}(n)\}$ in $\mathcal{P}^{ik}$ by performing at time $\hat{\tau}^{ik}(n+1)$, $n \ge 0$, the iteration:

$$\begin{aligned} \tilde{q}^{ik}(n+1) &= \Gamma^{ik} \Big( \tilde{q}^{ik}(n) + b(n)\mathrm{Diag}(\tilde{q}^{ik}(n))(q^{ik}(n)^T \tilde{x}^{ik}(\hat{\tau}^{ik}(n+1))\mathbf{1} \\ &\quad - \hat{x}^{ik}(\hat{\tau}^{ik}(n+1))) + \xi^{ik}(n) \Big), \end{aligned} \tag{4}$$

**(i)** Above, $\tilde{q}^{ik}(0) = \left[\frac{1}{d(i,k)}, \ldots, \frac{1}{d(i,k)}\right]^T$,

**(ii)** $\{b(n)\}$ is a nonincreasing sequence in $(0, \infty)$ satisfying: $\sum_n b(n) = \infty$, $\sum_n b(n)^2 < \infty$, and

$$\sum_n \left(\frac{b(n)}{a(n)}\right)^s < \infty \tag{5}$$

for some $s \ge 1$, implying, in particular that

$$\frac{b(n)}{a(n)} \;\to\; 0. \tag{6}$$

(As for $\{a(n)\}$, the nonincreasing requirement on $\{b(n)\}$ can be relaxed to: $\sup_{n,m} \frac{b(n+m)}{b(n)} < \infty$).

**(iii)** $\{\xi^{ik}(n)\}$ are i.i.d. random variables distributed uniformly on the unit ball of $\mathcal{R}^{d(i,k)-1}$. The presence of $\{\xi^{ik}(n)\}$ makes this a randomized scheme. This is essential in order to escape some undesirable potential limit sets for (4), as we shall see later.

**Step 2:** Set

$$q_j^{ik}(n+1) = \tilde{q}_j^{ik}(n+1), 1 \in j < d(i,k), \tag{7}$$

$$q_{d(i,k)}^{ik}(n+1) = 1 - \sum_{\ell=1}^{d(i,k)-1} \tilde{q}_j^{ik}(n+1). \tag{8}$$

Let $\gamma^{ik}$ denote the uniform distribution on $\mathcal{N}(i,k)$ and $1 > \epsilon > 0$ a small number. Set

$$p^{ik}(n) \quad = \quad (1-\epsilon)q^{ik}(n) + \epsilon\gamma^{ik}.$$

Note that this ensures that (A1) holds.

**Remark:** The $d(i)-$th component is not really "discriminated against". The $\{q^{ik}_{d(i,k)}(n)\}$ is seen to satisfy an iteration similar to (4) with $\xi^{ik}_{d(i,k)}(n) = 1 - \sum_{\ell<d(i,k)} \xi^{ik}_{\ell}(n)$ and a suitably redefined $\Gamma^{ik}$ with range $\mathcal{P}^{ik}$. Then $\{[\xi^{ik}_1(n), \ldots, \xi^{ik}_{d(i,k)}(n)]\}$ are i.i.d. uniformly distributed on the unit ball in the subspace $\{x = [x_1, \ldots, x_{d(i,k)}] : \sum x_i = 1\}$, thus showing that the $d(i)$-th component is not really "discriminated against." The advantage of iterating in $Q^{ik}$ instead of $\mathcal{P}^{ik}$ is that the latter has an empty interior, while the former doesn't, making the application of the projection less frequent.

Note that DES is a standard averaging scheme, whereas FAS is a "reinforcement learning" mechanism that incrementally adjusts $i$'s probability of sending a $k$-packet to $j$ downwards if the observed delay via $j$ exceeds the current estimate of average delay over all of $\mathcal{N}(i,k)$, and upwards if it is less than the latter, leaving it unaltered if the two coincide.

Condition (6) ensures that FAS moves on a slower scale than DES, as in the "two time scale stochastic approximation" of [5]. As in [5], DES will then see FAS as quasi-static, while FAS sees DES as quasi-equilibrated. This will be made precise in the analysis that follows in the next section. Many of the difficulties in the analysis arise because both DES and FAS are event-driven, and therefore, when viewed across the entire system, are completely asynchronous.

# 5 Convergence analysis: Preliminaries

To analyze DES and FAS, we shall use a variant of the Ordinary Differential Equation (ODE) approach of [18]. We will show that the discrete asynchronous FAS follows the ODE:

$$\dot{y}^{ik}_j(t) \quad = \quad y^{ik}_j(t) \left[ \sum_{\ell=1}^{d(i,k)} y_\ell(t) D^{ik}_\ell(y(t)) - D^{ik}_j(y(t)) \right], 1 \le i \le M, k \in J(i), j \in \mathcal{N}(i,k)/\{d(i,k)\}.$$

Then we shall identify candidates for the Cesaro limit points of the empirical measures of the ODE. This will turn out to be a set which contained in the so called Cesaro-Wardrop equilibria which we introduce later. Next we will make critical use of the randomization $\{\xi^{ik}(n)\}$, and show by a direct analysis of the stochastic recursion that all non-Wardrop points are unstable. In this and the next section we pursue the first goal, that of establishing the ODE.

## 5.1 Unscrambling truncated delays

To employ the ODE approach will require us to show the following for the DES: If

$$\alpha(n,t) \quad = \quad \min\{\ell \geq n : \sum_{k=n}^{\ell} a(k) \geq t\}, t > 0,$$

denotes the number of steps of adjustment beyond $n$ for which the "gain" amounts to $t$, then, asymptotically, we have

$$\sum_{m=n}^{\alpha(n,t)} a(m)\widetilde{\Delta}_j^{ik}(m) \quad \approx \quad tE_{\tilde{p}(\widetilde{\tau}_j^{ik}(n))}[\Delta_j^{ik}(\ell)],$$

where for $p \in P$, $E_p[\cdot]$ denotes the expectation with respect to the unique limiting stationary law under $\tilde{p}(\cdot) \equiv p$ stipulated in (A5). (In particular, by Lemma 3.1, $p \to E_p[\Delta_j^{ik}(\ell)]$ will be continuous.) Keeping this aim in mind and observing that $\sum_{m=n}^{\alpha(n,t)} a(m) \approx t$ by construction, we seek to analyze the limiting behavior of

$$\frac{\sum_{m=n}^{\alpha(n,t)} a(m)\widetilde{\Delta}_j^{ik}(m)}{\sum_{m=n}^{\alpha(n,t)} a(m)} \tag{9}$$

as $n \to \infty$. Recall that $\{\widetilde{\Delta}_j^{ik}(m)\}$ are the truncated delays ordered in the sequence in which their acknowledgments arrive and not in the sequence in which the corresponding packets were sent. Thus our first task is to unscramble these, i.e., show that (9) has the same a.s. limit behavior as

$$\frac{\sum_{m=n}^{\alpha(n,t)} a(m)\Delta_j^{ik}(m)}{\sum_{m=n}^{\alpha(n,t)} a(m)}. \tag{10}$$

**Lemma 5.1.**

$$\lim_{n\to\infty} \frac{\sum_{m=n}^{\alpha(n,t)} a(m)\left(\widetilde{\Delta}_j^{ik}(m) - \Delta_j^{ik}(m)\right)}{\sum_{m=n}^{\alpha(n,t)} a(m)} \quad = \quad 0 \text{ a.s.}$$

**Proof:** The summation in the numerator will have "error terms" of three kinds.

**(i)** *Terms in which a particular $\Delta_j^{ik}(m)$ appears in both (9) and (10), but in different places:* Let $\widetilde{\Delta}_j^{ik}(\ell) = \Delta_j^{ik}(m)$ (say). Set $\tilde{a}(m) = a(\ell)$. The net contribution of such terms is bounded by

$$\frac{\sum_{m=n}^{\alpha(n,t)} |a(m) - \tilde{a}(m)|}{\sum_{m=n}^{\alpha(n,t)} a(m)} \quad \leq \quad \frac{\sum_{m=n}^{\alpha(n,t)}(a(m) - a(m + A_m)) + \sum_{m=n}^{\alpha(n,t)}(a(m - B_m) - a(m))}{\sum_{m=n}^{\alpha(n,t)} a(m)}$$

where $A_m$, $B_m$ denote, respectively, the number of $k$-packets routed by $i$ through $j$ in the intervals $[\sigma_j^{ik}(m), \sigma_j^{ik}(m)+N]$ and $[\sigma_j^{ik}(m)-N, \sigma_j^{ik}(m)]$. The expression on the right side above will tend to zero if both

$$\frac{a(n+A_n)}{a(n)}, \frac{a(n-B_n)}{a(n)} \quad \rightarrow \quad 1 \text{ a.s.}$$

Since the proof of either case is similar, consider only the former. For $\delta > 0$ and $r > 0$ as in (3),

$$\begin{aligned}
\sum_n P\left(\frac{a(n)-a(n+A_n)}{a(n)} \geq \delta\right) &\leq \delta^{-r} \sum_n E\left[\left(\frac{a(n)-a(n+A_n)}{a(n)}\right)^r\right] \\
&\leq \delta^{-r} \sum_n \sum_m P(A_n = m) \left(\frac{a(n)-a(n+m)}{a(n)}\right)^r \\
&= \delta^{-r} \sum_m \sum_n P(A_n = m) \left(\frac{a(n)-a(n+m)}{a(n)}\right)^r \\
&\leq K\delta^{-r} \sum_m \sum_n \frac{1}{m^3} \left(\frac{a(n)-a(n+m)}{a(n)}\right)^r \\
&< \infty,
\end{aligned}$$

where $K > 0$ is a suitable constant and the penultimate inequality follows from (A4) and the Chebyshev inequality. The desired claim now follows from the Borel-Cantelli lemma.

**(ii)** *Terms in which a particular* $\widetilde{\Delta}_j^{ik}(m)$ *appears in (9) but not in (10)*:

Such an $m$ must lie in $\{n - B_n, n - B_n + 1, \ldots, n-1, \alpha(n,t)+1, \ldots, \alpha(n,t) + A_{\alpha(n,t)}\}$. The contribution of such terms cannot exceed $\frac{(A_{\alpha(n,t)}+B_n)a(n-B_n)}{\sum_{m=n}^{\alpha(n,t)} a(m)}$, which, since $\frac{a(n-B_n)}{a(n)} \rightarrow 1$ a.s., will converge to zero a.s. if $a(n)(A_{\alpha(n,t)} + B_n) \rightarrow 0$ a.s. But for $\Delta > 0$,

$$\begin{aligned}
\sum_n P(a(n)(A_{\alpha(n,t)} + B_n) \geq \delta) &\leq \delta^{-2} \sum_n E[(A_{\alpha(n,t)} + B_n)^2] a(n)^2 \\
&\leq K \sum_n a(n)^2 < \infty,
\end{aligned}$$

for a suitable $K > 0$, in view of (A4). The desired claim follows from the Borel-Cantelli lemma.

**(iii)** *Terms in which a particular* $\Delta_j^{ik}(m)$ *appears in (10), but not in (9)*:

Such an $m$ must lie in $\{n, n+1, \ldots, n+A_n, \alpha(n,t) - B_{\alpha(n,t)}, \ldots, \alpha(n,t)\}$. Argue as above. □

## 5.2 Interfacing the time scales

We now study the behavior of FAS over the interval $I_n = [\sigma_j^{ik}(n), \sigma_j^{ik}(\alpha(n,t))]$. (Recall that $\sum_{m=n}^{\alpha(n,t)} a(m) \approx t$). Also recall the random times $\{\theta^{ik}(m)\}$ which were defined as the successive times at which either a $k$-packet was sent by $i$ or an acknowledgment for the same was received by $i$ (where we also count "time-outs" as a part of the latter). Thus suppose that $\sigma_j^{ik}(n) = \theta^{ik}(\tilde{n})$ and $\sigma_j^{ik}(\alpha(n,t)) = \theta^{ik}(\tilde{n}_h)$ with $\tilde{n} = \tilde{n}_0 < \tilde{n}_1 < \ldots < \tilde{n}_h$ denoting the integers for which $\sigma^{ik}(n+\ell) = \theta^{ik}(\tilde{n}_\ell)$, $0 \le \ell \le h$, i.e., those $\theta^{ik}(m)$'s in $I_n$ at which $k$-packets are being sent from $i$. Let $Y_\ell$ = the number of acknowledgments (including time-outs) received by $i$ during the interval

$$[\sigma_j^{ik}(n+\ell), \sigma_j^{ik}(n+\ell+1)], 0 \le \ell < h.$$

**Lemma 5.2.** $E[(Y_\ell)^m] < \infty$, $1 \le m \le s$.

**Proof:** Clearly $Y_\ell$ is bounded by the number of $k$-packets sent by $i$ in the interval $[\sigma_j^{ik}(n+\ell) - N, \sigma_j^{ik}(n+\ell+1)]$. Let $Y_{1\ell}, Y_{2\ell}$ denote, respectively, the $k$-packets sent by $i$ in the intervals $[\sigma_j^{ik}(n+\ell) - N, \sigma_j^{ik}(n+\ell)]$ and $[\sigma_j^{ik}(n+\ell), \sigma_j^{ik}(n+\ell+1)]$. By (A4), $E[(Y_{1\ell})^m] < \infty$, $1 \le m \le s$. By (A1),

$$E[(Y_{2\ell})^m] \quad \le \quad K \sum_\beta \beta^m (1-\overline{\epsilon})^\beta < \infty, \quad 1 \le m \le s,$$

for a suitable $K > 0$. The claim follows. □

**Lemma 5.3.** $\sup_{t \in I_n} \|\tilde{p}(t) - \tilde{p}(\sigma_j^{ik}(n))\| \to 0$ *a.s. as* $n \to \infty$.

**Proof:** As before, any reference to "acknowledgments" will also include the time-outs. Let $\hat{n}_\ell$ denote the largest integer $m$ for which $\hat{\tau}^{ik}(m) \le \theta^{ik}(\tilde{n}_\ell)$. Thus $b(\hat{n}_\ell)$ will be the most recent stepsize used by FAS at time $\sigma_j^{ik}(n+\ell)$, due, in turn, to the most recent acknowledgment for a $k$-packet received by $i$. Also, let $G_\ell$ denote the set of consecutive integers $m$ such that $\sigma_j^{ik}(n+\ell) \le \hat{\tau}^{ik}(m) < \sigma_j^{ik}(n+\ell+1)$. That is, $G_\ell$ is the set of $m$ such that the $m$-th update of FAS takes place during the interval $[\sigma_j^{ik}(n+\ell), \sigma_j^{ik}(n+\ell+1)]$, which marks the time between two successive transmissions of $k$-packets by $i$ via $j$. In particular, note that $|G_\ell| = Y_\ell$. Furthermore, since $\{b(m)\}$ is nonincreasing, we have $b(\hat{n}_\ell) \ge b(m)$ for all $m \in G_\ell$. Furthermore, FAS updates on all acknowledgments of $k$-packets by $i$, whereas DES updates only for those corresponding to the $k$-packets sent via a prescribed $j$. Therefore FAS has at least as many updates in any given time interval as DES. In particular, $\hat{n}_\ell \ge n + \ell$, implying $b(\hat{n}_\ell) \le b(n+\ell)$. Hence,

$$\sup_{t \in I_n} \|\tilde{p}(t) - \tilde{p}(\sigma_j^{ik}(n))\| \quad \le \quad C_1 \sum_{\ell=0}^{h} \sum_{m \in G_\ell} b(m) \text{ (for a suitable constant } C_1 > 0)$$

$$
\begin{aligned}
&\leq \quad C_1 \sum_{\ell=0}^{h} Y_\ell b(\hat{n}_\ell) \\
&= \quad C_1 \sum_{\ell=0}^{h} a(n+\ell) Y_\ell \left( \frac{b(\hat{n}_\ell)}{a(n+\ell)} \right) \\
&\leq \quad C_1 \sum_{\ell=0}^{h} a(n+\ell) Y_\ell \left( \frac{b(n+\ell)}{a(n+\ell)} \right) \\
&\leq \quad C_2 t \left( \frac{\sum_{\ell=0}^{h} a(n+\ell) Y_\ell \left( \frac{b(n+\ell)}{a(n+\ell)} \right)}{\sum_{\ell=0}^{h} a(n+\ell)} \right) \text{ (for a suitable constant } C_2 > 0\text{),}
\end{aligned}
$$

where we use the fact that $\sum_{\ell=0}^{h} a(n+\ell) = \sum_{\ell=n}^{\alpha(n,t)} a(n) \in [t, t + a(\alpha(n,t))]$ and choose $C_2 \geq C_1(t + a(1))/t$. The right side of the above sequence of inequalities will tend to zero a.s. if we have $Y_\ell \left( \frac{b(n+1)}{a(n+1)} \right) \to 0$ a.s. But for any $\delta > 0$ and $s \geq 1$ as in (5), and a suitable constant $K > 0$,

$$
\sum_\ell P\left( Y_\ell \left( \frac{b(n+\ell)}{a(n+\ell)} \right) \geq \delta \right) \quad \leq \quad \delta^{-s} K \sum_\ell \left( \frac{b(\ell)}{a(\ell)} \right)^s < \infty,
$$

from (5) and Lemma 5.2. The desired claim now follows from the Borel-Cantelli lemma. □

## 5.3 Quasi-stationarity

Returning to (10), our final "preliminary" will be to show that a closely related weighted average

$$
\frac{\sum_{m=n}^{\alpha(n,t)} a(m) f(Z(\sigma_j^{ik}(m)))}{\sum_{m=n}^{\alpha(n,t)} a(m)} \tag{11}
$$

for some $f \in C(\overline{S})$ is asymptotically quasi-stationary in the sense that it tracks the stationary average with respect to a control policy that is slowly varying. Let $P(\overline{S})$ denote the space of probability measures on $\overline{S}$ with the Prohorov topology ([4], Chap. II) and define $\nu^n \in \mathcal{P}(\overline{S})$ by

$$
\int g d\nu^n \quad = \quad \frac{\sum_{m=n}^{\alpha(n,t)} a(m) g(Z(\sigma_j^{ik}(m)))}{\sum_{m=n}^{\alpha(n,t)} a(m)}, \quad g \in C(\overline{S}).
$$

Let $\hat{p}(n) \triangleq \tilde{p}(\tilde{\tau}_j^{ik}(n)) \in \mathcal{P}$, which is compact. Let $\{n(\ell)\} \subset \{n\}$ be a subsequence along which $\hat{p}(n(\ell)) \to p^* \in \mathcal{P}$ (say). Under $p^*$, $\{Z(\sigma_j^{ik}(n))\}$ will be a time-homogeneous Markov chain because of the strong Markov property of $Z(\cdot)$. Let $q^*(x, dy)$ denote its transition kernel. Let $\nu^*$ denote its unique limiting (hence invariant) distribution in $\mathcal{P}(\overline{S})$ stipulated in (A5).

**Lemma 5.4.** *For all sample points outside a zero probability set and for $\{\nu^n\}$, $\{n(\ell)\}$, $\nu^*$ defined as above, $\nu^{n(\ell)} \to \nu^*$.*

**Proof:** Let $\{\mathcal{F}_t\}$ be defined as in (A6) and set $\hat{\mathcal{F}}_n = \mathcal{F}_{\sigma_j^{ik}(n)}$, $n \geq 1$. For $g \in C(\overline{S})$, the sum

$$\sum_{m=1}^{n} a(m) \left(g(Z(\sigma_j^{ik}(m))) - E[g(Z(\sigma_j^{ik}(m)))/\hat{\mathcal{F}}_{m-1}]\right), n \geq 1,$$

is seen to be a zero mean martingale. Since $\overline{S}$ is compact, $g$ is bounded. It is then easily verified that the quadratic variation process for this martingale is $O\left(\sum_{m=1}^{n} a(m)^2\right)$ and therefore convergent. By Theorem 3.3.4, p. 53, of [4], the above martingale converges a.s. Since $\sum_{m=n}^{\alpha(n,t)} a(m) \approx t$, we have

$$\lim_{n\to\infty} \left( \frac{\sum_{m=n}^{\alpha(n,t)} a(m) g(Z(\sigma_j^{ik}(m)))}{\sum_{m=n}^{\alpha(n,t)} a(m)} - \frac{\sum_{m=n}^{\alpha(n,t)} a(m) E[g(Z(\sigma_j^{ik}(m)))/\hat{\mathcal{F}}_{m-1}]}{\sum_{m=n}^{\alpha(n,t)} a(m)} \right) = 0 \text{ a.s.} \quad (12)$$

The set of sample points where either Lemma 5.3 fails or (12) fails for a $g$ in some prescribed countable dense subset of $C(\overline{S})$, has zero probability. Fix a sample point outside this set and consider $\{n(k)\}$, $p^*$, $q^*$ defined as above. By Lemma 5.3,

$$\sup_{t \in I_{n(\ell)}} \|\tilde{p}(t) - p^*\| \rightarrow 0.$$

By continuity arguments,

$$\sup_{n(\ell) \leq m \leq \alpha(n(\ell),t)} |E[g(Z(\sigma_j^{ik}(m)))/\hat{\mathcal{F}}_{m-1}] - \int g(y) q^*(Z(\sigma_j^{ik}(m-1)), dy)| \rightarrow 0 \text{ a.s.}$$

for $g \in C(\overline{S})$. By (12) and (2), we then have

$$\int g d\nu^{n(\ell)} - \int \int g(y) q^*(x, dy) \nu^{n(\ell)}(dx) \rightarrow 0$$

for $g$ in a countable dense subset of $C(\overline{S})$. Since the latter is a convergence determining class for $\mathcal{P}(\overline{S})$ ([4], Chapter III), it follows that $\nu^{n(\ell)} \to \nu^*$ a.s. $\square$

Recall the notation $E_p[\cdot]$ for expectation with respect to the unique limiting stationary law under $\tilde{p}(\cdot) \equiv p$, as in (A5).

**Corollary 5.1.**

$$\frac{\sum_{m=n}^{\alpha(n,t)} a(m) f(Z(\sigma_j^{ik}(m)))}{\sum_{m=n}^{\alpha(n,t)} a(m)} - E_{\hat{p}(n)}[f(Z(\sigma_j^{ik}(m)))] \rightarrow 0 \text{ a.s. }, \quad f \in C(\overline{S}).$$

**Proof:** Combine the foregoing with continuity of the map $p \to E_p[f(Z(\sigma_j^{ik}(m)))]$ (cf. (A5)). □

**Corollary 5.2.**

$$\frac{\sum_{m=n}^{\alpha(n,t)} a(m)\varphi_{\hat{p}(n)}(Z(\sigma_j^i(m)))}{\sum_{m=n}^{\alpha(n,t)} a(m)} - \widetilde{E}_{\hat{p}(n)}[\varphi_{\hat{p}(n)}(Z(\sigma_j^{ik}(m)))] \quad \to \quad 0 \text{ a.s. },$$

*where* $\widetilde{E}_{\hat{p}(n)}[\varphi_{\hat{p}(n)}(Z(\sigma_j^{ik}(m)))]$ *denotes the expectation with respect to the unique limiting stationary law for* $\{Z(\sigma_j^{ik}(m))\}$ *under* $\hat{p}(n)$ *(cf. (A5)), treating* $\hat{p}(n)$ *as a parameter.*

**Proof:** Use the above in conjunction with Lemma 3.1. □

# 6 The analysis of the Delay Estimation Scheme

First we conduct an analysis of the DES, and then the FAS. We begin by analyzing (10), more or less in continuation of Section 5.3.

**Lemma 6.1.** *Almost surely,*

$$\lim_{n\to\infty}\left(\frac{\sum_{\ell=n}^{\alpha(n,t)} a(\ell)\Delta_j^{ik}(\sigma_j^{ik}(\ell))}{\sum_{\ell=n}^{\alpha(n,t)} a(\ell)} - \frac{\sum_{\ell=n}^{\alpha(n,t)} a(\ell)E[\Delta_j^{ik}(\sigma_j^{ik}(\ell))/\hat{\mathcal{F}}_\ell]}{\sum_{\ell=n}^{\alpha(n,t)} a(\ell)}\right) \quad = \quad 0.$$

**Proof:** Fix $T > N$ and let $t(0) = 0$, $t(n) = \sum_{m=1}^{n} a(i)$, $n \geq 1$.

Define $\{n(m)\} \subset \{n\}$ by: $n(0) = 0$ and

$$n(m+1) \quad = \quad \min\{k : t(k) \geq t(n(m)) + T\}, m \geq 0.$$

Using the fact that truncated delays are bounded by $N < \infty$, a successive conditioning argument leads to

$$E\left[\left(\sum_{\ell=n(m)}^{n(m+1)-1} a(\ell)\Delta_j^{ik}(\sigma_j^{ik}(\ell)) - E\left[\sum_{\ell=n(m)}^{n(m+1)-1} a(\ell)\Delta_j^{ik}(\sigma_j^{ik}(\ell))/\hat{\mathcal{F}}_{n(m)}\right]\right)^2 /\hat{\mathcal{F}}_{n(m)}\right]$$
$$= O\left(\sum_{\ell=n(m)}^{n(m+1)-1} a(\ell)^2\right).$$

Therefore, as in the proof of Lemma 5.4.,

$$\sum_{\substack{m=0 \\ m \text{ odd}}}^{h}\left(\sum_{\ell=n(m)}^{n(m+1)-1} a(\ell)\Delta_j^{ik}(\sigma_j^{ik}(\ell)) - E\left[\sum_{\ell=n(m)}^{n(m+1)-1} a(\ell)\Delta_j^{ik}(\sigma_j^{ik}(\ell))/\hat{\mathcal{F}}_{n(m)}\right]\right)$$

and

$$\sum_{\substack{m=0\\ m\text{ even}}}^{h}\left(\sum_{\ell=n(m)}^{n(m+1)-1} a(\ell)\Delta_j^{ik}(\sigma_j^{ik}(\ell)) - E\left[\sum_{\ell=n(m)}^{n(m+1)-1} a(\ell)\Delta_j^{ik}(\sigma_j^{ik}(\ell))/\hat{\mathcal{F}}_{n(m)}\right]\right)$$

are zero mean martingales (the fact that $T > N$ plays a role here) with a.s. convergent quadratic variation processes and therefore a.s. convergent themselves. Then their sum

$$\sum_{m=0}^{h}\left(\sum_{\ell=n(m)}^{n(m+1)-1} a(\ell)\Delta_j^{ik}(\sigma_j^{ik}(\ell)) - E\left[\sum_{\ell=n(m)}^{n(m+1)-1} a(\ell)\Delta_j^{ik}(\sigma_j^{ik}(\ell))/\hat{\mathcal{F}}_{n(m)}\right]\right) \tag{13}$$

is also a.s. convergent. A similar argument shows that

$$\sum_{m=0}^{h}\left(\sum_{\ell=n(m)}^{n(m+1)-1} a(\ell)E\left[\Delta_j^{ik}(\sigma_j^{ik}(\ell))/\hat{\mathcal{F}}_\ell\right] - E\left[\sum_{\ell=n(m)}^{n(m+1)-1} a(\ell)E\left[\Delta_j^{ik}(\sigma_j^{ik}(\ell))/\hat{\mathcal{F}}_\ell\right]/\hat{\mathcal{F}}_{n(m)}\right]\right) \tag{14}$$

is also a.s. convergent. A successive conditioning argument shows that the conditional expectation with respect to $\hat{\mathcal{F}}_{n(m)}$ in the summands of (13) and (14) is the same, implying that

$$\sum_{m=0}^{h} a(m)(\Delta_j^{ik}(\sigma_j^{ik}(m)) - E[\Delta_j^{ik}(\sigma_j^{ik}(m))/\hat{\mathcal{F}}_m])$$

is a.s. convergent. Thus

$$\lim_{n\to\infty} \frac{\sum_{\ell=n}^{\alpha(n,t)} a(\ell)\Delta_j^{ik}(\sigma_j^{ik}(\ell)) - \sum_{\ell=n}^{\alpha(n,t)} a(\ell)E[\Delta_j^{ik}(\sigma_j^{ik}(\ell))/\hat{\mathcal{F}}_\ell]}{\sum_{\ell=n}^{\alpha(n,t)} a(\ell)} = 0 \text{ a.s.}$$

□

This leads to:

**Lemma 6.2.** *Almost surely,*

$$\lim_{n\to\infty}\left(\frac{\sum_{m=n}^{\alpha(n,t)} a(m)\Delta_j^{ik}(\sigma_j^{ik}(m))}{\sum_{m=n}^{\alpha(n,t)} a(m)} - \widetilde{E}_{\hat{p}(n)}[\varphi_{\hat{p}(n)}(Z(\sigma_j^{ik}(m)))]\right) = 0.$$

**Proof:** Fix a sample point outside the zero probability set where any of the Lemmas 5.3, 5.4, 6.1, Corollaries 5.1, and 5.2 fail. Pick a sequence $\{n(\ell)\}$ of $\{n\}$ such that $\hat{p}(n(\ell)) \to p^*$ (say) in $\mathcal{P}$. Let $\psi_n$ denote the regular conditional law of $(Z(\sigma_j^{ik}(n)+t), \xi(\sigma_j^{ik}(n)+t))$, $t \in [0, N]$, give $\hat{\mathcal{F}}_n$ for $n \geq 1$. Let $\psi^*(x, dy) : S \to \mathcal{P}(D([0,N]; \overline{S}) \times D([0,N]; \mathcal{N}))$ denote its regular conditional law given $Z(\sigma_j^{ik}(n))$ (and hence given $\hat{\mathcal{F}}_n$) if $\tilde{p}(\cdot) \equiv p^*$ were used. Since $Z(\cdot)$ is

Feller under $p^*$, we may take $\psi^*(\cdot, dy)$ to be continuous. By Lemma 5.3, for all bounded $\Phi \in C(D([0,N];\overline{S}) \times D([0,N];\mathcal{N}))$,

$$\int \Phi d\psi_{n(\ell)} - \int \Phi(y)\psi^*(Z(\sigma_j^{ik}(n(\ell))), dy) \quad \rightarrow \quad 0.$$

(An appropriate choice of the versions of the regular conditional laws used is implicit here). Since $\overline{S}$ is compact and $\psi^*(\cdot, dy)$ continuous, we may take a further subsequence of $\{n(\ell)\}$ along which

$$Z(\sigma_j^{ik}(n(\ell))) \rightarrow \bar{z}, \psi^*(Z(\sigma_j^{ik}(n(\ell)), dy) \rightarrow \psi^*(\bar{z}, dy))$$

and therefore $\psi_{n(\ell)} \rightarrow \psi^*(\bar{z}, dy)$. Then along this subsequence,

$$\int \Phi d\psi_{n(\ell)} \rightarrow \int \Phi(y)\psi^*(\bar{z}, dy),$$

$$\int \Phi(y)\psi^*(Z(\sigma_j^{ik}(n(\ell))), dy) \quad \rightarrow \quad \int \Phi(y)\psi^*(\bar{z}, dy),$$

for *all* $\Phi : D([0,N];\overline{S}) \times D([0,N];\mathcal{N}) \rightarrow R$ that are bounded and a.s. continuous with respect to $\psi^*(\bar{z}, dy)$ (cf. the proof of Lemma 3.1). Thus along this sequence,

$$E[\Delta_j^{ik}(\sigma_j^{ik}(n(\ell)))/\hat{\mathcal{F}}_{n(\ell)}] - E^*[\Delta_j^{ik}(\sigma_j^{ik}(n(\ell)))/Z(\sigma_j^{ik}(n(\ell)))] \quad \rightarrow \quad 0,$$

where $E^*[\cdot|\cdot]$ denotes the conditional expectation under $\tilde{p}(\cdot) \equiv p^*$ and the $\psi^*(\bar{z}, dy) -$ a.s. continuity of the truncated delay follows from (A6) as in the proof of Lemma 3.1. A similar argument shows that

$$\varphi_{\hat{p}(n(\ell))}(Z(\sigma_j^{ik}(n(\ell)))) - E^*[\Delta_j^{ik}(n(\ell))/Z(\sigma_j^{ik}(n(\ell)))] \quad \rightarrow \quad 0$$

along this subsequence. It follows that

$$E[\Delta_j^{ik}(\sigma_j^{ik}(n))/\hat{\mathcal{F}}_n] - \varphi_{\hat{p}(n)}(Z(\sigma_j^{ik}(n))) \quad \rightarrow \quad 0$$

along the above sequence. Then it must hold true along $\{n\}$ as $n \rightarrow \infty$. The claim now follows on combining this observation with Lemma 6.1. □

Returning to DES, define "interpolated iterations" $\bar{x}(t)$, $t \geq 0$, corresponding to (1) as follows: Define $\{t(n)\}$ as above and set $\bar{x}(t(n)) = X_j^{ik}(n)$ with linear interpolation on $[t(n), t(n+1)]$ for $n \geq 0$. Also, define $\overline{D}_j^{ik}(p) = E_p[\Delta_j^{ik}(n)]$. By Lemma 3.1, $\overline{D}(\cdot)$ is continuous.

**Theorem 6.1.** *For any* $T > 0$,

$$\lim_{n\rightarrow\infty} \sup_{t(n)\leq t\leq t(n)+T} |\overline{D}_j^{ik}(\hat{p}(n)) - \bar{x}(t)| \quad = \quad 0 \text{ a.s.}$$

**Proof:** Rewrite (1) as

$$\begin{aligned}\bar{x}(t(n+\ell)) &= \bar{x}(n(\ell)) + \sum_{m=t(n)}^{t(n+\ell-1)} a(m)[\Delta_j^{ik}(\tilde{\sigma}_j^{ik}(m+1)) - \bar{x}(t(m))] \\ &= \bar{x}(t(n)) + \int_{t(n)}^{t(n+\ell)} (\overline{D}_j^{ik}(\tilde{p}(\tilde{\tau}_j^{ik}(n))) - \bar{x}(t))dt + o(1),\end{aligned}$$

for $t(n) \leq t(n+\ell) \leq t(n)+T$. This along with Lemma 5.3 yields the desired result by a standard argument based on Gronwall Lemma (as in, e.g., [5]). □

It is worth noting that the DES being a standard averaging scheme, the actual convergence proof is quite standard and short, most of the work has been for the bookkeeping required to manage the multiple time scales, unscrambling of events, etc. brought about by the fact that these are asynchronous, event-driven iterations.

# 7 The ODE analysis of the Flow Adjustment Scheme

In this section we use the ODE method to analyze the Flow Adjustment Scheme. First we define the delays in terms of $q$'s rather than the $p$'s. For $p^{ik} = (1-\epsilon)q^{ik} + \epsilon\gamma^{ik} \in \mathcal{P}^{ik}$, $1 \leq i \leq M$, $k \in J(i)$, define $D_j^{ik}(q) = \overline{D}_j^{ik}(p)$ for $\overline{D}_j^{ik}(\cdot)$ defined as above.

Our main result of this section is to identify that, in the limit, the delays are equalized along all paths of positive flow "on the average." For this purpose, define $\hat{b}(n) = \sum_{m=0}^{n-1} b(m), \hat{b}(0) = 0$, and define the $\mathcal{P}-$valued "interpolated trajectory" $\bar{q}(\cdot)$ by: $\bar{q}_j^{ik}(\hat{b}(n)) = q_j^{ik}(n)$, $n \geq 0$, with linear interpolation on each $[\hat{b}(n), \hat{b}(n+1)]$. Consider the $\mathcal{P}-$valued ODE

$$\begin{aligned}\dot{y}_j^{ik}(t) &= y_j^{ik}(t)\left[\sum_{\ell=1}^{d(i,k)} y_\ell^{ik}(t) D_\ell^{ik}(y(t)) - D_j^{ik}(y(t))\right], 1 \leq i \leq M, \\ & k \in J(i), j \in \mathcal{N}(i,k). \qquad (15)\end{aligned}$$

**Lemma 7.1.** *$\bar{q}(\cdot)$ tracks (15) in the sense that for any $T > 0$,*

$$lim_{t\to\infty} sup_{s\in[t,t+T]} ||\bar{q}(s) - q^t(s)|| = 0,$$

*where $q^t(s), s \in [t, t+T]$, is the trajectory of (15) with $q^t(t) = \bar{q}(t)$ for $t \geq 0$.*

**Proof:** Rewrite FAS as

$$\tilde{q}^{ik}(n+1) = \Gamma^{ik}(\tilde{q}^{ik}(n) + b(n)[Diag(\tilde{q}^{ik}(n))\left(\sum_{\ell=1}^{d(i,k)} q_\ell^{ik}(n)D_\ell^{ik}(q(n))\mathbf{1} - D_j^{ik}(q(n))\right)$$
$$+\xi^{ik}(n) + o(1)]).$$

Standard "ODE" approach to projected stochastic approximation algorithms (see, e.g., Theorem 2.1, page 95 of [17]) then implies the claim for the case when the index $j$ in (15) is restricted to $\mathcal{N}(i,k)\backslash\{d(i,k)\}$. Note that despite the presence of the projection operator $\Gamma^{ik}(\cdot)$, this o.d.e. does not contain any "boundary correction terms". This is because the driving vector field of (15) is never directed outwards at the boundary, as can be easily verified. Now set $y_{d(i,k)}^{ik} = 1 - \sum_{\ell=1}^{d(i,k)-1} y_\ell^{ik}(t)$. Direct verification shows that (15) extends to include $j = d(i,k)$ and so does the claim. □

Let $\mathcal{P}^*$ denote the space of probability measures on $\mathcal{P}$ with Prohorov topology.

**Definition** A pair $(q, \nu^*)$ in $\mathcal{P} \times \mathcal{P}^*$ will be said to be a Cesaro-Wardrop equilibrium if:

1. $\nu^*$ is invariant under (15), and,

2. If $q_j^{ik} > 0$, then

$$E^*[D_j^{ik}(p)] = min_m E^*[D_m^{ik}(p)],$$

   where $E^*[\cdot]$ denotes the expectation w.r.t. $\nu^*$ (i.e., $E^*[D_m^{ik}(p)] = \int D_m^{ik}(y)\nu^*(dy)$).

Define empirical measures $\nu(t), t > 0$, by:

$$\int_{\mathcal{P}} f d\nu(t) \triangleq \frac{1}{t}\int_0^t f(\bar{q}(s))ds,$$

for $f \in C(\mathcal{P})$. Our main result, proved via the subsequent lemmas, asserts that, in the limit, a link is assigned a positive flow only if the (properly normalized) time-averaged delay along the link is minimal:

**Theorem 7.1.** *Almost surely, $(\bar{q}(t), \nu(t))$ converges to the set of Cesaro-Wardrop equilibria as $t \to \infty$.*

**Lemma 7.2.** *Almost surely, every limit point of $\nu(t)$ in $\mathcal{P}^*$ as $t \to \infty$ is invariant under (15).*

**Proof:** Let $t_n \to \infty, \nu(t_n) \to \tilde{\nu}$ in $\mathcal{P}^*$. We need to show that $\tilde{\nu}$ is invariant under (15). For $f \in C(\mathcal{P})$,

$$\frac{1}{t_n}\int_0^{t_n} f(\bar{q}(s))ds \to \int f d\tilde{\nu}.$$

Let $\Phi_t : \mathcal{P} \to \mathcal{P}$ denote the "time $t$" map of (15), i.e., the map that maps $y(0)$ to $y(t)$ for $t > 0$. Thus in view of Lemma 7.1, for $t > 0$,

$$\begin{aligned} lim_{n\to\infty}\frac{1}{t_n}\int f(\bar{q}(s))ds &= \int f d\tilde{\nu} = \\ lim_{n\to\infty}\frac{1}{t_n}\int_t^{t_n+t} f(\bar{q}(s))ds &= \\ lim_{n\to\infty}\frac{1}{t_n}\int_0^{t_n} f\circ\Phi_t(\bar{q}(s))ds &= \int f\circ\Phi_t d\tilde{\nu}. \end{aligned}$$

Since $t > 0$ was arbitrary, the claim follows. □

Let $(q^*, \tilde{\nu})$ be a limit point of $(\bar{q}(t), \nu(t))$ as $t \to \infty$. By the above lemma, $\tilde{\nu}$ is invariant under (15), thus satisfying the first condition in the definition of the Cesaro-Wardrop equilibrium. Let $\tilde{E}[\cdot]$ denote the expectation under $\tilde{\nu}$.

**Lemma 7.3.** *For all $m$ such that $(q^*)^{ik}_m > 0$, $\tilde{E}[D^{ik}_m(p)]$ has the same value.*

**Proof:** Let $t_n \to \infty, \bar{q}(t_n) \to q^*, \nu(t_n) \to \tilde{\nu}$ in $\mathcal{P}^*$. Let $\eta \in \mathcal{P}$ be such that $\eta^{ik}_j > 0$ iff $(q^*)^{ik}_j > 0$. Then by using (4) to approximate (15) as in Lemma 7.1, we have,

$$\sum_\ell \eta^{ik}_\ell \ell n(\frac{\bar{q}^{ik}_\ell(t_n)}{\bar{q}^{ik}_\ell(0)}) = \int_0^{t_n}((\bar{q}^{ik})^T(s)D^{ik}(\bar{q}(s)) - (\eta^{ik})^T D^{ik}(\bar{q}(s)))ds + o(t_n),$$

where the term $o(t_n)$ collects together the usual asymptotically negligible error terms. (We are implicitly assuming that $\bar{q}^{ik}_\ell(0) > 0\ \forall i, k, \ell$ and thus the ratio in parentheses on the l.h.s. is well-defined. There is no loss of generality, as if it were not so initially, it will be so eventually because of our additional noise $\{\xi^{ik}(n)\}$ in (4). Thus the above can be ensured by shifting the time origin if necessary.) As $n \to \infty$, the r.h.s. is

$$t_n(\tilde{E}[(p^{ik})^T D^{ik}(p)] - (\eta^{ik})^T\tilde{E}[D^{ik}(p)]) + o(t_n).$$

The l.h.s. in the above equation then remains bounded by our choice of $\eta$. Dividing both sides by $t_n$ and letting $t_n \to \infty$, we get

$$\tilde{E}[(p^{ik})^T D^{ik}(p)] - (\eta^{ik})^T\tilde{E}[D^{ik}(p)] = 0, \qquad (16)$$

where $\tilde{E}[\,\cdot\,]$ denotes the expectation w.r.t. $\tilde{\nu}$. This is true for all $\eta$ mutually absolutely continuous w.r.t. $q^*$, which is possible if and only if $\tilde{E}[D^{ik}_m(p)]$ is independent of $m$ for all $m$ such that $(q^*)^{ik}_m > 0$. The claim follows. □

**Proof of Theorem 7.1:** We only need to prove that in the foregoing, for $\ell$ such that $(q^*)^{ik}_\ell = 0$,

$$\tilde{E}[D^{ik}_\ell(p)] \geq min_y \tilde{E}[D^{ik}_y(p)] = \tilde{E}[D^{ik}_j(p)]$$

for $j \in \{r : (q^*)^{ik}_r > 0\}$. If not, there exists an $m$ such that $(q^*)^{ik}_m = 0$ and $\tilde{E}[D^{ik}_m(p)] < \tilde{E}[D^{ik}_j(p)] - \gamma$ for some $\gamma > 0$, with $j$ as above. Let $\{t_n\}$ be as in the proof of the above lemma. Let $\eta = \frac{1}{2}(q^* + \delta_m)$, where $\delta_m$ is a point mass at $m$. As in Lemma 7.3, we have,

$$\sum_\ell \eta^{ik}_\ell \ell n(\frac{\bar{q}^{ik}_\ell(t_n)}{\bar{q}^{ik}_\ell(0)}) = \int_0^{t_n}((\bar{q}^{ik})^T(s)D^{ik}(\bar{q}(s)) - (\eta^{ik})^T D^{ik}(\bar{q}(s)))ds + o(t_n),$$

where the term $o(t_n)$ collects together the usual asymptotically negligible error terms. (Again we are implicitly assuming that $\bar{q}^{ik}_\ell(0) > 0 \; \forall i, k, \ell$ and thus the ratio in parentheses on the l.h.s. is well-defined.) As $n \to \infty$, the r.h.s. is

$$t_n(\tilde{E}[(p^{ik})^T D^{ik}(p)] - (\eta^{ik})^T \tilde{E}[D^{ik}(p)]) + o(t_n),$$

which is positive for sufficiently large $n$ in view of (16) and our choice of $\eta$, in fact it is $O(\gamma)t_n$, and thus increasing to $+\infty$. But the l.h.s. is going to $-\infty$ because $\bar{q}^{ik}_m(t_n) \to (q^*)^{ik}_m = 0$, which yields a contradiction. The claim follows. □

Note that there is an implicit time scaling involved in the passage from $\{n\}$ to $\{t(n)\}$, thus in the passage from the FAS to the interpolated trajectory $\bar{q}(\cdot)$. Our empirical measures are defined w.r.t. the latter timescale, not the 'natural' one on which FAS runs. In the latter, these would be weighted empirical measures with the stepsizes serving as weights. This is natural for diminishing stepsize stochastic approximation algorithms, where the timescaling is intrinsic to the o.d.e. approximation.

In Section 9 below, a special case is discussed where this result can be strengthened.

## 8 Avoidance of non-Wardrop solutions

Let $H \stackrel{\Delta}{=} \{y : y_\ell^{ik} > 0 \Rightarrow D_\ell^{ik}(y) = \sum_m y_m^{ik} D_m^{ik}(y)\}$. Note that the Wardrop equilibria form the following subset of $H$:

$$H_s \;:=\; \{y \in H : y_\ell^{ik} > 0 \Rightarrow D_\ell^{ik}(y) = \sum_m y_m^{ik} D_m^{ik}(y) = \min_m D_m^{ik}(y)\},$$

We will now show that all the non-Wardrop equilibria in $H \backslash H_s$ are rendered unstable by our addition of the Lebesgue-continuous noise $\{\xi^{ik}(n)\}$ to FAS.

**Theorem 8.1.** *$P(q(n) \to H \backslash H_s$ along a subsequence$) = 0$.*

**Proof:** We begin with some new notation. To start with, rewrite (15) with $1 \leq j < d(i,k)$ as $\dot{y}^{ik}(t) = f^{ik}(y(t))$ for appropriately defined $f^{ik}(\cdot)$ and the entire system as $\dot{y}(t) = f(y(t))$ for $f(\cdot) = [f^{ik}(\cdot)]$, suitably ordered.

Recall that an equilibrium $\bar{q}$ of (15) is characterized by the fact that $\bar{q}_\ell^{ik} > 0$ implies $D_\ell^{ik}(\bar{q}) = \sum_m \bar{q}_m^{ik} D_m^{ik}(\bar{q})$, and will be an unstable equilibrium in $H \backslash H_s$ if for some $i \in \{i, \ldots, M\}$, $k \in J(i)$, $j \in \mathcal{N}(i,k)$, we have $\bar{q}_j^{ik} = 0$ and $D_j^{ik}(\bar{q}) < \sum_m \bar{q}_m^{ik} D_m^{ik}(\bar{q})$. Consider a connected set $\tilde{C}$ of unstable equilibria in $H \backslash H_s$ and let $C$ be its image under the projection $\Pi_{\ell,m} Q^{\ell m} \to Q^{ik}$, sitting in the relative interior of some face $F = \{q^{ik} : q_{j_\ell}^{ik} = 0, 1 \leq l \leq m\}$ of $Q^{ik}$, and assumed to be unstable in the direction of a unit vector $\theta$ normal to $F$ and directed inwards (i.e., either towards the interior of $Q^{ik}$ or towards the relative interior of a face thereof). That is,

$$f^{ik}(q) \cdot \theta \;>\; 0, \tag{17}$$

for all $q$ sufficiently close to $C$, but not on $F$. It is easy to see that if this is not feasible for our original $C$, we can always make it so by replacing $C$ by a small piece thereof. It is certainly possible for a single point $\bar{q}^{ik} \in C$, for we can take $\theta$ any inward normal with strictly positive components along the directions of increase of $q_j^{ik}$ for which $D_j^{ik}(\bar{q}) < \sum_m \bar{q}_m^{ik} D_m^{ik}(\bar{q})$ and $\bar{q}_j^{ik} = 0$. Thus it will be true in a small neighborhood of $\bar{q}^{ik}$ in $C$, which we take to be our new $C$.

We further stipulate that for a sufficiently small $\delta > 0$, (17) holds for all $q$ in the $\delta$-neighborhood $C^\delta$ of $C$ in $Q^{ik}$ that are in $H \backslash H_s$. Once again, it is clear that a connected set of unstable equilibria in $F$ can be written as a countable union of sets that satisfy the foregoing *and* admit a neighborhood as above (though possibly with different values of $\delta$). Thus for our purposes it will suffice to show that FAS a.s. avoids a $C$ satisfying all of above.

Furthermore, for $z \in F$,

$$((\Gamma^{ik}(z+y)-z)\cdot\theta)^- \quad = \quad 0$$

for any "small" $y$. Thus given our assumptions on $\{\xi^{ik}(n)\}$, one sees that for a suitable $a > 0$,

$$E[(\Gamma^{ik}(z+b(n)\xi^{ik}(n))-z)\cdot\theta] \quad \geq \quad b(n)a, z \in F. \tag{18}$$

For $\delta$ above sufficiently small, this will also hold for $z \in C^\delta$, with $a$ replaced by, say, $\frac{a}{2}$.

Returning to (4), define

$$\hat{q}^{ik}_j(n+1) \quad = \quad q^{ik}_j(n) + b(n)q^{ik}_j(n)[q^{ik}(n)^T\tilde{x}^{ik}(\hat{\tau}^{ik}(n+1)) - \tilde{x}^{ik}_j(\hat{\tau}^{ik}(n+1))].$$

Direct verification shows that $\sum_{j=1}^{d(i,k)} \hat{q}^{ik}_j(n+1) = 1$. Thus $\hat{q}^{ik}(n+1) = [\hat{q}^{ik}_1(n+1), \ldots, \hat{q}^{ik}_{d(i,k)-1}(n+1)]^T \in Q^{ik}$. Since $\tilde{q}^{ik}(n+1) = \Gamma^{ik}(\hat{q}^{ik}(n+1) + b(n)\xi^{ik}(n))$, (4) can be written as

$$\tilde{q}^{ik}(n+1) \quad = \quad \hat{q}^{ik}(n+1) + b(n)\bar{\xi}^{ik}(n)$$

where

$$\bar{\xi}^{ik}(n) \quad = \quad \Gamma^{ik}(\hat{q}^{ik}(n+1) + b(n)\xi^{ik}(n)) - \hat{q}^{ik}(n+1))/b(n).$$

Suppose that $P(\tilde{q}^{ik}(n) \to C) > 0$. Using Egorov's theorem, we can find an $n_0 \geq 1$ such that

$$P(\tilde{q}^{ik}(n) \to C, \tilde{q}^{ik}(m) \in C^\delta \text{ for } m \geq n_0) > 0.$$

By Theorem 6.1 and Egorov's theorem, we may increase $n_0$ if necessary and suppose that $P(A_1 \cap A_2) > 0$ where

$$\begin{aligned} A_1 &= \{\tilde{q}^{ik}(n) \to C, \tilde{q}^{ik}(m) \in C^\delta \text{ for } m \geq n_0\} \\ A_2 &= \{\|\text{diag }(\tilde{q}^{ik}(m))[\tilde{q}^{ik}(m)\cdot(\tilde{x}^{ik}(\hat{\tau}(m+1)) - D^{ik}(\tilde{q}(m))) \\ &\quad -(\tilde{x}^{ik}(\hat{\tau}(m+1)) - D^{ik}(\tilde{q}(m)))]\| < \frac{a}{4}, m \geq n_0\} \end{aligned}$$

where $a > 0$ is as in (17). Define

$$Y(m) := \theta\cdot(\tilde{q}^{ik}(m) - z), z \in C, m \geq n_0.$$

Notice that the definition is independent of the choice of $z \in C$ because $\theta \perp F$. Then on $A_1 \cap A_2$, $Y(m) \to 0$. But $Y(m+1) = Y(m) + b(m)(f^{ik}(\tilde{q}^{ik}(m)) + \eta(m) + \bar{\xi}^{ik}(m)) \cdot \theta$ where $\|\eta(m)\| < \frac{a}{4}$ on $A_1 \cap A_2$. Iterating,

$$\begin{aligned} Y(m+1) = Y(n_0) \quad &+ \sum_{m=n_0}^{n} b(m) f^{ik}(\tilde{q}^{ik}(m)) \cdot \theta + \sum_{m=n_0}^{n} b(m)\eta(m) \cdot \theta \\ &+ \sum_{m=n_0}^{n} b(m)(\bar{\xi}^{ik}(m) - E[\bar{\xi}^{ik}(m)/\hat{\mathcal{F}}_{m-1}]) \cdot \theta \\ &+ \sum_{m=n_0}^{n} b(m) E[\bar{\xi}^{ik}(m)/\bar{\mathcal{F}}_{m-1}] \cdot \theta, \end{aligned} \tag{19}$$

(20)

where $\bar{\mathcal{F}}_m \triangleq \mathcal{F}_{\widehat{\tau}^{ik}(m)}$. Note that

$$Z_n \quad = \quad \sum_{m=n_0}^{n} b(m) \left( \bar{\xi}^{ik}(m) - E[\bar{\xi}^{ik}(m)/\bar{\mathcal{F}}_{m-1}] \right) \cdot \theta$$

is a zero mean martingale with increments bounded by a constant times $b(m)$, therefore with a quadratic variation process that is $O\left(\sum_{m=n_0}^{n} b(m)^2\right)$ and hence convergent. Thus the martingale itself converges a.s. to some random variable $\hat{\xi}$ (Theorem 3.3.4, p. 53, of [4]). Using (17), the remark following (18), and the facts that $Y(m) \to 0$, $|\eta(m) \cdot \theta| \leq \frac{a}{4}$ on $A_1 \cap A_2$, we can let $m \to \infty$ in (19) to obtain on $A_1 \cap A_2$,

$$0 \quad \geq \quad Y(n_0) + \hat{\xi} + \lim_{n\to\infty} \frac{a}{4} \sum_{m=n_0}^{n} b(m) = \infty \text{ a.s.},$$

a contradiction. Thus $P(A_1 \cap A_2) = 0$, leading to the conclusion $P(\tilde{q}^{ik}(n) \to C) = 0$.

Now suppose that for $A_3 = \{\tilde{q}^{ik}(n) \to C \text{ along a subsequence}\}$, we have $P(A_3) > 0$. Pick $n_0 \geq 1$ and $A_2$ as before so that $P(A_3 \cap A_2) > 0$. Consider a sample point in $A_3 \cap A_2$ that lies outside the set of zero probability on which either $\{Z_n\}$ fails to converge, or $\tilde{q}^{ik}(n) \to C$, or both. Then for a $\delta > 0$ sufficiently small, $\tilde{q}^{ik}(n) \notin C^\delta$ infinitely often. By decreasing $\delta$ further if necessary, we may suppose that (17) and (18) hold. Let $K = d(i,k)(2N+1)$. Pick $0 < \epsilon < \frac{a\delta}{16K}$. Let $Z_\infty = \lim_{n\to\infty} Z_n$ and take $\bar{n}_0 \geq n_0$ sufficiently large so that $|Y(n)| + |Z_{m+n} - Z_n| < \frac{\epsilon}{2}$ for $n \geq \bar{n}_0$ and $m \geq 0$. By our choice of the sample point, there must be infinitely many segments $\{\tilde{q}^{ik}(n), n_1 \leq n \leq n_2\}$ of the algorithm such that $\tilde{q}^{ik}(n) \in C^\delta$ for some $n_1 \leq n \leq n_2, \tilde{q}^{ik}(n_1) \notin C^{\delta/2}$, $\tilde{q}^{ik}(n) \notin C^{\frac{\delta}{4}}$ for $n_1 \leq n < n_2$ and $\tilde{q}^{ik}(n_2) \in C^{\frac{\delta}{4}}$. (That is, the algorithm, while remaining inside $C^\delta$, moves from outside of $C^{\frac{\delta}{2}}$ to just inside $C^{\frac{\delta}{4}}$). We may take $n_1$ sufficiently large so that $n_1 \geq \bar{n}_0$ and $b(n_1) < \frac{\delta}{8K}$. Since the size of the $n$-th increment of (15) is at most $b(n)K$, it follows that $\tilde{q}^{ik}(n) \notin C^\alpha$ for $0 < \alpha < \frac{\delta}{8}$ and $n_1 \leq n \leq n_2$. Take $\alpha$ sufficiently small so that for $q \in C^\alpha$, $z \in F$, $|\theta \cdot (q - z)| < \frac{\epsilon}{4}$. Now, between $n = n_1$ and $n = n_2$, the algorithm has traversed a distance of at least $\frac{\delta}{4}$. Once again using the fact that the size of its $n$-th increment is at most $b(n)K$, we conclude that

$$\sum_{n=n_1}^{n_2} b(n) \quad \geq \quad \frac{\delta}{4K} > \frac{4\epsilon}{a}.$$

Then by arguments similar to the ones we used earlier, for all $n > n_2$ such that $\tilde{q}^{ik}(m)$, $n_1 \leq m \leq n$, is in $C^\delta$, we have

$$\begin{aligned} Y(n) \quad &\geq \quad -\frac{\epsilon}{2} + \frac{a}{4} \sum_{m=n_1}^{n_2} b(m) + \frac{a}{4} \sum_{m=n_2+1}^{n} b(m) \\ &\geq \quad \frac{\epsilon}{2} + \frac{a}{4} \sum_{m=n_2+1}^{n} b(m). \end{aligned}$$

Hence, for $n > n_2$, $Y(n)$ is bounded from below by a monotone increasing sequence of positive numbers exceeding $\frac{\epsilon}{2}$, till $\tilde{q}^{ik}(n)$ exits from $C^\delta$. In particular, $\tilde{q}^{ik}(n), n > n_1$, must exit $C^\delta$ without entering $C^\alpha$. Since this is true for all such patches with $n_1$ sufficiently large, $\{\tilde{q}^{ik}(n)\}$ cannot converge to $C$ along a subsequence, which contradicts our choice of the sample point. Thus $P(A_3) = 0$. □

**Corollary 8.1.** $P(p(n) \overset{a.s.}{\to} \{p : p^{ik} = (1-\epsilon)q^{ik} + \epsilon\gamma^{ik}, q \in H \backslash H_s\}$ *along a subsequence*$) = 0$.

The use of $p^{ik}$'s instead of $q^{ik}$'s in the algorithm represents a standard trade-off in adaptive control. An adaptive control policy has a two-fold task: To probe (or "explore") the system and to optimize (or "exploit") it. Our decision to ensure a probability of at least $\bar{\epsilon} > 0$ for transmission of a packet along any contending channel ensures that all actions are being probed often enough to "learn" their delay characteristics. Increasing $\bar{\epsilon}$ would typically speed up this learning (because possible state-action pairs are being sampled more frequently) while degrading the performance (because this would mean further departure from the "optimal" choice of $q^{ik}$'s). One can also conceive of alternative schemes that start with a higher $\bar{\epsilon}$ to facilitate rapid learning in the initial stages and then decrease it very slowly, possibly to zero, which would ensure Cesaro convergence to the set of exact Cesaro-Wardrop equilibria.

In the context of our results of the preceding section, the foregoing has the following important implication: The invariant measures appearing in the Cesaro-Wardrop equilibria obtained in Theorem 7.1 will necessarily assign zero probability to unstable invariant sets.

We may further let $N \uparrow \infty$ to conclude the Wardrop property for actual (untruncated) delays in the case where the stationary expectation with respect to the aforementioned stationary law, of the roundtrip delays (between transmission of a packet and its acknowledgment) remains bounded. This is a further justification of our use of the term "approximate Wardrop equilibria."

# 9 On uniqueness of the Wardrop equilibrium

We begin with some remarks on the uniqueness of the Wardrop equilibrium, in our context where routing decisions at a node are independent of the origin of a packet. The first is essentially a restatement of Theorem 6.1 of [12] and is well known in trasportation literature [20]. Define vectors $p, D(p)$ by arranging $p_j^{ik}, D_j^{ik}(p)$ lexicographically.

**Theorem 9.1.** *The Wardrop equilibrium is unique if $D(\cdot)$ is monotone in the sense that*

$$(q_1 - q_2) \cdot (D(q_1) - D(q_2)) \quad > \quad 0 \text{ whenever } q_1 \neq q_2. \tag{21}$$

**Proof:** Let $q_1$, $q_2$ be distinct Wardrop equilibria. Then $q_1 \cdot D(q_1) \leq q_2 \cdot D(q_1)$, and $q_2 \cdot D(q_2) \leq q_1 \cdot D(q_2)$ leading to the contradiction $(q_1 - q_2) \cdot (D(q_1) - D(q_2)) \leq 0$. $\square$

We note that Theorem 9.1 can be relaxed so that (21) need only hold for $q_1$ and $q_2$ which are Wardrop equilibria. A similar argument to the above shows that if (21) is relaxed to $(q_1 - q_2) \cdot (D(q_1) - D(q_2)) \geq 0$, then for any two Wardrop equilibria $q_1$, $q_2$, we have support $(q_1^{ik}) \cup$ support $(q_2^{ik}) \subset$ Argmin $(D_\cdot^{ik}(q_1)) \cap$ Argmin $(D_\cdot^{ik}(q_2))$, for $1 \leq i \leq M$, $k \in J(i)$.

Under the conditions of Theorem 9.1, the conclusions of Theorem 7.1 can be strengthened:

**Theorem 9.2.** *If $D(\cdot)$ is monotone, then $\bar{q}(\cdot)$ (and hence $\{q(n)\}$) converge a.s. to the unique Wardrop equilibrium.*

**Proof:** Let $p^*$ denote the unique Wardrop equilibrium and for $q \in \mathcal{P}$, set $F(q) = -\sum_B (p^*)_j^{ik} \ell n(q_j^{ik})$, then for $y(\cdot)$ given by the above ODE,

$$\frac{d}{dt} F(y(t)) \quad = \quad -(y(t) - p^*) \cdot D(y(t))$$

$$\leq -(y(t) - p^*) \cdot D(y(t)) + (y(t) - p^*) \cdot D(p^*)$$
$$= -(y(t) - p^*) \cdot (D(y(t)) - D(p^*))$$
$$< 0, \text{for } y(t) \neq p^*.$$

(The first inequality uses the fact that $p^*$ is a Wardrop equilibrium.) Thus $F$ serves as a Lyapunov function for the ODE establishing that $p^*$ is the unique globally asymptotically stable equilibrium thereof. The claim follows from a standard argument based on Lemma 7.1 [17]. □

One cannot, however, expect this result to hold in general, as examples of nonconvergent replicator dynamics abound [27].

# 10 Concluding remarks

The problem of routing packets in a network can be regarded as a problem of control of a large scale system based on partial observations. We have proposed a distributed, asynchronous, adaptive algorithm for seeking Cesaro-Wardrop equilibria in networks. The algorithm is a modification of that in [11] which includes a probing term to shake off unstable non-Wardrop equilibria and unstable invariant sets in general. The network model allows for wireline networks where delays are caused by flows long links, as well as wireless networks where delays are caused by other flows in the vicinity of nodes. A proof of Cesaro convergence to the set of Cesaro-Wardrop equilibria is given in a fully dynamic context.

To render this a fully acceptable algorithm, some features need to be put in which will ensure that no packet ever traverses the same node twice.

# References


[1] E. Altman, H. Kameda, " Equilibria for Multi-class routing problems in multi-agent networks", *40th IEEE Conf. on Decision and Control*, pp. 604–609, Orlando, Florida, Dec. 2001.

[2] N. G. Beans, F. P. Kelly and P. G. Taylor, "Braess's paradox in a loss network", *J. Applied Probability* vol. 34, pp. 155–159, 1997.

[3] D. Bertsekas and R. Gallager, *Data Networks*, Prentice Hall, Englewood Cliffs, 1987.

[4] V. S. Borkar, *Probability Theory–An Advanced Course*, Springer-Verlag, New York, 1995.

[5] V. S. Borkar, "Stochastic Approximation with Two Time Scales," *Systems and Control Letters*, vol. 29, pp. 291–294, 1997.

[6] M. Bowling and M. Veloso, "Multiagent learning using a variable learning rate", *Artificial Intelligence* vol. 136, 2002, pp. 215–250.

[7] E. Codina and J. Barcelo, "Dynamic Traffic Assignment: Considerations on Some Deterministic Modelling Approaches," *Annals of Operations Research*, to appear.

[8] S. Ethier and T. G. Kurtz, *Markov Processes: Characterization and Convergence*, John Wiley, New York, 1986.

[9] M. Florian and D. Hearn, "Network Equilibrium Models and Algorithms," in *Handbook of OR and MS*, vol. 8:Network Routing, M.O. Ball, et al (eds)., North Holland, Amsterdam, 1995, pp. 485–550.

[10] D. Fudenberg and D. Levine, *Theory of Learning in Games*, MIT Press, Cambridge MA, 1998.

[11] P. Gupta and P.R. Kumar, "A system and traffic dependent adaptive routing algorithm for ad hoc networks," *Proceedings of the 36th IEEE Conference on Decision and Control*, pp. 2375–2380, San Diego, Dec. 1997.

[12] P. Gupta, "Design and Performance Analysis of Wireless Networks," PhD Thesis, Coordinated Science Laboratory, University of Illinois at Urbana-Champaign, 2000.

[13] A. Haurie and P. Marcotte, "A Game Theoretic Approach to Network Equilibrium," *Math. Programming Study*, vol. 26, pp. 252–255, 1986.

[14] A. Haurie and P. Marcotte, "On the Relationship between Nash-Cournot and Wardrop Equilibria," *Networks*, vol. 15, pp. 295–308, 1985.

[15] H. Kameda, E. Altman, T. Kozawa and Y. Hosokawa, "Braess-like paradoxes of Nash equilibrium for load balancing in distributed computer systems," *IEEE Trans. on Automatic Control*, vol. 45, pp. 1687–1691, 2000.

[16] Y. A. Korilis and A. Orda, "Incentive compatible pricing strategies for QoS switching", *IEEE INFOCOM*, pp. 891–899, 1999.

[17] H. J. Kushner and G. Yin, *Stochastic Approximation Algorithms and Applications*, Springer Verlag, New York, 1997.

[18] L. Ljung, "Analysis of Recursive Stochastic Algorithms," *IEEE Transactions on Automatic Control*, vol. AC-22, pp. 551–575, 1977.

[19] P. Marcotte, S. Nguyen (eds)., *Equilibrium and Advanced Transportation Modelling*, Kluwer Academic, Boston, MA, 1998

[20] M. Patriksson, "The Traffic Assignment Problem: Models and Methods," *VSP BV*, The Netherlands, 1994.

[21] W. H. Sandholm, "Evolutionary implementation and congestion pricing", *Review of Economic Studies*, to appear.

[22] W. H. Sandholm, "Evolutionary implementation and congestion pricing", *Journal of Economic Theory* vol. 97, pp. 81–108, 2001.

[23] S. P. Singh, M. Kearns and Y. Mansour, "Nash convergence of gradient dynamics in general sum games", *Proc. of the 16th Conf. on Uncertainty in Artificial Intelligence*, Stanford, CA, 2000, pp. 541–548.

[24] W. K. Tsai, G. Huang, J. K. Antonio and W. T. Tsai, "Distributed iterative aggregation algorithms for box-constrained minimization problems of optimal routing in data networks," *IEEE Transactions on Automatic Control* vol. AC-34, pp. 20-33, 1989.

[25] F. Vega Redondo, "Evolution, Games, and Economic Behavior," Oxford University Press, Oxford, UK, 1996.

[26] J. G. Wardrop, "Some Theoretical Aspects of Road Traffic Research," *Proc. Inst. Civil Engineers, Part 2*, pp. 325–378, 1952.

[27] J. Weibull, *Evolutionary Game Theory*, MIT Press, Cambridge, MA 1995.

[28] H. P. Young, *Individual Strategy and Social Structure*, Princeton University Press, Princeton, NJ, 1998.