Proportional fairness in heavy traffic: process limit convergence and insensitivity of the limit process

Maria Vlasiou*, Jiheng Zhang†, and Bert Zwart‡

*Eindhoven University of Technology
†The Hong Kong University of Science and Technology
‡Centrum Wiskunde & Informatica, Amsterdam

May 1, 2017

Abstract

Proportional fairness is a popular service allocation mechanism to describe and analyze the performance of data networks at flow level. Recently, several authors have shown that the invariant distribution of networks operating according to proportional fairness admits a product form distribution under critical loading. They focus however on exponential job size distributions, leaving the case of general job size distributions as an open question. Motivated by this, we consider a network operating under proportional fairness where the job size distributions belong to a dense class of distributions. We establish a heavy-traffic process limit theorem and show that the invariant distribution of the limit process is determined by the first moments of the job sizes. Our analysis relies on a uniform convergence result for a fluid model, which is of independent interest.

AMS subject classification: 60K25, 68M20, 90B15

Keywords: Brownian approximations, Lyapunov functions, network utility maximization

1 Introduction

A popular way to model congestion of data traffic is to consider such traffic at a level where files or jobs are represented by continuous flows rather than discrete packets. This gives rise to bandwidth-sharing networks, as introduced in Massoulié and Roberts (1999). Such networks model the dynamic interaction among flows that compete for bandwidth along their source-destination paths. Apart from offering insight into the complex behavior of computer-communication networks, they have also been suggested recently as a model to analyze road-traffic congestion (see for instance Kelly and Williams (2010)). The analysis of bandwidth-sharing networks is challenging, requiring tools from both optimization and stochastics.

The most important bandwidth-allocation mechanism that has been considered so far is perhaps *proportional fairness*. In a static setting, this policy can be implemented in a distributed fashion, simultaneously maximizing users' utility; cf. Kelly (1997), Yi and Chiang (2008). In addition, proportional fairness is known to be the only policy that satisfies the four axioms of

the *Nash bargaining theory* (Mazumdar et al. (1991); Ştefănescu and Ştefănescu (1984)). These are desirable properties in a static setting. Furthermore, proportional fairness has attractive dynamic properties: while being a greedy policy, proportional fairness has also been shown to optimize some long term cost objectives, at least in a heavy traffic environment (Ye and Yao (2012)). In particular, it is known to be stable in internet flow-level models (Massoulié (2007)), assuming phase-type document size distributions. Recently, proportional fairness has been suggested as an attractive alternative to maximum-pressure policies in Walton (2014b).

In some special cases detailed below, a bandwidth-sharing network operating under proportional fairness admits a (computable) invariant distribution of the number of users. As these cases are rather restrictive, it is natural to obtain insight in the performance of proportional fairness for more general network topologies. In Kang et al. (2009), it is shown, assuming exponential job size distributions, that the performance of proportional fairness is still tractable if the network is heavily loaded. Under a heavy traffic assumption, a limit theorem is developed yielding an approximating semimartingale reflected Brownian motion (SRBM), of which the invariant distribution is shown to have a product form. A restrictive assumption in Kang et al. (2009) is the so-called 'local traffic assumption', stating that each link in the network serves a route consisting only of that link. This was removed in Ye and Yao (2012) by using elegant geometric arguments. While Ye and Yao (2012) allows for generally distributed flow sizes, it assumes that the service policy within a class is first-in-first-out (FIFO). This is well-suited for packet-level models Walton (2014a), but not for flow-level models, which has been the main focus in the literature. In the present paper, we focus on flow-level models, where the per-class discipline is Processor Sharing (PS). This discipline is harder to analyze than FIFO and corresponds to the original open question posed in Kang et al. (2009). A recent survey on these developments can be found in Williams (2015).

While the Poisson arrival assumption can often be justified to some degree in practice, the same cannot be said for exponential job size distributions. As such, it is desirable for the performance of a network to be insensitive to fluctuations in higher moments of the job size distribution. There is overwhelming statistical evidence that the variance of file sizes is in fact infinite (Resnick (1997)). As perfectly stated in Bonald and Proutière (2003): "the practical value of insensitivity is best illustrated by the enduring success of Erlang's loss formula in telephone networks". In Bonald and Proutière (2003), it is shown that proportional fairness is the only utility-maximizing policy that yields this insensitivity property, provided that the network topology has a hypercube structure and that all servers work at the same speed. Given these limitations on the insensitivity of proportional fairness, some related allocation mechanisms have been suggested that yield insensitivity for arbitrary networks topologies. One such suggestion, based on connections with Whittle networks, is *balanced fairness* (Bonald and Proutière (2003)). Another suggestion (Massoulié (2007)) is modified proportional fairness. However, neither of these two policies are utility maximizing.

Though proportional fairness itself may not be always insensitive, it remains a key allocation mechanism for the reasons mentioned above. In fact, the key question addressed but left open in both Kang et al. (2009) and Ye and Yao (2012), is whether the product form property of their heavy traffic approximation, derived for exponential job sizes, still holds for more general job size distributions, yielding insensitivity of proportional fairness in heavy traffic. Informally, the main conjecture is whether the vector N of the number of users along each route in steady

state can be approximated as follows:

$$N \approx \text{diag}\,(\rho)\, A^T E_s. \tag{1.1}$$

Here, $\text{diag}\,(\rho)$ is a diagonal matrix having the load of each route on the diagonal, A is a 0-1 matrix encoding which server (link) is used by which route, and E_s is a vector of independent exponential random variables. Each random variable corresponds to a server and has as parameter the slack of that resource; i.e., if c is the vector of service speeds, then $s = c - A\rho$. The random variables E_s can be interpreted as equilibrium values of the Lagrange multipliers associated with the resources. In Walton (2014a), this property is called *product form resource pooling*. In addition, Jonckheere and López (2014) establish insensitivity of large deviation rate functions assuming the network has a tree topology. Other recent developments of proportional fairness are described in Harrison et al. (2014).

Though we are not resolving the conjecture (1.1) in this paper, we develop several results that support this conjecture. In particular, our main results are Theorems 4.1, 6.1 and 7.1 below. To derive these results, we adapt the state-space collapse approach of Bramson (1998); Williams (1998); Stolyar (2004) to our setting, building also on Bramson (1996); Kang et al. (2009); Massoulié (2007); Ye and Yao (2012). Specifically, we first investigate a fluid model assuming the system is critically loaded, and define a critical fluid model, significantly extending and simplifying the treatment of a Lyapounov function that was introduced by Massoulié (2007) in the subcritical case. Adapting and extending techniques from Ye and Yao (2012) and Kang et al. (2009), we investigate the set of invariant points of the fluid model. Our set-up is related to Ye and Yao (2012) in terms of the assumptions on the network topology. However, unlike Ye and Yao (2012), we do not need to assume that the service discipline within a class is FIFO. Instead, we consider Processor Sharing, as in all other works in this domain. We are able to extend the geometric ideas in Ye and Yao (2012) to deal with some form of routing in the network, which is sufficiently general to deal with phase-type distributions. Our results can be seen as extensions to networks of heavy-traffic limits for single-node single class Processor Sharing queues, as derived in Gromoll (2004); Puha and Williams (2004).

The main technical challenge of this paper is to show that fluid model solutions converge uniformly and at an exponential rate to an invariant point, which is Theorem 4.1. Ideas from Bramson (1996) and Massoulié (2007) form a useful starting point, but the analysis pertaining to our setting demands significant additional work. In particular, though we use the same candidate Lyapounov function that Massoulie Massoulié (2007) used in the analysis of the sub-critical regime, the analysis in the critical regime is much harder. Our main idea is a novel application of a rearrangement inequality, significantly simplifying Massoulié (2007). The resulting upper bound on the derivative of this function is then bounded further using properties like the utility-maximizing nature of proportional fairness. The fact that the proportionally fair bandwidth-allocation function may be discontinuous at the boundary complicates the analysis. The analysis of the fluid model is valid for general Markovian routing; we expect the convergence result to be useful beyond its present application, though we need to assume that all external arrival rates are positive. With the uniform convergence of fluid model solutions in place, the remaining steps follow arguments similar to Ye and Yao (2012), using in particular some of their intermediate results. This yields the diffusion limit in Theorem 6.1.

In our analysis, we additionally assume that job size distributions have a particular phase-

type structure, which is non-restrictive in the sense that any distribution with nonnegative support can be approximated arbitrary closely by such a phase-type distribution. This assumption is technically convenient as it allows for a finite-dimensional Markovian description of the system. Extending our results to more general distributions requires a measure-valued state descriptor and is beyond the scope of the techniques developed in this paper. Note that this would still not cover the practically relevant case of job sizes with infinite variance, which has not even been resolved even in the single-node single-class case, cf. Lambert et al. (2013). In the present paper, second moments show up in the description of the process limit, but cancel out against one another while computing the invariant distribution of the SRBM, using the skew symmetric condition developed by Harrison and Williams (1987). In particular, we characterize and simplify the invariant distribution in Theorem 7.1 using, among others, renewal-theoretic arguments in the computations.

To provide further context to our results, we note that Conjecture (1.1) relies on the assumption that a link in the network is work-conserving. When individual users have additional constraints on their individual access rates, (1.1) no longer holds and the distribution of N is better approximated by a multivariate normal, cf. Reed and Zwart (2014). When relaxing the assumption of proportional fairness to other utility-maximizing bandwidth-allocation policies, the theory becomes much harder and is still partly conjectural, as the resulting SRBMs no longer live in polyhedral domains, cf. Kang and Williams (2007); Kang et al. (2009). In this case, the simple approximation (1.1) cannot be expected to hold. Another assumption is that A is of full row rank. In Kelly et al. (2009), it is shown that (1.1) may not hold in in general if A is not of full row rank. Extensions to multi-path routing, of which its nature and importance is described in Kang et al. (2009), require the elements of A to be nonnegative rather than 0-1. The methodology developed in the present paper can deal with this more general case. To prove (1.1) for phase-type distributions, an interchange of the heavy traffic and steady state limits is required. In the case of exponential job sizes, this interchange is established in Shah et al. (2014) (using the process limit that was derived in Kang et al. (2009); Ye and Yao (2012)). Unfortunately, we were not able to utilize the existing methods and techniques from Gamarnik and Zeevi (2006); Budhiraja and Lee (2009); Gurvich (2014); Braverman et al. (2015); Ye and Yao (2016) to resolve this interchange problem, and we (have to) leave this question open. We hope our work stimulates more research in this direction.

The paper is organized as follows. The network model and some assumptions are introduced in Section 2. In Section 3, we give a detailed description of the dynamics of our model. An auxiliary fluid model with general Markovian routing is introduced and analyzed in detail in Section 4. In Section 5, we develop a suitable decomposition of our process. The diffusion limit is given in Section 6, and its invariant distribution is computed in Section 7.

2 The network model

In this section, we provide a detailed model description. As we make heavy use of results from Ye and Yao (2012), we follow their notation whenever possible. All vectors are column vectors. Throughout the paper, e is a column vector with all elements equal to 1 and I denotes the identity matrix. The dimensions of e and I should be clear from the context.

Network structure. The network consists of a set of routes $\mathcal{R} = \{1, \ldots, R\}$, which are typically indexed by r. Each route traverses several links, which are indexed by l, $l \in \mathcal{L} = \{1, \ldots, L\}$. Each link has a service capacity c_l. Let A denote the *link-route* matrix of dimension $L \times R$. We set $A_{l,r} = 1$ if route r needs 1 unit of capacity from link l and 0 otherwise. Assume A has full row rank; hence $L \leq R$. We note that all arguments in the paper remain valid if A is a nonnegative matrix of full row rank.

Stochastic assumptions. Next, we introduce the arrival process and service time assumptions. We assume for convenience that arrival processes are Poisson with rate λ_r. Service times at route r follow a phase type distribution with F_r phases. The set $\mathcal{F}_r = \{1, \ldots, F_r\}$ contains all phases for jobs on route r. As is commonplace (cf. Asmussen (2003)), a phase-type random variable is the lifetime of an absorbing Markov chain with initial distribution $\boldsymbol{a}_r = (\boldsymbol{a}_{r,1}, \ldots, \boldsymbol{a}_{r,F_r})^T \in \mathbb{R}_+^{F_r}$, sub-stochastic transition matrix $P^r \in \mathbb{R}^{F_r} \times \mathbb{R}^{F_r}$ with $P^r_{i,j}$ being the transition probability from state i to state j and $1 - \sum_j P^r_{i,j}$ being the transition probability from state i to the absorbing state (which corresponds to service completion). Furthermore, it is assumed that the service time in phase f is exponentially distributed with rate $\boldsymbol{\mu}_{r,f}$, and we write $\boldsymbol{\mu}_r = (\boldsymbol{\mu}_{r,1}, \ldots, \boldsymbol{\mu}_{r,F_r})^T \in \mathbb{R}_+^{F_r}$. In particular, the mean service time in phase f on route r is $\boldsymbol{m}_{r,f} = \frac{1}{\boldsymbol{\mu}_{r,f}}$, and $\boldsymbol{m}_r = (\boldsymbol{m}_{r,1}, \ldots, \boldsymbol{m}_{r,F_r})^T \in \mathbb{R}_+^{F_r}$. We assume

$$\lambda_r \boldsymbol{a}_{r,f} > 0 \text{ for all } f \in \mathcal{F}_r \text{ and } r \in \mathcal{R}, \tag{2.1}$$

$$(I - P^r) \text{ is invertible.} \tag{2.2}$$

The first assumption, that all routes have arrivals for each phase, is non-standard, and required in our analysis in Section 4. It is non-restrictive in the sense that an inspection of the proof of (Asmussen, 2003, Theorem III.4.2) shows that the resulting class of distributions is still dense in the class of all distributions with nonnegative support. Let $P^{r,T}$, $\boldsymbol{a}_r^T$ and $\boldsymbol{m}_r^T$ denote the transpose of P^r, $\boldsymbol{a}_r$ and $\boldsymbol{m}_r$. Then the mean service requirement β_r at route r is

$$\beta_r = \boldsymbol{m}_r^T (I - P^{r,T})^{-1} \boldsymbol{a}_r. \tag{2.3}$$

State-space description. Denote the R-dimensional vector of jobs on each route by $n = (n_1, \ldots, n_R)^T$, with n_r being the number of jobs on route $r \in \mathcal{R}$. To obtain a Markovian description of our network, it is useful to introduce a more detailed state-space descriptor:

$$\boldsymbol{n} = \left(\boldsymbol{n}_{1,1}, \ldots, \boldsymbol{n}_{1,F_1}, \ \ldots\ldots \ , \boldsymbol{n}_{R,1}, \ldots, \boldsymbol{n}_{R,F_R}\right)^T, \tag{2.4}$$

with $\boldsymbol{n}_{r,f}$ denoting the number of jobs on phase f at route r. It is clear that $\boldsymbol{n}$ is a $\sum_{r \in \mathcal{R}} F_r$-dimensional vector and $n_r = \sum_{f \in \mathcal{F}_r} \boldsymbol{n}_{r,f}$. We also need a *link-phase* matrix, denoted by $\boldsymbol{A}$, which is of dimension $L \times \sum_{r \in \mathcal{R}} F_r$. For a link $l \in \mathcal{L}$ and for all $f = \sum_{r'=1}^{r-1} F_{r'} + 1, \ldots, \sum_{r'=1}^{r} F_{r'}$, we have

$$\boldsymbol{A}_{l,f} = A_{l,r}, \quad r \in \mathcal{R}. \tag{2.5}$$

Thus, $\boldsymbol{A}$ is obtained by taking the rth column of A and repeating it for F_r times. From now on, when we make a distinction between routes and phases, we speak of 'route level' and 'phase level'. The associated notation will be distinguished by using boldface.

Traffic load. The route-level traffic load for each $r \in \mathcal{R}$ is

$$\rho_r = \lambda_r \beta_r. \tag{2.6}$$

Denote $\rho = (\rho_1, \ldots, \rho_R)^T \in \mathbb{R}_+^R$ and $c = (c_1, \ldots, c_L)^T \in \mathbb{R}_+^L$, then

$$A\rho = c. \tag{2.7}$$

Let A_l be the lth row of A. A link l is said to be a bottleneck if $A_l \rho = c_l$. For convenience, we assume that all links are a bottleneck. This assumption can be removed along the lines of the electronic companion of Ye and Yao (2012). Note however that we assume (2.7) for our limiting process. Later on, we introduce a sequence of processes, indexed by $k = 1, 2, 3, \ldots$, for which $c - A\rho^k$ is of the order $1/k$.

Let $\mathrm{diag}\,(x)$ be a square matrix with the diagonal being equal to the vector x and all off-diagonal elements being equal to 0. The traffic load for each route r at the phase level is defined as

$$\boldsymbol{\rho}_r = \lambda_r [\mathrm{diag}\,(\boldsymbol{m}_r)\,(I - P^{r,T})^{-1} \boldsymbol{a}_r]. \tag{2.8}$$

In other words, $\boldsymbol{\rho}_r = (\boldsymbol{\rho}_{r,1}, \ldots, \boldsymbol{\rho}_{r,F_r})^T \in \mathbb{R}_+^{F_r}$. It is clear from (2.3) and (2.6) that the aggregated load for each route r is

$$\rho_r = \sum_{f \in \mathcal{F}_r} \boldsymbol{\rho}_{r,f}. \tag{2.9}$$

Proportional fairness allocation. Denote by $\Lambda_r(n)$, $r \in \mathcal{R}$, the capacity allocated to route r jobs when the network status is n. Let Γ denote the set of all feasible allocations, i.e.

$$\Gamma = \left\{ \gamma \in \mathbb{R}^R : A\gamma \leq c, \gamma \geq 0 \right\}. \tag{2.10}$$

The proportional fair allocation $\Lambda(n)$ is the unique solution to the optimization problem

$$\max_{\gamma \in \Gamma} \sum_{r \in \mathcal{R}} n_r \log(\gamma_r), \tag{2.11}$$

with $\Lambda_r(n) = 0$ if $n_r = 0$. According to the optimality condition, this optimal solution to (2.11) satisfies

$$\frac{n_r}{\Lambda_r(n)} = \sum_{l \in \mathcal{L}} A_{l,r} \eta_l, \quad r \in \mathcal{R}, \tag{2.12}$$

for some $\eta = (\eta_l)_l \in \mathbb{R}_+^L$. It is known that Λ is directionally differentiable on $(0, \infty)^R$ by Reed and Zwart (2014) (earlier Kelly and Williams (2004) established continuity). In addition, Λ is radially homogeneous, i.e. $\Lambda(yn) = \Lambda(n)$ for $y > 0$, cf. Kelly and Williams (2004).

The allocation to each phase f on route r is $\boldsymbol{\Lambda}_{r,f}(\boldsymbol{n}) = \frac{\boldsymbol{n}_{r,f}}{n_r} \Lambda_r(n)$, where we make the convention throughout the paper that $0/0 = 0 \times \infty = 0$. This is consistent with the fact that $\boldsymbol{\Lambda}(\boldsymbol{n})$, as a $\sum_{r \in \mathcal{R}} F_r$-dimensional vector, is the optimal solution to

$$\max_{\gamma \in \boldsymbol{\Gamma}} \sum_{r \in \mathcal{R}, f \in \mathcal{F}_r} \boldsymbol{n}_{r,f} \log(\boldsymbol{\gamma}_{r,f}), \tag{2.13}$$

where

$$\boldsymbol{\Gamma} = \left\{ \boldsymbol{\gamma} \in \mathbb{R}^{\sum_{r \in \mathcal{R}} F_r} : \boldsymbol{A}\boldsymbol{\gamma} \leq c, \boldsymbol{\gamma} \geq 0 \right\}.$$

The extended vector $\boldsymbol{\gamma}$, together with $\boldsymbol{\mu}$, $\boldsymbol{m}$ and $\boldsymbol{\rho}$, is interpreted in the same way as (2.4). Extending (2.12) to phase level yields

$$\frac{\boldsymbol{n}_{r,f}}{\boldsymbol{\Lambda}_{r,f}(\boldsymbol{n})} = \sum_{l\in\mathcal{L}} A_{l,r}\eta_l, \quad r \in \mathcal{R}, \tag{2.14}$$

for some $\eta = (\eta_l)_l \in \mathbb{R}_+^L$.

3 System dynamics

Consider a sequence of systems indexed by $k = 1, 2, \ldots$. For the kth system, let $\boldsymbol{N}_{r,f}^k(t)$ denote the number of jobs on route r at phase f; then, $N_r^k(t) = \sum_{f\in\mathcal{F}_r} \boldsymbol{N}_{r,f}^k(t)$ denotes the total number of jobs on route r. Set the column vector $\boldsymbol{N}^k(t) = (\boldsymbol{N}_{r,f}^k(t))$. The resource allocated to phase f on route r at time t is $\boldsymbol{\Lambda}_{r,f}(\boldsymbol{N}^k(t))$ according to (2.13).

For convenience, set $P_{f,0}^r = 1 - \sum_{f'\in\mathcal{F}_r} P_{f,f'}^r$. Let $\boldsymbol{E}_{r,f}(\cdot)$ and $\boldsymbol{S}_{r,f,f'}(\cdot)$, $r \in \mathcal{R}$, $f \in \mathcal{F}_r$, $f' \in \mathcal{F}_r \cup \{0\}$, denote independent unit rate Poisson processes. In the remainder, symbols without subscript denote the column vector with the corresponding component; e.g., $\boldsymbol{E}(\cdot) = (\boldsymbol{E}_{r,f}(\cdot))$. The dynamics of $\boldsymbol{N}^k(t)$ can be written as

$$\begin{aligned}\boldsymbol{N}_{r,f}^k(t) = \boldsymbol{N}_{r,f}^k(0) + \boldsymbol{E}_{r,f}(\lambda_r^k \boldsymbol{a}_{r,f} t) + \sum_{f'\in\mathcal{F}_r} \boldsymbol{S}_{r,f',f}\Big(\boldsymbol{\mu}_{r,f'} P_{f',f}^r \boldsymbol{D}_{r,f'}(t)\Big) \\ - \sum_{f'\in\mathcal{F}_r\cup\{0\}} \boldsymbol{S}_{r,f,f'}\Big(\boldsymbol{\mu}_{r,f} P_{f,f'}^r \boldsymbol{D}_{r,f}^k(t)\Big),\end{aligned} \tag{3.1}$$

where

$$\boldsymbol{D}_{r,f}^k(t) = \int_0^t \boldsymbol{\Lambda}_{r,f}(\boldsymbol{N}^k(s))ds. \tag{3.2}$$

As users at a given route and phase may not leave the network immediately, we define a phase-based workload $\boldsymbol{W}^k(t)$ as a $\sum_{r\in\mathcal{R}} F_r$-dimensional vector interpreted as in (2.4). In particular, setting

$$\boldsymbol{P} = \begin{pmatrix} P^1 & 0 & 0 & 0 \\ 0 & P^2 & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & P^R \end{pmatrix}, \tag{3.3}$$

the phase-base workload is now defined as

$$\boldsymbol{W}^k(t) = \operatorname{diag}(\boldsymbol{m})\,(I - \boldsymbol{P}^T)^{-1}\boldsymbol{N}^k(t). \tag{3.4}$$

This is not the true workload, but a convenient proxy and a customary choice in heavy-traffic analysis; see Harrison (2000) for background.

Define the centered processes

$$\breve{\boldsymbol{E}}_{r,f}^k(t) = \boldsymbol{E}_{r,f}(\lambda_r^k \boldsymbol{a}_{r,f} t) - \lambda_r^k \boldsymbol{a}_{r,f} t, \tag{3.5}$$

$$\breve{\boldsymbol{S}}_{r,f,f'}^k(t) = \boldsymbol{S}_{r,f,f'}^k(\boldsymbol{\mu}_{r,f} P_{f,f'}^r t) - \boldsymbol{\mu}_{r,f} P_{f,f'}^r t, \tag{3.6}$$

and

$$\left(\breve{\boldsymbol{S}}^k_+(\boldsymbol{D}^k(t))\right)_{r,f} = \sum_{f'\in\mathcal{F}_r} \breve{\boldsymbol{S}}^k_{r,f',f}(D^k_{r,f'}(t)), \tag{3.7}$$

$$\left(\breve{\boldsymbol{S}}^k_-(\boldsymbol{D}^k(t))\right)_{r,f} = \sum_{f'\in\mathcal{F}_r\cup\{0\}} \breve{\boldsymbol{S}}^k_{r,f,f'}(D^k_{r,f}(t)). \tag{3.8}$$

Define the vector $\boldsymbol{\rho}^k$ as in (2.8) with λ_r replaced by λ^k_r. It follows from (3.1) and (3.2) that

$$\boldsymbol{W}^k(t) = \boldsymbol{W}^k(0) + \boldsymbol{X}^k(t) + \int_0^t [\boldsymbol{\rho} - \boldsymbol{\Lambda}(\boldsymbol{N}^k(s))]ds, \tag{3.9}$$

where

$$\boldsymbol{X}^k(t) = (\boldsymbol{\rho}^k - \boldsymbol{\rho})t + \mathrm{diag}\,(\boldsymbol{m})\,(I - \boldsymbol{P}^T)^{-1}\left[\breve{\boldsymbol{E}}^k(t) + \breve{\boldsymbol{S}}^k_+(\boldsymbol{D}^k(t)) - \breve{\boldsymbol{S}}^k_-(\boldsymbol{D}^k(t))\right]. \tag{3.10}$$

Let

$$Y^k(t) = \boldsymbol{A}\int_0^t [\boldsymbol{\rho} - \boldsymbol{\Lambda}(\boldsymbol{N}^k(s)]ds. \tag{3.11}$$

It is easily seen that

$$Y^k(t) = \int_0^t [c - \boldsymbol{A}\boldsymbol{\Lambda}(\boldsymbol{N}^k(s))]ds = \int_0^t [c - A\Lambda(N^k(s))]ds.$$

The vector $Y^k(t)$ can be interpreted as cumulative service capacities that have not been used in $[0,t]$.

4 A fluid model and its convergence to equilibrium

The goal of this self-contained section is to introduce and analyze a fluid model. We consider a more general setting: rather than analyzing the model at the phase level, we assume there is a general routing matrix P between different routes. The precise condition P needs to satisfy is stated in (4.1). Completed jobs from route r have probability $P_{r,r'}$ to be routed to route r'. It is clear that this setting is more general than the phase-type model introduced in Section 2, where routing is only restricted within phases of each route. This also allows us to simplify the notation in this section.

The overview of the present section is as follows.

1. We introduce a fluid model for a model with general routing, which is related to the fluid model in Massoulié (2007) – in fact, we add another requirement to the definition of Massoulié (2007), so that a function which is a fluid model in our sense, also satisfies the requirements in Massoulié (2007). We introduce an entropy-like function, which was shown in Massoulié (2007) to be a Lyapunov function in the sub-critically loaded case.

2. We show that the entropy-like function remains a Lyapunov function under critical loading. This requires a careful analysis, as also stipulated in Bramson (1996), who considered subcritical and critical fluid models of head of the line PS systems. As in Massoulié (2007), we use classical rearrangement inequalities, but we do so in an entirely different way: we show that the derivative of the Lyapunov function can be rewritten as the expected value

of a path functional of a terminating Markov chain, for which we obtain pathwise bounds (see proof of Lemma 4.2). Our arguments can provide a substantial simplification of the subcritical case treated in Massoulié (2007).

3. Using the bound of the derivative of the Lyapunov function, we then proceed to prove uniform convergence of fluid model solutions towards the invariant manifold, leading to Theorem 4.1. On a high level, our approach is similar to that of Bramson (1996):

   (a) Find a function L that is a Lyapunov function; i.e., show that $f(t) = L(n(t))$ has negative derivative bounded by $-g(n(t))$, with g a nonnegative function.

   (b) Show that $f(t) \leq |n(t)|Kg(n(t))$ for some constant K independent of $n(t)$.

   (c) The two inequalities combined give $f'(t) \leq -f(t)/(K|n(t)|)$. By bounding $|n(t)|$ in terms of $|n(0)|$, we get uniform rates of convergence of $f(t)$ to 0, leading to uniform convergence of $n(t)$ for all fluid models starting in a compact set.

   On a more detailed level, our arguments are different. Apart from simplifying and extending ideas from Massoulié (2007), we develop and use several additional properties of proportional fairness in the process.

In this section, we use lower case for fluid model quantities, such as $n(t)$.

4.1 A fluid model

Consider a network model with Poisson arrival vector λ, exponential service rate vector μ, and general routing matrix P that is no longer block-diagonal. Define $\rho = \text{diag}\left(\frac{1}{\mu}\right)(1 - P^T)^{-1}\lambda$. Later, in Section 6, we apply results obtained in this section by specializing the routing matrix to the block-diagonal $\boldsymbol{P}$ defined in (3.3). For notational simplicity, we still use $\mathcal{R}$ to denote the set of routes. The two assumptions we invoke are

$$I - P^T \text{ is invertible,} \tag{4.1}$$

$$\lambda_r > 0 \text{ for all } r \in \mathcal{R}. \tag{4.2}$$

The latter assumption is required for the analysis in this section. Recall that $\Lambda_r(n(t))$ solves the problem (2.11). In what follows, we mainly follow the notation of Massoulié (2007). We can now present our definition of a fluid model.

Definition 4.1 (Fluid Model)**.** *A fluid model solution is a vector-valued function* $\{n(t), t \geq 0\}$ *where for* $t \geq 0$, $n(t) = (n_r(t))_{r\in\mathcal{R}}$ *satisfies the following two conditions: 1) For each* $r \in \mathcal{R}$, $n_r(\cdot)$ *is a nonnegative function that is absolutely continuous with respect to the Lebesgue measure and for almost every* t

$$\dot{n}_r(t) = \lambda_r - \mu_r\Phi_r(n(t)) + \sum_{s\in\mathcal{R}} P_{s,r}\mu_s\Phi_s(n(t)), \tag{4.3}$$

where

$$\Phi_r(n(t)) \begin{cases} = \Lambda_r(n(t)), & \text{if } n_r(t) > 0, \\ \in [0, \limsup_{y\to n(t)} \Lambda_r(y)], & \text{if } n_r(t) = 0. \end{cases} \tag{4.4}$$

2) For almost every $t \geq 0$,

$$\sum_{r \in \mathcal{R}} A_{l,r} \Phi_r(n(t)) \leq c_l \quad \textit{for all} \;\; l \in \mathcal{L}. \tag{4.5}$$

Last, define the auxiliary functions $w(\cdot)$ and $y(\cdot)$ by

$$w(t) = \mathrm{diag}\,(m)\,(I - P^T)^{-1} n(t), \tag{4.6}$$

$$y(t) = A \int_0^t [\rho - \Phi(n(s))] ds. \tag{4.7}$$

This definition of a fluid model solution is essentially the same as the one in Massoulié (2007), though we also require (4.5). As our fluid model solutions also are fluid model solutions in the sense of Massoulié (2007), we can exploit properties developed in that work. We call any t for which (4.3)–(4.5) are satisfied for all routes $r \in \mathcal{R}$, a regular point. If t is regular, we will often say that the associated state $n(t)$ is regular. We now provide a more explicit representation for $\Phi_r(n(t))$ for any regular t. For each $t \geq 0$, introduce

$$\mathcal{R}_0(t) = \{r \in \mathcal{R} : n_r(t) = 0\} \;\; \text{and} \;\; \mathcal{R}_+(t) = \{r \in \mathcal{R} : n_r(t) > 0\}. \tag{4.8}$$

Since any fluid model solution is nonnegative, (4.8) gives a partition of all the routes. Note that $\mathcal{R}_0$ and $\mathcal{R}_+$ depends on t. We drop the notation t in the following when the context is clear. It is immediate that for any regular t, $n_r(t) = \dot{n}_r(t) = 0$ for $r \in \mathcal{R}_0(t)$ because $n_r(s) \geq 0$ for all $s \geq 0$ and $r \in \mathcal{R}$. This implies that for each regular t

$$\lambda_r - \mu_r \Phi_r(n(t)) + \sum_{s \in \mathcal{R}_0} \mu_s \Phi_s(n(t)) P_{s,r} + \sum_{s \in \mathcal{R}_+} \mu_s \Lambda_s(n(t)) P_{s,r} = 0, \quad r \in \mathcal{R}_0. \tag{4.9}$$

This gives an affine relationship between $(\Phi_r)_{r \in \mathcal{R}_0}$ and $(\Lambda_r)_{r \in \mathcal{R}_+}$. Such an affine relationship depends on the set $\mathcal{R}_+$, which can take only finitely many different values. Thus, we can derive the scalability of Φ from that of Λ; i.e., for any fluid model solution $n(t)$ at a regular t and a scalar $y > 0$,

$$\Phi(n(t)) = \Phi(y n(t)). \tag{4.10}$$

The main goal of this section is to give a proof of the following result:

Theorem 4.1. *Assume* (4.1) *and* (4.2). *Let $n(\cdot)$ be a fluid model solution. If $|n(0)| < M$ for some constant $M > 0$, then for all $\epsilon > 0$, there exists a time $T_{M,\epsilon}$ (not depending on $n(\cdot)$) and a state $n(\infty)$, such that*

$$|n(t) - n(\infty)| < \epsilon \quad \textit{for all} \;\; t > T_{M,\epsilon}.$$

This theorem will be a key tool in the derivation of the diffusion limit later on. The remainder of the current section is devoted to its proof.

4.2 A Lyapunov function

Introduce

$$L(n(t)) = \sum_{r \in \mathcal{R}} n_r(t) \log \left(\frac{\Phi_r(n(t))}{\rho_r} \right). \tag{4.11}$$

Note that $0 \log 0$ is set to be 0. For convenience, denote $f(t) = L(n(t))$. We follow Lemma 5 of Massoulié (2007), which we copy almost verbatim.

Lemma 4.1 (Basic characterizations from Massoulié (2007)). *Let $n(\cdot)$ be a fluid model solution and let $\mathcal{R}_0(\cdot)$ and $\mathcal{R}_+(\cdot)$ be as defined in* (4.8).

(i) There exists a constant M, such that for all $t \geq 0$:

$$\limsup_{h \downarrow 0} \frac{f(t+h) - f(t)}{h} \leq M.$$

With some abuse of notation, define for all regular $t > 0$

$$\dot{f}(t) := \sum_{r \in \mathcal{R}_+(t)} \dot{n}_r(t) \log\left(\frac{\Lambda_r(n(t))}{\rho_r}\right). \tag{4.12}$$

Then for every regular $t > 0$,

$$\limsup_{h \downarrow 0} \frac{f(t+h) - f(t)}{h} \leq \dot{f}(t).$$

(ii) For every regular $t > 0$, there exist modified arrival rates $(\tilde{\lambda}_r)_{r \in \mathcal{R}_+(t)}$ and modified routing probabilities $(\tilde{P}_{r,s})_{r,s \in \mathcal{R}_+(t)}$ that depend only on the set $\mathcal{R}_0(t)$, such that the matrix $(\tilde{P}_{r,s})_{r,s \in \mathcal{R}_+(t)}$ is sub-stochastic with spectral radius strictly less than 1. The identity

$$(\lambda_r)_{r \in \mathcal{R}_+(t)} = (I - \tilde{P}^T)^{-1} \tilde{\lambda}$$

holds, and in addition

$$\dot{n}_r(t) = \begin{cases} \tilde{\lambda}_r + \sum_{s \in \mathcal{R}_+(t)} \mu_s \tilde{P}_{r,s} \Lambda_s(n(t)) - \mu_r \Lambda_r(n(t)), & r \in \mathcal{R}_+(t), \\ 0, & r \in \mathcal{R}_0(t). \end{cases} \tag{4.13}$$

(iii) For a regular $t > 0$, let $u_r(t) = \log\left(\frac{\Lambda_r(n(t))}{\rho_r}\right)$ for all $r \in \mathcal{R}_+(t)$. We then have

$$\dot{f}(t) = -\sum_r \lambda_r \sum_{k=0}^{\infty} \sum_{s \in \mathcal{R}_+(t)} \tilde{P}_{r,s}^k (e^{u_s(t)} - 1) \left[u_s(t) - \sum_{s' \in \mathcal{R}_+(t)} \tilde{P}_{s,s'} u_{s'}(t) \right]. \tag{4.14}$$

Proof. Properties (i) and (ii) follow from Lemma 5 of Massoulié (2007) and property (iii) follows from the arguments on page 821 of Massoulié (2007). □

In Massoulié (2007), an elaborated argument is followed to show that $\dot{f}(t) < 0$ in the subcritically loaded case. In this paper, we study the critical loaded case (i.e., $A\rho = c$). The arguments in these two cases are quite different (cf. the difference in complexity between the convergence of subcritical and critical fluid models, as exhibited in Bramson (1996)). From this moment on, our analysis and the analysis in Massoulié (2007) follow separate ways.

4.3 Bounding the derivative of the Lyapunov function

Proposition 4.1. *Suppose that $n(\cdot)$ is a fluid model solution. For any regular $t > 0$,*

$$\dot{f}(t) \leq -\sum_{r \in \mathcal{R}_+(t)} \lambda_r \left(\frac{\Lambda_r(n(t))}{\rho_r} - 1\right) \log\left(\frac{\Lambda_r(n(t))}{\rho_r}\right).$$

Furthermore, there exists an $\epsilon > 0$ such that

$$\dot{f}(t) \leq -\epsilon \sum_{r \in \mathcal{R}_+(t)} (\Lambda_r(n(t)) - \rho_r) \log\left(\frac{\Phi_r(n(t))}{\rho_r}\right).$$

The proof follows directly from (4.2), (4.14) and the following lemma that is based on a rearrangement inequality, which is of independent interest.

Lemma 4.2. *Let $\{u_r\}_{r \in \mathcal{R}_+}$ be arbitrary real numbers, where $\mathcal{R}_+$ is any subset of positive integers. Let $\hat{P}$ be an $|\mathcal{R}_+|$-dimensional sub-stochastic matrix such that $1 - \hat{P}$ is invertible. Define*

$$h_r = \sum_{k=0}^{\infty} \sum_{s \in \mathcal{R}_+} \hat{P}_{r,s}^k (e^{u_s} - 1) \left[u_s - \sum_{s' \in \mathcal{R}_+} \hat{P}_{s,s'} u_{s'} \right].$$

Then

$$h_r \geq u_r(e^{u_r} - 1).$$

Proof. Let X_k be a Markov chain on $\mathcal{R}_+ \cup \{0\}$ starting from $X_0 = r$ and evolving according to the transition matrix $\hat{P}$ with 0 as absorbing state. Note that $\hat{P}$ is extended with one row and one column so as to add 0 as an additional state making $\hat{P}$ stochastic. Set $h_0 = 0$ and $u_0 = 0$. Note that

$$\mathbb{P}(X_k = s, X_{k+1} = s' \mid X_0 = r) = \hat{P}_{r,s}^k \hat{P}_{s,s'}.$$

Let $\mathbb{E}_r[\cdot]$ denote the conditional expectation given that $X_0 = r$. Set $v_r = e^{u_r} - 1$ for all $r \in \mathcal{R}_+ \cup \{0\}$, since $u_0 = 0$ we have

$$\begin{aligned} h_r &= \sum_{k=0}^{\infty} \sum_{s \in \mathcal{R}_+} \hat{P}_{r,s}^k v_s (u_s - \sum_{s' \in \mathcal{R}_+ \cup \{0\}} \hat{P}_{s,s'} u_{s'}) = \sum_{k=0}^{\infty} \sum_{s \in \mathcal{R}_+} \sum_{s' \in \mathcal{R}_+ \cup \{0\}} \hat{P}_{r,s}^k \hat{P}_{s,s'} v_s (u_s - u_{s'}) \\ &= \sum_{k=0}^{\infty} \mathbb{E}_r \left[v_{X_k} (u_{X_k} - u_{X_{k+1}}) \right]. \end{aligned}$$

Let $k_0 = \inf\{k : X_k = 0\}$, then

$$h_r = \mathbb{E}_r \left[\sum_{k=0}^{k_0 - 1} v_{X_k} (u_{X_k} - u_{X_{k+1}}) \right] = \mathbb{E}_r \left[\sum_{k=0}^{k_0 - 1} v_{X_k} (u_{X_k} \mathbb{1}_{\{k>0\}} - u_{X_{k+1}}) \right] + v_r u_r.$$

We claim that, a.s.,

$$\sum_{k=0}^{k_0 - 1} v_{X_k} u_{X_k} \mathbb{1}_{\{k>0\}} \geq \sum_{k=0}^{k_0 - 1} v_{X_k} u_{X_{k+1}}. \tag{4.15}$$

This follows from a classical rearrangement inequality in Hardy et al. (1988) stating that if (a_k) and (b_k) are two nondecreasing finite sequences and $(b_k^{[p]})$ is a permutation of (b_k), then $\sum_k a_k b_k \geq \sum_k a_k b_k^{[p]}$. Note that $u_j \geq u_i$ if and only if $v_j \geq v_i$. So the left hand side of (4.15) is the same as the sum if we order the sequences $\{v_i\}$ and $\{u_i\}$ from small to large, while the right hand side is the sum of a rearrangement since $0 = u_{X_0} \mathbb{1}_{\{0>0\}} = u_{X_{k_0}}$. Thus, $h_r \geq v_r u_r$ and the lemma is proven. □

4.4 Bounding the Lyapunov function in terms of its derivative

Having established an upper bound for $\dot{f}(t)$, our next task is to connect this bound to $f(t)$, which is establish in the next proposition.

Proposition 4.2. *Suppose that $n(\cdot)$ is a fluid model solution. Let ϵ be given by Proposition 4.1. There exists $0 < \zeta^* < \infty$ such that for any regular point $t > 0$,*

$$\dot{f}(t) \leq -\frac{\epsilon}{\zeta^*} \frac{1}{|n(t)|} f(t).$$

The proposition will be proven after developing some necessary background. Given a fluid model solution $n(\cdot)$, define $p(t) = \frac{n(t)}{|n(t)|}$ with the convention that $0/0 = 0$. By the scalability of Φ in (4.10), we have that $\Phi_r(n(t)) = \Phi_r(p(t))$. Let $(\eta_l(p(t)))_{l \in \mathcal{L}}$ be the Lagrange multipliers satisfying the Karush-Kuhn-Tucker (KKT) conditions (cf. Section 5.5.3 in Boyd and Vandenberghe (2004)) associated with the optimization problem (2.11), and define for any vector $p \geq 0$

$$\zeta_r(p) = \sum_l A_{l,r} \eta_l(p).$$

Lemma 4.3. *(i) For any r,*

$$\sup_{\{p\}} \zeta_r(p) < \infty.$$

Here the supremum is taken over all p for which we can write $p = n(t)/|n(t)|$ for a regular time t, which implies that $|p| = 1 and that A\Phi(p) \leq c$.
(ii) In particular, if for a regular time t, $p_r(t) = 0$, then $\zeta_r(p(t)) = 0$.

Proof. It follows from (4.9) and condition (4.2) that $\Phi_r(p) \geq \frac{\lambda_r}{\mu_r} > 0$, for all $r \in \mathcal{R}_0$. For all regular $t > 0$, define

$$\mathcal{L}_0 = \{l \in \mathcal{L} : A_{l,r} > 0 \text{ for some } r \in \mathcal{R}_0\}.$$

Since regularity implies $A\Phi(p) \leq c$ we obtain

$$\sum_{r \in \mathcal{R}_+} A_{l,r} \Phi_r(p) \leq c_l - \sum_{r \in \mathcal{R}_0} \frac{\lambda_r}{\mu_r} < c_l \tag{4.16}$$

for all $l \in \mathcal{L}_0$. We can see that $\eta_l(p) = 0$ for all $l \in \mathcal{L}_0$ since the lth constraint in (2.10) is not binding due to (4.16). This implies that $\zeta_r(p) = 0$ for any regular p with $p_r = 0$, and leads to statement (ii) of the lemma.

For statement (i) we also need to handle cases where $p_r > 0$. We use Lagrange duality. For fixed t, let

$$\mathcal{L}_+ = \{l \in \mathcal{L} : A_{l,r} > 0 \text{ for some } r \in \mathcal{R}_+\}.$$

We see that $\mathcal{L}_+ \neq \emptyset$. Note that though this set depends on t, we drop this for convenience. The Lagrangian dual $H(\eta) : \mathbb{R}^{|\mathcal{L}_+|} \to \mathbb{R}$ for any fixed $(\eta_l)_{l \in \mathcal{L}_+}$ of the optimization problem (2.11) with feasible region (2.10) can be written as

$$H(\eta) = \max_{\gamma_r, r \in \mathcal{R}_+} \sum_{r \in \mathcal{R}_+} p_r \log \gamma_r - \sum_{l \in \mathcal{L}_+} \eta_l \left(\sum_{r \in \mathcal{R}_+} A_{l,r} \gamma_r - c_l \right). \tag{4.17}$$

In this representation we have removed the constraints for the lines $l \in \mathcal{L} \setminus \mathcal{L}_+$, which are irrelevant for this t.

The optimal solution $(\gamma_r)_{r\in\mathcal{R}}$ satisfies the optimality condition $\frac{p_r}{\gamma_r} = \sum_{l\in\mathcal{L}_+} A_{l,r}\eta_l$, for all $r \in \mathcal{R}_+$. So we obtain

$$\sum_{l\in\mathcal{L}_+} \eta_l \sum_{r\in\mathcal{R}_+} A_{l,r}\gamma_r = \sum_{l\in\mathcal{L}_+} \eta_l \sum_{r\in\mathcal{R}_+} A_{l,r} \frac{p_r}{\sum_{l'\in\mathcal{L}_+} \eta_{l'} A_{l',r}} = \sum_{r\in\mathcal{R}_+} p_r \frac{\sum_{l\in\mathcal{L}_+} \eta_l A_{l,r}}{\sum_{l'\in\mathcal{L}_+} \eta_{l'} A_{l',r}} = 1.$$

Thus, (4.17) can be simplified as

$$H(\eta) = \sum_{r\in\mathcal{R}_+} p_r \log p_r - \sum_{r\in\mathcal{R}_+} p_r \log\Big(\sum_{l\in\mathcal{L}_+} \eta_l A_{l,r}\Big) + \sum_{l\in\mathcal{L}_+} \eta_l c_l - 1.$$

By duality, $(\eta_l(p))_{l\in\mathcal{L}_+}$ solves the optimization problem

$$\inf_{\eta\geq 0} \left(\sum_{l\in\mathcal{L}_+} \eta_l c_l - \sum_{r\in\mathcal{R}_+} p_r \log\Big(\sum_{l\in\mathcal{L}_+} \eta_l A_{l,r}\Big) \right),$$

which, by recalling that $\sum_{r\in\mathcal{R}_+} p_r = 1$, is equivalent to

$$\sup_{\eta\geq 0} \sum_{r\in\mathcal{R}_+} p_r \left[\log\Big(\sum_{l\in\mathcal{L}_+} \eta_l A_{l,r}\Big) - \sum_{l\in\mathcal{L}_+} \eta_l c_l \right].$$

Since $c_l > 0$ for $l \in \mathcal{L}_+$, we have that $\left[\log(\sum_{l\in\mathcal{L}_+} \eta_l A_{l,r}) - \sum_{l\in\mathcal{L}_+} \eta_l c_l\right]$ is negative when η is outside a compact set. This implies that $\eta(p)$ is necessarily uniformly bounded in p for any fixed $\mathcal{L}_+$. Since there are only finite choices (2^L) for $\mathcal{L}_+$, we must have $\sup_{p:|p|=1} |\eta(p)| < \infty$. □

Proof of Proposition 4.2. Let t be a regular point. By Lemma 4.3, let $\zeta^* < \infty$ be an upper bound of $\zeta_r(p(t))$ for all $p(t)$ such that $|p(t)| = 1$ and $A\Phi(p(t)) \leq c$. Using (4.10) and Proposition 4.1, we have

$$\begin{aligned}
\dot{f}(t) &\leq -\epsilon \sum_{r\in\mathcal{R}_+(t)} (\Phi_r(p(t)) - \rho_r) \log\left(\frac{\Phi_r(p(t))}{\rho_r}\right) \\
&\leq -\frac{\epsilon}{\zeta^*} \sum_{r\in\mathcal{R}_+(t)} \zeta_r(p(t))(\Phi_r(p(t)) - \rho_r) \log\left(\frac{\Phi_r(p(t))}{\rho_r}\right). \qquad (4.18)
\end{aligned}$$

By the KKT conditions, $p_r(t) = \zeta_r(p(t))\Phi_r(p(t))$ for all $r \in \mathcal{R}_+(t)$. Define $q_r(t) = \zeta_r(p(t))\rho_r$ for all $r \in \mathcal{R}$. Observe that $q_r(t) = 0$ for all $r \in \mathcal{R}_0(t)$ due to Lemma 4.3 (ii). Then (4.18) becomes

$$\dot{f}(t) \leq -\frac{\epsilon}{\zeta^*} \sum_{r\in\mathcal{R}_+(t)} (p_r(t) - q_r(t)) \log\left(\frac{\Phi_r(p(t))}{\rho_r}\right). \qquad (4.19)$$

Consider now the allocation $\Lambda(q)$, which is the solution to the program $\max_\gamma \sum_{r\in\mathcal{R}} q_r \log\gamma_r$ subject to $A\gamma \leq c$ and $\gamma_r = 0$ if $q_r = 0$. The KKT conditions then read $q_r/\Lambda_r(q) = \sum_{l\in\mathcal{L}} A_{l,r}\eta_l(q)$, $\eta(q)(A\Lambda(q) - c) = 0$ for some $\eta(q) \geq 0$. Since the network is critically loaded, i.e., $A\rho = c$, a feasible solution to these KKT equations is to take $\eta(q) = \eta(p)$ and $\Lambda_r(q) = \rho_r$ if $q_r > 0$.

From this, it follows that since $\Phi(q(t))$ is maximizing the function $\sum_{r\in\mathcal{R}_+(t)} q_r(t)\log\gamma_r$ over all feasible γ,

$$\sum_{r\in\mathcal{R}_+(t)} q_r(t)\log\Phi_r(p(t)) \le \sum_{r\in\mathcal{R}_+(t)} q_r(t)\log\Phi_r(q(t)) = \sum_{r\in\mathcal{R}_+} q_r(t)\log\rho_r.$$

This together with (4.19) implies

$$\dot{f}(t) \le -\frac{\epsilon}{\zeta^*}\sum_{r\in\mathcal{R}} p_r(t)\log\left(\frac{\Phi_r(p(t))}{\rho_r}\right) = -\frac{\epsilon}{\zeta^*}\frac{1}{|n(t)|}f(t).$$

□

4.5 Compactness and convergence to the invariant manifold

We first derive some additional properties of f in order to be able to use some ideas developed in Bramson (1996).

Proposition 4.3. *The following inequalities hold.*

$$f(0) = \sum_{r\in\mathcal{R}} n_r(0)\log\left(\frac{\Phi_r(n(0))}{\rho_r}\right) \le |n(0)|\log\left(\frac{\max_l c_l}{\min_r \rho_r}\right),$$
$$\dot{f}(t) \le -\epsilon\sum_{r\in\mathcal{R}}\left(\frac{\Phi_r(n(t))}{\rho_r} - 1\right)^2,$$

for some $\epsilon > 0$ and almost every t.

Proof. The first inequality is trivial. The second inequality is derived in two steps. Let t be a regular point. We first note that

$$\dot{f}(t) \le -\epsilon\sum_{r\in\mathcal{R}_+}\left(\frac{\Phi_r(n(t))}{\rho_r} - 1\right)^2, \tag{4.20}$$

which follows from Proposition 4.1 and the inequality $(a-b)\log(a/b) \ge (a-b)^2/\max\{a,b\}$. Again, the exact value of ϵ may change from step to step, but it will always be strictly positive. The challenge is to extend this to the entire index set r, a task the rest of this proof is devoted to.

Let $(\nu_r)_{r\in\mathcal{R}}$ be the solution to the following traffic equation

$$\nu_r = \lambda_r + \sum_{r'\in\mathcal{R}} P_{r',r}\nu_{r'} \quad \text{for all } \ r\in\mathcal{R}.$$

Then $\rho_r = \nu_r/\mu_r$. Set $d_r(t) = \mu_r\Phi_r(n(t))$. We see that

$$\dot{f}(t) \le -\epsilon\sum_{r\in\mathcal{R}_+}\left(d_r(t) - \nu_r\right)^2.$$

Note that

$$d_r(t) = \lambda_r + \sum_{r'\in\mathcal{R}} P_{r',r}d_{r'}(t) \quad \text{for all } \ r\in\mathcal{R}_0(t).$$

So for all $r \in \mathcal{R}_0(t)$,

$$d_r(t) - \nu_r = \sum_{r' \in \mathcal{R}} P_{r',r}(d_{r'}(t) - \nu_{r'}).$$

We use this expression to derive properties of the vector $(d(t) - \nu)|_{\mathcal{R}_0(t)}$, which is formed by the coordinates of the vector $d(t) - \nu$ corresponding to those coordinates $r \in \mathcal{R}_0(t)$. Let $P^{0,0}$ be the matrix built up from all routing probabilities from $\mathcal{R}_0(t)$ to $\mathcal{R}_0(t)$ and let $P^{+,0}$ be the matrix consisting of routing probabilities from states $\mathcal{R}_+(t)$ to $\mathcal{R}_0(t)$. Then

$$(d(t) - \nu)|_{\mathcal{R}_0} = P^{0,0}(d(t) - \nu)|_{\mathcal{R}_0} + P^{+,0}(d(t) - \nu)|_{\mathcal{R}_+}.$$

Since $I - P$ is invertible, so is $I - P^{0,0}$ (where I is of appropriate dimension) and we see that

$$(d(t) - \nu)|_{\mathcal{R}_0} = (I - P^{0,0})^{-1}P^{+,0}(d(t) - \nu)|_{\mathcal{R}_+} =: P^{\dagger}(d(t) - \nu)|_{\mathcal{R}_+}.$$

The matrix $P^{\dagger}$ consists of nonnegative elements. We conclude that for $r \in \mathcal{R}_0$,

$$d_r(t) - \nu_r = \sum_{r' \in \mathcal{R}_+} P^{\dagger}_{r'r}(d_{r'}(t) - \nu_{r'}). \tag{4.21}$$

The Cauchy-Schwarz inequality yields

$$(d_r(t) - \nu_r)^2 \leq \sum_{r' \in \mathcal{R}_+} {P^{\dagger}_{r'r}}^2 (d_{r'}(t) - \nu_{r'})^2 \leq \|P^{\dagger}\|_{\infty}^2 \sum_{r' \in \mathcal{R}_+} (d_{r'}(t) - \nu_{r'})^2,$$

where $\|P^{\dagger}\|_{\infty}$ denotes the largest element in the matrix $P^{\dagger}$. Summing up over $r \in \mathcal{R}_0(t)$ yields

$$\sum_{r' \in \mathcal{R}_0(t)} (d_{r'}(t) - \nu_{r'})^2 \leq \|P^{\dagger}\|_{\infty}^2 R \sum_{r' \in \mathcal{R}_+(t)} (d_{r'}(t) - \nu_{r'})^2.$$

Combining the above inequality and (4.20) leads to the second inequality of this proposition. □

Proof of Theorem 4.1. Bramson's proof of his Proposition 6.1 also applies to our setting if we set $d_r(t) = \left(\frac{\Phi_r(n(t))}{\rho_r} - 1\right)$. The same holds for his Proposition 6.2, using Proposition 5.3 at various points in his line of argument. We omit the details. This guarantees the existence of a constant B such that for all $t \geq 0$,

$$|n(t)| \leq B|n(0)|.$$

Combining this with Proposition 4.2 and recalling that $f(t)$ is nonincreasing, we see that there exists an $\epsilon > 0$ such that for all $t \geq 0$ and all $\delta \in (0, t)$

$$f(t) - f(t - \delta) \leq -\epsilon\delta f(t)/|n(0)|.$$

Thus, assuming t/δ is an integer, we derive by iteration that

$$f(t) \leq f(t - \delta)/(1 + \delta\epsilon/|n(0)|) \leq \cdots \leq f(0)\left(\frac{1}{1 + \delta\epsilon/|n(0)|}\right)^{t/\delta}.$$

Since this is true for every δ such that t/δ is an integer, we obtain by letting $\delta \downarrow 0$ along an appropriate sequence that

$$f(t) \leq f(0)\exp\{-\epsilon t/|n(0)|\}. \tag{4.22}$$

From (6.26)–(6.28) of Bramson (1996), we then obtain that for all $t' \geq t \geq 0$

$$|n(t) - n(t')| \leq B|n(0)| \exp\{-\epsilon t/|n(0)|\},$$

for appropriate constants ϵ, B. Consequently, $n(t)$ satisfies the Cauchy criterion for convergence, and thus converges to some $n(\infty)$. The last equation implies that for all $t \geq 0$

$$|n(t) - n(\infty)| \leq B|n(0)| \exp\{-\epsilon t/|n(0)|\}.$$

In other words, convergence is exponentially fast, u.o.c. in $|n(0)|$. Recall the definition of $L(n(t))$ in (4.12). Since $f(x)$ is lower semi-continuous (cf. Theorem 1 in Massoulié (2007)), we see that

$$0 \leq L(n(\infty)) \leq \liminf_{t\to\infty} L(n(t)) = \lim_{t\to\infty} f(t) = 0.$$

Consequently,

$$0 = \sum_r n_r(\infty) \log(\Phi_r(n(\infty))/\rho_r) = \sum_{r:n_r(\infty)>0} n_r(\infty) \log(\Lambda_r(n(\infty))/\rho_r).$$

Furthermore, $\sum_r n_r(\infty) \log(\Lambda_r(n(\infty))) \geq \sum_r n_r(\infty) \log(\Lambda'_r)$ for any feasible Λ', since $\Lambda(n(\infty))$ is the unique optimum of the PF utility maximization problem. It then follows that $\Phi_r(n(\infty)) = \Lambda_r(n(\infty)) = \rho_r$ if $n_r(\infty) > 0$. If $n_r(\infty) = 0$, an additional argument is needed to show that $\Phi_r(n(\infty)) = \rho_r$.

Observe that $n(\infty)$ is an invariant point, since $n(t)$ and $n(t+s)$ both converge to $n(\infty)$ for every fixed s as $t \to \infty$, and $(n(t+s))_s$ can be seen as time-shifted fluid model with starting point $n(t)$. Since fluid model solutions are regular almost everywhere, a fluid model solution with starting position $n(\infty)$ is regular everywhere. This enables us to apply equation (4.21) with $t = \infty$ to conclude that $\Phi_r(n(\infty)) = \rho_r$ when $n_r(\infty) = 0$. Consequently, $n(\infty)$ is on the invariant manifold. □

5 Geometry of the fixed-point state space

In this section, we determine and analyze the set of points $\boldsymbol{n}$ for which $\boldsymbol{\Lambda}(\boldsymbol{n}) = \boldsymbol{\rho}$. As we have seen in the previous section, this equality (practically saying that on average, work is flowing out and in at the same rate) determines the invariant points of our fluid model. A main technical task is to show that only such states $\boldsymbol{n}$ show up in the heavy-traffic limit. The analysis in this section is inspired by (Ye and Yao, 2012, Section 3), though our situation is different, as we need to deal with routing.

Let $\boldsymbol{B}^\dagger = \mathrm{diag}\,(\boldsymbol{m})\,(I - \boldsymbol{P}^T)^{-1}\mathrm{diag}\,(\boldsymbol{\rho})$. Define

$$\mathcal{W} := \{w = \boldsymbol{B}^\dagger \boldsymbol{A}^T \pi : \pi = (\pi_l)_{l\in\mathcal{L}} \geq 0\}. \tag{5.1}$$

The following lemma shows that $\mathcal{W}$ arises from the so-called *invariant manifold*, or *fixed-point state space* for the fluid model introduced in Definition 4.1.

Lemma 5.1. *For any state $\boldsymbol{n} \geq 0$ satisfying $\boldsymbol{\Lambda}_{r,f}(\boldsymbol{n}) = \boldsymbol{\rho}_{r,f}$ for all $r \in \mathcal{R}$ and $l \in \mathcal{L}$, its associated workload, as defined by* (3.4)*, must be in $\mathcal{W}$.*

Proof. Suppose $\boldsymbol{\gamma}_{r,f} = \boldsymbol{\rho}_{r,f}$, for all $r \in \mathcal{R}$ and $l \in \mathcal{L}$, is the optimal solution to (2.13) and let $\pi_l = \eta_l \geq 0$ be the corresponding Lagrangian multiplier. According to (3.4), the workload of phase f on route r is

$$\boldsymbol{w}_{r,f} = \boldsymbol{m}_{r,f} \sum_{f'} [(1 - P^{r,T})^{-1}]_{f,f'} \boldsymbol{n}_{r,f'} = \boldsymbol{m}_{r,f} \sum_{f'} [(1 - P^{r,T})^{-1}]_{f,f'} \boldsymbol{\rho}_{r,f'} \sum_{l \in \mathcal{L}} A_{l,r} \pi_l.$$

In matrix form, $\boldsymbol{w} = \boldsymbol{B}\boldsymbol{A}^T \pi$. □

The routing in our model is only "local" in the sense that phase transitions only happen within each route. This means that $\boldsymbol{P}$ is a block-diagonal matrix and so is $\boldsymbol{B}^\dagger$. We now develop some notions that enable us to connect this part of the analysis to work done in Ye and Yao (2012). Let $\boldsymbol{C}$ be an $R \times \sum_{r \in \mathcal{R}} F_r$ matrix with the first F_1 columns all being the R-dimensional vector $(1, 0, \ldots, 0)^T$, the next F_2 columns all being $(0, 1, \ldots, 0)^T$ and so on. In particular,

$$\boldsymbol{C} = \begin{pmatrix} 1 & \ldots & 1 & 0 & \ldots & 0 & \ldots & 0 & \ldots & 0 \\ 0 & \ldots & 0 & 1 & \ldots & 1 & \ldots & 0 & \ldots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \ldots & 0 & 0 & \ldots & 0 & \ldots & 1 & \ldots & 1 \end{pmatrix}.$$

Then, we have that

$$\boldsymbol{A} = A\boldsymbol{C}. \tag{5.2}$$

Define now the diagonal matrix

$$\boldsymbol{B} := \operatorname{diag}(\boldsymbol{m}) \operatorname{diag}\left((I - \boldsymbol{P}^T)^{-1} \boldsymbol{\rho}\right). \tag{5.3}$$

Due to the structure of $\boldsymbol{A}$ (repeating the rth column of A for F_r times; see (2.5)), we have

$$\boldsymbol{B}^\dagger \boldsymbol{A}^T = \boldsymbol{B}\boldsymbol{A}^T. \tag{5.4}$$

The key difference between our model and that of Ye and Yao (2012) is in the definition of the workload in (3.4). Since some of our work is rerouted, we need to carefully handle the indirect work that an arrival brings to the system. As a consequence, the matrix $\boldsymbol{B}^\dagger$ is not a diagonal matrix, as required in the geometric analysis in Ye and Yao (2012). However, due to the special structure of our routing matrix (3.3), we can replace $\boldsymbol{B}^\dagger$ with $\boldsymbol{B}$ (cf. (5.4)) and the structure of $\mathcal{W}$ coincides with that of the similar manifold introduced in Ye and Yao (2012). Thus, all the analysis in Ye and Yao (2012) applies to our situation; this would no longer be the case if we consider more general routing schemes as considered in Section 4. We now briefly cite some relevant results from Ye and Yao (2012).

Workload decomposition. Let Δ be the left null space of $\boldsymbol{A}^T$, i.e. the kernel of A:

$$\Delta := \{\delta \in \mathbb{R}^{\sum_{r \in \mathcal{R}} F_r} : \boldsymbol{A}\delta = 0\},$$

as A is of full row rank. We assume without loss of generality that $\sum_{r \in \mathcal{R}} F_r > L$; if equality were true, then this would actually simplify the analysis, as $\mathcal{W}$ would be the positive orthant in this case. Δ is of dimension $\sum_{r \in \mathcal{R}} F_r - L$. Since $\boldsymbol{B}$ is diagonal, and thus of full rank, then for any basis H (which is of dimension $\sum_{r \in \mathcal{R}} F_r \times (\sum_{r \in \mathcal{R}} F_r - L)$) of Δ, $\boldsymbol{B}H$ is also a basis and

$$\boldsymbol{A}\boldsymbol{B}H = 0. \tag{5.5}$$

Moreover, as $\boldsymbol{B}$ is symmetric, one can chose the basis H such that

$$H^T \boldsymbol{B} H = I.$$

The null space Δ can now be expressed as

$$\Delta = \{\boldsymbol{B} H z : z \in \mathbb{R}^{\sum_{r\in\mathcal{R}} F_r - L}\}. \tag{5.6}$$

So any $\sum_{r\in\mathcal{R}} F_r$-dimensional real-valued vector w can be decomposed into two linearly independent vectors, one belonging to $\mathcal{W}$ and one belonging to Δ:

$$w = \boldsymbol{B}\boldsymbol{A}^T \pi + \boldsymbol{B} H z, \tag{5.7}$$

with π and z as specified in (5.1) and (5.6). Note that because $\boldsymbol{A}$ and $\boldsymbol{B}$ are both full rank and $\boldsymbol{A}\boldsymbol{B}$ is surjective, $\boldsymbol{A}\boldsymbol{B}\boldsymbol{A}^T$ is invertible. Then, set $G = \boldsymbol{A}^T(\boldsymbol{A}\boldsymbol{B}\boldsymbol{A}^T)^{-1}$ and observe that

$$G^T \boldsymbol{B}^T G = G^T \boldsymbol{B} G = (\boldsymbol{A}\boldsymbol{B}\boldsymbol{A}^T)^{-1}, \;\; G^T \boldsymbol{B} H = 0, \;\; G^T \boldsymbol{B}^T \boldsymbol{A}^T = G^T \boldsymbol{B}\boldsymbol{A}^T = \boldsymbol{A}\boldsymbol{B} G = I. \tag{5.8}$$

In other words, g_l, the lth column of G, is perpendicular to Bh_m, with h_m the mth column of H. (Keep in mind that $\boldsymbol{B}$ is diagonal.) Let $\mathcal{W}_l := \{w \in \mathcal{W} : \pi_l = 0\}$ denote the lth facet of $\mathcal{W}$; we see that g_l is perpendicular to $\mathcal{W}_l$. The $\sum_{r\in\mathcal{R}} F_r$-dimensional matrix (G, H) is invertible (cf. Ye and Yao (2012)); hence, we can decompose the $\sum_{r\in\mathcal{R}} F_r$-dimensional vector w as

$$w = \boldsymbol{B} G y + \boldsymbol{B} H z. \tag{5.9}$$

It follows from (5.5), (5.7) and (5.8) that

$$H^T w = z \;\text{ and }\; G^T w = \pi. \tag{5.10}$$

Dynamic complementarity problem. Consider the following dynamic complementarity problem (DCP), also known as Skorokhod problem.

$$\boldsymbol{w}(t) = \boldsymbol{w}(0) + \boldsymbol{x}(t) + \boldsymbol{B} G y(t) + \boldsymbol{B} H z(t) \geq 0, \tag{5.11}$$

$$G^T \boldsymbol{w}(t) \geq 0, \tag{5.12}$$

$$y_l(t) \text{ is nondecreasing in } t \text{ with } y(0) = 0, \tag{5.13}$$

$$\int_0^\infty \boldsymbol{w}(t)^T G dy(t) = 0, \tag{5.14}$$

$$H^T \boldsymbol{w}(t) = 0. \tag{5.15}$$

If we multiply (5.11) by H^T from the left, we have $z(t) = -H^T x(t)$ due to (5.5), (5.8) and (5.15). Also note that (5.7) and (5.10) imply that

$$\boldsymbol{B}\boldsymbol{A}^T G^T + \boldsymbol{B} H H^T = I.$$

Therefore, we can eliminate $z(t)$ in (5.11) to obtain

$$\boldsymbol{w}(t) = \boldsymbol{w}(0) + \boldsymbol{B}\boldsymbol{A}^T G^T \boldsymbol{x}(t) + \boldsymbol{B} G y(t) \geq 0. \tag{5.16}$$

It is pointed out in Ye and Yao (2012) that the DCP characterized by (5.16) and (5.12)–(5.15) can be transformed to a standard Skorohod problem (e.g., Williams (1998)) if we consider $\boldsymbol{w}_G(t) = G^T \boldsymbol{w}(t)$. Let $\Psi : \mathcal{D} \to \mathcal{D}^3$ denote the solution to the DCP (5.11)–(5.15), i.e.,

$$(\boldsymbol{w}, y, z) = \Psi(\boldsymbol{x}).$$

The results in this section are required to derive the diffusion limit in Section 6.

Reflection on the boundary. To connect this DCP with our workload process, observe that by applying the decomposition (5.9) to $\int_0^t [\boldsymbol{\rho} - \boldsymbol{\Lambda}(\boldsymbol{N}^k(s))]ds$, the dynamics for the stochastic workload process (3.9) can be written as

$$\boldsymbol{W}^k(t) = \boldsymbol{W}^k(0) + \boldsymbol{X}^k(t) + \boldsymbol{B}GY^k(t) + \boldsymbol{B}HZ^k(t), \tag{5.17}$$

where $Y^k(t)$ is defined in (3.11) and

$$Z^k(t) = H^T \int_0^t [\boldsymbol{\rho} - \boldsymbol{\Lambda}(\boldsymbol{N}^k(s))]ds.$$

We see that (5.11) and (5.13) are valid, while in general (5.12), (5.14) and (5.15) are not. A main technical challenge of the paper is to show that they are approximately valid for large k under a heavy traffic assumption.

The condition (5.15) says $\boldsymbol{w}$ lives in $\mathcal{W}$. This is not the case in the pre-limit, but if $\boldsymbol{w}$ is close to $\mathcal{W}$ and there is backlog at link l, then that link is working at full capacity, which is approximately (5.14). To make this formal, define the distance from any state $\boldsymbol{w}$ to $\mathcal{W}$ as

$$\mathbf{d}^{fp}(\boldsymbol{w}) = \sum_{l \in \mathcal{L}} (-g_l^T \boldsymbol{w})^+ + \sum_{m=1}^{\sum_{r \in \mathcal{R}} F_r - L} |h_m^T \boldsymbol{w}|.$$

The intuition behind this definition is that, following from (5.10), $\boldsymbol{w}$ is an invariant point if and only if $z = H^T \boldsymbol{w} = 0$ and $\pi \equiv G^T \boldsymbol{w} \geq 0$. A key lemma is (Ye and Yao, 2012, Lemma 2).

Lemma 5.2 (Ye and Yao (2012))**.** *Let $M > 0$ and $\epsilon > 0$ be given. There exists a constant $\sigma = \sigma(M, \epsilon) > 0$ (sufficiently small) such that the following implication holds for any $l \in \mathcal{L}$:*

$$g_l^T \boldsymbol{w} > \epsilon \Rightarrow \boldsymbol{A}_l \boldsymbol{\Lambda}(\boldsymbol{n}) = c_l$$

if both $|\boldsymbol{w}| \leq M$ and $\mathbf{d}^{fp}(\boldsymbol{w}) \leq \sigma$.

6 Diffusion approximations

The main objective of this section is to study the network in heavy traffic in order to establish the diffusion approximation, which is stated in Theorem 6.1 below. The main difficulty is that the DCP in Section 5 does not hold for the stochastic system; however, it holds only asymptotically in the heavy traffic regime, in a sense we make precise later on. To this end, we establish state space collapse (SSC) in Section 6.2, which shows that the diffusion-scaled workload process will be close to the invariant manifold and the DCP is satisfied asymptotically (Proposition 6.2(ii)). Using the framework of Bramson (1998), we prove SSC using a uniform fluid approximation shown in Section 6.1 and the convergence to the invariant state of the fluid model, as we have shown in Section 4.

Our heavy-traffic assumption is, as $k \to \infty$,

$$\lambda^k \to \lambda, \tag{6.1}$$

$$k(\rho - \rho^k) \to \theta, \tag{6.2}$$

for some λ and $\theta \in \mathbb{R}_+^R$. By (2.8), this implies $k(\boldsymbol{\rho}_{r,f} - \boldsymbol{\rho}_{r,f}^k) \to \boldsymbol{\theta}_{r,f}$ for some $\boldsymbol{\theta}_{r,f} \geq 0$ as $k \to \infty$. The diffusion scaling is defined as

$$\hat{\boldsymbol{N}}^k(t) = \frac{1}{k}\boldsymbol{N}^k(k^2 t), \quad \hat{\boldsymbol{W}}^k(t) = \frac{1}{k}\boldsymbol{W}^k(k^2 t),$$

and the diffusion scaling for the process quantities is defined as

$$\hat{\boldsymbol{E}}^k(t) = \frac{1}{k}\breve{\boldsymbol{E}}^k(k^2 t), \quad \hat{\boldsymbol{S}}^k(t) = \frac{1}{k}\breve{\boldsymbol{S}}^k(k^2 t).$$

The definitions of the scaling for the corresponding route-level quantities are defined in exactly the same way. Following the above definition, we have the following diffusion scaling

$$\begin{aligned}
\hat{\boldsymbol{X}}^k(t) &= \operatorname{diag}(\boldsymbol{m})\,(I - \boldsymbol{P}^T)^{-1}\hat{\boldsymbol{E}}^k(t) + k(\boldsymbol{\rho}^k - \boldsymbol{\rho})t \\
&\quad + \operatorname{diag}(\boldsymbol{m})\,(I - \boldsymbol{P}^T)^{-1}\left[\hat{\boldsymbol{S}}_+^k\big(\tilde{\boldsymbol{D}}^k(t)\big) - \hat{\boldsymbol{S}}_-^k\big(\tilde{\boldsymbol{D}}^k(t)\big)\right], \qquad (6.3)\\
\hat{\boldsymbol{Y}}^k(t) &= \frac{1}{k}\boldsymbol{A}\int_0^{k^2 t}[\boldsymbol{\rho} - \boldsymbol{\Lambda}(\boldsymbol{N}^k(s)]ds,\\
\hat{\boldsymbol{Z}}^k(t) &= \frac{1}{k}\boldsymbol{H}^T\int_0^{k^2 t}[\boldsymbol{\rho} - \boldsymbol{\Lambda}(\boldsymbol{N}^k(s)]ds,
\end{aligned}$$

where $\tilde{\boldsymbol{D}}^k(t) = \boldsymbol{D}^k(k^2 t)/k^2$. The above diffusion-scaled processes still satisfiy the dynamic equation (5.17). We do not copy it, but later refer to it as the diffusion-scaled version of (5.17).

Theorem 6.1. *Assume that* (2.1), (2.2), (2.7), (6.1), *and* (6.2) *hold. In addition, assume that the diffusion-scaled initial state converges weakly as* $k \to \infty$*:*

$$\hat{\boldsymbol{W}}^k(0) \Rightarrow \chi_0 \in \mathcal{W}. \qquad (6.4)$$

The stochastic processes under the proportional-fairness allocation policy converge weakly as $k \to \infty$

$$\left(\hat{\boldsymbol{X}}^k(\cdot), \hat{\boldsymbol{W}}^k(\cdot), \hat{Y}^k(\cdot), \hat{Z}^k(\cdot)\right) \Rightarrow \left(\hat{\boldsymbol{X}}(\cdot), \Psi(\hat{\boldsymbol{X}})(\cdot)\right),$$

where Ψ *is defined after the DCP in Section 5 and* $\hat{\boldsymbol{X}}(\cdot)$ *is a Brownian motion with drift* $-\boldsymbol{\theta}$ *and covariance matrix*

$$\Sigma_X = \operatorname{diag}(\boldsymbol{m})\,(I - \boldsymbol{P}^T)^{-1}\left(\operatorname{diag}(\lambda\boldsymbol{a}) + \Sigma_U\right)(I - \boldsymbol{P})^{-1}\operatorname{diag}(\boldsymbol{m}), \qquad (6.5)$$

with

$$\Sigma_U = \operatorname{diag}\big((I + \boldsymbol{P}^T)(\boldsymbol{\rho}\cdot\boldsymbol{\mu})\big) - \boldsymbol{P}^T\operatorname{diag}(\boldsymbol{\rho}\cdot\boldsymbol{\mu}) - \operatorname{diag}(\boldsymbol{\rho}\cdot\boldsymbol{\mu})\,\boldsymbol{P}, \qquad (6.6)$$

$(\lambda\boldsymbol{a})_{r,f} = \lambda_r \boldsymbol{a}_{r,f}$, *and* $(\boldsymbol{\rho}\cdot\boldsymbol{\mu})_{r,f} = \boldsymbol{\rho}_{r,f}\boldsymbol{\mu}_{r,f}$.

The proof of this theorem is postponed to the end of this section. The strategy of proving this result follows that of Williams (1998) where the key step is to prove the state spaces collapse (Proposition 6.2) using the method developed by Bramson (1998).

6.1 Uniform fluid approximations

We follow the approach and terminology of Bramson (1998). The shifted fluid scaling for "status" quantities (the state *at* a particular time) is defined as

$$\bar{U}^{k,j}(t) = \frac{1}{k} U^k(kj + kt),$$

where U^k could be any of the processes $\boldsymbol{N}^k$, $\boldsymbol{W}^k$, $\boldsymbol{D}^k$ and Y^k. The shifted fluid scaling for "process" quantities (the accumulative number of events *by* a particular time) is defined as

$$\bar{U}^{k,j}(t) = \frac{1}{k}[U^k(kj + kt) - U^k(kj)],$$

where U^k could be any of the processes $\breve{\boldsymbol{E}}^k$ and $\breve{\boldsymbol{S}}^k$. To connect the shifted fluid scaling and diffusion scaling, consider the diffusion-scaled process on the interval $[0, T]$, which corresponds to the interval $[0, k^2T]$ for the unscaled process. Fix a constant $L > 1$; then, the interval will be covered by the $\lfloor kT \rfloor + 1$ overlapping intervals

$$[kj, kj + kL], \quad j = 1, 2, \ldots, \lfloor kT \rfloor.$$

For each $t \in [0, T]$, there exists a $j \in \{0, ..., \lfloor kT \rfloor\}$ and $s \in [0, L]$ (which may not be unique) such that $k^2 t = kj + ks$. Thus,

$$\hat{X}^k(t) = \bar{X}^{k,j}(s). \tag{6.7}$$

To utilize the shifted fluid-scaled processes to analyze the diffusion-scaled processes, we present a uniform fluid approximation, which is the same as (Ye and Yao, 2012, Lemma 12).

Proposition 6.1. *Assume* (6.1) *and the existence of* $M > 0$ *such that the initial state* $|\bar{\boldsymbol{N}}^{k,j_k}(0)| < M$ *for all* k*, where* j_k *is an integer in* $[0, kT]$*. For any subsequence of* $\{k\}_0^\infty$*, there exists a subsequence* $\mathcal{K}$ *along which* $(\bar{\boldsymbol{N}}^{k,j_k}(\cdot), \bar{\boldsymbol{W}}^{k,j_k}(\cdot), \bar{\boldsymbol{D}}^{k,j_k}(\cdot), \bar{Y}^{k,j_k}(\cdot))$ *converges with probability 1* u.o.c. *to a fluid model solution* $(\bar{\boldsymbol{N}}(\cdot), \bar{\boldsymbol{W}}(\cdot), \bar{\boldsymbol{D}}(\cdot), \bar{Y}(\cdot))$ *that satisfies the fluid model equations* (4.3)–(4.7)*.*

Proof. Following (Bramson, 1998, Proposition 4.2) and (Stolyar, 2004, Appendix A.2), using Chebyshev's inequality and the Borel-Cantelli lemma, we have that, as $k \to \infty$,

$$\sup_{s \in [0, kT]} \sup_{t \in [0, L]} |\frac{1}{k} \breve{\boldsymbol{E}}^k(ks + kt)| \to 0,$$

$$\sup_{s \in [0, kT]} \sup_{t \in [0, L]} |\frac{1}{k} \breve{\boldsymbol{S}}^k(ks + kt)| \to 0,$$

a.s. (almost surely) for any fixed $T > 0$ and $L > 0$. This implies that a.s. as $k \to \infty$,

$$\max_{j \in kT} \sup_{t \in [0, L]} \big(\bar{\breve{\boldsymbol{E}}}^{k,j}(t), \bar{\breve{\boldsymbol{S}}}^{k,j}(t)\big) \to (\boldsymbol{0}, \boldsymbol{0}).$$

From this point, we can apply exactly the same approach as in (Massoulié, 2007, Appendix A.1) to obtain convergence. Applying the shifted fluid scaling to the dynamics equations (3.1) and (3.2) and the scalability of $\boldsymbol{\Lambda}_{r,f}(\cdot)$ (see the comment after (2.12)), we have

$$\bar{\boldsymbol{N}}^{k,j}_{r,f}(t) = \bar{\boldsymbol{N}}^{k,j}_{r,f}(0) + \lambda_r \boldsymbol{a}_{r,f} t + \sum_{f' \in \mathcal{F}_r} \boldsymbol{\mu}_{r,f'} P^r_{f',f} \int_0^t \boldsymbol{\Lambda}_{r,f'}\big(\bar{\boldsymbol{N}}^{k,j}(s)\big) ds$$

$$- \boldsymbol{\mu}_{r,f} \int_0^t \boldsymbol{\Lambda}_{r,f}\big(\bar{\boldsymbol{N}}^{k,j}(s)\big) ds + \bar{\epsilon}^k_{r,f}(t),$$

where, utilizing the notations defined in (3.5)–(3.8),

$$\sup_{t\in[0,L]}|\bar{\epsilon}^k_{r,f}(t)| \le \sup_{s\in[0,kT]}\sup_{t\in[0,L]}\frac{1}{k}|\breve{\boldsymbol{E}}^k_{r,f}(ks+kt)| + \sup_{s\in[0,kT]}\sup_{t\in[0,L]}\frac{1}{k}\sum_{f'\in\mathcal{F}\cup\{0\}}|\breve{\boldsymbol{S}}^k_{r,f,f'}(ks+kt)| + \sup_{s\in[0,kT]}\sup_{t\in[0,L]}\frac{1}{k}\sum_{f'\in\mathcal{F}}|\breve{\boldsymbol{S}}^k_{r,f',f}(ks+kt)|.$$

This implies $\sup_{t\in[0,L]}|\bar{\epsilon}^k_{r,f}(t)|\to 0$ a.s. as $k\to\infty$. Moreover, $|\bar{\boldsymbol{N}}^{k,j}(0)|<M$ for all j,k by our assumption. So we have verified the two conditions of a variation of the Arzela-Ascoli theorem (see (Ye et al., 2005, Lemma 6.3)). Thus, for any subsequence, there exists a further subsequence such that, as $k\to\infty$, almost surely,

$$\int_0^{\cdot}\boldsymbol{\Lambda}_{r,f}\big(\bar{\boldsymbol{N}}^{k,j}(s)\big)ds\to\bar{\boldsymbol{D}}_{r,f'}(\cdot)\quad u.o.c.\ \text{ on } [0,L], \qquad \bar{\boldsymbol{N}}^{k,j}_{r,f}(\cdot)\to\bar{N}(\cdot)\quad u.o.c.\ \text{ on } [0,L], \tag{6.8}$$

where

$$\bar{\boldsymbol{N}}_{r,f}(t)=\bar{\boldsymbol{N}}_{r,f}(0)+\lambda_r\boldsymbol{a}_{r,f}t+\sum_{f'\in\mathcal{F}_r}\boldsymbol{\mu}_{r,f'}P^r_{f',f}\bar{\boldsymbol{D}}_{r,f}(t)-\boldsymbol{\mu}_{r,f}\bar{\boldsymbol{D}}_{r,f}(t).$$

To avoid complicating the notation, we still use k to index the subsequence. By Rademacher's theorem, $\bar{\boldsymbol{D}}_{r,f}(t)$ is differentiable almost everywhere on $[0,L]$. For any differentiable point t, if $\bar{\boldsymbol{N}}_{r,f}(t)>0$, then $\Lambda_{r,f}(\cdot)$ is continuous at $\bar{\boldsymbol{N}}(t)$ according to (Ye et al., 2005, Lemma 6.2(b)). Thus, there exists an $h>0$ such that $\bar{\boldsymbol{N}}_{r,f}(s)>0$ for all $s\in[t,t+h]$ and as $k\to\infty$,

$$\int_t^{t+h}\boldsymbol{\Lambda}_{r,f}\big(\bar{\boldsymbol{N}}^{k,j}(s)\big)ds\to\int_t^{t+h}\boldsymbol{\Lambda}_{r,f}\big(\bar{\boldsymbol{N}}(s)\big)ds.$$

If $\bar{\boldsymbol{N}}_{r,f}(t)=0$, then by Fatou's lemma,

$$\lim_{k\to\infty}\int_t^{t+h}\boldsymbol{\Lambda}_{r,f}\big(\bar{\boldsymbol{N}}^{k,j}(s)\big)ds\le\int_t^{t+h}\limsup_{y\to\bar{\boldsymbol{N}}(s)}\boldsymbol{\Lambda}_{r,f}(y)ds.$$

On the other hand, the function $x\to\limsup_{y\to x}\boldsymbol{\Lambda}_{r,f}(y)$ is upper semi-continuous, thus

$$\limsup_{s\to t}\limsup_{y\to\bar{\boldsymbol{N}}(s)}\boldsymbol{\Lambda}_{r,f}(y)\le\limsup_{y\to\bar{\boldsymbol{N}}(t)}\boldsymbol{\Lambda}_{r,f}(y).$$

This implies that the derivative of $\bar{\boldsymbol{D}}_{r,f}(t)$ at t must lie in the interval $[0,\limsup_{y\to\bar{\boldsymbol{N}}(t)}\boldsymbol{\Lambda}_{r,f}(y)]$. This is why we construct $\Phi(\cdot)$ (see (4.4)) as the extension of $\Lambda(\cdot)$ in Definition 4.1. It remains to be shown that for all regular $t\ge 0$,

$$\boldsymbol{A}\Phi(\bar{\boldsymbol{N}}(t))\le c. \tag{6.9}$$

Observing that $\boldsymbol{A}\boldsymbol{\Lambda}(\boldsymbol{n})\le c$ for any regular state $\boldsymbol{n}$ due to the allocation policy (2.11), we conclude for the pre-limit process $\bar{\boldsymbol{N}}(\cdot)$ that

$$\int_t^{t+h}\boldsymbol{A}\boldsymbol{\Lambda}_{r,f}\big(\bar{\boldsymbol{N}}^{k,j}(s)\big)ds\le ch.$$

By the convergence (6.8), we must have (6.9). □

6.2 State space collapse and asymptotic complementarity

There are two key properties leading to the proof of Theorem 6.1. Note that the diffusion-scaled stochastic process $(\hat{\boldsymbol{X}}^k(\cdot), \hat{\boldsymbol{W}}^k(\cdot), \hat{Y}^k(\cdot), \hat{Z}^k(\cdot))$ only satisfies equations (5.11) and (5.13) of the DCP, but does not satisfy equations (5.12), (5.14) and (5.15). We will show in the following proposition that it satisfies these equations in an approximation sense. The approximate satisfaction of (5.12) and (5.15) is called *state space collapse*, meaning that the diffusion-scaled workload process $\hat{\boldsymbol{W}}^k(\cdot)$ gets close to the invariant manifold $\mathcal{W}$ as k grows large; The approximate satisfaction of (5.14) is called *Asymptotic Complementarity* and is instrumental in establishing tightness. The latter is formalized by the following proposition.

Proposition 6.2. *Pick a sample-path dependent constant C such that*

$$\sup_{s,t\in[0,T]} |\hat{\boldsymbol{X}}^k(t) - \hat{\boldsymbol{X}}^k(s)| \leq C, \tag{6.10}$$

and any $\epsilon > 0$. Under condition (6.4), there exists a constant $k(\epsilon)$ such that for all $k > k(\epsilon)$ the following properties hold:

1. *State space collapse:*
$$\mathbf{d}^{fp}\big(\hat{\boldsymbol{W}}^k(t)\big) \leq \epsilon, \quad \textit{for all } t \in [0,T];$$

2. *Asymptotic complementarity:*
$$\hat{Y}_l^k(t) \textit{ can not increase at time } t \textit{ if } g_l^T \hat{\boldsymbol{W}}^k(t) > 2\epsilon, \quad \textit{for all } t \in [0,T];$$

3. *Boundedness: There exists $M > 0$, depending on C and network parameters, such that*
$$|\hat{\boldsymbol{W}}^k(t)| \leq M, \quad \textit{for all } t \in [0,T].$$

Proof. Due to the relationship (6.7) between the diffusion and fluid-scaled processes, we just need prove these three results for the shifted fluid-scaled processes, i.e.,

$$\mathbf{d}^{fp}\big(\bar{\boldsymbol{W}}^{k,j}(s)\big) \leq \epsilon, \tag{6.11}$$

$$\bar{Y}_l^{k,j}(s) = \bar{Y}_l^{k,j}(0) \text{ if } \sup_{s'\in[0,L]} g_l^T \bar{\boldsymbol{W}}^{k,j}(s') > 2\epsilon, \tag{6.12}$$

$$|\bar{\boldsymbol{W}}^{k,j}(s)| \leq M, \tag{6.13}$$

for all $j = 0, 1, \ldots, \lfloor kT \rfloor$ and $s \in [0, L]$. We choose $L > T_{M,\min(\epsilon/4,\sigma/2)} + 1$ with $T_{M,\min(\epsilon/4,\sigma/2)}$ specified in Theorem 4.1 and use induction. First, we show (6.11)–(6.13) hold for $j = 0$. It follows from the initial condition (6.4), Proposition 6.1 and Theorem 4.1 that

$$\bar{\boldsymbol{W}}^{k,0}(s) \to \chi \quad \textit{u.o.c.} \text{ on } [0,L],$$

for some $\chi \in \mathcal{W}$. Though the above convergence should be interpreted as for any subsequence there is a further convergent subsequence, an easy proof by contradiction can show this is enough to prove results for all sufficiently large k. Thus, we omit the complication of introducing notation for subsequences. Thus (6.11) and (6.13) hold for $j = 0$ and all sufficiently large k. Moreover,

$$|g_l^T\big(\bar{\boldsymbol{W}}^{k,0}(s) - \chi\big)| < \min(\epsilon/4, \sigma/2),$$

for all $s \in [0, L]$. This implies that

$$\begin{aligned} & |g_l^T \bar{\boldsymbol{W}}^{k,0}(s) - g_l^T \bar{\boldsymbol{W}}^{k,0}(s')| \\ & \leq |g_l^T \big(\bar{\boldsymbol{W}}^{k,0}(s) - \chi\big)| + |g_l^T \big(\bar{\boldsymbol{W}}^{k,0}(s') - \chi\big)| \\ & \leq \epsilon. \end{aligned} \tag{6.14}$$

So if $\sup_{s' \in [0,L]} g_l^T \bar{\boldsymbol{W}}^{k,0}(s') > 2\epsilon$ for some link l, then $\inf_{s' \in [0,L]} g_l^T \bar{\boldsymbol{W}}^{k,0}(s') > \epsilon$ due to the triangle inequality

$$g_l^T \bar{\boldsymbol{W}}^{k,0}(s) \geq g_l^T \bar{\boldsymbol{W}}^{k,0}(s') - |g_l^T \bar{\boldsymbol{W}}^{k,0}(s) - g_l^T \bar{\boldsymbol{W}}^{k,0}(s')|.$$

Applying Lemma 5.2, we have

$$\bar{Y}_l^{k,0}(t) - \bar{Y}_l^{k,0}(0) = \int_0^t \big(c_l - \boldsymbol{A}_l \boldsymbol{\Lambda}(\hat{\boldsymbol{N}}^k(s))\big) ds = 0. \tag{6.15}$$

Thus (6.12) is proved for $j = 0$.

Now assume for each k there exits j_k such that (6.11)–(6.13) hold for all $j = 0, 1, \ldots, j_k - 1$ for all sufficiently large k. Note that

$$\bar{\boldsymbol{W}}^{k,j_k}(s) = \bar{\boldsymbol{W}}^{k,j_k-1}(1 + s). \tag{6.16}$$

Since $L > 1$, due to overlapping, (6.11)–(6.13) hold for $j = j_k$ on $[0, L - 1]$. We just need to extend the result from $[0, L - 1]$ to $[0, L]$. By Proposition 6.1 (again we omit the technicality of subsequences), as $k \to \infty$

$$\bar{\boldsymbol{W}}^{k,j_k}(s) \to \bar{\boldsymbol{W}}(s) \quad u.o.c. \text{ on } [0, L], \tag{6.17}$$

for some fluid limit $\bar{\boldsymbol{W}}(\cdot)$. Due to (6.16), we readily have $|\bar{\boldsymbol{W}}^{k,j_k}(0)| \leq M$. This implies that $|\bar{\boldsymbol{W}}(0)| \leq M$. So, by applying Theorem 4.1, we have for all $s \geq L - 1 \geq T_{M,\epsilon/4}$

$$\mathbf{d}^{fp}\big(\bar{\boldsymbol{W}}(s)\big) < \min(\epsilon/4, \sigma/2), \tag{6.18}$$

$$|g_l^T \big(\bar{\boldsymbol{W}}(s) - \chi\big)| < \min(\epsilon/4, \sigma/2), \tag{6.19}$$

for some $\chi \in \mathcal{W}$. Moreover, (6.17) and (6.18) imply that (6.11) holds for $j = j_k$ and $s \in [L-1, L]$. Further, (6.17) and (6.19) imply that

$$|g_l^T \big(\bar{\boldsymbol{W}}^{k,j}(s) - \chi\big)| \leq |g_l^T \big(\bar{\boldsymbol{W}}^{k,j}(s) - \bar{\boldsymbol{W}}(s)| + |g_l^T \big(\bar{\boldsymbol{W}}(s) - \chi\big)| \leq \min(\epsilon/2, \sigma),$$

for all $s \in [L - 1, L]$. So (6.14) and (6.15) also hold for $j = j_k$ on $[L - 1, L]$. By Lemma 5.2, (6.12) is proved for $j = j_k$ and $s \in [L - 1, L]$. The proof of boundedness (6.13) relies on the asymptotic complementarity (6.12). Introduce the oscillation of a function on the interval $[a, b]$

$$\mathrm{Osc}(f, [a, b]) = \sup_{a \leq s \leq t \leq b} |f(t) - f(s)|.$$

It follows from (Ye and Yao, 2012, Lemma 13) (also see (Kang et al., 2009, Proposition 7)) that (6.12) implies that

$$\begin{aligned} \mathrm{Osc}(G^T \hat{\boldsymbol{W}}^k, [0, \frac{j_k + L}{k}]) & \leq \kappa_c \mathrm{Osc}(\hat{\boldsymbol{X}}^k, [0, \frac{j_k + L}{k}]) + \kappa_c \epsilon \\ & \leq \kappa_c (C + \epsilon) \end{aligned} \tag{6.20}$$

by condition (6.10). Recall the definition $G = \boldsymbol{A}^T(\boldsymbol{A}\boldsymbol{B}\boldsymbol{A}^T)^{-1}$, and observe that we have

$$\boldsymbol{A}\hat{\boldsymbol{W}}^k(t) = (\boldsymbol{A}\boldsymbol{B}\boldsymbol{A}^T)G^T\hat{\boldsymbol{W}}^k(t). \tag{6.21}$$

So there exists another constant κ_a, which only depends on $(\boldsymbol{A}\boldsymbol{B}\boldsymbol{A}^T)$, such that

$$|\boldsymbol{A}\hat{\boldsymbol{W}}^k(t)| \le |\boldsymbol{A}\hat{\boldsymbol{W}}^k(0)| + \mathrm{Osc}(\boldsymbol{A}\hat{\boldsymbol{W}}^k, [0,t]) \le |\boldsymbol{A}\chi_0| + \epsilon + \kappa_a\kappa_c(C+\epsilon)$$

for all $t \le (j_k + L)/k$ and all sufficiently large k, where the last inequality is due to the initial condition (6.4) and (6.21). Choose

$$M = \frac{|\boldsymbol{A}\chi_0| + \epsilon + \kappa_a\kappa_c(C+\epsilon)}{\min_{l,r}\{A_{l,r} : A_{l,r} > 0\}}.$$

Thus, $|\bar{\boldsymbol{W}}^{k,j_k}(s)| \le M$ for all $s \in [0, L]$ due to (6.7), and (6.13) holds for $j = j_k$. □

Proof of Theorem 6.1. According to the functional central limit theorem (e.g., Chapter 5 of Chen and Yao (2001)), as $k \to \infty$,

$$\hat{\boldsymbol{E}}^k(\cdot) \Rightarrow \hat{\boldsymbol{E}}(\cdot) \quad \text{and} \quad \hat{\boldsymbol{S}}^k(\cdot) \Rightarrow \hat{\boldsymbol{S}}(\cdot), \tag{6.22}$$

where $\hat{\boldsymbol{E}}_{r,f}(\cdot)$ and $\hat{\boldsymbol{S}}_{r,f}(\cdot)$ are standard Brownian motions independent of each other. Using the Skorohod representation theorem, we can map all random objects to the same probability space on which the above convergence, as well as the convergence (6.4), holds a.s. This enables us to employ sample-path arguments for the rest of this proof.

We first show the convergence of $\hat{\boldsymbol{X}}^k(\cdot)$. Define the scaling $\tilde{\boldsymbol{W}}^k(t) := \boldsymbol{W}^k(k^2t)/k^2$. Such a scaling is essentially same as the fluid scaling we introduced in Section 6.1, with the only difference being that k^2 is used to scale the time and space instead of k. Without symbolically repeating the proof, we claim that the fluid approximation result in Proposition 6.1 still holds for $\tilde{\boldsymbol{W}}^k(\cdot)$. Note that by condition (6.4), as $k \to \infty$,

$$\tilde{\boldsymbol{W}}^k(0) = \frac{1}{k}\hat{\boldsymbol{W}}^k(0) \to 0 \in \mathcal{W}.$$

This implies, by Theorem 4.1, that, as $k \to \infty$,

$$\tilde{\boldsymbol{D}}^k_{r,f}(\cdot) \to \boldsymbol{\rho}_{r,f}\cdot, \quad u.o.c. \ \text{ on } [0,\infty). \tag{6.23}$$

The convergence (6.23), together with the a.s. version of (6.22), implies that

$$\hat{\boldsymbol{S}}^k_{r,f,f'}\big(\tilde{\boldsymbol{D}}_{r,f}(\cdot)\big) \to \hat{\boldsymbol{S}}_{r,f,f'}(\boldsymbol{\rho}_{r,f}\cdot), \quad u.o.c. \ \text{ on } [0,\infty). \tag{6.24}$$

Let

$$\hat{\boldsymbol{U}}(t) = \sum_{f' \in \mathcal{F}_r} \hat{\boldsymbol{S}}_{r,f',f}(\boldsymbol{\rho}_{r,f'}t) - \sum_{f' \in \mathcal{F}_r \cup \{0\}} \hat{\boldsymbol{S}}_{r,f,f'}(\boldsymbol{\rho}_{r,f}t).$$

Recall (6.3), the diffusion-scaled version of the system dynamics (3.10). From the above convergence (6.22)–(6.24), we can conclude that, *u.o.c.* on $[0,\infty)$,

$$\hat{\boldsymbol{X}}^k(\cdot) \to \hat{\boldsymbol{X}}(\cdot), \tag{6.25}$$

where $\hat{\boldsymbol{X}}(t) = -\boldsymbol{\theta}t + \mathrm{diag}\,(\boldsymbol{m})\,(I - \boldsymbol{P}^T)^{-1}(\hat{\boldsymbol{E}}(\boldsymbol{a}t) + \hat{\boldsymbol{U}}(t))$. Clearly, this is a Brownian motion with drift $-\boldsymbol{\theta}$. We now show that the covariance matrix is (6.5). The covariance matrix of

$\hat{\boldsymbol{E}}(\boldsymbol{a}\cdot)$ is diag $(\lambda \boldsymbol{a})$. To compute the covariance matrix of $\hat{\boldsymbol{U}}(\cdot)$, we only need to do that for each fixed $r \in \mathcal{R}$. Note that each $\hat{\boldsymbol{S}}^k_{r,f,f'}(\boldsymbol{\rho}_{r,f}\cdot)$, $f \in \mathcal{F}_r$, $f' \in \mathcal{F}_r \cup \{0\}$, is an independent Brownian motion with variance $\boldsymbol{\rho}_{r,f}\boldsymbol{\mu}_{r,f}P^r_{f,f'}$. Observe that for any $t \geq 0$,

$$\begin{aligned}
&\mathbb{E}[\hat{\boldsymbol{U}}_{f_0}(t)\hat{\boldsymbol{U}}_{f_1}(t)] \\
&= \mathbb{E}\bigg[\Big(\sum_{g\in\mathcal{F}_r} \hat{\boldsymbol{S}}_{r,g,f_0}(\boldsymbol{\rho}_{r,g}t) - \sum_{g\in\mathcal{F}_r\cup\{0\}} \hat{\boldsymbol{S}}_{r,f_0,g}(\boldsymbol{\rho}_{r,f_0}t)\Big)\Big(\sum_{g\in\mathcal{F}_r} \hat{\boldsymbol{S}}_{r,g,f_1}(\boldsymbol{\rho}_{r,g}t) - \sum_{g\in\mathcal{F}_r\cup\{0\}} \hat{\boldsymbol{S}}_{r,f_1,g}(\boldsymbol{\rho}_{r,f_1}t)\Big)\bigg]
\end{aligned}$$

Writing out this product we get an expression of the form $I - II - III + IV$. We compute each term separately. Let $\mathbb{1}_{\{\cdot\}}$ be the indicator function.

$$\begin{aligned}
I &= \mathbb{1}_{\{f_0=f_1\}} \sum_{g\in\mathcal{F}_r} \mathbb{E}\Big[\hat{\boldsymbol{S}}^2_{r,g,f_0}(\boldsymbol{\rho}_{r,g}t)\Big] = \mathbb{1}_{\{f_0=f_1\}} \sum_{g\in\mathcal{F}_r} \boldsymbol{\mu}_{r,g}\boldsymbol{\rho}_g P_{gf_0}, \\
IV &= \mathbb{1}_{\{f_0=f_1\}}\boldsymbol{\mu}_{r,f_0}\boldsymbol{\rho}_{r,f_0}, \qquad II = \boldsymbol{\rho}_{r,f_1}\boldsymbol{\mu}_{r,f_1}P^r_{f_1,f_0}, \qquad III = \boldsymbol{\rho}_{r,f_0}\boldsymbol{\mu}_{r,f_0}P^r_{f_0,f_1}.
\end{aligned}$$

Thus, the covariance matrix of $\hat{\boldsymbol{U}}$ is given by (6.6), from which we obtain (6.5).

Second, we study the convergence of $\hat{Z}^k(\cdot)$. By Proposition 6.2 (a), as $k \to \infty$,

$$|h_m^T \hat{\boldsymbol{W}}^k(\cdot)| \to 0, \quad u.o.c. \text{ on } [0,\infty).$$

Multiplying both sides of the diffusion-scaled version of (5.17), we have for all $t \geq 0$,

$$h_m^T \hat{\boldsymbol{W}}^k(t) = h_m^T \hat{\boldsymbol{W}}^k(0) + h_m^T \hat{\boldsymbol{X}}^k(t) + \hat{Z}^k(t).$$

Thus, as $k \to \infty$,

$$\hat{Z}^k(\cdot) \to \hat{Z}(\cdot) := -h_m^T \hat{\boldsymbol{X}}(\cdot), \quad u.o.c. \text{ on } [0,\infty). \tag{6.26}$$

Next, we study the convergence of $\hat{Y}^k(\cdot)$. It follows from Proposition 6.2 (c) that $\hat{Y}^k(\cdot)$ is also uniformly bounded on the interval $[0,T]$. Hence, according to Helly's selection theorem (e.g., (Billingsley, 1995, p. 336)), for any subsequence of $\hat{Y}^k(t)$, there exists a further subsequence $\mathcal{K}$ along which as $k \to \infty$,

$$\hat{Y}^k(t) \to \hat{Y}(t), \tag{6.27}$$

for a nondecreasing function $\hat{Y}(t)$ which is continuous almost everywhere. The above convergence holds for all time $t \in [0,T]$ at which $\hat{Y}(t)$ is continuous.

Summarizing (6.25)–(6.27), by (5.17), we have along the subsequence $\mathcal{K}$ as $k \to \infty$,

$$\hat{\boldsymbol{W}}^k(t) \to \hat{\boldsymbol{W}}(t) = \hat{\boldsymbol{W}}(0) + \hat{\boldsymbol{X}}(t) + \boldsymbol{B}G\hat{Y}(t) + \boldsymbol{B}H\hat{Z}(t)$$

for almost all $t \in [0,L]$ (those t at which $\hat{Y}(t)$ is continuous). Note that $\hat{Y}(\cdot)$ can be chosen to be right continuous with left limits since it is continuous almost everywhere. Thus, $\hat{\boldsymbol{W}}(\cdot)$ is also right continuous with left limits. By Proposition 6.2, the $\big(\hat{\boldsymbol{W}}^k(\cdot), \hat{\boldsymbol{X}}^k(\cdot), \hat{Y}^k(\cdot), \hat{Z}^k(\cdot)\big)$ satisfies the DCP (5.11)–(5.15). It follows from the oscillation bound (6.20) that the limit $\hat{\boldsymbol{W}}(\cdot)$ is continuous, and so is the process $\hat{Y}(\cdot)$. By the uniqueness of the solution to the DCP (e.g., (Ye and Yao, 2012, Proposition 4)), the convergence along the subsequence $\mathcal{K}$ implies the convergence along the original sequence. □

7 The invariant distribution: insensitivity and product form

In this section we analyze the SRBM $\hat{\boldsymbol{N}}(t)$, $t \geq 0$; the limit of our queue-length process. Define

$$\hat{W}_G(t) = G^T \hat{\boldsymbol{W}}(t), t \geq 0.$$

It follows from (5.4) and (5.8) (in particular $G^T \boldsymbol{B} \boldsymbol{A}^T = I$) that

$$\hat{\boldsymbol{W}}(t) = \boldsymbol{B}\boldsymbol{A}^T \hat{W}_G(t) = \boldsymbol{B}^\dagger \boldsymbol{A}^T \hat{W}_G(t), t \geq 0.$$

Letting $1/\boldsymbol{m}$ be the vector with each component being the reciprocal of the corresponding one of $\boldsymbol{m}$, we obtain by (3.4),

$$\hat{\boldsymbol{N}}(t) = (I - \boldsymbol{P}^T)\text{diag}\,(1/\boldsymbol{m})\,\hat{\boldsymbol{W}}(t), t \geq 0.$$

According to the definition of $\boldsymbol{B}^\dagger$ (given above (4.1)),

$$\hat{\boldsymbol{N}}(t) = \text{diag}\,(\boldsymbol{\rho})\,\boldsymbol{A}^T \hat{W}_G(t), t \geq 0.$$

Since by definition $N_r(t) = \sum_{f \in \mathcal{F}_r} \boldsymbol{N}_{r,f}(t)$, the same relation holds for the diffusion limits $\hat{N}_r(t)$, $r \in \mathcal{R}$, $t \geq 0$ and $\hat{\boldsymbol{N}_{r,f}}(t)$, $r \in route$, $f \in \mathcal{F}_r$, $t \geq 0$. By invoking (5.2), the limiting queue length process at the route level, given by $\hat{N}(t) = (\hat{N}_r(t), r \in \mathcal{R}), t \geq 0$, satisfies

$$\hat{N}(t) = \text{diag}\,(\rho)\,A^T \hat{W}_G(t), t \geq 0. \tag{7.1}$$

The main result of this section is an expression for the invariant distribution of $\hat{W}_G(t), t \geq 0$:

Theorem 7.1. *Assume $\theta > 0$. As $t \to \infty$, $\hat{W}_G(t) \to \hat{W}_G(\infty)$ in distribution, where the random variable $\hat{W}_G(\infty)$ is a vector of independent exponential distributions with rate θ.*

Observe that this result, combined with (7.1), is consistent with conjecture (1.1).

We prove Theorem 7.1 by checking a sufficient condition for product form, due to Harrison and Williams (1987). A version of this result, suitable for our purposes, is stated in Section 7.1. The condition involves a relationship between the covariance matrix and the reflection matrix, which are analyzed in Section 7.2 and 7.3. All insights are combined in Section 7.4.

7.1 Sufficient condition for product form

The results in Sections 5 and 6 imply that W_G, $t \geq 0$, is a semi-martingale reflected Brownian motion as defined in, e.g. Harrison and Williams (1987). A SRBM is characterized by the drift $-\theta$, covariance matrix Γ of the driving Brownian motion, and reflection matrix R. In our setting, the SRBM has a stationary distribution as we assume $\theta > 0$. Harrison and Williams (1987) shows when this stationary distribution is of product form assuming a normalized form of R. For our purposes the version presented as Theorem 7.12 in Chen and Yao (2001) is most convenient, and we follow that verbatim here.

Suppose that $R^{-1}\theta > 0$. Let Γ_d be a diagonal matrix containing the diagonal elements of Γ, and let R_d be a diagonal matrix containing the diagonal elements of R. If

$$2\Gamma = R R_d^{-1} \Gamma_d + \Gamma_d R_d^{-1} R^T, \tag{7.2}$$

the density of the stationary distribution is given by

$$f(z) = \prod_{r \in \mathcal{R}} \sigma_r e^{-\sigma_r z},$$

where for all r, σ_r are elements of the R-dimensional vector $\sigma = 2\Gamma_d^{-1} R_d \theta$. We need to verify this in our situation. From the discussion of the reflection mapping in Section 5, in particular (5.16), we have for $t \geq 0$,

$$\begin{aligned} \hat{W}_G(t) &= \hat{W}_G(0) + G^T \hat{\boldsymbol{X}}(t) + G^T \boldsymbol{B} G Y(t) \\ &= \hat{W}_G(0) + (\boldsymbol{A}\boldsymbol{B}\boldsymbol{A}^T)^{-1} \boldsymbol{A} \hat{\boldsymbol{X}}(t) + (\boldsymbol{A}\boldsymbol{B}\boldsymbol{A}^T)^{-1} Y(t), \end{aligned}$$

where the last inequality follows from the definition of G (recall $G = \boldsymbol{A}^T(\boldsymbol{A}\boldsymbol{B}\boldsymbol{A}^T)^{-1}$) and (5.8). So, the reflection matrix is given by $R = (\boldsymbol{A}\boldsymbol{B}\boldsymbol{A}^T)^{-1}$. Since in our case the reflection matrix is symmetric, the sufficient condition (7.2) becomes

$$\Gamma = R R_d^{-1} \Gamma_d. \tag{7.3}$$

The remainder of this section is devoted to the verification of (7.3). In Section 7.2, we derive an expression for the covariance matrix of $A\hat{\boldsymbol{X}}(t)$, $t \geq 0$. Then in Section 7.3, we simplify the reflection matrix R. Together, these insights yield the covariance matrix of $G^T \hat{\boldsymbol{X}}(t) = R\boldsymbol{A}\hat{\boldsymbol{X}}(t)$, $t \geq 0$, allowing us to verify (7.3) in Section 7.4.

7.2 The covariance matrix

The covariance matrix of $\boldsymbol{A}\hat{\boldsymbol{X}}(t)$, $t \geq 0$, is

$$\boldsymbol{A}\Sigma_X \boldsymbol{A}^T = AC\Sigma_X C^T A^T, \tag{7.4}$$

by Theorem 6.1 and (5.2). We use the representation (6.5) of Σ_X and set $\boldsymbol{\tau}_r = (I - P^r)^{-1}\boldsymbol{m}_r$. In other words, $\boldsymbol{\tau}_r = (\boldsymbol{\tau}_{r,1}, \ldots, \boldsymbol{\tau}_{r,F_r})^T$ where $\boldsymbol{\tau}_{r,f}$ can be interpreted as the residual service requirement of a job at phase f on route r. Note that by (6.5) and (6.6), Σ_X is a block diagonal matrix with the rth block being the F_r-dimensional matrix

$$\Sigma_X^r = \lambda_r \mathrm{diag}\,(\boldsymbol{a}_r) + \Sigma_U^r,$$

where

$$\Sigma_U^r = \mathrm{diag}\left((I + P^{r,T})(\boldsymbol{\rho}_r \cdot \boldsymbol{\mu}_r)\right) - P^{r,T}\mathrm{diag}\,(\boldsymbol{\rho}_r \cdot \boldsymbol{\mu}_r) - \mathrm{diag}\,(\boldsymbol{\rho}_r \cdot \boldsymbol{\mu}_r)\, P^r. \tag{7.5}$$

Due to the structure of $\boldsymbol{C}$ (cf. Section 5), the matrix $\boldsymbol{C}\Sigma_X\boldsymbol{C}^T$ is an $R \times R$ diagonal matrix, with on each diagonal entry an expression of the form

$$\boldsymbol{\tau}_r^T \left(\lambda_r \mathrm{diag}\,(\boldsymbol{a}_r) + \Sigma_U^r\right) \boldsymbol{\tau}_r. \tag{7.6}$$

To simplify this expression, we first need to simplify Σ_U^r. Note that, by (2.8)

$$\boldsymbol{\rho}_r = \lambda_r \mathrm{diag}\,(\boldsymbol{m}_r)\,(I - P^{r,T})^{-1}\boldsymbol{a}_r. \tag{7.7}$$

Thus, we see that

$$\boldsymbol{\rho}_r \cdot \boldsymbol{\mu}_r = \lambda_r (I - P^{r,T})^{-1}\boldsymbol{a}_r.$$

Consequently, we have

$$(I + P^{r,T})(\boldsymbol{\rho}_r \cdot \boldsymbol{\mu}_r) = \lambda_r (I + P^{r,T})(I - P^{r,T})^{-1} \boldsymbol{a}_r = \lambda_r [2(I - P^{r,T})^{-1} - I] \boldsymbol{a}_r.$$

So the first term on the right hand side of (7.5) can be transformed into $2\lambda_r \text{diag}\left((I - P^{r,T})^{-1}\boldsymbol{a}_r\right) - \lambda_r \text{diag}\left(\boldsymbol{a}_r\right)$. The second and the third terms on the right hand side of (7.5) are just transpose of each other, thus they play the same role in computing the quadratic form (7.6). This implies that (7.6) can be written as

$$\begin{aligned} 2\lambda_r \boldsymbol{\tau}_r^T \left(\text{diag}\left(\boldsymbol{a}_r^T (I - P^r)^{-1}\right)(I - P^r)\right) \boldsymbol{\tau}_r &= 2\lambda_r \boldsymbol{m}_r^T (I - P^{r,T})^{-1} \text{diag}\left(\boldsymbol{a}_r^T (I - P^r)^{-1}\right) \boldsymbol{m}_r \\ &= 2\lambda_r \boldsymbol{m}_r^T (I - P^{r,T})^{-1} \boldsymbol{\rho}_r, \end{aligned} \tag{7.8}$$

where the first equality is due to the definition of $\boldsymbol{\tau}_r$ in the above. Let $\beta_r^{(2)}$ be the second moment of the phase-type distribution specified by $\boldsymbol{a}_r$ and P^r. We now show that (7.8) equals $\lambda_r \beta_r^{(2)}$. The normalized load vector $\boldsymbol{\rho}_r/\rho_r$ has a renewal-theoretic interpretation: for a renewal process with phase-type inter-renewal times, $\boldsymbol{\rho}_{r,f}/\rho_r$ contains the probability that the renewal process is in phase f in stationarity. Using renewal theory, and recalling (2.3), we see that

$$\frac{\beta_r^{(2)}}{2\beta_r} = \frac{\boldsymbol{\tau}_r^T \boldsymbol{\rho}_r}{\rho_r} = \frac{\boldsymbol{m}_r^T (I - P^{r,T})^{-1} \text{diag}\left(\boldsymbol{m}_r\right) (I - P^{r,T})^{-1} \boldsymbol{a}_r}{\beta_r} = \frac{\boldsymbol{m}_r^T (I - P^{r,T})^{-1} \boldsymbol{\rho}_r}{\beta_r},$$

where the last equality is due to (7.7). Consequently,

$$\beta_r^{(2)} = 2\boldsymbol{m}_r^T (I - P^{r,T})^{-1} \boldsymbol{\rho}_r.$$

In view of (7.6)–(7.8), the rth element of the diagonal matrix $C\Sigma_X C^T$ is $\lambda_r \beta_r^{(2)}$. Thus, setting $\beta^{(2)} = (\beta_1^{(2)}, \ldots, \beta_R^{(2)})$,

$$C\Sigma_X C^T = \text{diag}\left(\lambda \cdot \beta^{(2)}\right).$$

By (7.4), we conclude that the covariance matrix of $\boldsymbol{AX}(t)$ is $A\text{diag}\left(\lambda \cdot \beta^{(2)}\right) A^T$.

7.3 The reflection matrix

By (5.2), the reflection mapping can be written as

$$R = (\boldsymbol{A}\boldsymbol{B}\boldsymbol{A}^T)^{-1} = (A\boldsymbol{C}\boldsymbol{B}\boldsymbol{C}^T A^T)^{-1}.$$

According to (5.3), $\boldsymbol{B}$ is a $\sum_{r\in\mathcal{R}} F_r$-dimensional diagonal matrix. Due to the structure of $\boldsymbol{C}$ (see Section 5), $\boldsymbol{C}\boldsymbol{B}\boldsymbol{C}^T$ is an R-dimensional diagonal matrix, with the rth element being the sum of the all the elements on the diagonal of the rth block of B. Thus, by (2.8), the rth diagonal element of $\boldsymbol{C}\boldsymbol{B}\boldsymbol{C}^T$ is

$$\lambda_r \boldsymbol{m}_r^T (I - P^{r,T})^{-1} \text{diag}\left(\boldsymbol{m}_r\right) (I - P^{r,T})^{-1} \boldsymbol{a}_r = \lambda_r \boldsymbol{m}_r^T (I - P^{r,T})^{-1} \boldsymbol{\rho}_r = \lambda_r \beta_r^{(2)}/2.$$

So we have $R = \frac{1}{2}(A\text{diag}\left(\lambda \cdot \beta^{(2)}\right) A^T)^{-1}$.

7.4 Verification of the skew symmetry condition

We are now in a position to verify the product-form condition (7.3).

Proof of Theorem 7.1. Set $D = \operatorname{diag}(\lambda)\operatorname{diag}\left(\beta^{(2)}\right)$. In the previous two sections, we derived for the reflection matrix $R = (ADA^T)^{-1}/2$ and for the covariance matrix $\Sigma = G^T A \Sigma_X A^T G$, which equals $RADA^T R$. This implies that $\Sigma = 2R$. The product form condition (7.3), which is $R^{-1}\Sigma = \Sigma_d R_d^{-1}$, is equivalent to $R^{-1}\Sigma = \Sigma_d R_d^{-1}$, which is now trivial: both sides equal 2. The vector σ is given by $\sigma = 2\Gamma_d^{-1} R_d \theta = \theta$; see also Harrison and Williams (1987) and Chen and Yao (2001). □

Acknowledgments

This research is made possible by grants from the 'Joint Research Scheme' program, sponsored by the Netherlands Organization of Scientific Research (NWO) and the Research Grants Council of Hong Kong (RGC) through projects 649.000.005 and D-HK007/11T, respectively. MV is also affiliated with CWI and is supported by a MEERVOUD grant from NWO. JZ is supported by the GRF Project No. 16501015 from Hong Kong RGC. BZ is also affiliated with Eindhoven University of Technology and is supported by an NWO VICI grant.

References

Asmussen, S. (2003). *Applied probability and queues* (Second ed.), Volume 51 of *Applications of Mathematics*. New York: Springer-Verlag.

Billingsley, P. (1995). *Probability and measure* (Third ed.). Wiley Series in Probability and Mathematical Statistics. New York: John Wiley & Sons Inc.

Bonald, T. and A. Proutière (2003). Insensitive bandwidth sharing in data networks. *Queueing Syst. 44*(1), 69–100.

Boyd, S. and L. Vandenberghe (2004). *Convex optimization*. Cambridge: Cambridge University Press.

Bramson, M. (1996). Convergence to equilibria for fluid models of head-of-the-line proportional processor sharing queueing networks. *Queueing Syst. 23*(3-4), 1–26.

Bramson, M. (1998). State space collapse with application to heavy traffic limits for multiclass queueing networks. *Queueing Syst. 30*(1-2), 89–148.

Braverman, A., J. Dai, and M. Miyazawa (2015). Heavy traffic approximation for the stationary distribution of a generalized jackson network: the bar approach. Technical report, Cornell University.

Budhiraja, A. and C. Lee (2009). Stationary distribution convergence for generalized jackson networks in heavy traffic. *Math. Oper. Res. 34*(1), 45–56.

Chen, H. and D. D. Yao (2001). *Fundamentals of queueing networks*, Volume 46 of *Applications of Mathematics (New York)*. New York: Springer-Verlag.

Gamarnik, D. and A. Zeevi (2006). Validity of heavy traffic steady-state approximation in generalized Jackson networks. *Ann. Appl. Probab. 16*(1), 56–90.

Gromoll, H. C. (2004). Diffusion approximation for a processor sharing queue in heavy traffic. *Ann. Appl. Probab. 14*(2), 555–611.

Gurvich, I. (2014, 12). Diffusion models and steady-state approximations for exponentially ergodic markovian queues. *Ann. Appl. Probab. 24*(6), 2527–2559.

Hardy, G., J. Littlewood, and G. Pólya (1988). *Inequalities* (2nd ed.). Cambridge Mathematical Library. Cambridge University Press.

Harrison, J. M. (2000). Brownian models of open processing networks: canonical representation of workload. *Ann. Appl. Probab. 10*(1), 75–103.

Harrison, J. M., C. Mandayam, D. Shah, and Y. Yang (2014). Resource sharing networks: overview and an open problem. *Stochastic Systems 4*, 524–555.

Harrison, J. M. and R. J. Williams (1987). Multidimensional reflected Brownian motions having exponential stationary distributions. *Ann. Probab. 15*(1), 115–137.

Jonckheere, M. and S. López (2014). Large deviations for the stationary measure of networks under proportional fair allocations. *Math. Oper. Res. 39*(2), 418–431.

Kang, W., F. P. Kelly, N. H. Lee, and R. J. Williams (2009). State space collapse and diffusion approximation for a network operating under a fair bandwidth sharing policy. *Ann. Appl. Probab. 19*(5), 1719–1780.

Kang, W. and R. J. Williams (2007). An invariance principle for semimartingale reflecting brownian motions in domains with piecewise smooth boundaries. *Ann. Appl. Probab. 17*(2), 741–779.

Kelly, F. (1997). Charging and rate control for elastic traffic. *European Transactions on Telecommunications 8*(1), 33–37.

Kelly, F. P., L. Massoulié, and N. S. Walton (2009). Resource pooling in congested networks: proportional fairness and product form. *Queueing Syst. 63*(1-4), 165–194.

Kelly, F. P. and R. J. Williams (2004). Fluid model for a network operating under a fair bandwidth-sharing policy. *Ann. Appl. Probab. 14*(3), 1055–1083.

Kelly, F. P. and R. J. Williams (2010). Heavy traffic on a controlled motorway. In *Probability and mathematical genetics*, Volume 378 of *London Math. Soc. Lecture Note Ser.*, pp. 416–445. Cambridge Univ. Press, Cambridge.

Lambert, A., F. Simatos, and B. Zwart (2013). Scaling limits via excursion theory: interplay between Crump-Mode-Jagers branching processes and processor-sharing queues. *Ann. Appl. Probab. 23*(6), 2357–2381.

Massoulié, L. (2007). Structural properties of proportional fairness: stability and insensitivity. *Ann. Appl. Probab. 17*(3), 809–839.

Massoulié, L. and J. Roberts (1999). Bandwidth sharing: objectives & algorithms. In *IEEE Infocom 1999*, pp. 1395–1403.

Mazumdar, R., L. Mason, and C. Douligeris (1991). Fairness in network optimal flow control: optimality of product forms. *Communications, IEEE Transactions on 39*(5), 775–782.

Puha, A. L. and R. J. Williams (2004). Invariant states and rates of convergence for a critical fluid model of a processor sharing queue. *Ann. Appl. Probab. 14*(2), 517–554.

Reed, J. E. and B. Zwart (2014). Limit theorems for bandwidth sharing networks with rate constraints. *Oper. Res. 62*(6), 1453–1466.

Resnick, S. I. (1997). Heavy tail modeling and teletraffic data. *Ann. Statist. 25*(5), 1805–1869.

Shah, D., J. N. Tsitsiklis, and Y. Zhong (2014). Qualitative properties of α-fair policies in bandwidth-sharing networks. *Ann. Appl. Probab. 24*(1), 76–113.

Ştefănescu, A. and M. V. Ştefănescu (1984). The arbitrated solution for multi-objective convex programming. *Rev. Roumaine Math. Pures Appl. 29*(7), 593–598.

Stolyar, A. L. (2004). Maxweight scheduling in a generalized switch: state space collapse and workload minimization in heavy traffic. *Ann. Appl. Probab. 14*(1), 1–53.

Walton, N. (2014a). Store-forward and its implications for proportional scheduling. In *Communication, Control, and Computing (Allerton), 2014 52nd Annual Allerton Conference on*, pp. 1174–1181.

Walton, N. S. (2014b). Concave switching in single and multihop networks. In *The 2014 ACM International Conference on Measurement and Modeling of Computer Systems*, SIGMETRICS '14, New York, NY, USA, pp. 139–151. ACM.

Williams, R. J. (1998). Diffusion approximations for open multiclass queueing networks: sufficient conditions involving state space collapse. *Queueing Syst. 30*(1-2), 27–88.

Williams, R. J. (2015). Stochastic processing networks. *Annual Review of Statistics and Its Application 3*, 323–345.

Ye, H., J. Ou, and X.-M. Yuan (2005). Stability of data networks: Stationary and bursty models. *Oper. Res. 53*(1), 107–125.

Ye, H. and D. D. Yao (2012). A stochastic network under proportional fair resource control – diffusion limit with multiple bottlenecks. *Oper. Res. 60*(3), 716–738.

Ye, H. and D. D. Yao (2016). Diffusion limit of fair resource control – stationarity and interchange of limits. *Math. Oper. Res. 41*(4), 1161–1207.

Yi, Y. and M. Chiang (2008). Stochastic network utility maximisation – a tribute to Kelly's paper published in this journal a decade ago. *European Transactions on Telecommunications 19*(4), 421–442.