On the Flow-level Dynamics of a Packet-switched Network

Ciamac Moallemi
Graduate School of Business
Columbia University
ciamac@gsb.columbia.edu

Devavrat Shah*
LIDS, EECS
Massachusetts Institute of Technology
devavrat@mit.edu

ABSTRACT

The packet is the fundamental unit of transportation in modern communication networks such as the Internet. Physical layer scheduling decisions are made at the level of packets, and packet-level models with exogenous arrival processes have long been employed to study network performance, as well as design scheduling policies that more efficiently utilize network resources. On the other hand, a user of the network is more concerned with end-to-end bandwidth, which is allocated through congestion control policies such as TCP. Utility-based flow-level models have played an important role in understanding congestion control protocols. In summary, these two classes of models have provided separate insights for flow-level and packet-level dynamics of a network.

In this paper, we wish to study these two dynamics together. We propose a joint flow-level and packet-level stochastic model for the dynamics of a network, and an associated policy for congestion control and packet scheduling that is based on α-weighted policies from the literature. We provide a fluid analysis for the model that establishes the throughput optimality of the proposed policy, thus validating prior insights based on separate packet-level and flow-level models. By analyzing a critically scaled fluid model under the proposed policy, we provide constant factor performance bounds on the delay performance and characterize the invariant states of the system.

Categories and Subject Descriptors

G [**Mathematics of Computing**]: PROBABILITY AND STATISTICS—*Queueing theory, Markov processes, Stochastic processes*

General Terms

Algorithms, Performance, Theory

*This work was supported in parts by NSF projects CCF 0728554 and CNS 0546590.

SIGMETRICS'10, June 14–18, 2010, New York, New York, USA.

1. INTRODUCTION

The optimal control of a modern, packet-switched data network can be considered from two distinct vantage points. From the first point of view, the atomic unit of the network is the *packet*. In a packet-level model, the limited resources of a network are allocated via the decisions on the scheduling of packets. Scheduling policies for packet-based networks have been studied across a long line of literature (e.g., [26, 23, 20]). The insights from this literature have enabled the design of scheduling policies that allow for the efficient utilization of the resources of a network, in the sense of maximizing the throughput of packets across the network, while minimizing the delay incurred by packets, or, equivalently, the size of the buffers needed to queue packets in the network.

Packet-level models accurately describe the mechanics of a network at a low level. However, they model the arrival of new packets to the network exogenously. In reality, the arrival of new packets is also under the control of the network designer, via rate allocation or congestion control decisions. Moreover, while efficient utilization of network resources is a reasonable objective, a network designer may also be concerned with the satisfaction of end users of the network. Such objectives cannot directly be addressed in a packet-level model.

Flow-level models (cf. [9, 1]) provide a different point of view by considering the network at a higher level of abstraction or, alternatively, over a longer time horizon. In a flow-level model, the atomic unit of the network is a *flow*, or user, who wishes to transmit data from a source to a destination. Resource allocation decisions are made via the allocation of a transmission rate to each flow. Each flow generates utility as a function of its rate allocation, and rate allocation decisions may be made so as to maximize a global utility function. In this way, a network designer can address end user concerns such as fairness.

Flow-level models typically make two simplifying assumptions. The first assumption is that, as the number of flows evolves stochastically over time, the rates allocated to flows are updated *instantaneously*. The rate allocation decision for a particular flow is made in a manner that requires immediate knowledge of the demands of other flows for the limited transmission resources along the flow's entire path. This assumption, referred to as *time-scale separation*, is based on the idea that flows arrive and depart according to much slower processes than the mechanisms of the rate control algorithm. The second assumption is that, once the rate allocation decision is made, each flow can transmit data *in-*

stantaneously across the network at its given rate. In reality, each flow generates discrete packets, and these packets must travel through queues to traverse the network. Moreover, the packet scheduling decisions within the network must be made in a manner that is consistent with and can sustain the transmission rates allocated to each flow, and the induced packet arrival process must not result in the inefficient allocation of low level network resources.

In this paper, our goal is to develop a stochastic model that jointly captures the packet-level and flow-level dynamics of a network, without any assumption of time-scale separation. The contributions of this paper are as follows:

1. We present a joint model where the dynamic evolution of flows and packets is simultaneous. In our model, it is possible to simultaneously seek efficient allocation of low level network resources (buffers) while maximizing the high-level metric of end-user utility.
2. For our network model, we propose packet scheduling and rate allocation policies where decisions are made via myopic algorithms that combine the distinct insights of prior packet- and flow-level models. Packets are scheduled according to a maximum weight policy. The rate allocation decisions are completely local and distributed. Further, in long term (i.e., under fluid scaling), the rate control policy exhibits the behavior of a primal algorithm for an appropriate utility maximization problem.
3. We provide a fluid analysis of the joint packet- and flow-level model. This analysis allows us to establish stability of the joint model and the throughput optimality of our proposed control policy.
4. We establish, using a fluid model under critical loading, a performance bound on our control policy under the metric of minimizing the outstanding number of packets and flows in the network (or, in other words, minimizing delay). We demonstrate that, for a class of *balanced* networks, our control policy performs to within a constant factor of any other control policy.
5. Under critical loading, we characterize the invariant manifold of the fluid model of our control policy, as well as establishing convergence to this manifold starting from any initial state. These results, along with the method of Bramson [2], lead to the characterization of multiplicative state space collapse under *heavy traffic scaling*. Further, we establish that the invariant states of the fluid model are asymptotically optimal under a limiting control policy.

In summary, our work provides a joint dynamic flow- and packet-level model that captures the microscopic (packet) and macroscopic (fluid, flow) behavior of large packet-based communications network faithfully. The performance analysis of our rate control and scheduling algorithm suggests that the separate insights obtained for dynamic flow-level models [9, 1] and for packet-level models [26, 23, 20] indeed continue to hold in the combined model.

The balance of the paper is organized as follows. In Section 1.1, we survey the related literature on flow- and packet-level models. In Section 2, we introduce our network model. Our network control policy, which combines features of maximum weight scheduling and utility-based rate allocation is described in Section 3. A fluid model is derived in Section 4. Stability (or, throughput optimality) of the network control policy is established in Section 5. The critically scaled fluid model is described in Section 6. In Section 7, we provide performance guarantees for balanced networks. The invariant states of the critically scaled fluid model are described in Section 8. Finally, in Section 9, we conclude.

1.1 Literature Review

The literature on scheduling in packet-level networks begins with Tassiulas and Ephremides [26], who proposed a class of ‘maximum weight’ (MW) or ‘back-pressure’ policies. Such policies assign a weight to every schedule, which is computed by summing the number of packets queued at links that the schedule will serve. At each instant of time, the schedule with the maximum weight will be selected. Tassiulas and Ephremides [26] establish that, in the context of multi-hop wireless networks, MW is throughput optimal. That is, the stability region of MW contains the stability region of any other scheduling algorithm. This work was subsequently extended to a much broader class of queueing networks by others (e.g., [14, 3, 25, 4]).

By allowing for a broader class of weight functions, the MW algorithm can be generalized to the family of so-called MW-α scheduling algorithms. These algorithms are parameterized by a scalar $\alpha \in (0, \infty)$. MW-α can be shown to inherit the throughput optimality of MW [11, 18] for all values of $\alpha \in (0, \infty)$. However, it has been observed experimentally that the average queue length (or, ‘delay’) under MW-α decreases as $\alpha \to 0^+$ [11]. Certain delay properties of this class of algorithms have been subsequently established under a heavy traffic scaling [23, 4, 20].

Flow-level models have received significant recent attention in the literature, beginning with the work of Kelly, Maulloo, and Tan [9]. This work developed rate-control algorithms as decentralized solutions to a deterministic utility maximization problem. This optimization problem seeks to maximize the utility generated by a rate allocation, subject to capacity constraints that define a set of feasible rates. This work was subsequently generalized to settings where flows stochastically depart and arrive [13, 6, 1], addressing the question of the stability of the resulting control policies. Fluid and diffusion approximations of the resulting systems have been subsequently developed [10, 8, 28]. Under these stochastic models, flows are assumed to be allocated rate as per the optimal solution of the utility maximization problem instantaneously. Essentially, this *time-scale separation* assumption captures the intuition that the dynamics of the arrivals and departures of flows happens on a much slower time-scale than the dynamics of rate control algorithm.

In reality, flow arrivals/departures and rate control happen on the *same* time-scale. Various authors have considered this issue, in the context of understanding the stability of the stochastic flow level models without the time-scale separation assumption [12, 7, 24, 17, 22]. Lin, Schroff, and Srikant [12] assume a stochastic model of flow arrivals and departures as well as the operation of a primal-dual algorithm for rate allocation. However, there are no packet dynamics present. Other work [7, 24, 17] has assumed that rate control for each type of flow is a function of a local Lagrange multiplier; and a separate Lagrange multiplier is associated with each link in the network. These multipliers are updated using a maximum weight-type policy. In this line of work, Lagrange multipliers are interpreted as queue lengths, but there are no actual packet-level dynamics present. Further, these models lack flow-level dynamics as well. Thus, while overall this collection of work is closest to the results of this paper, it stops short of offering a complete characterization of a joint flow- and packet-level dynamic model.

Finally, we take note of recent work by Walton [27], which presents a simple but insightful model for joint flow- and packet-level dynamics. In this model, each source generates packets by reacting to the acknowledgements from its destination, and at each time instant, each source has at most one packet in flight. Under a many-source scaling for a specific network topology, it is shown that the network operates with rate allocation as per the proportional fair criteria. This work provides important intuition about the relationship between utility maximization and the rate allocation resulting from the packet-level dynamics in a large network. However, it is far from providing a comprehensive joint flow- and packet-level dynamic model as well as efficient control policy.

2. NETWORK MODEL

In this section, we introduce our network model. This model captures both the flow-level and the packet-level aspects of a network, and will allow us to study the interplay between the dynamics at these two levels. In a nutshell, flows of various types arrive according to an exogenous process and seek to transmit some amount of data through the network. As in the standard congestion control algorithm, TCP, the flows generate packets at their ingress to the network. The packets travel to their respective destinations along links in the network, queueing in buffers at intermediate locations. As a packet travels along its route, it is subject to physical layer constraints, like medium access constraints, switching constraints, or constraints due to limited link capacity. A flow departs once all of its packets have been sent.

In this section, we describe the mechanics of the network that are independent of the network control policy. In Section 3, we will propose a specific network control policy to be applied in this context.

2.1 Network Structure

Consider a network consisting of a set $\mathcal{V}$ of destination nodes, a set $\mathcal{L}$ of links, and a set $\mathcal{F}$ the set of flow types. Each flow type identified by a fixed given route starting at the source link $s(f) \in \mathcal{L}$ and ending at the destination node $d(f) \in \mathcal{V}$. At a given time, multiple flows of a given type exist in the network, each flow injects packets into the network.

The network maintains buffers for packets that are in transit across the network. At each link, there is a separate queue for the packets corresponding to each possible destination. Let $\mathcal{E} = \mathcal{L} \times V$ denote the set of all such queues, with each $e = (\ell, v)$ being the queue at link ℓ for final destination v. Traffic in each queue is transmitted to the next hop along the route to the destination, and leaves the network when it reaches the destination. We define the routing matrix $R \in \{0,1\}^{\mathcal{E}\times\mathcal{E}}$ by setting $R_{ee'} \triangleq 1$ if the next hop for queue e is queue e', and $R_{ee'} \triangleq 0$ otherwise. Note that traffic for a flow of type f enters the network in the queue $\iota(f) \triangleq \big(s(f), d(f)\big) \in \mathcal{E}$. Define the matrix $\Gamma \in \{0,1\}^{\mathcal{E}\times\mathcal{F}}$ by setting $\Gamma_{ef} \triangleq 1$ if $\iota(f) = e$, and $\Gamma_{ef} \triangleq 0$ otherwise. We will assume that the routes are acyclic. In this case, we can define the matrix

$$\Xi \triangleq \left(I - R^\top\right)^{-1} = I + R^\top + (R^\top)^2 + \cdots . \qquad (1)$$

Under the acyclic routing assumption, $\Xi_{e'e} = 1$ if and only if a packet arriving at queue e subsequently eventually passes through queue e'.

2.2 Dynamics: Flow-Level

In this section, we will describe in detail the stochastic model for dynamics of flows in the network. The system evolves in continuous time, with $t \in [0, \infty)$ denoting time, starting at $t = 0$. For each flow type $f \in \mathcal{F}$, Let $N_f(t)$ denote the number of flows of type f active at time t. Flows of type f arrive according to an independent Poisson process of rate ν_f. Flows of type f receive an *aggregate* rate of service $X_f(t) \in [0, C]$ at time t. Here, $C > 0$ is the maximal the rate of service that can be provided to any flow type. This service is divided equally amongst the $N_f(t)$ flows. As flows are serviced, packets are generated. The evolution of packets and flows proceeds according to:

- Packets are generated by all the flows of type f, in aggregate, as a time varying Poisson process of rate $X_f(t)$ at time t. If $N_f(t) = 0$, then $X_f(t) = 0$.
- When a packet is generated by a flow of type f, it joins the ingress queue $\iota(f) \in \mathcal{E}$.
- When a packet is generated by a flow of type f, the flow departs from the network with a probability of $0 < \mu_f < 1$, independent of everything else.

Thus, each flow of type f requires an amount of service according to an independent exponential random variable with mean[1] $1/\mu_f$, and the flow departure process for flows of type f is a Poisson process of rate $\mu_f X_f(t)$ at time t. We can summarize the flow count process $N_f(\cdot)$ by the transitions

$$N_f(t) \to \begin{cases} N_f(t) + 1 & \text{at rate } \nu_f, \\ N_f(t) - 1 & \text{at rate } \mu_f X_f(t). \end{cases}$$

Define the *offered load* vector $\rho \in \mathbb{R}_+^{\mathcal{F}}$ by $\rho_f \triangleq \nu_f/\mu_f$, for each flow type f. Without loss of generality, we will make the following assumptions:[2]

- $\rho > \mathbf{0}$, i.e., we restrict attention to flows with a non-trivial loading.
- $\rho < C\mathbf{1}$, i.e., we assume that the maximal service rate C is sufficient for the load generated by any single flow type.
- $\Xi\Gamma\rho > 0$, i.e., we restrict attention to queues with a non-trivial loadings.

Denote by $\mathsf{A}_f(t)$ the cumulative number of flows of type f that have arrived in the time interval $[0, t]$. By definition,

[1]The assumption that $\mu_f < 1$ is without loss of generality. This is because, by thinning the flow arrival process, we can restrict attention to flow types f that are expected to generate at least 1, in other words, flow types with mean service time satisfying $1/\mu_f > 1$.

[2]In what follows, inequalities between vectors are to be interpreted component-wise. The vector $\mathbf{0}$ (resp., $\mathbf{1}$) is the vector where every component is 0 (resp., 1), and whose dimension can be inferred from the context.

$\mathsf{A}_f(\cdot)$ is a Poisson process of rate ν_f. Denote the cumulative number of packets generated by flow type f in the time interval $[0,t]$ by $A_f(t)$. By definition $A_f(\cdot)$ is a Poisson process with a time-varying rate given by $X_f(\cdot)$. We suggest that the reader take note of difference between $\mathsf{A}_f(\cdot)$ and $A_f(\cdot)$. Let $D_f(t)$ denote the cumulative number of flows of type f that have departed in the time interval $[0,t]$. $D_f(\cdot)$ is a Poisson process with time varying rate $\mu_f X_f(\cdot)$. The evolution of flow count for flow type f over time can be written as

$$N_f(t) = N_f(0) + \mathsf{A}_f(t) - D_f(t). \tag{2}$$

2.3 Dynamics: Packet-Level

As we have just described, flows generate packets which are injected into the network. These packets must traverse the links of the network from source to destination. In this section, we describe the dynamics of packets in the network.

We assume that each queue in the network is capable of transmitting at most 1 data packet per unit time. However, the collection of queues that can simultaneously transmit is restricted by a set of scheduling constraints. These scheduling constraints are meant to capture any limitations of the network due to scarce resources (e.g., limited wireless bandwidth, limited link capacity, etc.).

Formally, the scheduling constraints are described by the set $\mathcal{S} \subset \{0,1\}^{\mathcal{E}}$. Under a permissible schedule $\pi \in \mathcal{S}$, a packet will be transmitted from a queue $e \in \mathcal{E}$ if and only if $\pi_e = 1$. We assume that $\mathbf{0} \in \mathcal{S}$. Further, we assume that $\mathcal{S}$ is *monotone*: if $\sigma \in \mathcal{S}$ and $\sigma' \in \{0,1\}^{\mathcal{E}}$ such that $\sigma'_e \leq \sigma_e$ for every queue e, then $\sigma' \in \mathcal{S}$. Finally, denote by $\Pi \in \{0,1\}^{\mathcal{E}\times\mathcal{S}}$ the matrix with columns consisting of the elements of $\mathcal{S}$.

We assume that the scheduling of packets happens at every integer time. At a time $\tau \in \mathbb{Z}_+$, let $\pi(\tau) \in \mathcal{S}$ denote the scheduled queues for the time interval $[\tau, \tau+1)$. For each queue e, denote by $Q_e(\tau^-)$ the length of the queue e immediately prior to the time τ (i.e., before scheduling happens). The queue length evolves, for times $t \in [\tau, \tau+1)$ according to[3]

$$Q_e(t) \triangleq Q_e(\tau^-) - \pi_e(\tau)\mathbb{I}_{\left\{Q_e(\tau^-)>0\right\}} + \sum_{f\in\mathcal{F}} \Gamma_{ef}\big(A_f(t) - A_f(\tau^-)\big) + \sum_{e'\in\mathcal{E}} R_{e'e}\pi_{e'}(\tau)\mathbb{I}_{\left\{Q_{e'}(\tau^-)>0\right\}}.$$

Here, for each flow type f, $A_f(\tau^-)$ is the cumulative number of packets generated by flows of type f in the time interval $[0,\tau)$. The term $\pi_e(\tau)\mathbb{I}_{\{Q_e(\tau)>0\}}$ enforces an idling constraint, i.e., if queue e is scheduled but empty, no packet departs. Note that, over a time interval $[\tau, \tau+1)$, we assume the transmission of packets already present in the network occurs instantly at time τ, while the arrival of new packets to the network occurs continuously throughout the entire time interval.

Finally, let $S_\pi(\tau)$ denote the cumulative number of time slots during which the schedule π was employed up to and including time τ. Let $Z_e(\tau)$ denote the cumulative idling time for queue e up to an including time τ. That is,

$$Z_e(\tau) \triangleq \sum_{s=0}^{\tau}\sum_{\pi\in\mathcal{S}} \pi_e\big(S_\pi(s) - S_\pi(s-1)\big)\mathbb{I}_{\{Q_e(s)=0\}}.$$

[3] $\mathbb{I}_{\{\cdot\}}$ denotes the indicator function.

Then, the overall queue length evolution can be written in vector form as

$$\begin{aligned} Q(\tau+1) = Q(0) &- \left(I - R^\top\right)\Pi S(\tau) \\ &+ \left(I - R^\top\right) Z(\tau) + \Gamma A(\tau+1), \end{aligned} \tag{3}$$

where we define the vectors

$$Q(t) \triangleq \left[Q_e(t)\right]_{e\in\mathcal{E}}, \qquad A(t) \triangleq \left[A_f(t)\right]_{f\in\mathcal{F}},$$
$$S(\tau) \triangleq \left[S_\pi(\tau)\right]_{\pi\in\mathcal{S}}, \qquad Z(\tau) \triangleq \left[Z_e(t)\right]_{e\in\mathcal{E}}.$$

3. MWUM CONTROL POLICY

A network control policy is a rule that, at each point in time, provides two types of decisions: (a) the rate of service provided to each flow, and (b) the scheduling of packets subject to the physical constraints in the network. In Section 2, we described the stochastic evolution of flows and packets in the network, taking as given the network control policy. In this section, we describe a control policy called the *maximum weight utility maximization-α* (MWUM-α) policy. MWUM-α takes as a parameter a scalar $\alpha \in (0,\infty)\setminus\{1\}$.

The MWUM-α policy is myopic and based only on local information. Specifically, a flow generates packets at rate that is based on the queue length at its ingress, and the scheduling of packets is decided as a function of the effected queue lengths.

At the flow-level, rate allocation decisions are made according to a per flow utility maximization problem. Each flow chooses a rate so as to myopically maximize its utility as a function of rate consumption, subject to a linear penalty (or 'price') for consuming limited network resources. As in the case of α-fair rate allocation, the utility function is assumed to have a constant relative risk aversion of α. The price charged is a function of the number of packets queued at the ingress queue associated with the flow, raised to the α power.

At the packet-level, packets are scheduled according to a maximum weight-α scheduling algorithm. In particular, each queue is assigned a weight equal to the number of queued packets to the α power, and a schedule is picked which maximizes the total weight of all scheduled queues.

3.1 Control: Rate Allocation

The first control decision we shall consider is that of *rate allocation*, or, the determination of the aggregate rate of service $X_f(t)$, at time t, for each flow type f. We will assume our network is governed by a variant of an α-fair rate allocation policy. This is as follows:

Assume that each flow of type f is allocated a rate $Y_f(t) \geq 0$ at time t by maximizing a (per flow) utility function that depends on the allocated rate, subject to a linear penalty, or cost, for consuming resources from the limited capacity of the network. In particular, we will assume a utility function given a rate allocation of $y \geq 0$ to an individual flow of type f of the form $V_f(y) \triangleq y^{1-\alpha}/(1-\alpha)$, for some $\alpha \in (0,\infty)\setminus\{1\}$. This utility function is popularly known as α-fair in the congestion control literature [15], and has a constant relative risk aversion of α. The individual flow will then be assigned capacity according to

$$Y_f(t) \in \operatorname*{argmax}_{y\geq 0}\; V_f(y) - Q^\alpha_{\iota(f)}(t)y.$$

Here, $Q^\alpha_{\iota(f)}(t)$ represents a 'price' or 'congestion signal'. Intuitively, a flow reacts to the congestion (or lack of it) through the length of the ingress or 'first-hop' queue $\iota(f)$. Then, if $N_f(t) > 0$, the aggregate rate $X_f(t)$ allocated to all flows of type f at time t is determined according to

$$X_f(t) = N_f(t)Y_f(t) = \operatorname*{argmax}_{x\geq 0} \frac{x^{1-\alpha}N_f^\alpha(t)}{1-\alpha} - Q^\alpha_{\iota(f)}(t)\,x.$$

If $N_f(t) = 0$, we require that $X_f(t) = 0$. Further, we will constrain the overall rate allocated to flows of type f by the constant C. Thus, rate allocation is determined by the equation

$$X_f(t) = \begin{cases} \operatorname*{argmax}_{x\in[0,C]} \dfrac{x^{1-\alpha}N_f^\alpha(t)}{1-\alpha} - Q^\alpha_{\iota(f)}(t)\,x & \text{if } N_f(t) > 0, \\ 0 & \text{otherwise.} \end{cases}$$

Given the strictly concave nature of the objective in this optimization program, it is clear that the maximizer is unique and $X_f(t)$ is well-defined.

Denote by $\bar{X}_f(t)$ the cumulative rate allocation to flows of type f in the time interval $[0,t]$, i.e.,

$$\bar{X}_f(t) \triangleq \int_0^t X_f(s)\,ds.$$

$\bar{X}_f(\cdot)$ is Lipschitz continuous and differentiable, since $X_f(\cdot)$ is always bounded by C.

3.2 Control: Scheduling

The second control decision that must be specified is that of scheduling. We will assume the following variation of the 'maximum weight' or 'back-pressure' policies introduced by Tassiulas and Ephremides [26].

At the beginning of each discrete-time slot $\tau \in \mathbb{Z}_+$, a schedule $\pi(\tau) \in \mathcal{S}$ is chosen according to the optimization problem

$$\begin{aligned} \pi(\tau) &\in \operatorname*{argmax}_{\pi\in\mathcal{S}} \sum_{e\in\mathcal{E}} \pi_e \left[Q^\alpha_e(\tau^-) - \sum_{e'\in\mathcal{E}} R_{ee'}Q^\alpha_{e'}(\tau^-) \right] \\ &= \operatorname*{argmax}_{\pi\in\mathcal{S}} \pi^\top (I-R)Q^\alpha(\tau^-), \end{aligned} \tag{4}$$

where $Q^\alpha(\tau^-) \triangleq \left[Q^\alpha_e(\tau^-)\right]_{e\in\mathcal{E}}$ is a vector of componentwise powers of queue lengths, immediately prior to time τ. In other words, the schedule $\pi(\tau)$ is such that it maximizes the summation of weights of queues served, where weight of queue $e \in \mathcal{E}$ is given by $\left[(I-R)Q^\alpha(\tau^-)\right]_e$. Given that $\mathcal{S}$ is monotone, there exists $\pi \in \mathcal{S}$ that maximizes this weight and is such that $\pi_e = 0$ if $Q_e(\tau^-) = 0$. We will restrict our attention to such schedules only. From this, it follows that the objective value of the optimization program in (4) is always non-negative.

By the discussion above, it is clear that the following invariants are satisfied:

1. For any schedule π and time τ, $S_\pi(\tau) = S_\pi(\tau-1)$ if $\pi^\top(I-R)Q^\alpha(\tau^-) < \sigma^\top(I-R)Q^\alpha(\tau^-)$, for some $\sigma \in \mathcal{S}$.
2. For any queue e and time τ, $Z_e(\tau) = 0$. In other words, there is no idling.

4. FLUID MODEL

We introduce the fluid model of the above described system. As we shall see, the evolution of rate allocation to flows in the fluid model resembles rate allocation of a 'flow-level' model that has been popular in the literature [13, 1]. In that sense, our model on original time-scale operates at the packet-level granularity and on the fluid or long time scale operates at the flow-level granularity.

4.1 Fluid Scaling and Fluid Model Equations

In order to introduce the fluid model of our network, we will consider the scaled version of the system. To this end, denote the overall system state at a time $t \geq 0$ by

$$\mathcal{Z}(t) \triangleq \Big(Q(t), Z(\lfloor t\rfloor), N(t), S(\lfloor t\rfloor), \bar{X}(t), \mathsf{A}(t), D(t), A(t)\Big).$$

Here, the components of the state $\mathcal{Z}(t)$ are the primitives introduced in Sections 2.2 and 2.3. That is, at times $t \in \mathbb{R}_+$, where we have, $Q(t) \triangleq \left[Q_e(t)\right]_{e\in\mathcal{E}}$ with $Q_e(t)$ being the length of queue e; $N(t) \triangleq \left[N_f(t)\right]_{f\in\mathcal{F}}$, with $N_f(t)$ being the number of flows of type f; $\bar{X}(t) \triangleq \left[\bar{X}_f(t)\right]_{f\in\mathcal{F}}$, with $\bar{X}_f(t)$ being the cumulative rate allocated to flow type f; $\mathsf{A}(t) \triangleq \left[\mathsf{A}_f(t)\right]_{f\in\mathcal{F}}$, with $\mathsf{A}_f(t)$ being the cumulative arrival count of flow type f; $D(t) \triangleq \left[D_f(t)\right]_{f\in\mathcal{F}}$, with $D_f(t)$ being the cumulative departure count of flow type f; $A(t) \triangleq \left[A_f(t)\right]_{f\in\mathcal{F}}$, with $A_f(t)$ being the cumulative packet arrival count of flow type f; and, at times $\tau \in \mathbb{Z}_+$, we have $Z(\tau) \triangleq \left[Z_e(\tau)\right]_{e\in\mathcal{E}}$ with $Z_e(\tau)$ being the cumulative idleness for queue e; $S(\tau) \triangleq \left[S_\pi(\tau)\right]_{\pi\in\mathcal{S}}$, with $S_\pi(\tau)$ being the cumulative time schedule π is employed.

For scaling parameter $r \in \mathbb{R}, r \geq 1$, define the scaled system state as

$$\begin{aligned} \mathcal{Z}^{(r)}(t) \triangleq \Big(&Q^{(r)}(t), Z^{(r)}(t), N^{(r)}(t), S^{(r)}(t), \\ &\bar{X}^{(r)}(t), \mathsf{A}^{(r)}(t), D^{(r)}(t), A^{(r)}(t)\Big). \end{aligned}$$

Here, the scaled components are defined as

$$\begin{aligned} &Q^{(r)}(t) \triangleq r^{-1}Q(rt), \quad N^{(r)}(t) \triangleq r^{-1}N(rt), \\ &\bar{X}^{(r)}(t) \triangleq r^{-1}\bar{X}(rt), \quad \mathsf{A}^{(r)}(t) \triangleq r^{-1}\mathsf{A}(rt), \\ &D^{(r)}(t) \triangleq r^{-1}D(rt), \quad A^{(r)}(t) \triangleq r^{-1}A(rt), \\ &Z^{(r)}(t) \triangleq r^{-1}\big[(rt-\lfloor rt\rfloor)Z(\lceil rt\rceil) + (\lceil rt\rceil - rt)Z(\lfloor rt\rfloor)\big], \\ &S^{(r)}(t) \triangleq r^{-1}\big[(rt-\lfloor rt\rfloor)S(\lceil rt\rceil) + (\lceil rt\rceil - rt)S(\lfloor rt\rfloor)\big]. \end{aligned}$$

Note that, in above we have 'linearized' the components $Z(\cdot)$ and $S(\cdot)$ merely for technical convenience.

Our interest is in understanding the behavior of $\mathcal{Z}^{(r)}(\cdot)$ as $r \to \infty$. Roughly speaking, in this limiting system the trajectories will satisfy certain deterministic equations, called fluid model equations. And, solutions to these equations will be denoted as fluid model solution defined below. The formal result is stated in Theorem 1.

Definition 1 (**Fluid Model Solution**). *Given fixed initial conditions $q(0) \in \mathbb{R}^{\mathcal{E}}_+$ and $n(0) \in \mathbb{R}^{\mathcal{F}}_+$, for every time horizon $T > 0$, let $\mathsf{FMS}(T)$ denote the set of all trajectories*

$$\begin{aligned} &\mathfrak{z}(t) \triangleq \big(q(t), z(t), n(t), s(t), \bar{x}(t), \mathsf{a}(t), d(t), a(t)\big) \\ &\in \mathcal{X} \triangleq \mathbb{R}^{\mathcal{E}}_+ \times \mathbb{R}^{\mathcal{E}}_+ \times \mathbb{R}^{\mathcal{F}}_+ \times \mathbb{R}^{\mathcal{S}}_+ \times \mathbb{R}^{\mathcal{F}}_+ \times \mathbb{R}^{\mathcal{F}}_+ \times \mathbb{R}^{\mathcal{F}}_+ \times \mathbb{R}^{\mathcal{F}}_+, \end{aligned} \tag{5}$$

over the time interval $[0, T]$ *such that:*

(F1) All components of $\mathfrak{z}(t)$ *are Lipschitz continuous and thus differentiable almost everywhere.*

(F2) For all $t \in [0, T]$, $n(t) = n(0) + \mathsf{a}(t) - d(t)$.

(F3) For all $t \in [0, T]$, $\mathsf{a}(t) = \nu t$.

(F4) For all $t \in [0, T]$, $d(t) = \text{diag}(\mu)\bar{x}(t)$.

(F5) For all $t \in [0, T]$,

$$q(t) = q(0) - \left(I - R^\top\right) \Pi s(t) + \left(I - R^\top\right) z(t) + \Gamma a(t).$$

(F6) For all $t \in [0, T]$, $a(t) = \bar{x}(t)$.

(F7) For all $t \in [0, T]$, $\mathbf{1}^\top s(t) = t$.

(F8) Each component of $z(\cdot), s(\cdot)$ *&* $\bar{x}(\cdot)$ *is non-decreasing.*

In addition, define the set $\mathsf{FMS}^\alpha(T)$ *to be the subset of trajectories in* $\mathsf{FMS}(T)$ *that also satisfy:*

(F9) If $t \in [0, T]$ *is a regular point, then for all* $f \in \mathcal{F}$,

$$x_f(t) = \begin{cases} \underset{x \in [0,C]}{\text{argmax}} \ \dfrac{x^{1-\alpha} n_f^\alpha(t)}{1-\alpha} - q_{\iota(f)}^\alpha(t)\, x & \text{if } n_f(t) > 0, \\ \nu_f/\mu_f \ (= \rho_f) & \text{otherwise,} \end{cases}$$

where $x_f(t) \triangleq \dot{\bar{x}}_f(t)$.

(F10) If $t \in [0, T]$ *is a regular point, then for all* $\pi \in \mathcal{S}$, $\dot{s}_\pi(t) = 0$, *if*

$$\pi^\top (I - R) q^\alpha(t) < \max_{\sigma \in \mathcal{S}} \ \sigma^\top (I - R) q^\alpha(t).$$

(F11) If $t \in [0, T]$ *is a regular point and* $n_f(t) = 0$ *for some* $f \in \mathcal{F}$, *then* $q_{\iota(f)}(t) = 0$.

(F12) For all $t \in [0, T]$, $z(t) = \mathbf{0}$.

Note that (F1)–(F8) correspond to fluid model equations that must be satisfied under *any* scheduling policy, and, hence, are *algorithm independent* fluid model equations. On the other hand, (F9)–(F12) are particular the networks controlled under the MWUM-α policy. (F9) captures the long-term effect of the rate allocation mechanism through the α-fair utility maximization based policy. Indeed, in a static resource allocation model, the (F9) can be thought of as the *primal* update in a in an algorithm for solving the maximization problem that tries to allocate rates to maximize the net α-fair utility of flows subject to capacity constraints. (F10) captures the effect of short-term packet-level behavior induced by the scheduling algorithm. Specifically, the characteristics of the maximum weight scheduling algorithm are captured by this equation.

4.2 Formal Statement

We wish to establish fluid model solutions as limit of the scaled system state process $\mathcal{Z}^{(r)}(\cdot)$ as $r \to \infty$. To this end, fix a time horizon $T > 0$. Let $\mathbf{D}[0, T]$ denote the space of all functions from $[0, T]$ to $\mathcal{X}$, as in (5), that are right continuous with left limits (RCLL). We will assume that this space is equipped with the Skorohod metric, which we denote by $\mathbf{d}(\cdot, \cdot)$. Given a fixed scaling parameter r, consider the scaled system dynamics over interval $[0, T]$. Each sample path $\{\mathcal{Z}^{(r)}(t),\ t \in [0, T]\}$ of the system state is RCLL, and hence is contained in the space $\mathbf{D}[0, T]$.

The following theorem formally establishes the convergence of the scaled system process to a fluid model solution of the form specified in Definition 1.

Theorem 1. *Given a fixed time horizon* $T > 0$, *consider a sequence of scaled system state processes* $\{\mathcal{Z}^{(r)}(t),\ t \in [0, T]\} \subset \mathbf{D}[0, T]$, *for* $r \geq 1$ *evolving under an arbitrary control policy. Suppose the initial conditions*

$$\lim_{r \to \infty} Q^{(r)}(0) = q(0), \quad \lim_{r \to \infty} N^{(r)}(0) = n(0), \quad a.s., \qquad (6)$$

are satisfied. Then, for any $\varepsilon > 0$,

$$\liminf_{r \to \infty} \ \mathsf{P}\left[\mathcal{Z}^{(r)}(\cdot) \in \mathsf{FMS}_\varepsilon(T)\right] = 1,$$

where

$$\mathsf{FMS}_\varepsilon(T) \triangleq \{\mathbf{x} \in \mathbf{D}[0, T] \ : \ \mathbf{d}(\mathbf{x}, \mathbf{y}) < \varepsilon,\ \mathbf{y} \in \mathsf{FMS}(T)\}.$$

Additionally, under the MWUM-α *control policy we have that*

$$\liminf_{r \to \infty} \mathsf{P}\left[\mathcal{Z}^{(r)}(\cdot) \in \mathsf{FMS}_\varepsilon^\alpha(T)\right] = 1,$$

where

$$\mathsf{FMS}_\varepsilon^\alpha(T) \triangleq \{\mathbf{x} \in \mathbf{D}[0, T] \ : \ \mathbf{d}(\mathbf{x}, \mathbf{y}) < \varepsilon,\ \mathbf{y} \in \mathsf{FMS}^\alpha(T)\}.$$

Proof Sketch. The result can be established by following a standard sequence of arguments (cf. [2, 21, 10]). First, the sequence of measures corresponding to the sequence of random processes $\{\mathcal{Z}^{(r)}(\cdot)\}$ is shown to be tight. This establishes that limit points must exist. Next, it is established that each limit point must satisfy the conditions of a fluid solution with probability 1. Details can be found in the longer version of this paper [16]. □

5. SYSTEM STABILITY

In this section, we characterize the stability of a network under the MWUM-α policy. In particular, we shall see that the network Markov process is positive recurrent as long as the system is *underloaded*, or the system is maximally stable. In order to construct the stability region under the MWUM-α policy, first define a set $\Lambda \subset \mathbb{R}_+^{\mathcal{E}}$ of *per queue* arrival rates by

$$\Lambda \triangleq \left\{\lambda \in \mathbb{R}_+^{\mathcal{E}} \ : \ \exists\ s \in \mathbb{R}_+^{\mathcal{S}} \text{ with } \lambda \leq \Pi s,\ \mathbf{1}^\top s \leq 1\right\}. \qquad (7)$$

Imagine that the network has no packet arrivals from flows, but instead has packets arriving according to exogenous processes. Suppose that $\lambda \in \mathbb{R}_+^{\mathcal{E}}$ is the vector of exogenous arrival rates, so that packets arrived to each queue e at rate λ_e. Then, it is not difficult to see that the network would not be stable under *any* scheduling policy if $\lambda \notin \Lambda$. Otherwise, there is at least one queue in the network that is loaded beyond its capacity. Hence, the set Λ represents the raw *scheduling capacity* of the network.

The set Λ can alternatively be described as follows: Given a vector $\lambda \in \mathbb{R}_+^{\mathcal{E}}$, consider the linear program

$$\begin{aligned} \mathsf{PRIMAL}(\lambda) \triangleq \underset{s}{\text{minimize}} \quad & \mathbf{1}^\top s \\ \text{subject to} \quad & \lambda \leq \Pi s, \\ & s \in \mathbb{R}_+^{\mathcal{S}}. \end{aligned}$$

Clearly $\lambda \in \Lambda$ if and only if $\mathsf{PRIMAL}(\lambda) \leq 1$. The quantity $\mathsf{PRIMAL}(\lambda)$ is called the *effective load* of a system with exogenous arrivals of rate λ.

Now, in our model, packets arrive to the network not through an exogenous process, but rather, they are generated by flows. As discussed in Section 2.2, each flow type $f \in \mathcal{F}$ generates packets at according to an offered load of

ρ_f. The generated packets subsequently traverse through the network along pre-determined paths specified by the routing matrix Γ. Let $\lambda \in \mathbb{R}_+^{\mathcal{E}}$ be the vector of *implied loads* on the scheduling network due to this. It seems reasonable to relate λ and the vector $\rho \in \mathbb{R}_+^{\mathcal{F}}$ of offered loads according to $\lambda = \Gamma\rho + R^\top\lambda$. Equivalently, we define $\lambda \triangleq \Xi\Gamma\rho$, where Ξ is from (1). We define the *effective load* $\mathsf{L}(\rho)$ of our network by $\mathsf{L}(\rho) \triangleq \mathsf{PRIMAL}(\Xi\Gamma\rho)$.

Given the above discussion, it seems natural to suspect that the network's scheduling capacity allows it to operate effectively as long as $\mathsf{L}(\rho) \leq 1$. This motivates the following definition:

DEFINITION 2 (**Admissibility**). *We call a vector $\rho \in \mathbb{R}_+^{\mathcal{F}}$ **admissible** if $\mathsf{L}(\rho) \leq 1$. ρ is strictly admissible if $\mathsf{L}(\rho) < 1$. Finally, ρ is critically admissible if $\mathsf{L}(\rho) = 1$.*

We establish system stability, or, formally, positive recurrence, when arrival process is *strictly admissible*. To this end, recall that the system is completely described by the $\mathcal{Z}(\cdot)$ process. Under the MWUM-α policy, the evolution of all the components of $\mathcal{Z}(t)$ is entirely determined by $\big(N(t), Q(t)\big)$. Further, the changes in $\big(N(t), Q(t)\big)$ occur at times specified by the arrivals of a (time-varying) Poisson process. Therefore, tuple $\big(N(\cdot), Q(\cdot)\big)$ forms a continuous-time Markov chain. The following is the main result of this section:

THEOREM 2. *Consider a network system with strictly admissible ρ operating under the MWUM-α policy. Then, the Markov chain $\big(Q(\cdot), N(\cdot)\big)$ is positive recurrent.*

It is worth noting that if $\mathsf{L}(\rho) > 1$, then at least of the queues in the network must be, on average, loaded beyond its capacity. Hence the network Markov process can not be positive recurrent or stable.

The proof of Theorem 2, which can be found in the longer version of this paper [16], uses the fluid model based approach pioneered by Dai [5]. A crucial step in this procedure is to establish the stability of the fluid model. To this end, consider the Lyapunov function L_α defined over the vector of the number of flows, $n = [n_f] \in \mathbb{R}_+^{\mathcal{F}}$, and the vector of queue lengths, $q = [q_e] \in \mathbb{R}_+^{\mathcal{E}}$, by

$$L_\alpha(n, q) \triangleq \sum_{f\in\mathcal{F}} \frac{n_f^{1+\alpha}}{\mu_f \rho_f^\alpha} + \sum_{e\in\mathcal{E}} q_e^{1+\alpha}. \tag{8}$$

The following lemma, whose proof is found in the longer version [16], demonstrates stability of the fluid model.

LEMMA 3. *Let $\big(n(\cdot), q(\cdot)\big)$ be, respectively, the flow count process and the queue length process of a fluid model solution in the set $\mathsf{FMS}^\alpha(T)$. If $\mathsf{L}(\rho) \leq 1$, then for every regular point $t \in [0, T]$,*

$$\frac{d}{dt} L_\alpha\big(n(t), q(t)\big) \leq 0.$$

Suppose further that $\mathsf{L}(\rho) < 1$, and that the initial conditions $\big(n(0), q(0)\big)$ satisfy

$$L_\alpha\big(n(0), q(0)\big) = 1.$$

Then, for T sufficiently large, there exist $\delta > 0$ and $\tau > T$ such that, for all $t \geq \tau$,

$$L_\alpha\big(n(t), q(t)\big) \leq 1 - \delta.$$

6. CRITICAL LOADING

We have established the throughput optimality of the system under the MWUM-α control policy, for any $\alpha \in (0, \infty) \setminus \{1\}$. Thus, this entire family of policies possesses good *first order* characteristics. Further, there may be many other throughput optimal policies outside the class of MWUM-α policies. This naturally raises the question of whether there is a 'best' choice of α, and how the resulting MWUM-α policy might compare to the universe of all other policies.

In order to answer these questions, we desire a more refined analysis of policy performance than throughput optimality. One way to obtain such an analysis is via the study of a *critically loaded system*, i.e., a system with critically admissible arrival rates. Under a critical loading, fluid model solutions take non-trivial values over entire horizon. In contrast, for strictly admissible systems under throughput optimal policies, all fluid trajectories go to 0 (cf. Lemma 3). We will employ the study of the fluid model solutions of critically loaded systems as a tool for the comparative analysis of network control policies.

In particular, given a vector of flow counts, $n = [n_f] \in \mathbb{R}_+^{\mathcal{F}}$, and the vector of queue lengths, $q = [q_e] \in \mathbb{R}_+^{\mathcal{E}}$, consider the linear cost function

$$c(n, q) \triangleq \sum_{f\in\mathcal{F}} \frac{n_f}{\mu_f} + \sum_{e\in\mathcal{E}} q_e = \mathbf{1}^\top \left[\Gamma \operatorname{diag}(\mu^{-1}) n + q\right]. \tag{9}$$

This cost function is analogous to a 'minimum delay' objective in a packet-level queueing network: a cost is incurred for each queued packet, and a cost is incurred for each outstanding flow proportional to the number of packets that it will generate.

In this section, we establish fundamental lower bounds that apply to the cost incurred in a critically loaded fluid model under *any* scheduling policies. In Sections 7 and 8, we will compare these with the cost incurred by MUWM-α control policies. We shall find that as $\alpha \to 0^+$, the cost induced by the MUWM-α algorithms improves and becomes close to the algorithm independent lower bound we establish.

6.1 Virtual Resources and Workload

We start with some definitions. First, consider the dual of the LP $\mathsf{PRIMAL}(\lambda)$,

$$\begin{aligned} \mathsf{DUAL}(\lambda) \triangleq \underset{\zeta}{\text{maximize}} \quad & \lambda^\top \zeta_e \\ \text{subject to} \quad & \Pi^\top \zeta \leq \mathbf{1}, \\ & \zeta \in \mathcal{R}_+^{\mathcal{E}}. \end{aligned}$$

Since there is no duality gap, the solution of $\mathsf{PRIMAL}(\lambda)$ is equal to the solution of $\mathsf{DUAL}(\lambda)$.

DEFINITION 3 (**Virtual Resource**). *We will call any feasible solution $\zeta \in \mathbb{R}_+^{\mathcal{E}}$ of dual optimization problem $\mathsf{DUAL}(\lambda)$ a **virtual resource**. Suppose the system is critically loaded, i.e., the offered load vector ρ satisfies*

$$\mathsf{L}(\rho) = \mathsf{PRIMAL}(\Xi\Gamma\rho) = \mathsf{DUAL}(\Xi\Gamma\rho) = 1.$$

*Then, we call a virtual resource that is an optimal solution of $\mathsf{DUAL}(\Xi\Gamma\rho)$ a **critical virtual resource**.*

For a critically loaded system with offered load vector ρ, let $\mathsf{CR}(\rho)$ denote the set of all critical virtual resources. Note that $\mathsf{CR}(\rho)$ is a bounded polytope and hence possesses finitely many extreme points. Let $\mathsf{CR}^*(\rho)$ denote the finite set of extreme points of $\mathsf{CR}(\rho)$.

The following definition captures the amount of 'work' associated with a critical resource, as a function of the current state of the system.

DEFINITION 4 (**Workload**). *Consider a critically loaded system with an offered load vector ρ and a critical virtual resource $\zeta \in \mathsf{CR}(\rho)$. If the flow count and queue length vectors are given by (n, q), the* ***workload*** *associated the resource ζ is defined to be*

$$w_\zeta(n, q) \triangleq \zeta^\top \Xi \left[q + \Gamma \operatorname{diag}(\mu)^{-1} n\right].$$

6.2 A Lower Bound on Fluid Trajectories

Given the critically loaded system with offered load vector ρ. We claim the following fundamental lower bound the fluid trajectory under *any* algorithm. This bound can be thought of as minimal work-conservation.

LEMMA 4. *Consider the fluid model trajectory of system under any scheduling and rate allocation algorithm, with flow count and queue length processes given by $\big(n(\cdot), q(\cdot)\big)$. Then, for any time $t \geq 0$ and any critical virtual resource $\zeta \in \mathsf{CR}(\rho)$,*

$$w_\zeta\big(n(0), q(0)\big) \leq w_\zeta\big(n(t), q(t)\big). \tag{10}$$

PROOF. Given a time interval $[0, T]$, for any $T > 0$, consider the fluid model trajectory $\mathfrak{z}(t)$ of the form (5). By Theorem 1, this fluid trajectory must satisfy the algorithm independent fluid model equations, (F1)–(F8) in Definition 1. By (F1), the trajectory is Lipschitz continuous and differentiable for almost all $t \in [0, T]$. For any such regular t, by (F2)–(F4), we have $\dot{n}(t) = \nu - \operatorname{diag}(\mu)\dot{\bar{x}}(t)$. Thus,

$$\Gamma \operatorname{diag}(\mu^{-1})\dot{n}(t) = \Gamma\rho - \Gamma\dot{\bar{x}}(t). \tag{11}$$

From (F5)–(F6), we obtain

$$\dot{q}(t) = \left(I - R^\top\right)\dot{z}(t) - \left(I - R^\top\right)\Pi\dot{s}(t) + \Gamma\dot{\bar{x}}(t) \tag{12}$$

Adding (11) and (12), we obtain

$$\dot{q}(t) + \Gamma \operatorname{diag}(\mu^{-1})\dot{n}(t) = \Gamma\rho + \left(I - R^\top\right)\left(\dot{z}(t) - \Pi\dot{s}(t)\right).$$

Now, multiplying both sides by $\Xi \triangleq \left(I - R^\top\right)^{-1}$, we obtain

$$\Xi\left[\dot{q}(t) + \Gamma \operatorname{diag}(\mu^{-1})\dot{n}(t)\right] = \Xi\Gamma\rho + \dot{z}(t) - \Pi\dot{s}(t). \tag{13}$$

Now, consider a critical virtual resource $\zeta \in \mathsf{CR}(\rho)$. Since ζ is $\mathsf{DUAL}(\Xi\Gamma\rho)$ optimal, $\zeta^\top\Xi\Gamma\rho = 1$. Taking an inner product of (13) with $\zeta^\top$, we obtain

$$\zeta^\top\Xi\left[\dot{q}(t) + \Gamma \operatorname{diag}(\mu^{-1})\dot{n}(t)\right] = 1 + \zeta^\top\dot{z}(t) - \zeta^\top\Pi\dot{s}(t). \tag{14}$$

Now, by (F8), $z(t)$ is non-decreasing, i.e., $\dot{z}(t)$ is non-negative. Since ζ is also non-negative, $\zeta^\top\dot{z}(t) \geq 0$. By (F8), $\dot{s}(t)$ is non-negative. Since ζ is $\mathsf{DUAL}(\Xi\Gamma\rho)$ feasible and from (F7), it follows that $\zeta^\top\Pi\dot{s}(t) \leq \mathbf{1}^\top\dot{s}(t) = 1$. Applying these observations to (14), it follows that

$$\frac{d}{dt}w_\zeta\big(n(t), q(t)\big) = \zeta^\top\Xi\left[\dot{q}(t) + \Gamma \operatorname{diag}(\mu^{-1})\dot{n}(t)\right] \geq 0.$$

Given that $\big(n(\cdot), q(\cdot)\big)$ are Lipschitz continuous, the desired result follows immediately. □

Lemma 4 guarantees the conservation of workload under any policy. This motivates the *effective cost* of a state $(n, q) \in \mathbb{R}_+^{\mathcal{F}} \times \mathbb{R}_+^{\mathcal{E}}$, defined by the linear program

$$\begin{aligned} c^*(n, q) \triangleq \underset{n', q'}{\text{minimize}} \quad & c(n', q') \\ \text{subject to} \quad & w_\zeta(n', q') \geq w_\zeta(n, q), \\ & \forall\, \zeta \in \mathsf{CR}^*(\rho), \\ & n \in \mathbb{R}_+^{\mathcal{F}}, \quad q \in \mathbb{R}_+^{\mathcal{E}}. \end{aligned} \tag{15}$$

The effective cost is the lowest cost of any state with at least as much workload at (n, q). We have the following lower bound on the cost achieved under any fluid trajectory:

THEOREM 5. *Consider fluid model trajectory of system under any scheduling and rate allocation algorithm, with flow count and queue length processes given by $\big(n(\cdot), q(\cdot)\big)$. Then, for any time $t \geq 0$, the instantaneous cost $c\big(n(t), q(t)\big)$ is bounded below according to*

$$c^*\big(n(0), q(0)\big) \leq c\big(n(t), q(t)\big). \tag{16}$$

PROOF. By Lemma 4, if the initial condition of a fluid trajectory satisfies $\big(n(0), q(0)\big) = (n, q)$, then $\big(n(t), q(t)\big)$ is feasible for (15) for every $t \geq 0$. The result immediately follows. □

7. BALANCED SYSTEMS

In this section, we will develop a bound on the cost achieved in a fluid model solution under the MWUM-α policy. In particular, we will establish that this cost, at any instant of time, is within a constant factor of the cost achievable under *any* policy. The constant factor is uniform across the entire fluid trajectory, and relates to a notion of balance on the critical resources of the network, which we will describe shortly.

We begin with a preliminary lemma, that provides an upper bound on the cost under the MWUM-α policy. This upper bound is closely related to the Lyapunov function introduced earlier for studying the system stability.

LEMMA 6. *Consider fluid model trajectory of system under the MWUM-α policy, where $\alpha \in (0, \infty) \setminus \{1\}$, and denote the flow count and queue length processes by $\big(n(\cdot), q(\cdot)\big)$. Suppose that the offered load vector ρ satisfies $\mathsf{L}(\rho) \leq 1$. Then, at any time $t \geq 0$, it must be that*

$$c\big(n(t), q(t)\big) \leq \big(1 + \beta(\alpha)\big)c\big(n(0), q(0)\big), \tag{17}$$

where $\beta(\alpha) \to 0$ as $\alpha \to 0^+$.

PROOF. Recall the Lyapunov function L_α from (8). It follows from Lemma 3 that, so long as $\mathsf{L}(\rho) \leq 1$,

$$L_\alpha\big(n(t), q(t)\big) \leq L_\alpha\big(n(0), q(0)\big). \tag{18}$$

Now, the standard norm inequality suggests that for any $v \in \mathbb{R}_+^d$ and $\alpha > 0$,

$$d^{-\frac{\alpha}{1+\alpha}}\|v\|_1 \leq \|v\|_{1+\alpha} \leq \|v\|_1. \tag{19}$$

Combining this with (18), it follows that, if $d \triangleq |\mathcal{E}| + |\mathcal{F}|$,

$$\begin{aligned} &\sum_{f \in \mathcal{F}} \frac{n_f(t)}{\mu_f}\left(\frac{1}{\nu_f}\right)^{\frac{\alpha}{1+\alpha}} + \sum_{e \in \mathcal{E}} q_e(t) \\ &\qquad \leq d^{\frac{\alpha}{1+\alpha}}\left[\sum_{f \in \mathcal{F}} \frac{n_f(0)}{\mu_f}\left(\frac{1}{\nu_f}\right)^{\frac{\alpha}{1+\alpha}} + \sum_{e \in \mathcal{E}} q_e(0)\right]. \end{aligned}$$

Now, as $\alpha \to 0^+$, $d^{\frac{\alpha}{1+\alpha}} \to 1$. Also,

$$\left(\frac{1}{\nu^*}\right)^{\frac{\alpha}{1+\alpha}} \le \left(\frac{1}{\nu_f}\right)^{\frac{\alpha}{1+\alpha}} \le \left(\frac{1}{\nu_*}\right)^{\frac{\alpha}{1+\alpha}}, \tag{20}$$

where $\nu_* \triangleq \min_f \nu_f$ and $\nu^* \triangleq \max_f \nu_f$. Thus, as $\alpha \to 0^+$, $1/\nu_f \to 1$ uniformly over f. The result then follows. □

The following definition is central to our performance guarantee:

Definition 5 (**Balance Factor**). *Given a system that is critically loaded with offered load vector ρ, define the* ***balance factor*** *as the value of the optimization problem*

$$\begin{aligned}\gamma(\rho) \triangleq \underset{n,q,n',q'}{\text{minimize}} \quad & c(n',q') \\ \text{subject to} \quad & w_\zeta(n',q') \ge w_\zeta(n,q),\ \forall\ \zeta \in \mathsf{CR}^*(\rho), \\ & c(n,q) = 1, \\ & n,n' \in \mathbb{R}_+^{\mathcal{F}}, \quad q,q' \in \mathbb{R}_+^{\mathcal{E}}.\end{aligned}$$

It is clear that $\gamma(\rho) \ge 0$, since $n', q' \ge \mathbf{0}$. Since there are feasible solutions with $(n,q) = (n',q')$, it is also true that $\gamma(\rho) \le 1$. In order to interpret $\gamma(\rho)$, assume for the moment that there is only a single critical extreme resource $\zeta \in \mathsf{CR}^*(\rho)$. If we define $v \triangleq \Xi^\top \zeta$, then the constraint that $w_\zeta(n',q') \ge w_\zeta(n,q)$ is equivalent to

$$v^\top \left[\Gamma \operatorname{diag}(\mu)^{-1} n' + q'\right] \ge v^\top \left[\Gamma \operatorname{diag}(\mu)^{-1} n + q\right].$$

In this case, it is clear that the solution to the LP defining $\gamma(\rho)$ is given by $\gamma(\rho) = (\min_e v_e)/(\max_e v_e)$. Hence, $\gamma(\rho)$ is the measure of the degree of 'balance' of the critical resource ζ across buffers in the network.

In the more general case (i.e., $|\mathsf{CR}^*(\rho)| \ge 1$), define the set $\mathcal{V} \triangleq \operatorname{span}\{\Xi^\top\zeta,\ \zeta \in \mathsf{CR}^*(\rho)\}$. It is not difficult to see that $\gamma(\rho) > 0$ if and only if, for each queue $e \in \mathcal{V}$, there exists $v \in \mathcal{V}$ with $v_e > 0$. In other words, if every queue is influenced by some critical resource. We call such networks *balanced*. In the extreme, if $\mathbf{1} \in \mathcal{V}$, then $\gamma(\rho) = 1$.

The following is the main theorem of this section. It offers a bound on the cost incurred under the MWUM-α policy, relative to *any* other policy, which is a function of the balance factor.

Theorem 7. *Consider fluid model trajectory of a critically loaded system under the MWUM-α policy, where $\alpha \in (0,\infty)\backslash\{1\}$, and denote the flow count and queue length processes by $\big(n(\cdot), q(\cdot)\big)$. Suppose that $\gamma(\rho) > 0$. Let $\big(n'(\cdot), q'(\cdot)\big)$ be the flow count and queue length policies under any other policy, given the same initial conditions, i.e., $n(0) = n'(0)$ and $q(0) = q'(0)$. Then, at any time $t \ge 0$, it must be that*

$$c\big(n(t), q(t)\big) \le \frac{1+\beta(\alpha)}{\gamma(\rho)} c\big(n'(t), q'(t)\big), \tag{21}$$

where $\beta(\alpha) \to 0$ as $\alpha \to 0^+$.

Proof. First, note that if $\big(n(0), q(0)\big) = \mathbf{0}$, i.e., the system is empty, then this holds for all $t \ge 0$ (cf. Theorem 9). In this case, (21) is immediate. Otherwise, fix $t \ge 0$, set $\bar{c} \triangleq c\big(n(0), q(0)\big) > 0$. Define

$$(n',q') \triangleq \big(n'(t), q'(t)\big)/\bar{c}, \quad (n,q) \triangleq \big(n(0), q'(0)\big)/\bar{c}.$$

Using Lemma 4, it is clear that (n,q,n',q') is feasible for the LP defining $\gamma(\rho)$. Thus,

$$c\big(n(0), q(0)\big) \le \frac{1}{\gamma(\rho)} c\big(n'(t), q'(t)\big).$$

The result then follows by applying Lemma 6. □

8. INVARIANT MANIFOLD

In Section 7, we proved a constant factor guarantee on the cost of the MWUM-α policy, relative to the cost achieved under any other policy. Our bound held point-wise, at *every* instant of time. However, the constant factor of the bound depended on the balance factor, which could be significantly large or potentially infinite.

In this section, we consider a different type of analysis. Instead of considering the evolution of the fluid model for *every* time t, we instead examine the asymptotically limiting states of the fluid model as $t \to \infty$. In particular, we characterize these *invariant* states as fixed points in the solution space of an optimization problem. We shall also show that these fixed points are attractive, i.e., starting from any initial condition, the fluid trajectory reaches an invariant state. We will quantify time to converge to the invariant manifold as a function of the initial conditions of the fluid trajectory.

This characterization is of invariant states is key towards establishing the *state space collapse* property of the system under a heavy traffic limit [2]. Moreover, we shall see that these invariant states are cost optimal as $\alpha \to 0^+$. In other words, the cost of an invariant state cannot be improved by *any* policy.

8.1 Optimization Problems

We start with two useful optimization problems that will be useful in characterizing invariant states of the fluid trajectory. To this end, consider a critically loaded system, i.e., a system where the offered load ρ is such that $\mathsf{L}(\rho) = 1$.

Suppose we are given a state $(n,q) \in \mathbb{R}_+^{\mathcal{F}} \times \mathbb{R}_+^{\mathcal{E}}$ of, respectively, flow counts and queue lengths. Define the optimization problem

$$\begin{aligned}\mathsf{ALGP}(n,q) \triangleq \underset{n',q',t,x,\sigma}{\text{minimize}} \quad & L_\alpha(n',q') \\ \text{subject to} \quad & n' = n + t\big[\nu - \operatorname{diag}(\mu)x\big], \\ & q' = q + t\left[\Gamma x - \left(I - R^\top\right)\sigma\right], \\ & n' \in \mathbb{R}_+^{\mathcal{F}}, \quad q' \in \mathbb{R}_+^{\mathcal{E}}, \quad t \in \mathbb{R}_+, \\ & x \in [0,C]^{\mathcal{F}},\ \sigma \in \Lambda.\end{aligned}$$

Here, recall that Λ is the scheduling capacity region of the network, defined by (7). Similarly, define the optimization problem

$$\begin{aligned}\mathsf{ALGD}(n,q) \triangleq \underset{n',q'}{\text{minimize}} \quad & L_\alpha(n',q') \\ \text{subject to} \quad & w_\zeta(n',q') \ge w_\zeta(n,q), \\ & \quad \forall\ \zeta \in \mathsf{CR}^*(\rho), \\ & n' \in \mathbb{R}_+^{\mathcal{F}}, \quad q' \in \mathbb{R}_+^{\mathcal{E}}.\end{aligned}$$

Intuitively, give a state (n,q), $\mathsf{ALGP}(n,q)$ finds a state (n',q') which minimizes the Lyapunov function L_α and can be reach starting from (n,q), using feasible scheduling and rate allocation decisions. $\mathsf{ALGP}(n,q)$, on the other hand, finds a state (n',q') which minimizes the Lyapunov function and has at least as much workload as (n,q). The following result states that both $\mathsf{ALGP}(n,q)$ and $\mathsf{ALGD}(n,q)$ are equivalent optimization problems:

Lemma 8. *A state $(n',q') \in \mathbb{R}_+^{\mathcal{F}} \times \mathbb{R}_+^{\mathcal{E}}$ is feasible for the optimization problem $\mathsf{ALGP}(n,q)$ if and only if it is feasible for the optimization problem $\mathsf{ALGD}(n,q)$.*

PROOF. First, consider any (n', q', t, x, σ) that is feasible for $\mathsf{ALGP}(n,q)$. Note that feasibility for $\mathsf{ALGP}(n,q)$ implies that

$$\Gamma \operatorname{diag}(\mu)^{-1} n' + q' \geq \Gamma \operatorname{diag}(\mu)^{-1} n + q + t\Big[\Gamma \operatorname{diag}(\mu)^{-1}\nu - \Gamma x\Big] + t\left[\Gamma x - \Big(I - R^\top\Big)\sigma\right].$$

Therefore, if $\zeta \in \mathsf{CR}^*(\rho)$, we have that

$$w_\zeta(n', q') = w_\zeta(n,q) + t\left[\zeta^\top \Xi\Gamma\rho - \zeta^\top\sigma\right].$$

Since $\sigma \in \Lambda$ and ζ is feasible for $\mathsf{DUAL}(\Xi\Gamma\rho)$, we have $\zeta^\top\sigma \leq 1$. Since $\zeta \in \mathsf{CR}^*(\rho)$, we have $\zeta^\top\Xi\Gamma\rho = 1$. Therefore, as $t \geq 0$, it follows that

$$w_\zeta(n', q') \geq w_\zeta(n,q).$$

That is, (n', q') is $\mathsf{ALGD}(n,q)$ feasible.

Next, assume that (n', q') is feasible for $\mathsf{ALGD}(n,q)$. Given some $t \geq 0$, define

$$x \triangleq \operatorname{diag}(\mu)^{-1}\left[\nu - t^{-1}(n' - n)\right], \quad \sigma \triangleq \Xi\left[\Gamma x - t^{-1}(q' - q)\right].$$

With these definitions, if we establish existence of $t \geq 0$ so that $\mathbf{0} \leq x \leq C\mathbf{1}$ and $\sigma \in \Lambda$, then (n', q', t, x, σ) is feasible for $\mathsf{ALGP}(n,q)$ feasible.

Note that as $t \to \infty$, $x \to \rho$. By assumption, $\rho_f > 0$ and $\rho_f < C$ for all $f \in \mathcal{F}$. Therefore, for t sufficiently large, $\mathbf{0} \leq x \leq C\mathbf{1}$.

Next, we wish to show that, for t sufficiently large, $\sigma \in \Lambda$. This requirement is equivalent to demonstrating that $\mathsf{PRIMAL}(\sigma) \leq 1$ and that $\sigma \geq \mathbf{0}$. To show that $\mathsf{PRIMAL}(\sigma) \leq 1$, note that $\mathsf{PRIMAL}(\sigma) = \mathsf{DUAL}(\sigma)$ and suppose that ζ is feasible for $\mathsf{DUAL}(\sigma)$. Then,

$$\begin{aligned}\zeta^\top\sigma &= \zeta^\top\left[\Xi\Gamma x - t^{-1}(q' - q)\right] \\ &= \zeta^\top\left[\Xi\Gamma\rho - t^{-1}\Xi\Gamma\operatorname{diag}(\mu)^{-1}(n' - n) - t^{-1}(q' - q)\right] \\ &= \zeta^\top\Xi\Gamma\rho - t^{-1}\left[w_\zeta(n', q') - w_\zeta(n,q)\right].\end{aligned}$$

If $\zeta \in \mathsf{CR}(\rho)$, then

$$\zeta^\top\Xi\Gamma\rho = 1, \quad \text{and} \quad w_\zeta(n', q') - w_\zeta(n,q) \geq 0,$$

thus $\zeta^\top\sigma \leq 1$. On the other hand, if $\zeta \notin \mathsf{CR}(\rho)$, $\zeta^\top\Xi\Gamma\rho < 1$. Therefore, in any event, for t sufficiently large, $\mathsf{DUAL}(\rho) \leq 1$.

To show that $\sigma \geq \mathbf{0}$, note that

$$\begin{aligned}\sigma &= \Xi\left[\Gamma x - t^{-1}(q' - q)\right] \\ &= \Xi\Gamma\rho - t^{-1}\left[\Xi\Gamma\operatorname{diag}(\mu)^{-1}(n' - n) + \Xi(q' - q)\right].\end{aligned}$$

By assumption, $\Xi\Gamma\rho > \mathbf{0}$. Therefore, for t sufficiently large enough, $\sigma \geq \mathbf{0}$. □

8.2 Fixed Points: Characterization

Note that the optimization problem $\mathsf{ALGD}(n,q)$ is a convex minimization problem: it has a convex feasible set with strictly convex and coercive objective. from standard arguments from theory of convex optimization, it follows that an optimal solution exists and is unique. Hence, we can make the following definition:

DEFINITION 6 (**Lifting Map**). *Given a critically scaled system, we define the* ***lifting map*** $\Delta \colon \mathbb{R}_+^{\mathcal{F}} \times \mathbb{R}_+^{\mathcal{E}} \to \mathbb{R}_+^{\mathcal{F}} \times \mathbb{R}_+^{\mathcal{E}}$ *to be the function that maps a state (n,q) to the unique solution of the optimization problem $\mathsf{ALGD}(n,q)$.*

The main result of this section is to characterize the invariant states of fluid model as the fixed points of lifting map Δ.

THEOREM 9. *A state $(n,q) \in \mathbb{R}_+^{\mathcal{F}} \times \mathbb{R}_+^{\mathcal{E}}$ is an invariant state of a fluid model solution under the MWUM-α policy if and only if it is a fixed point of Δ, i.e.,*

$$(n,q) = \Delta(n,q).$$

PROOF. The proof follows by establishing equivalence of the following statements, for every state (n,q):

(i) $(n,q) = \Delta(n,q)$.

(ii) Any fluid model solution satisfying the initial condition $\big(n(0), q(0)\big) = (n,q)$ has $\big(n(t), q(t)\big) = (n,q)$ for all $t \geq 0$.

(iii) There exists a fluid model solution with $\big(n(t), q(t)\big) = (n,q)$ for all $t \geq 0$.

(iv) (n,q) satisfy

$$(\Gamma\rho)^\top q^\alpha = \max_{\pi \in \mathcal{S}} \pi^\top (I - R) q^\alpha, \tag{22}$$

$$n_f > 0 \ \Rightarrow\ \rho_f q_{\iota(f)} = n_f, \quad \forall\, f \in \mathcal{F}, \tag{23}$$

$$n_f = 0 \ \Rightarrow\ q_{\iota(f)} = 0, \quad \forall\, f \in \mathcal{F}. \tag{24}$$

(i) ⇒ (ii): If $(n,q) = \Delta(n,q)$, then it solves $\mathsf{ALGD}(n,q)$. Suppose that the fluid model $\mathfrak{z}(t)$ satisfies with initial state $\big(n(0), q(0)\big) = (n,q)$. By Lemma 3, it follows that, for all $t \geq 0$, $L_\alpha\big(n(t), q(t)\big) \leq L_\alpha(n,q)$. From the fluid model equations (F1)–(F12), $\big(n(t), q(t)\big)$ is $\mathsf{ALGP}(n,q)$ feasible. Therefore, it follows that $\big(n(t), q(t)\big)$ is an optimal solution of $\mathsf{ALGP}(n,q)$, and, by Lemma 8, of $\mathsf{ALGD}(n,q)$. Since $\mathsf{ALGD}(n,q)$ has unique solution, it follows that $\big(n(t), q(t)\big) = (n,q)$ for all $t \geq 0$.

(ii) ⇒ (iii): This follows in a straightforward manner by considering the arguments in Theorem 1 with initial conditions given by (n,q).

(iii) ⇒ (iv): Suppose that the fluid trajectory $\mathfrak{z}(t)$ satisfies $\big(n(t), q(t)\big) = (n,q)$, for all $t \geq 0$. Then, for any regular point $t \geq 0$, we have $\dot{n}(t) = \mathbf{0}$ and $\dot{q}(t) = \mathbf{0}$. Using (F2)–(F4), it follows that $x(t) \triangleq \frac{d}{dt}\bar{x}(t) = \rho$. For any $f \in \mathcal{F}$, if $n_f = n_f(t) > 0$ and $x_f(t) = \rho_f < C$, then by (F9) it must be that $x_f(t) = n_f(t)/q_f(t)$. Therefore, $\rho_f q_f = n_f$. Similarly, if $n_f = 0$, it must be that $q_{\iota(f)} = 0$ by (F11).

Now, define $H(t) \triangleq \mathbf{1}^\top q^{1+\alpha}(t)$. Since $q(\cdot)$ is constant, applying (F5), (F6), (F12), it must be that

$$0 = \dot{H}(t) = \dot{q}(t)\top q^\alpha(t) = \left[\Gamma\rho - \Big(I - R^\top\Big)\Pi\dot{s}(t)\right]^\top q^\alpha(t).$$

Applying (F7) and (F10),

$$0 = (\Gamma\rho)^\top q^\alpha - \max_{\pi \in \mathcal{S}}\ \pi^\top (I - R) q^\alpha.$$

(iv) ⇒ (i): Suppose (n,q) satisfy (22)–(23). Define $(n', q') \triangleq \Delta(n,q)$. Since (n', q') solves $\mathsf{ALGD}(n,q)$, by Lemma 8, there exist (t, x, σ) so that (n', q', t, x, σ) is an optimal solution for $\mathsf{ALGP}(n,q)$. This solution must satisfy

$$n' = n + t\big[\nu - \operatorname{diag}(\mu)x\big], \quad q' = q + t\left[\Gamma x - \Big(I - R^\top\Big)\sigma\right].$$

Now consider the trajectory

$$\big(n(\tau), q(\tau)\big) \triangleq (n, q) + \frac{\tau}{t}(n' - n, q' - q), \quad \forall\, \tau \in [0, t].$$

Define J to be the Lyapunov function L_α evaluated along this path, i.e., $J(\tau) \triangleq L_\alpha\big(n(\tau), q(\tau)\big)$. Then,

$$\begin{aligned}\frac{\dot{J}(0)}{1+\alpha} &= \sum_{f\in\mathcal{F}} \frac{n_f^\alpha(\nu_f - \mu_f x_f)}{\mu_f \rho_f^\alpha} + (\Gamma x)^\top q^\alpha - \sigma^\top (I-R) q^\alpha \\ &= \underbrace{\left(\sum_{f\in\mathcal{F}} \frac{n_f^\alpha(\nu_f - \mu_f x_f)}{\mu_f \rho_f^\alpha} + (\Gamma\delta)^\top q^\alpha\right)}_{(X)} \\ &\quad + \underbrace{\left((\Gamma\rho)^\top q^\alpha - \sigma^\top (I-R) q^\alpha\right)}_{(Y)},\end{aligned}$$

where $\delta \triangleq x - \rho$.

First, consider Y. Since $\sigma \in \Lambda$, there exists some $s \in \mathbb{R}_+^{\mathcal{S}}$ with $\mathbf{1}^\top s \le 1$ and $\sigma \le \Pi s$. From the monotonicity of $\mathcal{S}$, we can pick s so that $\sigma = \Pi s$. Therefore,

$$\sigma^\top (I-R) q^\alpha = s^\top \Pi^\top (I-R) q^\alpha \le \max_{\pi\in\mathcal{S}} \pi^\top (I-R) q^\alpha.$$

Then, by (22), it follows that $Y \ge 0$. Now, consider X, and note that $X = 0$ by (23)–(24) along with

$$X = \sum_{f\in\mathcal{F}} \left(\frac{n_f^\alpha(\rho_f - x_f)}{\rho_f^\alpha} + \delta_f q_{\iota(f)}^\alpha\right) = \sum_{f\in\mathcal{F}} \delta_f \left(q_{\iota(f)}^\alpha - \frac{n_f^\alpha}{\rho_f^\alpha}\right). \tag{25}$$

Thus, we have that $\dot{J}(0) \ge 0$. Since $J(\tau)$ is a convex function, this implies that $J(0) \le J(t)$, i.e., $L_\alpha(n, q) \le L_\alpha(n', q')$. Due to uniqueness of the optimal solution to $\mathsf{ALGD}(n, q)$, it follows that $(n', q') = (n, q)$. □

8.3 Fixed Points: Attractiveness

Now we establish the attractiveness of the space of fixed points. Specifically, we will show that starting from any initial state under fluid trajectory, the state reaches (arbitrarily close to) space of fixed points (in finite time).

Given $\varepsilon > 0$, define

$$\mathcal{J}_\varepsilon \triangleq \left\{(n, q) \in \mathbb{R}_+^{\mathcal{F}} \times \mathbb{R}_+^{\mathcal{E}} \;:\; \|(n, q) - \Delta(n, q)\|_1 < \varepsilon\right\}.$$

In other words, $\mathcal{J}_\varepsilon$ is the set of states (n, q) which are ε-approximate fixed points (in an ℓ_1-norm sense) of the lifting map. Given a fluid trajectory $\big(n(\cdot), q(\cdot)\big)$, define

$$h_\varepsilon\big(n(\cdot), q(\cdot)\big) \triangleq \inf\ \left\{t \ge 0 \;:\; \big(n(s), q(s)\big) \in \mathcal{J}_\varepsilon,\ \forall\, s \ge t\right\}.$$

In other words, $h_\varepsilon\big(n(\cdot), q(\cdot)\big)$ is the amount of time required for the trajectory $\big(n(\cdot), q(\cdot)\big)$ to reach and subsequently remain in the set $\mathcal{J}_\varepsilon$.

Theorem 10. *For any $\varepsilon > 0$, there exists $H_\varepsilon > 0$ so that if $\big(n(\cdot), q(\cdot)\big)$ is a fluid trajectory of the MWUM-α policy in a critically loaded system, with initial condition satisfying $\|(n(0), q(0))\|_\infty \le 1$, then*

$$h_\varepsilon\big(n(\cdot), q(\cdot)\big) \le H_\varepsilon.$$

Proof Sketch. Given $\delta > 0$,

$$\begin{aligned}\mathcal{D} &\triangleq \left\{(n, q) \in \mathbb{R}_+^{\mathcal{F}} \times \mathbb{R}_+^{\mathcal{E}} \;:\; L_\alpha(n, q) \le L_\alpha(\mathbf{1})\right\}, \\ \mathcal{I} &\triangleq \{(n, q) \in \mathcal{D} \;:\; (n, q) = \Delta(n, q)\}, \\ \mathcal{I}_\delta &\triangleq \{(n, q) \in \mathcal{D} \;:\; \|(n, q) - (n', q')\|_1 < \delta,\ (n', q') \in \mathcal{I}\}, \\ \mathcal{K}_\delta &\triangleq \{(n, q) \in \mathcal{D} \;:\; K(n, q) < K(n', q'),\ \forall\, (n', q') \in \mathcal{D} \setminus \mathcal{I}_\delta\}.\end{aligned}$$

where $K(n, q) \triangleq L_\alpha(n, q) - L_\alpha\big(\Delta(n, q)\big)$. The result can be established by showing that the following hold:

(i) $K\big(n(t), q(t)\big)$ is non-increasing in t.

(ii) For $\delta > 0$ sufficiently small, $\mathcal{I} \subset \mathcal{K}_\delta \subset \mathcal{I}_\delta \subset \mathcal{J}_\varepsilon$.

(iii) Starting from any initial condition in $\mathcal{D}$ (this includes all (n, q) with $\|(n, q)\|_\infty \le 1$), the time to hit $\mathcal{K}_\delta$ is bounded uniformly.

In particular, (iii) implies that starting from any state in $\mathcal{D}$, the fluid trajectory hits the set $\mathcal{K}_\delta$ in finite time. By (i), once the trajectory is in set $\mathcal{K}_\delta$, it remains in that set forever. By (ii), $\mathcal{K}_\delta \subset \mathcal{J}_\varepsilon$, and the result follows. We refer an interested reader to the longer version [16] for details. □

8.4 Fixed Points: Optimality

The following theorem characterizes the cost associated with an invariant state, relative to the effective cost. The effective cost represents the lowest cost achievable under *any* policy (cf. Theorem 5). Hence, this result implies that the invariant states of the MWUM-α policy are cost optimal, as $\alpha \to 0^+$.

Theorem 11. *Suppose (n^*, q^*) is an invariant state of a critically loaded system under the MWUM-α policy. Then,*

$$c(n^*, q^*) \le \big(1 + \phi(\alpha)\big) c^*(n^*, q^*), \tag{26}$$

where $\beta(\alpha) \to 0$ as $\alpha \to 0^+$.

Proof. Suppose (n^*, q^*) is an invariant state. Define (n', q') to be an optimal solution to the effective cost LP $c^*(n^*, q^*)$, defined by (15). Clearly

$$L_\alpha(n^*, q^*) \le L_\alpha(n, q),$$

since (n^*, q^*) is optimal for $\mathsf{ALGD}(n^*, q^*)$, and (n, q) is feasible for $\mathsf{ALGD}(n^*, q^*)$. Then, following the same argument as in Lemma 6,

$$c(n^*, q^*) \le \big(1 + \beta(\alpha)\big) c(n', q') = \big(1 + \beta(\alpha)\big) c^*(n^*, q^*),$$

where $\beta(\alpha) \to 0$ as $\alpha \to 0^+$. □

9. FUTURE DIRECTIONS

There are several interesting directions for future work. To start with, by characterizing the invariant manifold of the critically loaded fluid model and establishing its attractiveness, the work here should lead to the multiplicative state-space collapse property in a relatively straightforward manner following the method of Bramson [2]. As the next step, establishing the strong state-space collapse property would require bounding the the maximal deviation in the system state over certain time-horizon. We strongly believe that under MWUM-α control policy for $\alpha \ge 1$, this should follow from a recently developed Lyapunov function based maximal inequality by Shah, Tsitsiklis and Zhong [19]. However, further obtaining a complete characterization of the diffusion (heavy traffic) approximation seems to be far more non-trivial question. Finally, the results about path-wise constant factor optimality of critically loaded fluid model seem to suggest the possibility of such constant factor optimality of MWUM-α control policy under diffusion approximation.

10. REFERENCES

[1] T. Bonald and L. Massoulié. Impact of fairness on Internet performance. In *Proceedings of the 2001 ACM SIGMETRICS International Conference on Measurement and Modeling of Computer Systems*, pages 82–91, Cambridge, MA, June 2001.

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

[3] J. Dai and B. Prabhakar. The throughput of switches with and without speed-up. In *Proceedings of IEEE Infocom*, pages 556–564, 2000.

[4] J. G. Dai and W. Lin. Maximum pressure policies in stochastic processing networks. *Operations Research*, 53(2):197–218, March-April 2005.

[5] JG Dai. On positive Harris recurrence of multiclass queueing networks: a unified approach via fluid limit models. *The Annals of Applied Probability*, 5(1):49–77, 1995.

[6] G. de Veciana, T.-J. Lee, and T. Konstantopoulos. Stability and performance analysis of networks supporting elastic services. *IEEE/ACM Transactions on Networking*, 9(1):2–14, 2001.

[7] A. Eryilmaz and R. Srikant. Joint congestion control, routing and MAC for stability and fairness in wireless networks. *IEEE Journal on Selected Areas in Communications*, 24(8):1514–1524, August 2006.

[8] W. Kang, F. P. Kelly, N. H. Lee, and R. J. Williams. Fluid and brownian approximations for an internet congestion control model. In *Proceedings of the 43rd IEEE Conference on Decision and Control*, volume 4, pages 3938–3943, December 2004.

[9] F. P. Kelly, A. K. Maulloo, and D. K. H. Tan. Rate control for communication networks: shadow prices, proportional fairness and stability. *Journal of the Operational Research Society*, 49(3):237–252, March 1998.

[10] F. P. Kelly and R. J. Williams. Fluid model for a network operating under a fair bandwidth-sharing policy. *The Annals of Applied Probability*, 14(3):1055–1083, 2004.

[11] I. Keslassy and N. McKeown. Analysis of scheduling algorithms that provide 100% throughput in input-queued switches. In *Proceedings of Allerton Conference on Communication, Control, and Computing*, 2001.

[12] X. Lin, N. B. Shroff, and R. Srikant. On the connection-level stability of congestion-controlled communication networks. *IEEE Transactions on Information Theory*, 54(5):2317–2338, May 2008.

[13] L. Massoulié and J. W. Roberts. Bandwidth sharing and admission control for elastic traffic. *Telecommunication Systems*, 15(1–2):185–201, November 2000.

[14] N. McKeown, V. Anantharam, and J. Walrand. Achieving 100% throughput in an input-queued switch. In *Proceedings of IEEE Infocom*, pages 296–302, 1996.

[15] J. Mo and J. Walrand. Fair end-to-end window-based congestion control. *IEEE/ACM Transactions on Networking (ToN)*, 8(5):556–567, 2000.

[16] C. Moallemi and D. Shah. On the flow level dynamics of a packet-switched network. *Under submission, available at arxiv.org/abs/1003.0929*, 2010.

[17] M. J. Neely, E. Modiano, and C. Li. Fairness and optimal stochastic control for heterogeneous networks. *IEEE/ACM Transactions on Networking*, 16(2):396–409, April 2008.

[18] D. Shah. Analysis of scheduling algorithms that provide 100% throughput in input-queued switches. In *Proceedings of Allerton Conference on Communication, Control, and Computing*, 2001.

[19] D. Shah, J. N. Tsitsiklis, and Y. Zhong. Qualitative properties of alpha-weighted scheduling policies. In *ACM SIGMETRICS*, New York, USA, June 2010.

[20] D. Shah and D. J. Wischik. Optimal scheduling algorithm for input queued switch. In *Proceedings of IEEE INFOCOM*, 2006.

[21] D. Shah and D. J. Wischik. The teleology of scheduling algorithms for switched networks under light load, critical load, and overload. http://web.mit.edu/devavrat/www/shahwischik.pdf, 2007-09.

[22] S. Shakkottai and R. Srikant. *Network Optimization and Control.* Foundations and Trends in Networking. Now Publishers, Inc., Boston, January 2008.

[23] A. L. Stolyar. MaxWeight scheduling in a generalized switch: State space collapse and workload minimization in heavy traffic. *Annals of Applied Probability*, 14(1):1–53, 2004.

[24] A. L. Stolyar. Maximizing queueing network utility subject to stability: Greedy primal-dual algorithm. *Queueing Systems*, 50(4):401–457, 2005.

[25] L. Tassiulas and P. B. Bhattacharya. Allocation of interdependent resources for maximal throughput. *Stochastic Models*, 16:27–48, 2000.

[26] L. Tassiulas and A. Ephremides. Stability properties of constrained queueing systems and scheduling policies for maximum throughput in multihop radio networks. *IEEE Transactions on Automatic Control*, 37(12):1936–1948, December 1992.

[27] N. S. Walton. Proportional fairness and its relationship with multi-class queueing networks. To appear in *Annals of Applied Probability*, 2008.

[28] H. Ye and D. D. Yao. Heavy-traffic optimality of a stochastic network under utility-maximizing resource control. *Operations Research*, 56(2):453–470, March-April 2008.