Fluid and Brownian Approximations for an Internet Congestion Control Model

W. Kang, F. P. Kelly, N. H. Lee and R. J. Williams

***Abstract*— We consider a stochastic model of Internet congestion control that represents the randomly varying number of flows present in a network where bandwidth is shared fairly amongst elastic document transfers. We focus on the heavy traffic regime in which the average load placed on each resource is approximately equal to its capacity. We first describe a fluid model (or functional law of large numbers approximation) for the stochastic model. We use the long time behavior of the solutions of this fluid model to establish a property called (multiplicative) state space collapse, which shows that in diffusion scale the flow count process can be approximately recovered as a continuous lifting of the workload process. Under proportional fair sharing of bandwidth and a mild condition, we show how state space collapse can be combined with a new invariance principle to establish a Brownian model as a diffusion approximation for the workload process and hence to yield an approximation for the flow count process. The workload diffusion behaves like Brownian motion in the interior of a polyhedral cone and is confined to the cone by reflection at the boundary, where the direction of reflection is constant on any given boundary face. We illustrate this approximation result for a simple linear network. Here the diffusion lives in a wedge that is a strict subset of the positive quadrant. This geometrically illustrates the entrainment of resources, whereby congestion at one resource may prevent another resource from working at full capacity.**

I. INTRODUCTION

Roberts and Massoulié [14] have introduced and studied a flow-level model of Internet congestion control, that represents the randomly varying number of flows present in a network where bandwidth is dynamically shared between flows that correspond to continuous transfers of individual elastic documents. This model assumes a "separation of time scales" such that the time scale of the flow dynamics (i.e., of document arrivals and departures) is much longer than the time scale of the packet level dynamics on which rate control schemes such as TCP converge to equilibrium.

Assuming exponentially distributed document sizes, de Veciana, Lee and Konstantopoulos [6] and Bonald and Massoulié [2] studied stability of the flow-level model under weighted α-fair bandwidth sharing policies, where $\alpha \in (0, \infty)$ (cf. Mo and Walrand [13]). For $\alpha = 1$ and $\alpha \to \infty$ in [6], and for $\alpha \in (0, \infty)$ in [2], Lyapunov functions were constructed which yield positive recurrence of the Markov chain associated with the flow-level model when the average load on each resource is less than its capacity.

We are interested in using Brownian models (i.e., diffusion approximations) to explore the performance of flow-level models operating under weighted α-fair bandwidth sharing policies when the average load placed on each resource is approximately equal to its capacity, i.e., the system is heavily loaded. We are particularly interested in manifestations of the phenomenon of entrainment, whereby congestion at some resources may prevent other resources from working at their full capacity.

There are several motivations for our work. One source of motivation lies in fixed point approximations of network performance for TCP networks (cf. [3], [7], [15]). These approximations require, as input, information on the joint distribution of the numbers of flows present on different routes, where dependencies between these numbers may be induced by the bandwidth sharing mechanism. Similarly, an understanding of such joint distributions seems important if the performance models for a single bottleneck described by Ben Fredj *et al.* [1] are to be generalized to a network. Another motivation is that the flow-level model typically involves the simultaneous use of several resources. With exponential document sizes, this model can be equated (in distribution) with a stochastic processing network (SPN) as introduced by Harrison [8], [9]. Open multiclass queueing networks operating under head-of-the-line (HL) service disciplines are a special case of SPNs without simultaneous resource possession. For certain queueing networks of this type, it has been shown [4], [16] that suitable asymptotic behavior of critical fluid models implies a property called state space collapse, which in turn validates the use of Brownian model approximations for these networks in heavy traffic. For more general SPNs, investigation of the behavior of critical fluid models, of a related notion of state space collapse, and of the implications for diffusion approximations, are in the early stages of development. The analysis in this paper can be viewed as a contribution to such an investigation for models involving simultaneous resource possession. Finally, although we restrict to exponential document sizes in this paper, we would like to relax that assumption in future work. Although this involves a significantly more elaborate stochastic model to keep track of residual document sizes (because of the processor sharing nature of the bandwidth sharing policy), knowing the results for exponential document sizes is likely to be useful for such

The research of F.P. Kelly is supported in part by EPSRC grant GR/S86266/01. The research of R. J. Williams is supported in part by NSF Grant DMS-0305272.

Address of F. P. Kelly: Statistical Laboratory, University of Cambridge, Wilberforce Road, Cambridge CB3 0WB, UK F.P.Kelly@statslab.cam.ac.uk

Address of W. Kang, N. H. Lee and R. J. Williams: Department of Mathematics, UCSD, 9500 Gilman Drive, La Jolla, CA 92093-0112, USA williams@math.ucsd.edu

work.

In this paper, we consider the flow-level model with exponentially distributed document sizes operating under a weighted α-fair bandwidth sharing policy for $\alpha \in (0, \infty)$. We assume that all resources are heavily loaded. We first define a critical fluid model which is a formal functional law of large numbers approximation to the flow-level model. The asymptotic behavior of this critical fluid model was studied in a prior work [12] of two of the authors, Kelly and Williams. We show that this behavior can be used to prove a property called *(multiplicative) state space collapse.* Loosely speaking this says that in diffusion scale the flow count process can be approximately recovered by a continuous lifting of the lower dimensional workload process. Given the asymptotic behavior of the critical fluid model, our proof of state space collapse proceeds in a similar manner to that in Bramson [4], where open multiclass queueing networks operating under certain head-of-the-line (HL) service disciplines are treated. Finally, for the case of proportional fair sharing of bandwidth ($\alpha = 1$), we combine our state space collapse result with a new *invariance principle* [10] for semimartingale reflecting Brownian motions living in cones to yield a Brownian model as a diffusion approximation for the flow count process under a mild condition. This diffusion process lives in a polyhedral cone. It behaves like Brownian motion in the interior of the cone and is confined to the cone by reflection at the boundary where the direction of reflection is constant on any given boundary face. We illustrate this approximation result for a simple linear network. Here the diffusion lives in a wedge that is a strict subset of the positive quadrant. This geometrically illustrates the entrainment of resources, whereby congestion at one resource may prevent another resource from working at full capacity.

Ongoing work is directed towards establishing diffusion approximations for the flow level model when $\alpha \neq 1$, to considering the situation when only some resources are heavily loaded, and to relaxing the exponential document size assumption. The latter is a challenging problem, since, as in processor sharing queues with general service times, a useful state descriptor needs to keep track of all of the residual document sizes.

II. STOCHASTIC MODEL

We consider a network with finitely many *resources* labelled by $j \in \mathcal{J} \neq \emptyset$. A *route* i is a non-empty subset of $\mathcal{J}$ (interpreted as the set of resources used by route i). We are given a finite, non-empty set $\mathcal{I}$ of allowed routes. Let $\mathbf{J} = |\mathcal{J}|$, the total number of resources, and $\mathbf{I} = |\mathcal{I}|$, the total number of routes. Let A be the $\mathbf{J} \times \mathbf{I}$ *incidence matrix* which contains only zeros and ones and is defined such that $A_{ji} = 1$ if resource j is used by route i, and $A_{ji} = 0$ otherwise. We assume that A has rank $\mathbf{J}$, so that it has full row rank. We further assume that resource (bandwidth) *capacities* $(C_j : j \in \mathcal{J})$ are given and that these are all strictly positive and finite.

An active flow on route i corresponds to the continuous transmission of a document through the resources used by route i. Transmission is assumed to occur simultaneously through all resources on route i. It is assumed that a new document arrives to route i at each jump time of a Poisson process that has rate parameter $\nu_i > 0$ and that each such document has an exponentially distributed size with mean $1/\mu_i$ where $\mu_i \in (0, \infty)$. These document sizes are assumed to be independent of one another and to be independent of all arrival times of documents. The number of documents on route i at time zero is assumed to be independent of the remaining sizes of those documents and these sizes are assumed to be independent and exponentially distributed with mean $1/\mu_i$. Initial numbers and sizes of documents, arrival times of new documents and their sizes for different routes $i \in \mathcal{I}$ are assumed to be mutually independent.

Bandwidth capacity is allocated dynamically to the routes according to the following bandwidth sharing policy which was first introduced by Mo and Walrand [13]. The bandwidth for a route is shared equally amongst all of the documents currently being transmitted over that route. Given a fixed parameter $\alpha \in (0, \infty)$ and strictly positive weights $(\kappa_i : i \in \mathcal{I})$, if $N_i(t)$ denotes the (random) number of flows on route i at time t for each $i \in \mathcal{I}$, and $N(t) = (N_i(t) : i \in \mathcal{I})$, then the bandwidth allocated to route i at time t is given by $\Lambda_i(N(t))$ and this bandwidth is shared equally amongst all of the flows on route i, where the function $\Lambda(\,\cdot\,) = (\Lambda_i(\,\cdot\,) : i \in \mathcal{I})$ is defined as follows (we define it on all of $\mathbb{R}_+^{\mathbf{I}}$ as we shall later apply it to rescaled versions of N).

Let $\Lambda : \mathbb{R}_+^{\mathbf{I}} \to \mathbb{R}_+^{\mathbf{I}}$ be defined such that for each $n \in \mathbb{R}_+^{\mathbf{I}}$, $\Lambda_i(n) = 0$ for $i \in \mathcal{I}_0(n) \equiv \{l \in \mathcal{I} : n_l = 0\}$, and when $\mathcal{I}_+(n) \equiv \{l \in \mathcal{I} : n_l > 0\}$ is non-empty, $\Lambda^+(n) \equiv (\Lambda_i(n) : i \in \mathcal{I}_+(n))$ is the unique value of $\Lambda^+ = (\Lambda_i : i \in \mathcal{I}_+(n))$ that solves the optimization problem:

$$\begin{aligned} &\text{maximize} && G_n(\Lambda^+) \\ &\text{subject to} && \sum_{i \in \mathcal{I}_+(n)} A_{ji}\Lambda_i \leq C_j, \quad j \in \mathcal{J}, \\ &\text{over} && \Lambda_i \geq 0, \quad i \in \mathcal{I}_+(n), \end{aligned}$$

where for $n \in \mathbb{R}_+^{\mathbf{I}} \setminus \{0\}$ and $\Lambda^+ = (\Lambda_i : i \in \mathcal{I}_+(n)) \in \mathbb{R}_+^{|\mathcal{I}_+(n)|}$,

$$G_n(\Lambda^+) = \begin{cases} \sum_{i \in \mathcal{I}_+(n)} \kappa_i n_i^\alpha \frac{\Lambda_i^{1-\alpha}}{1-\alpha} & \text{if } \alpha \neq 1, \\ \sum\limits_{i \in \mathcal{I}_+(n)} \kappa_i n_i \log \Lambda_i & \text{if } \alpha = 1, \end{cases} \tag{1}$$

and the value of the right member above is taken to be $-\infty$ if $\alpha \in [1, \infty)$ and $\Lambda_i = 0$ for some $i \in \mathcal{I}_+(n)$. The resulting bandwidth allocation is called a *weighted α-fair allocation.*

From Lemma A.4 of Kelly and Williams [12], we know that for each fixed $n \in \mathbb{R}_+^{\mathbf{I}}$ there is $p \in \mathbb{R}_+^{\mathbf{J}}$ such that

$$p_j \left(C_j - \sum_{i \in \mathcal{I}} A_{ji}\Lambda_i(n) \right) = 0 \quad \text{for all } j \in \mathcal{J}, \tag{2}$$

$$n_i = \Lambda_i(n) \left(\frac{\sum_{j \in \mathcal{J}} p_j A_{ji}}{\kappa_i} \right)^{1/\alpha} \quad \text{for all } i \in \mathcal{I}. \quad (3)$$

The flow count process $N = (N_i : i \in \mathcal{I})$ is a Markov process with state space $\mathbb{Z}_+^{\mathbf{I}}$. We use the following (equivalent in distribution) representation for N and the cumulative unused capacity process $U = (U_j : j \in \mathcal{J})$:

$$N_i(t) = N_i(0) + E_i(t) - S_i(T_i(t)), \quad i \in \mathcal{I}, \quad (4)$$

$$U_j(t) = C_j t - \sum_{i \in \mathcal{I}} A_{ji} T_i(t), \quad j \in \mathcal{J}, \quad (5)$$

where E_i is a Poisson process with rate ν_i, S_i is a Poisson process with rate μ_i, $T_i(t)$ is the cumulative amount of bandwidth allocated to route i up to time t and

$$T_i(t) = \int_0^t \Lambda_i(N(s))\, ds. \quad (6)$$

It is assumed that $E_i, S_i, N_i(0),\ i \in \mathcal{I}$ are mutually independent. We define an (average) *workload process* by

$$W(t) = AM^{-1}N(t) \quad \text{for all } t \geq 0, \quad (7)$$

where $M = \text{diag}(\mu)$ is the $\mathbf{I} \times \mathbf{I}$ diagonal matrix with the entries of μ on its diagonal. Let $\rho_i = \nu_i/\mu_i$ for each $i \in \mathcal{I}$. We shall assume henceforth that the following *heavy traffic* condition holds.

Assumption 2.1:

$$A\rho = C. \quad (8)$$

Remark 2.1: In fact, one may consider a slightly more general situation in which a sequence of systems is considered with parameters converging to those satisfying the above heavy traffic condition. To simplify the exposition, we have not considered that more general situation here.

III. SEQUENCE OF SYSTEMS AND SCALING

Consider an increasing sequence of positive scale parameters $\{r_n\}$ which converges to infinity. To ease the notation, we shall simply write r in place of r_n, where it is understood that r increases to infinity through a sequence. We consider a sequence of stochastic systems indexed by r. Each member of the sequence is a stochastic system as described in the previous section. The processes E, S are kept fixed as r varies but $N(0)$ may vary with r. (Our results can be extended to the case where E, S also depend on r under additional assumptions. To simplify the exposition, we have chosen not to include that extension here.) We append a superscript of r to a process associated with the r^{th} system that depends on r. Thus, we have processes N^r, W^r, U^r, T^r, E, S.

We define fluid scaled processes $\overline{N}^r, \overline{W}^r, \overline{U}^r, \overline{T}^r, \overline{E}^r, \overline{S}^r$, as follows. For each r and $t \geq 0$, let

$$\begin{aligned} \overline{N}^r(t) &= N^r(rt)/r, & \overline{W}^r(t) &= W^r(rt)/r, \\ \overline{U}^r(t) &= U^r(rt)/r, & \overline{T}^r(t) &= T^r(rt)/r, \\ \overline{E}^r(t) &= E(rt)/r, & \overline{S}^r(t) &= S(rt)/r. \end{aligned}$$

We define diffusion scaled processes $\hat{N}^r, \hat{W}^r, \hat{U}^r$ as follows. For each r and $t \geq 0$, let

$$\begin{aligned} \hat{N}^r(t) &= \frac{N(r^2t)}{r}, \quad \hat{U}^r(t) = \frac{U(r^2t)}{r}, \\ \hat{W}^r(t) &= \frac{W(r^2t)}{r} = AM^{-1}\hat{N}^r(t), \\ \hat{E}^r(t) &= \frac{E(r^2t) - \nu r^2 t}{r}, \quad \hat{S}^r(t) = \frac{S(r^2t) - \mu r^2 t}{r}. \end{aligned}$$

For each $x \in \mathbb{R}^n$, let $|x|$ denote the Euclidean norm of x. For each r and $m = 0, 1, 2, \ldots$, let

$$\overline{n}_m^r = |\overline{N}^r(m)| \vee 1,$$

and for each $t \geq 0$, let

$$\begin{aligned} \overline{N}^{r,m}(t) &= \frac{\overline{N}^r(m + \overline{n}_m^r t)}{\overline{n}_m^r}, \\ \overline{W}^{r,m}(t) &= \frac{\overline{W}^r(m + \overline{n}_m^r t)}{\overline{n}_m^r} = AM^{-1}\overline{N}^{r,m}(t). \end{aligned}$$

IV. FLUID MODEL

A. Fluid Model Solution

A fluid model solution can be thought of as a formal limit of the sequence $\{\overline{N}^r\}$ as $r \to \infty$. The following notions are used in the definition below. A function $f = (f_1, \ldots, f_{\mathbf{I}}) : [0, \infty) \to \mathbb{R}_+^{\mathbf{I}}$ is absolutely continuous if each of its components $f_i : [0, \infty) \to \mathbb{R}_+$, $i = 1, \ldots, \mathbf{I}$, is absolutely continuous. A *regular point* for an absolutely continuous function $f : [0, \infty) \to \mathbb{R}_+^{\mathbf{I}}$ is a value of $t \in [0, \infty)$ at which each component of f is differentiable. (Since f is absolutely continuous, almost every time $t \in [0, \infty)$ is a regular point for f.)

Definition 4.1: A fluid model solution is an absolutely continuous function $n : [0, \infty) \to \mathbb{R}_+^{\mathbf{I}}$ such that at each regular point t for $n(\cdot)$ we have for each $i \in \mathcal{I}$,

$$\frac{d}{dt} n_i(t) = \begin{cases} \nu_i - \mu_i \Lambda_i(n(t)), & \text{if } n_i(t) > 0, \\ 0, & \text{if } n_i(t) = 0, \end{cases} \quad (9)$$

and for each $j \in \mathcal{J}$:

$$\sum_{i \in \mathcal{I}_+(n(t))} A_{ji} \Lambda_i(n(t)) + \sum_{i \in \mathcal{I}_0(n(t))} A_{ji} \rho_i \leq C_j, \quad (10)$$

where $\mathcal{I}_+(n(t)) = \{i \in \mathcal{I} : n_i(t) > 0\}$ and $\mathcal{I}_0(n(t)) = \{i \in \mathcal{I} : n_i(t) = 0\}$.

B. Invariant Manifold

Definition 4.2: A state $n_0 \in \mathbb{R}_+^{\mathbf{I}}$ is called invariant (for the fluid model) if there is a fluid model solution $n(\cdot)$ such that $n(t) = n_0$ for all $t \geq 0$. Let $\mathcal{M}_\alpha$ denote the set of all invariant states. We call $\mathcal{M}_\alpha$ the invariant manifold.

Various characterizations of the invariant states were given in [12]. We summarize these in Theorem 4.1 below. For this, we need the following definitions.

For each $n \in \mathbb{R}_+^{\mathbf{I}}$, define $w(n) = (w_j(n) : j \in \mathcal{J})$ to be given by

$$w_j(n) = \sum_{i \in \mathcal{I}} A_{ji} \frac{n_i}{\mu_i}, \quad j \in \mathcal{J}. \quad (11)$$

We call $w(n)$ the *workload* associated with n.

For each $w \in \mathbb{R}_+^{\mathbf{J}}$, define $\Delta(w)$ to be the unique value of $n \in \mathbb{R}_+^{\mathbf{I}}$ that solves the following optimization problem:

$$\begin{array}{ll} \text{minimize} & F(n) \\ \text{subject to} & \sum_{i\in\mathcal{I}} A_{ji}\frac{n_i}{\mu_i} \geq w_j, \quad j \in \mathcal{J}, \\ \text{over} & n_i \geq 0, \quad i \in \mathcal{I}, \end{array} \tag{12}$$

where

$$F(n) = \frac{1}{\alpha+1}\sum_{i\in\mathcal{I}} \nu_i\kappa_i\mu_i^{\alpha-1}\left(\frac{n_i}{\nu_i}\right)^{\alpha+1}, \quad n \in \mathbb{R}_+^{\mathbf{I}}. \tag{13}$$

The nonlinear function Δ has the two properties stated in the next proposition.

Proposition 4.1: The function $\Delta : \mathbb{R}_+^{\mathbf{J}} \to \mathbb{R}_+^{\mathbf{I}}$ is continuous. Furthermore, for each $w \in \mathbb{R}_+^{\mathbf{J}}$ and $c > 0$,

$$\Delta(cw) = c\Delta(w). \tag{14}$$

Proof: The first property is proved in [12] and the second is proved in [11]. ∎

Theorem 4.1: The following are equivalent.

(i) n is an invariant state (for the fluid model), i.e., $n \in \mathcal{M}_\alpha$,

(ii) $\Lambda_i(n) = \rho_i$ for all $i \in \mathcal{I}_+(n) = \{l \in \mathcal{I} : n_l > 0\}$,

(iii) there is $q \in \mathbb{R}_+^{\mathbf{J}}$ such that

$$n_i = \rho_i\left(\frac{\sum_{j\in\mathcal{J}} q_jA_{ji}}{\kappa_i}\right)^{1/\alpha} \quad \text{for all } i \in \mathcal{I}, \tag{15}$$

(iv) $n = \Delta(w(n))$.

Proof: This follows immediately from Lemma 5.1 and Theorems 5.1, 5.3 of [12]. ∎

The next three properties of fluid model solutions are used in the proof of Theorem 5.1. They follow from the statement and proof of Theorem 5.2 in [12] and the fact that for any fluid model solution $n(\cdot)$, $F(n(\cdot))$ is a non-increasing function.

Proposition 4.2: For each $R > 0$, there is a constant $D(R) \in [R, \infty)$ such that for any fluid model solution $n(\cdot)$ satisfying $|n(0)| \leq R$, we have $|n(t)| \leq D(R)$ for all $t \geq 0$.

The following theorem states that fluid model solutions converge uniformly to the invariant manifold $\mathcal{M}_\alpha$, where the uniformity applies across all fluid model solutions that start inside a compact subset of $\mathbb{R}_+^{\mathbf{I}}$. This result is critical to the proof of (multiplicative) state space collapse. For the statement of this theorem, we define the distance between $n \in \mathbb{R}_+^{\mathbf{I}}$ and $\mathcal{M}_\alpha$ by

$$d(n, \mathcal{M}_\alpha) = \inf\{|n - v| : v \in \mathcal{M}_\alpha\}. \tag{16}$$

Theorem 4.2: Fix $R \in (0, \infty)$ and $\varepsilon > 0$. There is a constant $T_{R,\varepsilon} \in [1, \infty)$ such that for each fluid model solution $n(\cdot)$ satisfying $|n(0)| \leq R$ we have

$$d(n(t), \mathcal{M}_\alpha) < \varepsilon \quad \text{for all } t \geq T_{R,\varepsilon}. \tag{17}$$

Proposition 4.3: For each $R \in (0, \infty)$ and $\varepsilon > 0$, there is $\delta > 0$ such that for any fluid model solution $n(\cdot)$ satisfying $|n(0)| \leq R$ and $d(n(0), \mathcal{M}_\alpha) < \delta$, we have $d(n(t), \mathcal{M}_\alpha) < \varepsilon$ for all $t \geq 0$.

V. (MULTIPLICATIVE) STATE SPACE COLLAPSE

Henceforth, for simplicity, we suppose that the systems start empty, i.e., $N^r(0) = 0$ for all r. (One can relax this condition to allow $N^r(0) \neq 0$ for some or all r, but then one needs to add an assumption so that state space collapse holds initially.)

For $0 \leq s < t < \infty$ and any bounded function $x : [s, t] \to \mathbb{R}_+^{\mathbf{I}}$, let $\|x(\cdot)\|_{[s,t]} = \sup_{u\in[s,t]} |x(u)|$ and when $s = 0$, let $\|x(\cdot)\|_t = \|x(\cdot)\|_{[0,t]}$.

The following result shows that (multiplicative) state space collapse holds. This property is critical to establishing a diffusion approximation result for $\hat{N}^r$.

Theorem 5.1: For each $T > 0$,

$$\frac{\left\|\hat{N}^r(\cdot) - \Delta\left(\hat{W}^r(\cdot)\right)\right\|_T}{\|\hat{N}^r(\cdot)\|_T \vee 1} \to 0 \tag{18}$$

in probability as $r \to \infty$.

Sketch of Proof. The proof of Theorem 5.1 proceeds in a similar manner to that in Bramson [4], where open multiclass queueing networks operating under certain head-of-the-line (HL) service disciplines are considered. A key ingredient in the proof is the fact that for each time $\overline{L} \geq 1$, there is a time $L > \overline{L}$ such that for each $\varepsilon > 0$, for all r sufficiently large, with probability at least $1 - \varepsilon$, the left member of (18) is dominated by

$$\left\|\overline{N}^{r,0}(\cdot) - \Delta\left(\overline{W}^{r,0}(\cdot)\right)\right\|_{[0,\overline{L}]} + \sup_{m=0}^{\lfloor rT\rfloor-1} \left\|\overline{N}^{r,m}(\cdot) - \Delta\left(\overline{W}^{r,m}(\cdot)\right)\right\|_{[\overline{L},L]}.$$

Here $\lfloor\cdot\rfloor$ denotes greatest integer part. In this way, study of a diffusion scaled quantity is reduced to study of $O(r)$ fluid scaled quantities. Indeed, with probability at least $1 - \varepsilon$, for r sufficiently large, the quantities $\{\overline{N}^{r,m}, m = 0, 1, \ldots, \lfloor rT\rfloor\}$ can be approximated by fluid model solutions over $[0, L]$. The results of Section IV on fluid model solutions, especially Theorem 4.2, can then be used to prove the desired result. For details of the proof, we refer the reader to [11].

VI. DIFFUSION APPROXIMATION

We are interested in obtaining a diffusion approximation for $\hat{W}^r$, where for each $t \geq 0$,

$$\begin{aligned} \hat{W}^r(t) &= \hat{X}^r(t) + \hat{U}^r(t), \\ \hat{X}^r(t) &= AM^{-1}\left(\hat{E}^r(t) - \hat{S}^r(\bar{\bar{T}}^r(t))\right), \\ \bar{\bar{T}}^r(t) &= \frac{T(r^2t)}{r^2}, \\ (\hat{S}^r(\bar{\bar{T}}^r(t)))_i &= \hat{S}_i^r(\bar{\bar{T}}_i^r(t)), \quad i \in \mathcal{I}. \end{aligned}$$

Formal manipulations using Theorem 5.1 and a representation of the bandwidth allocations in terms of Lagrange multipliers $p = (p_j : j \in \mathcal{J})$ suggest the following natural conjecture. To state this, we need the following notation. Let

$$\mathcal{W}_\alpha = AM^{-1}\mathcal{M}_\alpha \tag{19}$$

where

$$\mathcal{M}_\alpha = \Big\{ n \in \mathbb{R}_+^{\mathbf{I}} : n_i = \rho_i \left(\frac{\sum_{k\in\mathcal{J}} q_k A_{ki}}{\kappa_i} \right)^{\frac{1}{\alpha}} \text{ for all } i \in \mathcal{I} \text{ and some } q \in \mathbb{R}_+^{\mathbf{J}} \Big\}. \tag{20}$$

For each $j \in \mathcal{J}$, let $\mathcal{M}_\alpha^j$ be the subset of $\mathcal{M}_\alpha$ obtained by setting q_j equal to zero in (20), and let

$$\mathcal{B}_j = \{AM^{-1}n : n \in \mathcal{M}_\alpha^j\}. \tag{21}$$

Conjecture 6.1: $\{\hat{W}^r : r > 0\}$ converges in distribution to a diffusion process $\tilde{W}$ that is a semimartingale reflecting Brownian motion living in the cone $\mathcal{W}_\alpha$.

One difficulty in resolving this conjecture is that existence and uniqueness in law for the conjectured limit process $\tilde{W}$ is not known for all values of the parameters $\alpha, A, \nu, \mu, \kappa$. The case $\alpha = 1$ corresponds to proportional fair sharing of bandwidth and in this case the cone $\mathcal{W}_1$ is a polyhedral cone. Existence and uniqueness in law for this case is provided in Dai-Williams [5] and a suitable invariance principle which covers this case was recently established in Kang-Williams [10]. Consequently, for $\alpha = 1$ we are able to prove Theorem 6.1 below. Work in progress is directed towards establishing a similar result for other values of α.

We now define precisely what we mean by a semimartingale reflecting Brownian motion. For this, let Γ be the $\mathbf{J}\times\mathbf{J}$ covariance matrix given by

$$\Gamma = 2AM^{-1}\text{diag}(\nu)M^{-1}A', \tag{22}$$

where $'$ denotes transpose. Also, for each $j \in \mathcal{J}$, let e_j denote the unit vector that is parallel to the positive j^{th} coordinate direction in $\mathbb{R}_+^{\mathbf{J}}$.

Definition 6.1: A Semimartingale Reflecting Brownian Motion (abbreviated as SRBM) that lives in the cone $\mathcal{W}_1$, has direction of reflection e_j on the boundary face $\mathcal{B}_j$ for $j \in \mathcal{J}$, has covariance matrix Γ, and starts from the origin, is an $\{\mathcal{F}_t\}$-adapted, $\mathbf{J}$-dimensional process $\tilde{W}$ defined on some filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}, P)$ such that

- (i) P-a.s., $\tilde{W}(t) = \tilde{X}(t) + \tilde{Y}(t)$ for all $t \geq 0$,
- (ii) P-a.s., $\tilde{W}$ has continuous paths and $\tilde{W}(t) \in \mathcal{W}_1$ for all $t \geq 0$,
- (iii) under P,
  - (a) $\tilde{X}$ is a $\mathbf{J}$-dimensional Brownian motion starting from the origin with zero drift and covariance matrix Γ,
  - (b) $\{\tilde{X}(t), \mathcal{F}_t,\ t \geq 0\}$ is a martingale,
- (v) for each $j \in \mathcal{J}$, $\tilde{Y}_j$ is an $\{\mathcal{F}_t\}$-adapted, one-dimensional process such that P-a.s.,
  - (a) $\tilde{Y}_j(0) = 0$,
  - (b) $\tilde{Y}_j$ is continuous and non-decreasing,
  - (c) $\tilde{Y}_j(t) = \int_{(0,t]} 1_{\{\tilde{W}(s)\in\mathcal{B}_j\}} d\tilde{Y}_j(s)$.

The following condition is used in the next theorem. This can be interpreted as a *local traffic* assumption under which each resource has at least one route that only uses that resource.

Assumption 6.1: For each $j \in \mathcal{J}$ there is $i \in \mathcal{I}$ such that $A_{ji} = 1$ and $A_{ki} = 0$ for all $k \neq j$.

Theorem 6.1: Suppose that $\alpha = 1$ and that Assumption 6.1 holds. Then the sequence of processes $\{(\hat{W}^r, \hat{N}^r) : r > 0\}$ converges in distribution to a continuous process $(\tilde{W}, \tilde{N})$, where $\tilde{W}$ is a semimartingale reflecting Brownian motion that lives in the polyhedral cone $\mathcal{W}_1$ and $\tilde{N} = \Delta(\tilde{W})$. For each $j \in \mathcal{J}$, the direction of reflection for $\tilde{W}$ on the boundary face $\mathcal{B}_j$ is given by the unit vector e_j, the covariance matrix for $\tilde{W}$ is Γ, and $\tilde{W}$ starts from the origin.

Sketch of Proof. The proof uses a new invariance principle developed in [10] for semimartingale reflecting Brownian motions living in cones with piecewise constant reflection fields on the boundary. A key assumption for that theory to apply is that there is existence and uniqueness in law for the limit process $\tilde{W}$, which follows in the case $\alpha = 1$ from work of Dai and Williams [5]. The main difficulty in establishing the convergence is in showing for each $j \in \mathcal{J}$ that with probability tending to one as $r \to \infty$, the process $\hat{U}^r$ increases only when $\hat{W}^r$ is near the boundary portion $\mathcal{B}_j$. The local traffic assumption is used to verify this. For details of the proof, we refer the reader to [11].

VII. EXAMPLE

We illustrate Theorem 6.1 with a simple example. Suppose that $\mathcal{J} = \{1, 2\}$ and $\mathcal{I} = \{\{1\}, \{2\}, \{1, 2\}\}$, corresponding to a linear network with two resources and three routes (see Figure 1). Let $\alpha = 1$, $\kappa_i = \mu_i = 1$, for $i = 1, 2, 3$, $C_j = 1$ for $j = 1, 2$, and $\rho_1 + \rho_3 = \rho_2 + \rho_3 = 1$. Then the state space for the diffusion $\tilde{W}$ is the two-dimensional cone

$$\begin{aligned}\mathcal{W}_1 = \{&(w_1, w_2) : w_1 = \rho_1 q_1 + \rho_3(q_1 + q_2),\\ &w_2 = \rho_2 q_2 + \rho_3(q_1 + q_2),\ \text{for some } q \in \mathbb{R}_+^2\},\end{aligned}$$

which is the same as the cone

$$\{(w_1, w_2) : w_1 \geq 0,\ w_1\rho_3 \leq w_2 \leq w_1\rho_3^{-1}\}$$

pictured in Figure 2. Reflection occurs in the horizontal direction (corresponding to resource 1 incurring idleness) on the bounding face $w_1 = w_2\rho_3$. The interpretation of this is that although there is work for resource 1 within the system, congestion at resource 2 is preventing resource 1 from working at its full capacity. Similarly, vertical reflection (corresponding to resource 2 incurring idleness) on the bounding face $w_2 = w_1\rho_3$ is interpreted to mean that congestion at resource 1 is preventing resource 2 from working at its full capacity. In fact, for this example, with $\alpha \neq 1$, the workload cone $\mathcal{W}_\alpha$ is the same as for $\alpha = 1$ and Theorem 6.1 is also true in this case. However, in general, for higher dimensional workloads, i.e., $\mathbf{J} > 2$, the shape of the workload cone will depend on α, as well as on A, ν, μ, κ.

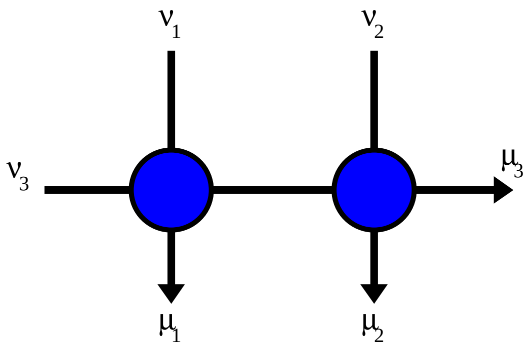

Fig. 1. A linear network with two resources and three routes.

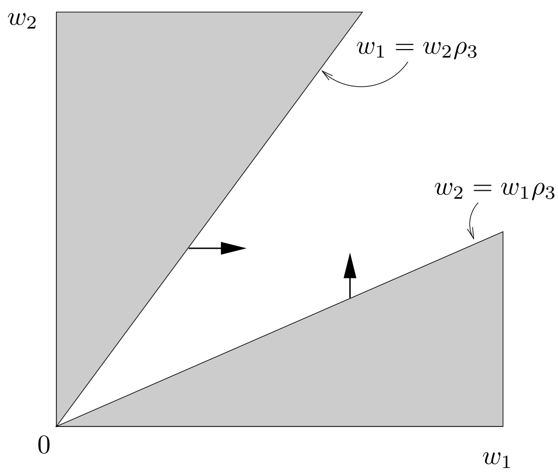

Fig. 2. The workload cone $\mathcal{W}_\alpha$ for a network with two resources, with workloads labelled w_1, w_2, and three routes, with traffic loads labelled ρ_1, ρ_2, ρ_3. Under the lifting map Δ, points (w_1, w_2) on the boundary $w_1 = w_2\rho_3$ are mapped to points (n_1, n_2, n_3) where $n_1 = 0$ (and the corresponding $q \in \mathbb{R}^2_+$ has $q_1 = 0$); similarly, points (w_1, w_2) on the boundary $w_2 = w_1\rho_3$ are mapped to points (n_1, n_2, n_3) where $n_2 = 0$ (and the corresponding $q \in \mathbb{R}^2_+$ has $q_2 = 0$).

For the linear network of Figure 1, consider an alternative bandwidth sharing policy that gives priority at each resource to flows on route 3 and allocates any remaining capacity at resource i to route i, for $i = 1, 2$. For this policy, the conjectured heavy traffic diffusion approximation $\tilde{W}$ has its state space equal to the positive quadrant and it has orthogonal reflection at the boundary. A resource is now idle only when there is no workload for it; moreover, a coupling can be constructed under which the sample paths for this diffusion are componentwise not greater than those for the earlier diffusion. Thus there exists a bandwidth sharing policy outside the class of weighted α-fair policies that avoids entrainment.

VIII. CONCLUDING REMARKS

A bandwidth sharing policy corresponds to a generalization of the notion of a processor sharing discipline from a single resource to a network with several shared resources. In particular, weighted α-fair policies provide a tractable theoretical abstraction of the bandwidth sharing effected by decentralized packet-based end-to-end congestion control algorithms such as TCP. It is known [2], [6] that, for flow-level models with exponentially distributed file sizes, weighted α-fair policies are stable when the average load on each resource is less than its capacity.

Weighted α-fair policies can nevertheless suffer from entrainment of resources, whereby congestion at some resources may prevent others from working at full capacity: this is manifested under diffusion scaling, where a Brownian model for the workload process lives in a cone which may be a strict subset of the positive orthant.

Under proportional fair sharing and a mild condition, this paper has shown how state space collapse can be combined with a new invariance principle to establish a Brownian model as a diffusion approximation for the flow level model.

Bandwidth sharing policies outside the class of weighted α-fair policies may avoid entrainment, although such policies may not be easy to effect via decentralized packet-based end-to-end congestion control algorithms.

REFERENCES

[1] S. Ben Fredj, T. Bonald, A. Proutiere, G. Regnie and J. Roberts (2001). Statistical bandwidth sharing: a study of congestion at flow level. *Proceedings of ACM Sigcomm 2001.*

[2] T. Bonald and L. Massoulié (2001). Impact of fairness on Internet performance. *Proceedings of ACM Sigmetrics 2001.*

[3] T. Bu and D. Towsley (2001). Fixed Point Approximation for TCP behavior in an AQM Network. *Proceedings of ACM Sigmetrics 2001.*

[4] M. Bramson (1998). State space collapse with application to heavy traffic limits for multiclass queueing networks. *Queueing Systems: Theory and Applications* **30** 89–148.

[5] J. G. Dai and R. J. Williams (1995). Existence and uniqueness of semimartingale reflecting Brownian motions in convex polyhedrons, *Theory of Probability and Its Applications* **40** 1–40.

[6] G. de Veciana, T.-J. Lee and T. Konstantopoulos (2001). Stability and performance analysis of networks supporting elastic services. *IEEE/ACM Transactions on Networking* **9** 2–14.

[7] R. J. Gibbens, S. K. Sargood, C. Van Eijl, F. P. Kelly, H. Azmoodeh, R. N. Macfadyen and N. W. Macfadyen (2000). Fixed-point models for the end-to-end performance analysis of IP networks. *13th ITC Specialist Seminar: IP Traffic Measurement, Modeling and Management,* Sept 2000, Monterey, California

[8] J. M. Harrison (2000). Brownian models of open processing networks: canonical representation of workload. *Annals of Applied Probability* **10** 75–103. Correction: **13** (2003), 390–393.

[9] J. M. Harrison (2003). A broader view of Brownian networks. *Annals of Applied Probability* **13** 1119–1150.

[10] W. Kang and R. J. Williams (2004). An invariance principle for semimartingale reflecting Brownian motions in cones with piecewise constant reflection fields, in preparation.

[11] W. Kang, F. P. Kelly, N. H. Lee and R. J. Williams (2004). State space collapse and diffusion approximation for a network operating under a fair bandwidth-sharing policy, in preparation.

[12] F. P. Kelly and R. J. Williams (2004). Fluid model for a network operating under a fair bandwidth-sharing policy. *Annals of Applied Probability* **14** 1055–1083.

[13] J. Mo and J. Walrand (2000). Fair end-to-end window-based congestion control. *IEEE/ACM Transactions on Networking* **8** 556–567.

[14] J. Roberts and L. Massoulié (1998). Bandwidth sharing and admission control for elastic traffic. *ITC specialists seminar,* Yokohama, 1998.

[15] M. Roughan, A. Erramilli and D. Veitch (2001). Network Performance for TCP Networks, Part I: Persistent Sources. *Proceeding ITC'17,* Brasil, September 2001.

[16] R. J. Williams (1998). Diffusion approximations for open multiclass queueing networks: sufficient conditions involving state space collapse. *Queueing Systems: Theory and Applications* **30** 27–88.