Heavy-Traffic Delay Insensitivity in Connection-Level Models of Data Transfer with Proportionally Fair Bandwidth Sharing

Weina Wang
CSL
University of Illinois at Urbana-Champaign
Urbana, IL 61801
weinaw@illinois.edu

Siva Theja Maguluri
ISyE
Georgia Institute of Technology
Atlanta, GA 30332
siva.theja@gatech.edu

R. Srikant
CSL & ECE
University of Illinois at Urbana-Champaign
Urbana, IL 61801
rsrikant@illinois.edu

Lei Ying
ECEE
Arizona State University
Tempe, AZ 85287
lei.ying.2@asu.edu

ABSTRACT

Motivated by the stringent requirements on delay performance in data center networks, we study a connection-level model for bandwidth sharing among data transfer flows, where file sizes have phase-type distributions and proportionally fair bandwidth allocation is used. We analyze the expected number of files in steady-state by setting the steady-state drift of an appropriately chosen Lyapunov function equal to zero. We consider the heavy-traffic regime and obtain asymptotically tight bounds on the expected number of files in the system. Our results show that the expected number of files under proportionally fair bandwidth allocation is insensitive in heavy traffic to file size distributions, thus complementing the diffusion approximation result of Vlasiou et al. [20].

1. INTRODUCTION

We consider the following resource allocation problem that stems from the transfer of data in communication networks, illustrated in Figure 1. Data transfer requests arrive to a network, and the transfer of each data file, also referred to as a *flow*, is through a predetermined route that consists of a set of consecutive links connecting the source node and destination node. Each link in the network has a finite bandwidth capacity, allocated to the flows on the link by a bandwidth allocation policy. The bandwidth/rate a flow receives determines the speed at which its data can be transferred, thus determines the delay of the file transfer, namely the time from when the file arrives until the completion of the transfer. We are interested in analyzing the delay as a performance metric.

This model was first proposed by Massoulié and Roberts [16] as a connection-level model for data transfer in the Internet. It has been used to study congestion control schemes (e.g., TCP Vegas) and inform new protocol designs. Now with data centers being the backbone of the ubiquitous data technology, data transfer in data centers has attracted great attention, motivating research studies to provide better understanding on fundamental performance limits.

An important bandwidth allocation policy that has been considered is *proportionally fair policy*, introduced by Kelly [10]. This policy has been studied in the *heavy-traffic* regime, where the load of a system is close to the boundary of system capacity. Heavy-traffic analysis is an approach that has been widely adopted to study queueing systems. It provides approximations on the performance of a system, and these approximations are also often found useful for other traffic regimes. Heavy-traffic analysis also gives insights into policy design since it examines a policy in the critical scenario of heavy load.

For exponentially distributed file sizes and Poisson arrivals, delay analysis of proportionally fair policy in heavy traffic was first studied by Kang et al. [9] using diffusion approximation, where they need a *local traffic assumption*. This assumption requires that for each link, there is a flow that uses this link only. The local traffic assumption is inappropriate in a data center context since a data transfer flow is from a server to another server, but links connect servers to switches or switches to switches. To justify that the diffusion approximation in [9] is a valid approximation, Shah et al. [18] established the so-called *interchange of limits*, also under the local traffic assumption. Ye and Yao [26] removed the local traffic assumption and replaced it with a much weaker assumption that requires the routing matrix to have full row-rank. We refer to this assumption as the *full-rank assumption*. Ye and Yao later established the corresponding interchange of limits in [27].

With the above results for proportionally fair policy assuming exponentially distributed file sizes, a question that is of great importance to both practice and theoretical research is whether this policy is insensitive. A bandwidth allocation policy is said to be *insensitive* if its performance does not depend on file size distributions. Insensitivity is a highly desirable property since file size distributions in practice may not be exponential, and they may change over time with the evolvement of application scenarios. For proportionally fair

IFIP WG 7.3 Performance 2017. Nov. 14-16, 2017, New York, NY USA

policy, Paganini et al. [17] showed that the natural stability condition that was proved to be sufficient for exponentially distributed file sizes in [6] is still a sufficient condition for generally distributed file sizes, by considering a fluid approximation to the original stochastic system. For certain network topologies, the stationary distribution of the number of flows on different routes was also shown to be insensitive [16, 3, 4]. Notably, Vlasiou et al. [20] recently showed that the diffusion approximation for proportionally fair policy is insensitive. However, it remains an open problem to prove interchange of limits for this diffusion approximation.

In this paper, we avoid the interchange-of-limits issue by directly working with the stationary distribution of the connection-level model. In particular, we analyze the expected number of flows in steady state and show that it is insensitive in heavy traffic in the following sense. We consider a class of phase-type distributions that can approximate any file size distribution arbitrarily closely. We prove the following main result assuming file size distributions in this class, Poisson arrivals and the full-rank assumption. Let n_r be the number of flows on a route r in steady state, and then $\sum_r n_r$ is the total number of flows on all the routes, which is also referred to as the *backlog*. We show that the expected backlog is bounded as follows:

$$\mathbb{E}\left[\sum_r n_r\right] = \frac{L}{\epsilon} + o\left(\frac{1}{\epsilon}\right), \tag{1}$$

where L is the number of links in the network, and $\epsilon > 0$ is the heavy-traffic parameter, depending only on the *mean* file sizes and representing how far away the traffic load is from the boundary of the system capacity. Since the dominant term, L/ϵ, does not depend on the specific file size distributions except for their means, we say that the expected number of flows is *insensitive in heavy traffic*. This result complements the diffusion approximation result of Vlasiou et al. [20] since it justifies the validity of the backlog bound given by diffusion approximation in steady state.

We remark that this backlog bound in (1) scales linearly with the *number of links*, while static planning for bandwidth allocation would result in a backlog that scales linearly with the *number of routes*. This scaling behavior of proportionally fair policy is very appealing to data centers and the Internet, since the number of links is typically several orders of magnitude smaller than the number of routes.

Our analysis is under the drift-based framework developed by Eryilmaz and Srikant [7] and [13], where the basic idea is to obtain bounds on expected backlog by setting the steady-state drift of an appropriately chosen Lyapunov function equal to zero. A key step in this approach is to establish a *state-space collapse* result in the following sense. The system is represented by a Markov chain, whose state space is a multi-dimensional vector space. Consider the steady state of this Markov chain and a lower-dimensional subspace of the state space. We say that the state-space collapses to this lower-dimensional subspace if the moments of the distance between the steady state and the lower-dimensional subspace are upper bounded by constants as the heavy-traffic parameter ϵ goes to 0. This intuitively means that the steady state concentrates around the lower-dimensional subspace in heavy traffic, hence the term collapse. In [7] and the papers [14, 22, 25] that apply this approach to different settings, the state-space collapses

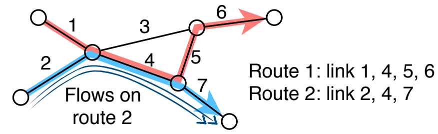

Figure 1: Bandwidth Sharing Network.

are to single-dimensional subspaces. Papers [13] and [15] generalized this approach to the case where the state space collapses to a multi-dimensional subspace, and resolved the open problem on the scaling behavior of backlog in a switch under the MaxWeight algorithm.

In this paper, our state-space collapse result is of a different type from the existing work above. Recall that the state-space collapse in the existing work above indicates that the distance between the steady state and a lower-dimensional subspace has constant moment upper bounds. In our state-space collapse result, these moment upper bounds are not constants: they grow to infinity as the heavy-traffic parameter ϵ goes to 0, but at a speed slower than the corresponding moments of the length of the state vector. Specifically, the m-th moment of this distance grows as $O\big((1/\sqrt{\epsilon})^m\big)$, while the m-th moment of the length of the state vector grows as $\Theta\big((1/\epsilon)^m\big)$. Therefore, the ratio between this distance and the length of the state vector still goes to 0. In this sense, the state-space collapse in this paper is of a *multiplicative* type, which has a similar flavor to the multiplicative state-space collapse in the diffusion approximation literature (see, e.g., [5, 24, 9]). We remark that a recent work [21] that studies switches with reconfiguration delay also deals with multiplicative type of state-space collapse, but does not give explicit moment upper bounds.

To establish the state-space collapse result and obtain backlog bounds, we construct an inner product that is different from the usual dot product in the state space, inspired by the Lyapunov function in [17]. This inner product rotates the space in a way such that the utilization of resources under proportionally fair policy is reflected by quantities with clear geometric meanings. This enables us to study the dynamics of geometric quantities such as the aforementioned distance between the state vector and a lower-dimensional subspace and the corresponding projection, which are needed in the drift-based approach. To show that the constructed inner product is well-defined and has desired properties, we make an interesting connection to the Popov-Belevitch-Hautus (PBH) test, which is a linear algebraic result well-known to control theorists.

2. SYSTEM MODEL

Basic Notation. Let $\mathbb{R}$ and $\mathbb{Z}_+$ denote the set of real numbers and positive integers, respectively. Let $[K]$ for a positive integer K denote the set $\{1, 2, \ldots, K\}$. We use $\mathbf{1}_{K\times 1}$ to denote an all-one vector with dimension $K \times 1$ for a positive integer K, and omit the subscript when it is clear from the context.

Bandwidth Sharing Network. We consider a network where nodes are connected by a set of links $\mathcal{L} = \{1, 2, \ldots, L\}$, illustrated in Figure 1. Data transfer requests arrive to the network, and the transfer of each data file, also referred to as a *flow*, is through a predetermined *route* that consists of a set of consecutive links connecting the source node and destination node. We consider a fixed set of routes $\mathcal{R} = \{1, 2, \ldots, R\}$. We write $\ell \in r$ if link ℓ is on route r. The relation between links and routes can be represented by the *routing matrix* $H = (h_{\ell r})_{\ell \in \mathcal{L}, r \in \mathcal{R}}$ with $h_{\ell r} = 1$ if $\ell \in r$, and $h_{\ell r} = 0$ otherwise. We assume that the routing matrix has full row-rank, referred to as the *full-rank assumption*.

The system is operated in continuous time. Each link ℓ in the network has a bandwidth capacity C_ℓ, allocated to the flows on the link by a *bandwidth allocation policy*. A bandwidth allocation policy specifies how much bandwidth/rate each flow receives according to the number of flows present on all the routes, subject to bandwidth capacity constraints. The rate a flow receives determines the speed at which the flow's data can be transferred. We are interested in the delay of a file transfer, namely the time from when the flow arrives until the completion of the transfer. Specifically, if we allocate a rate of $x(t)$ at time t to a flow that arrives at time A and has a file size F, then its delay D is given by the following equation:

$$\int_A^{A+D} x(t)dt = F, \tag{2}$$

i.e., the transfer completes when the accumulative rate equals to the file size. Therefore, the bandwidth allocation policy affects the delay by specifying the rates $x(t)$'s for the flows.

Proportionally Fair Policy. Let $N_r(t)$ be the total number of flows on route r at time t. The so-called proportionally fair policy allocates a rate of $x_r(t)$ to each flow on route r, where $(x_r(t))_{r\in\mathcal{R}}$ is the optimal solution of the following optimization problem with n_r equal to $N_r(t)$:

$$\max_{(x_1,\ldots,x_R)} \quad \sum_r n_r \log x_r \tag{3}$$

$$\text{subject to} \quad \sum_{r:\ell\in r} n_r x_r \leq C_\ell, \forall \ell, \tag{4}$$

$$x_r \geq 0, \forall r, \tag{5}$$

and $x_r(t) = 0$ when $n_r = 0$. The constraints in (4) are the bandwidth capacity constraints of the links, which indicates that the total rate allocated to the flows on the link should be within the link's capacity C_ℓ. Let p_ℓ denote the Lagrange multiplier for the capacity constraint of link ℓ. Then the rate allocation $(x_r(t))_{r\in\mathcal{R}}$ satisfies

$$x_r(t) = \begin{cases} \frac{1}{\sum_{l:l\in r} p_\ell} & \text{when } n_r > 0, \\ 0 & \text{otherwise.} \end{cases} \tag{6}$$

For simplicity, we will just write $x_r(t) = x_r$ in the remainder of this paper, but keep in mind that x_r implicitly depends on the flow counts at time t.

Arrivals and Service. Flows arrive at route r as a Poisson process with rate λ_r, and the arrival processes for different routes are independent. Let $\boldsymbol{\lambda} = [\lambda_1, \ldots, \lambda_R]^T$ denote the arrival rate vector. The file sizes of flows on route r are i.i.d. with a phase-type distribution given by the absorption time of a Markov chain specified as follows:

- The Markov chain has $K_r + 1$ states, where state 0 is an absorbing state and states $1, 2, \ldots, K_r$ are transient states (or *phases*).
- The initial distribution is $(\beta_0, \boldsymbol{\beta}_r)$ with $\beta_0 = 0$, where $\boldsymbol{\beta}_r$ is a $1 \times K_r$ vector.
- The transition rate matrix is

$$\begin{bmatrix} 0 & \mathbf{0} \\ \boldsymbol{s}_r & S_r \end{bmatrix}, \tag{7}$$

where $\boldsymbol{s}_r$ is a $K_r \times 1$ vector and S_r is a $K_r \times K_r$ matrix.

We further assume that this phase-type distribution belongs to a special class of phase-type distributions: (finite) mixtures of Erlang distributions [1] where these Erlangs have different rates. It can be proved that any probability distribution on $[0, \infty)$ can be approximated arbitrarily closely by a distribution in this class. With this assumption, each S_r is a block-diagonal matrix in the following form:

$$S_r = \begin{bmatrix} S_r^{(1)} & 0 & \cdots & 0 \\ 0 & S_r^{(2)} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & S_r^{(B_r)} \end{bmatrix}, \tag{8}$$

where each matrix $S_r^{(b)}$ with $1 \leq b \leq B_r$ has the following form:

$$S_r^{(b)} = \begin{bmatrix} -\mu_r^{(b)} & \mu_r^{(b)} & 0 & \cdots & 0 \\ 0 & -\mu_r^{(b)} & \mu_r^{(b)} & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & -\mu_r^{(b)} & \mu_r^{(b)} \\ 0 & \cdots & 0 & 0 & -\mu_r^{(b)} \end{bmatrix}. \tag{9}$$

The $\mu_r^{(b)}$'s are the rates of the Erlangs, so they are positive and distinct for different b's. The initial distribution $\boldsymbol{\beta}_r$ has the following structure: $\beta_{r,k} > 0$ for each phase k that corresponds to the first row of some $S_r^{(b)}$, and $\beta_{r,k} = 0$ for other phases.

For the above phase-type distribution with the parameter $(\boldsymbol{\beta}_r, S_r)$, the expected flow size is given by

$$\frac{1}{\overline{\mu}_r} = \boldsymbol{\beta}_r(-S_r)^{-1}\mathbf{1}_{K_r\times 1}. \tag{10}$$

We define the *load on route* r to be $\overline{\rho}_r = \lambda_r/\overline{\mu}_r$.

State Representation. With the above model for arrivals and service, a Markovian representation of the flow dynamics consists of the flow counts for every phase on every route. Let $N_{r,k}(t)$ denote the number of flows present on route r that are in phase k at time t. Let

$$\boldsymbol{N}_r(t) = [N_{r,1}(t), \ldots, N_{r,K_r}(t)]^T, \tag{11}$$

$$\boldsymbol{N}(t) = [(\boldsymbol{N}_1(t))^T, \ldots, (\boldsymbol{N}_R(t))^T]^T, \tag{12}$$

i.e., $\boldsymbol{N}_r(t)$ is a vector stacking together the $N_{r,k}(t)$'s, and $\boldsymbol{N}(t)$ is a vector concatenating the $\boldsymbol{N}_r(t)$'s. Note that the scaler $N_r(t)$, the total number of flows on route r at time t used in the proportionally fair policy, is given by $N_r(t) = \sum_{k\in[K_r]} N_{r,k}(t)$. Then the flow count vector $\boldsymbol{N}(t)$ is a vector in $\mathbb{R}^K$ with $K = \sum_r K_r$, and the flow count process $(\boldsymbol{N}(t) \colon t \geq 0)$ is a Markov chain.

Below we give the state transition rates for this Markov chain. Let $\boldsymbol{e}^{(r,k)} \in \mathbb{R}^K$ be a vector in the state space whose entry that corresponds to phase k of route r is equal to 1 and other entries are equal to 0. Then the transition rate $q_{\boldsymbol{n}\boldsymbol{n}'}$ from state $\boldsymbol{n}$ to state $\boldsymbol{n}' \neq \boldsymbol{n}$ is as follows:

$$q_{\boldsymbol{n}\boldsymbol{n}'} = \begin{cases} \lambda_r \beta_{r,k} & \text{if } \boldsymbol{n}' = \boldsymbol{n} + \boldsymbol{e}^{(r,k)}, \\ n_{r,k_1} x_r (S_r)_{k_1,k_2} & \text{if } \boldsymbol{n}' = \boldsymbol{n} - \boldsymbol{e}^{(r,k_1)} + \boldsymbol{e}^{(r,k_2)}, \\ n_{r,k} x_r \sum_{k'} (-S_r)_{k,k'} & \text{if } \boldsymbol{n}' = \boldsymbol{n} - \boldsymbol{e}^{(r,k)}, n_{r,k} > 0, \\ 0 & \text{otherwise.} \end{cases} \tag{13}$$

Heavy-Traffic Regime. We are interested in the stationary distribution of the flow count process in a heavy-traffic regime. Specifically, we consider a sequence of systems with the arrival rate vectors approaching the boundary of the capacity region. Let the systems be indexed by a nonnegative parameter ϵ, which represents how far away the arrival rate vector is from the boundary of the system capacity, with smaller ϵ being closer and $\epsilon = 0$ being on the boundary. We study the expected number of flows, i.e., *backlog*, in steady state for each system, and then look at how they scale in the heavy-traffic regime where ϵ is small.

For clarity, we append the superscript $^{(\epsilon)}$ to the quantities that depend on ϵ in the ϵ-th system. We say a quantity is a *constant* if it does not depend on either the system state or ϵ. Assume that the arrival rate vector is given by $\boldsymbol{\lambda}^{(\epsilon)} = (1-\epsilon)\boldsymbol{\lambda}^{(0)}$ for some $\boldsymbol{\lambda}^{(0)}$ on the boundary of the capacity region such that all the links are saturated. Recall that the load on route r is $\overline{\rho}_r^{(\epsilon)} = \lambda_r^{(\epsilon)} / \overline{\mu}_r$. Then

$$\sum_{r:\ell \in r} \overline{\rho}_r^{(0)} = C_\ell, \quad \text{for all } \ell. \tag{14}$$

We call $\sum_{r:\ell\in r} \overline{\rho}_r^{(\epsilon)}$ the *load on link* ℓ, which is equal to $(1-\epsilon)C_\ell$ in this heavy-traffic regime. Note that the results in this paper can be easily generalized to the heavy-traffic regime where only a subset of links are saturated, i.e., the regime where the equality (14) holds for only some ℓ's instead of all the ℓ's. But for ease of exposition, we only present the all-saturated regime.

An L-Dimensional Cone. We introduce an L-dimensional cone $\mathcal{K}$ in the state space $\mathbb{R}^K$, which is where the state space collapses to in heavy traffic. Note that $L \leq K$ due to the full-rank assumption. The cone $\mathcal{K}$ is finitely generated by a set of vectors $\{\boldsymbol{b}^{(\ell)}, \ell \in \mathcal{L}\}$, i.e.,

$$\mathcal{K} = \left\{ \boldsymbol{y} \in \mathbb{R}^K : \boldsymbol{y} = \sum_{\ell \in \mathcal{L}} \alpha_\ell \boldsymbol{b}^{(\ell)}, \alpha_\ell \geq 0 \text{ for all } \ell \in \mathcal{L} \right\}, \tag{15}$$

where the $\boldsymbol{b}^{(\ell)}$'s are defined below. For each route r, we define the load vector as follows:

$$\boldsymbol{\rho}_r^{(\epsilon)} = \lambda_r^{(\epsilon)} (-S_r)^{-T} \boldsymbol{\beta}_r^T. \tag{16}$$

The k-th entry of this vector, $\rho_{r,k}^{(\epsilon)}$, can be thought of as the load of phase k on route r, since the k-th entry of $(-S_r)^{-T}\boldsymbol{\beta}_r^T$ is the expected time a flow spends in phase k if given a unit of bandwidth. We can verify that the load on route r we have introduced, $\overline{\rho}_r^{(\epsilon)}$, is the sum of loads of phases on this route, i.e., $\overline{\rho}_r^{(\epsilon)} = \sum_{k\in[K_r]} \rho_{r,k}^{(\epsilon)}$. Let $\boldsymbol{\rho}^{(\epsilon)}$ be a vector concatenating the $\boldsymbol{\rho}_r^{(\epsilon)}$'s, i.e.,

$$\boldsymbol{\rho}^{(\epsilon)} = [(\boldsymbol{\rho}_1^{(\epsilon)})^T, \dots, (\boldsymbol{\rho}_R^{(\epsilon)})^T]^T. \tag{17}$$

Now we construct a $K \times 1$ vector $\boldsymbol{b}^{(\ell)}$ from $\boldsymbol{\rho}^{(0)}$ for each link ℓ: we index the entries of $\boldsymbol{b}^{(\ell)}$ using the route and phase (r,k), and let $b_{r,k}^{(\ell)} = \rho_{r,k}^{(0)} \mathbb{1}_{\{\ell \in r\}}$, where $\mathbb{1}_{\{\ell\in r\}}$ is equal to 1 when route r uses link ℓ and equal to 0 otherwise. That is, we keep the entries of $\boldsymbol{\rho}^{(0)}$ that correspond to phases of the routes that use link ℓ, and set other entries to zero. We give a concrete example below to explain this structure of $\boldsymbol{b}^{(\ell)}$'s.

EXAMPLE 1. *Consider the network illustrated in Figure 2.*

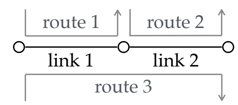

Figure 2: Example with two links and three routes.

The routing matrix is

$$H = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix}. \tag{18}$$

Suppose the service time distributions of routes 1, 2 *and* 3 *have* $K_1 = 2, K_2 = 1$ *and* $K_3 = 3$ *phases, respectively. Then*

$$\boldsymbol{\rho}^{(0)} = \begin{bmatrix} \rho_{1,1}^{(0)} & \rho_{1,2}^{(0)} & \rho_{2,1}^{(0)} & \rho_{3,1}^{(0)} & \rho_{3,2}^{(0)} & \rho_{3,3}^{(0)} \end{bmatrix}, \tag{19}$$

and

$$\boldsymbol{b}^{(1)} = \begin{bmatrix} \rho_{1,1}^{(0)} & \rho_{1,2}^{(0)} & 0 & \rho_{3,1}^{(0)} & \rho_{3,2}^{(0)} & \rho_{3,3}^{(0)} \end{bmatrix}^T, \tag{20}$$

$$\boldsymbol{b}^{(2)} = \begin{bmatrix} 0 & 0 & \rho_{2,1}^{(0)} & \rho_{3,1}^{(0)} & \rho_{3,2}^{(0)} & \rho_{3,3}^{(0)} \end{bmatrix}^T. \tag{21}$$

3. MAIN RESULTS

In this section, we present the two main results of this paper: state-space collapse and an asymptotically tight bound on backlog. The proofs are given in Sections 5 and 6, respectively.

3.1 State-Space Collapse

Our first main result is state-space collapse, which intuitively means that the steady state of the flow count process concentrates around a lower-dimensional subspace of the state space in heavy traffic. Specifically, the state space of the flow count process $(\boldsymbol{N}(t) : t \geq 0)$ is the K-dimensional space $\mathbb{R}^K$ with $K = \sum_r K_r$. We consider the L-dimensional cone $\mathcal{K}$ defined in (15), which lies in a lower dimensional subspace since $L \leq K$. Let $\overline{\boldsymbol{N}}^{(\epsilon)}$ denote a random vector whose distribution is the stationary distribution of the flow count process $(\boldsymbol{N}^{(\epsilon)}(t) : t \geq 0)$. We decompose $\overline{\boldsymbol{N}}^{(\epsilon)}$ into its projection onto the cone, referred to as the *parallel component* and denoted by $\overline{\boldsymbol{N}}_{\shortparallel}^{(\epsilon)}$, and the remainder, referred to as the *perpendicular component* and denoted by $\overline{\boldsymbol{N}}_{\perp}^{(\epsilon)}$ since it is perpendicular to the parallel term. Then

$$\overline{\boldsymbol{N}}^{(\epsilon)} = \overline{\boldsymbol{N}}_{\shortparallel}^{(\epsilon)} + \overline{\boldsymbol{N}}_{\perp}^{(\epsilon)}, \tag{22}$$

where $\overline{\boldsymbol{N}}_{\shortparallel}^{(\epsilon)} \in \mathcal{K}$ and $\|\overline{\boldsymbol{N}}_{\perp}^{(\epsilon)}\|$ is the distance between $\overline{\boldsymbol{N}}^{(\epsilon)}$ and $\mathcal{K}$. The state-space collapse indicates that as the arrival rate vector approaches the boundary of the capacity region, the perpendicular component $\overline{\boldsymbol{N}}_{\perp}^{(\epsilon)}$ becomes negligible compared to the parallel component $\overline{\boldsymbol{N}}_{\shortparallel}^{(\epsilon)}$. We formally state this result in terms of moments in the following theorem.

THEOREM 1. *Consider a sequence of bandwidth sharing networks under the proportionally fair policy, indexed by a parameter ϵ with $0 < \epsilon < 1$. The load on each link ℓ is $(1-\epsilon)C_\ell$, where C_ℓ is the bandwidth capacity of link ℓ. Let $\overline{\boldsymbol{N}}^{(\epsilon)}$ denote a random vector whose distribution is the stationary distribution of the flow count process $(\boldsymbol{N}^{(\epsilon)}(t) \colon t \geq 0)$. Then the m-th moment of $\|\overline{\boldsymbol{N}}_{\perp}^{(\epsilon)}\|$ can be bounded as follows:*

$$\mathbb{E}\Big[\Big\|\overline{\boldsymbol{N}}_{\perp}^{(\epsilon)}\Big\|^m\Big] = O\left(\left(\frac{1}{\sqrt{\epsilon}}\right)^m\right), \text{ for all } m \in \mathbb{Z}_+. \tag{23}$$

We remark that $\mathbb{E}[\|\overline{\boldsymbol{N}}_{\perp}^{(\epsilon)}\|]/\mathbb{E}[\|\overline{\boldsymbol{N}}_{\shortparallel}^{(\epsilon)}\|] \to 0$ as $\epsilon \to 0^+$ since it can be proved that $\mathbb{E}[\|\overline{\boldsymbol{N}}_{\shortparallel}^{(\epsilon)}\|] = \Theta(1/\epsilon)$. Therefore, our state-space collapse is of a *multiplicative* type. We obtain moment bounds based on Lyapunov drift using an approach similar to that in [2]. However, approaches such as [8, 2] require the drift to be negative whenever the value of the Lyapunov function is large enough. But for this system and the Lyapunov function $V(\boldsymbol{n}) = \|\boldsymbol{n}_\perp\|$, where $\boldsymbol{n}$ is a state of the flow count process, large $\|\boldsymbol{n}_\perp\|$ alone may not be enough to give a negative drift. We prove that the drift is negative under the additional condition that the ratio $\|\boldsymbol{n}_\perp\|/\|\boldsymbol{n}\|$ is also large enough, which leads to the multiplicative type of state-space collapse.

Note that we have not specified the inner product and the corresponding norm for the projection and the state-space collapse result. In fact, if we can obtain moment bounds for *a* norm, then we can actually have moment bounds in the same orders for *any* norm, since all the norms are equivalent in $\mathbb{R}^K$. However, we will see that the choice of inner product is crucial in obtaining proper drift bounds, and the inner product we choose is different from the usual dot product in $\mathbb{R}^K$. We defer the definition of the specific inner product we choose to Section 4.

3.2 Backlog Bound

Based on the state-space collapse result, we establish the following bound on the backlog, which is asymptotically tight in the heavy-traffic regime where ϵ becomes small. This bound is said to be insensitive in heavy traffic since the dominant term, L/ϵ, does not depend on the specific file size distributions except for their means.

THEOREM 2. *Consider a sequence of bandwidth sharing networks under the proportionally fair policy, indexed by a parameter ϵ with $0 < \epsilon < 1$. The load on each link ℓ is $(1-\epsilon)C_\ell$, where C_ℓ is the bandwidth capacity of link ℓ. Let $\overline{\boldsymbol{N}}^{(\epsilon)}$ denote a random vector whose distribution is the stationary distribution of the flow count process $(\boldsymbol{N}^{(\epsilon)}(t) \colon t \geq 0)$. Then*

$$\mathbb{E}\Big[\sum_{r,k} \overline{N}_{r,k}^{(\epsilon)}\Big] = \frac{L}{\epsilon} + o\left(\frac{1}{\epsilon}\right), \tag{24}$$

where L is the number of links in the network.

4. INNER PRODUCT

In this section, we present the inner product and its induced norm used throughout this paper. We prove that the constructed inner product satisfies two conditions that are essential for the proofs of the main results. The meaning of these two conditions will become clearer when we reach those proofs.

For the space $\mathbb{R}^K$ where the states of the flow count process lie in, we consider the following weighted inner product defined by a block-diagonal matrix M:

$$\langle \boldsymbol{y}, \boldsymbol{z} \rangle = \boldsymbol{y}^T M \boldsymbol{z}, \quad \boldsymbol{y}, \boldsymbol{z} \in \mathbb{R}^K, \tag{25}$$

where

$$M = \begin{bmatrix} M_1 & 0 & \cdots & 0 \\ 0 & M_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & M_R \end{bmatrix}, \tag{26}$$

and each M_r is a $K_r \times K_r$ matrix defined as follows:

$$M_r = \frac{1}{\lambda_r^{(0)}} \int_0^{+\infty} \frac{\exp(S_r\sigma)\mathbf{1}_{K_r\times 1}\mathbf{1}_{K_r\times 1}^T \exp(S_r^T\sigma)}{\boldsymbol{\beta}_r(-S_r)^{-1}\exp(S_r\sigma)\mathbf{1}_{K_r\times 1}} d\sigma. \tag{27}$$

Then the induced norm is defined as:

$$\|\boldsymbol{y}\| = \sqrt{\langle \boldsymbol{y}, \boldsymbol{y} \rangle} = \sqrt{\boldsymbol{y}^T M \boldsymbol{y}}, \quad \boldsymbol{y} \in \mathbb{R}^K. \tag{28}$$

We remark that the Lyapunov function $\|\boldsymbol{n}\|^2$ under this norm, where $\boldsymbol{n}$ is a state, is equivalent to the Lyapunov function in [17] with a parameter K there chosen to be 1. But this choice for K is not allowed in [17]. To study weighted delay in switches, [11] also considers a weighted norm, where the matrix that defines it is a diagonal matrix.

Below we justify the validity of the constructed inner product in Lemma 1 and give two properties of M_r in Lemmas 2 and 3, which will be used later to show that the inner product satisfies two desired conditions. The proofs of these lemmas are given in Appendix A.

LEMMA 1. *The inner product defined by the matrix M is well-defined, i.e., the matrix M is well-defined and positive definite.*

LEMMA 2. *For each route r,*

$$(\boldsymbol{\rho}_r^{(0)})^T M_r(-S_r^T) = \mathbf{1}_{K_r\times 1}^T. \tag{29}$$

LEMMA 3. *For each route r, there exists a constant $\kappa_r > 0$, such that the matrix $\frac{1}{2}M_r(-S_r^T) + \frac{1}{2}(-S_r)M_r - \kappa_r M_r$ is positive semi-definite.*

Next we identify the conditions on the inner product that are needed in the proofs of the main results. Recall that $\{\boldsymbol{b}^{(\ell)}, l \in \mathcal{L}\}$ are constructed from the load vector $\boldsymbol{\rho}^{(0)}$ such that $b_{r,k}^{(\ell)} = \rho_{r,k}^{(0)}\mathbb{1}_{\{l\in r\}}$. Similarly, we construct a set of vectors $\{\widehat{\boldsymbol{b}}^{(\ell)}, l \in \mathcal{L}\}$ from the current rate allocation as follows. Recall that a state $\boldsymbol{n}$ is a $K\times 1$ vector that has the form $\boldsymbol{n} = [\boldsymbol{n}_1^T, \ldots, \boldsymbol{n}_R^T]^T$ with $\boldsymbol{n}_r = [n_{r,1}, \ldots, n_{r,K_r}]^T$, and x_r is the bandwidth allocated to each flow on route r based on $\boldsymbol{n}$ by proportionally fair sharing. Let $\boldsymbol{nx} = [\boldsymbol{n}_1^T x_1, \ldots, \boldsymbol{n}_R^T x_R]^T$, whose (r,k)-th entry, $n_{r,k}x_r$, is the total bandwidth allocated to the flows in phase k on route r. Then $\widehat{\boldsymbol{b}}^{(\ell)}$ is defined by $\widehat{b}_{r,k}^{(\ell)} = n_{r,k}x_r\mathbb{1}_{\{l\in r\}}$. We claim that the constructed inner product satisfies the following two conditions, where the norm is the induced norm:

(C1) For each link ℓ,

$$\langle \boldsymbol{b}^{(\ell)}, (-S^T)(\boldsymbol{\rho}^{(0)} - \boldsymbol{nx}) \rangle = U_\ell,$$

where U_ℓ is the *unused bandwidth* on link ℓ, i.e., the amount of bandwidth that is not allocated to any flow.

(C2) For each link ℓ,

$$\langle \boldsymbol{b}^{(\ell)} - \widehat{\boldsymbol{b}}^{(\ell)}, (-S^T)(\boldsymbol{\rho}^{(0)} - \boldsymbol{n}\boldsymbol{x})\rangle \geq \kappa_{\min}\|\boldsymbol{b}^{(\ell)} - \widehat{\boldsymbol{b}}^{(\ell)}\|^2$$

for a positive constant $\kappa_{\min}$, where

$$S = \begin{bmatrix} S_1 & 0 & \cdots & 0 \\ 0 & S_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & S_R \end{bmatrix}.$$

We remark that both conditions are concerned with the difference between the load vector $\boldsymbol{\rho}^{(0)}$ and the bandwidth allocation vector $\boldsymbol{n}\boldsymbol{x}$, rotated and scaled by $-S^T$. Condition (C1) requires the projection of this altered difference $(-S^T)(\boldsymbol{\rho}^{(0)} - \boldsymbol{n}\boldsymbol{x})$ onto the vector $\boldsymbol{b}^{(\ell)}$ to be the unused bandwidth. For condition (C2), observe that under the regular dot product of Euclidean space, $\langle \boldsymbol{b}^{(\ell)} - \widehat{\boldsymbol{b}}^{(\ell)}, \boldsymbol{\rho}^{(0)} - \boldsymbol{n}\boldsymbol{x}\rangle_{\text{Euclidean}} = \|\boldsymbol{b}^{(\ell)} - \widehat{\boldsymbol{b}}^{(\ell)}\|^2_{\text{Euclidean}}$. Condition (C2) requires that this Euclidean inner product is not diminished by the matrix $(-S^T)$ under the constructed inner product.

Proof of Conditions (C1), (C2). We first prove (C1):

$$\begin{aligned}
&\langle \boldsymbol{b}^{(\ell)}, (-S)^T(\boldsymbol{\rho}^{(0)} - \boldsymbol{n}\boldsymbol{x})\rangle \\
&= (\boldsymbol{b}^{(\ell)})^T M(-S)^T(\boldsymbol{\rho}^{(0)} - \boldsymbol{n}\boldsymbol{x}) \qquad (30)\\
&= \sum_{r:\ell\in r} (\boldsymbol{\rho}_r^{(0)})^T M_r(-S_r)^T(\boldsymbol{\rho}_r^{(0)} - \boldsymbol{n}_r x_r) \qquad (31)\\
&= \sum_{r:\ell\in r} \mathbf{1}^T(\boldsymbol{\rho}_r^{(0)} - \boldsymbol{n}_r x_r) \qquad (32)\\
&= C_\ell - \sum_{r:\ell\in r}\sum_k n_{r,k}x_r \qquad (33)\\
&= U_\ell, \qquad (34)
\end{aligned}$$

where (32) follows from Lemma 2, and (33) follows from the heavy-traffic condition in (14).

Next we prove condition (C2). The inner product can be written in the following form:

$$\begin{aligned}
&\langle \boldsymbol{b}^{(\ell)} - \widehat{\boldsymbol{b}}^{(\ell)}, (-S^T)(\boldsymbol{\rho}^{(0)} - \boldsymbol{n}\boldsymbol{x})\rangle \\
&= (\boldsymbol{b}^{(\ell)} - \widehat{\boldsymbol{b}}^{(\ell)})^T M(-S^T)(\boldsymbol{\rho}^{(0)} - \boldsymbol{n}\boldsymbol{x}) \\
&= \sum_{r:\ell\in r} (\boldsymbol{\rho}_r^{(0)} - \boldsymbol{n}_r x_r)^T M_r(-S_r^T)(\boldsymbol{\rho}_r^{(0)} - \boldsymbol{n}_r x_r) \\
&= \frac{1}{2}\sum_{r:\ell\in r} (\boldsymbol{\rho}_r^{(0)} - \boldsymbol{n}_r x_r)^T (M_r(-S_r^T) + (-S_r)M_r)(\boldsymbol{\rho}_r^{(0)} - \boldsymbol{n}_r x_r).
\end{aligned}$$

Then by Lemma 3,

$$\begin{aligned}
&\frac{1}{2}\sum_{r:\ell\in r} (\boldsymbol{\rho}_r^{(0)} - \boldsymbol{n}_r x_r)^T (M_r(-S_r^T) + (-S_r)M_r)(\boldsymbol{\rho}_r^{(0)} - \boldsymbol{n}_r x_r) \\
&\geq \sum_{r:\ell\in r} \kappa_r(\boldsymbol{\rho}_r^{(0)} - \boldsymbol{n}_r x_r)^T M_r(\boldsymbol{\rho}_r^{(0)} - \boldsymbol{n}_r x_r) \\
&\geq \kappa_{\min}\|\widehat{\boldsymbol{b}}^{(\ell)} - \boldsymbol{b}^{(\ell)}\|^2,
\end{aligned}$$

where $\kappa_{\min} = \min_r\{\kappa_r\} > 0$. Therefore,

$$\langle \boldsymbol{b}^{(\ell)} - \widehat{\boldsymbol{b}}^{(\ell)}, (-S^T)(\boldsymbol{\rho}^{(0)} - \boldsymbol{n}\boldsymbol{x})\rangle \geq \kappa_{\min}\|\widehat{\boldsymbol{b}}^{(\ell)} - \boldsymbol{b}^{(\ell)}\|^2,$$

which completes the proof. □

5. STATE-SPACE COLLAPSE

In this section, we prove the state-space collapse result in Theorem 1. We divide the proof into three steps: (i) We establish a bound on the drift of the Lyapunov function $V(\boldsymbol{n}) = \|\boldsymbol{n}_\perp\|$, where $\boldsymbol{n}$ is a state of the flow count process $(\boldsymbol{N}^{(\epsilon)}(t) \colon t \geq 0)$; (ii) We bound the distribution tail of $\|\overline{\boldsymbol{N}}_\perp\|$ based on the drift bound using an approach similar to that in Bertsimas et al. [2]; (iii) We obtain the moment bounds using the tail bound.

(i) Drift Bound. For any Lyapunov function $V(\boldsymbol{n})$, the drift of V at a state $\boldsymbol{n}$ in the ϵ-th system is defined as

$$\Delta V(\boldsymbol{n}) = \sum_{\boldsymbol{n}':\boldsymbol{n}'\neq\boldsymbol{n}} q_{\boldsymbol{n}\boldsymbol{n}'}(V(\boldsymbol{n}') - V(\boldsymbol{n})). \qquad (35)$$

where $q_{\boldsymbol{n}\boldsymbol{n}'}$ is the transition rate from state $\boldsymbol{n}$ to $\boldsymbol{n}'$ of the flow count process $(\boldsymbol{N}^{(\epsilon)}(t) \colon t \geq 0)$. We establish the following drift bound for $V(\boldsymbol{n}) = \|\boldsymbol{n}_\perp\|$, the proof of which is given in Section 5.1.

Lemma 4. *In the ϵ-th system, the drift of the Lyapunov function $\|\boldsymbol{n}_\perp\|$ satisfies that*

$$\Delta\|\boldsymbol{n}_\perp\| \leq -\sqrt{\epsilon} \qquad (36)$$

when

$$\epsilon \leq \epsilon_{\max},\ \|\boldsymbol{n}_\perp\| \geq \frac{A_1}{2\xi_1\sqrt{\epsilon}},\ \frac{\|\boldsymbol{n}_\perp\|}{\sum_{r,k} n_{r,k}} \geq \frac{\xi_2\sqrt{\epsilon}}{\kappa_{\min}C_{\min}}, \qquad (37)$$

where $\epsilon_{\max}, A_1, \xi_1, \xi_2, \kappa_{\min}$ and $C_{\min}$ are positive constants.

We remark that the last condition on $\|\boldsymbol{n}_\perp\|/\sum_{r,k} n_{r,k}$ in (37) is equivalent to that $\|\boldsymbol{n}_\perp\|/\|\boldsymbol{n}\|$ is large enough, since all norms are equivalent in $\mathbb{R}^K$ and thus there exist positive constants a_1 and a_2 such that $a_1\|\boldsymbol{n}\| \leq \sum_{r,k} n_{r,k} \leq a_2\|\boldsymbol{n}\|$.

(ii) Tail Bound. Next we bound the distribution tail of $\|\overline{\boldsymbol{N}}_\perp\|$ based on the drift bound in Lemma 4. Bertsimas et al. [2] gave an exponential type upper bound on the distribution tail of a Lyapunov function when the Lyapunov function has a negative drift for large enough value of the Lyapunov function. However, their results do not directly apply here since the drift bound in Lemma 4 has an additional requirement on $\|\boldsymbol{n}_\perp\|/\sum_{r,k} n_{r,k}$. We use a similar approach and show the following tail bound, which has an additional term besides the exponential term. The proof is given in Section 5.2.

Lemma 5. *For any nonnegative $\epsilon \leq \epsilon_{\max}$, the tail distribution of $\|\overline{\boldsymbol{N}}_\perp\|$ is bounded by an exponential term plus an additional term as follows: for any nonnegative integer j,*

$$\mathbb{P}\left(\|\overline{\boldsymbol{N}}_\perp\| > \frac{A_1}{2\xi_1\sqrt{\epsilon}} + 2\nu_1 j\right) \leq \alpha^{j+1} + \xi_2(1-\alpha)\sum_{i=0}^{j}\alpha^i\left(\beta^{\theta/\sqrt{\epsilon}}\right)^{j-i}, \qquad (38)$$

where A_1, ξ_1, ξ_2 are the constants in Lemma 4, ν_1 and θ are positive constants, and

$$\alpha = \frac{a}{a+\sqrt{\epsilon}}, \quad \beta = \frac{b}{b+\epsilon}, \qquad (39)$$

for positive constants a and b.

(iii) Moment bounds. The tail bound in Lemma 5 is enough to give the $O\big((1/\sqrt{\epsilon})^m\big)$ bound on the m-th moment of $\|\overline{\boldsymbol{N}}_\perp\|$ in Theorem 1. The derivation of the moment bounds based on the tail bound is much intuitive and is similar to [13], so the proof is given in our technical report [23] due to space limit.

5.1 Proof of Lemma 4 (Drift Bound)

Proof. Recall that a state $\boldsymbol{n}$ is a $K \times 1$ vector that has the form $\boldsymbol{n} = [\boldsymbol{n}_1^T, \dots, \boldsymbol{n}_R^T]^T$ with $\boldsymbol{n}_r = [n_{r,1}, \dots, n_{r,K_r}]^T$, and x_r is the bandwidth allocated to each flow on route r based on $\boldsymbol{n}$ by proportionally fair sharing. Also recall that $\boldsymbol{nx}$ denotes a vector whose (r,k)-th entry is $n_{r,k}x_r$, which is the total bandwidth allocated to the flows in phase k on route r. We fix an $\epsilon > 0$ and omit the superscript $^{(\epsilon)}$ for conciseness.

We prove Lemma 4 by combining the following claim, which holds for any inner product and its induced norm, with the conditions (C1) and (C2) that are satisfied by the inner product we choose in Section 4. The proof of the following claim is given at the end of this proof.

Claim 1.

$$\Delta\|\boldsymbol{n}_\perp\| \le \frac{1}{\|\boldsymbol{n}_\perp\|}\langle \boldsymbol{n} - \boldsymbol{n}_{\shortparallel}, (-S^T)(\boldsymbol{\rho}^{(0)} - \boldsymbol{nx})\rangle + \epsilon\|(-S^T)\boldsymbol{\rho}^{(0)}\| + \frac{A_1}{2\|\boldsymbol{n}_\perp\|},$$

where A_1 is a constant.

Next we analyze the terms in Claim 1, utilizing conditions (C1) and (C2). We first consider the term $\langle \boldsymbol{n}, (-S^T)(\boldsymbol{\rho}^{(0)} - \boldsymbol{nx})\rangle$. By the proportionally fair sharing, $(x_r)_{r\in\mathcal{R}}$ satisfies

$$x_r = \begin{cases} \frac{1}{\sum_{l:l\in r} p_\ell} & \text{when } \sum_{k\in[K_r]} n_{r,k} > 0, \\ 0 & \text{otherwise,} \end{cases}$$

where p_ℓ is the Lagrange multiplier of the capacity constraint of link ℓ. Then $n_{r,k}$ can be written as

$$n_{r,k} = n_{r,k}x_r \sum_{l:l\in r} p_\ell,$$

i.e.,

$$\boldsymbol{n} = \sum_{l:l\in r} p_\ell \widehat{\boldsymbol{b}}^{(\ell)}.$$

Note that by condition (C1) and complementary slackness,

$$p_\ell\langle \boldsymbol{b}^{(\ell)}, (-S^T)(\boldsymbol{\rho}^{(0)} - \boldsymbol{nx})\rangle = p_\ell U_\ell = 0.$$

Thus

$$\begin{aligned}&\langle \boldsymbol{n}, (-S^T)(\boldsymbol{\rho}^{(0)} - \boldsymbol{nx})\rangle \\ &= \sum_\ell p_\ell \langle \widehat{\boldsymbol{b}}^{(\ell)} - \boldsymbol{b}^{(\ell)}, (-S^T)(\boldsymbol{\rho}^{(0)} - \boldsymbol{nx})\rangle \\ &\overset{(a)}{\le} -\kappa_{\min} \sum_\ell p_\ell \|\widehat{\boldsymbol{b}}^{(\ell)} - \boldsymbol{b}^{(\ell)}\|^2,\end{aligned}$$

where (a) follows from (C2). Note that

$$\begin{aligned}\|\boldsymbol{n}_\perp\|^2 &\overset{(a)}{\le} \Big\|\boldsymbol{n} - \sum_\ell p_\ell \boldsymbol{b}^{(\ell)}\Big\|^2 \\ &= \Big\|\sum_\ell p_\ell\Big(\widehat{\boldsymbol{b}}^{(\ell)} - \boldsymbol{b}^{(\ell)}\Big)\Big\|^2 \\ &\le \Big(\sum_\ell p_\ell \|\widehat{\boldsymbol{b}}^{(\ell)} - \boldsymbol{b}^{(\ell)}\|\Big)^2 \\ &\overset{(b)}{\le} \Big(\sum_\ell p_\ell\Big)\Big(\sum_\ell p_\ell \|\widehat{\boldsymbol{b}}^{(\ell)} - \boldsymbol{b}^{(\ell)}\|^2\Big) \\ &\overset{(c)}{\le} \frac{1}{C_{\min}}\Big(\sum_{r,k} n_{r,k}\Big)\Big(\sum_\ell p_\ell \|\widehat{\boldsymbol{b}}^{(\ell)} - \boldsymbol{b}^{(\ell)}\|^2\Big),\end{aligned}$$

where (a) follows from the definition of projection, (b) follows from Cauchy-Schwarz inequality, and (c) is due to the equality $\sum_\ell p_\ell C_\ell = \sum_{r,k} n_{r,k}$ derived from the proportionally fair sharing policy and $C_{\min} = \min_\ell C_\ell$. Then there holds

$$\langle \boldsymbol{n}, (-S^T)(\boldsymbol{\rho}^{(0)} - \boldsymbol{nx})\rangle \le -\kappa_{\min} C_{\min} \frac{\|\boldsymbol{n}_\perp\|^2}{\sum_{r,k} n_{r,k}}. \qquad (40)$$

We then consider the term $\langle \boldsymbol{n}_{\shortparallel}, (-S^T)(\boldsymbol{\rho}^{(0)} - \boldsymbol{nx})\rangle$. Since $\boldsymbol{n}_{\shortparallel} \in \mathcal{K}$, we can represent it by

$$\boldsymbol{n}_{\shortparallel} = \sum_\ell \alpha_\ell \boldsymbol{b}^{(\ell)}, \quad \alpha_\ell \ge 0 \text{ for all } \ell.$$

Then

$$\begin{aligned}\langle \boldsymbol{n}_{\shortparallel}, (-S^T)(\boldsymbol{\rho}^{(0)} - \boldsymbol{nx})\rangle &= \sum_\ell \alpha_\ell \langle \boldsymbol{b}^{(\ell)}, (-S^T)(\boldsymbol{\rho}^{(0)} - \boldsymbol{nx})\rangle \\ &\overset{(a)}{=} \sum_\ell \alpha_\ell U_\ell \\ &\ge 0, \qquad (41)\end{aligned}$$

where (a) follows from condition (C1).

Combining (40) and (41) yields

$$\Delta\|\boldsymbol{n}_\perp\| \le -\kappa_{\min} C_{\min} \frac{\|\boldsymbol{n}_\perp\|}{\sum_{r,k} n_{r,k}} + \epsilon\|(-S^T)\boldsymbol{\rho}^{(0)}\| + \frac{A_1}{2\|\boldsymbol{n}_\perp\|}.$$

We choose constants $\xi_1 > 0, \xi_2 > 0$ such that

$$\xi_2 - \xi_1 = 2. \qquad (42)$$

Then when

$$\epsilon \le \epsilon_{\max} \triangleq \frac{1}{\|(-S^T)\boldsymbol{\rho}^{(0)}\|^2}, \quad \|\boldsymbol{n}_\perp\| \ge \frac{A_1}{2\xi_1\sqrt{\epsilon}},$$
$$\frac{\|\boldsymbol{n}_\perp\|}{\sum_{r,k} n_{r,k}} \ge \frac{\xi_2\sqrt{\epsilon}}{\kappa_{\min} C_{\min}},$$

we have $\Delta\|\boldsymbol{n}_\perp\| \le -\xi_2\sqrt{\epsilon} + \sqrt{\epsilon} + \xi_1\sqrt{\epsilon} = -\sqrt{\epsilon}$, which is the drift bound in Lemma 4.

Lastly, we prove the Claim 1 at the beginning of this proof. We first bound $\Delta\|\boldsymbol{n}_\perp\|$ in the following form

$$\begin{aligned}\Delta\|\boldsymbol{n}_\perp\| &\le \frac{1}{2\|\boldsymbol{n}_\perp\|}\Delta\|\boldsymbol{n}_\perp\|^2 \\ &= \frac{1}{2\|\boldsymbol{n}_\perp\|}(\Delta\|\boldsymbol{n}\|^2 - \Delta\|\boldsymbol{n}_{\shortparallel}\|^2),\end{aligned}$$

where the inequality follows from the fact that $\|\boldsymbol{n}_\perp\| = \sqrt{\|\boldsymbol{n}_\perp\|^2}$ and the square-root function is concave. The drifts $\Delta\|\boldsymbol{n}\|^2$ and $\Delta\|\boldsymbol{n}_{\shortparallel}\|^2$ can be bounded as follows, where we omit the superscript $^{(\epsilon)}$ for conciseness.

$$\begin{aligned}&\Delta\|\boldsymbol{n}\|^2 \\ &\overset{(a)}{=} \sum_r \Bigg(\sum_{k\in[K_r]} \lambda_r \beta_{r,k}\Big(\|\boldsymbol{n} + \boldsymbol{e}^{(r,k)}\|^2 - \|\boldsymbol{n}\|^2\Big)\end{aligned}$$

$$
+ \sum_{\substack{k_1, k_2 \in [K_r] \\ k_1 \neq k_2}} n_{r,k_1} x_r (S_r)_{k_1,k_2} \cdot \Big(\|\boldsymbol{n} - \boldsymbol{e}^{(r,k_1)} + \boldsymbol{e}^{(r,k_2)}\|^2 - \|\boldsymbol{n}\|^2 \Big)
$$

$$
+ \sum_{k \in [K_r]} n_{r,k} x_r \sum_{k'} (-S_r)_{k,k'} \Big(\|\boldsymbol{n} - \boldsymbol{e}^{(r,k)}\|^2 - \|\boldsymbol{n}\|^2 \Big) \Bigg)
$$

$$
\overset{(b)}{\leq} 2\langle \boldsymbol{n}, \begin{bmatrix} \lambda_1 \boldsymbol{\beta}_1^T \\ \lambda_2 \boldsymbol{\beta}_2^T \\ \vdots \\ \lambda_R \boldsymbol{\beta}_R^T \end{bmatrix} - \begin{bmatrix} -S_1^T & 0 & \cdots & 0 \\ 0 & -S_2^T & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & -S_R^T \end{bmatrix} \begin{bmatrix} \boldsymbol{n}_1 x_1 \\ \boldsymbol{n}_2 x_2 \\ \vdots \\ \boldsymbol{n}_R x_R \end{bmatrix} \rangle + A_1,
$$

$$
\overset{(c)}{=} 2\langle \boldsymbol{n}, (-S^T)(\boldsymbol{\rho} - \boldsymbol{n}\boldsymbol{x}) \rangle + A_1
$$

$$
= 2\langle \boldsymbol{n}, (-S^T)(\boldsymbol{\rho}^{(0)} - \boldsymbol{n}\boldsymbol{x}) \rangle - 2\epsilon \langle \boldsymbol{n}, (-S^T)\boldsymbol{\rho}^{(0)} \rangle + A_1. \quad (43)
$$

where (a) follows from the transition rates of the flow count process $(\boldsymbol{N}(t)\colon \geq 0)$ under the proportionally fair sharing, (b) is obtained by expressing norms in terms of inner products and bounding the sum of the terms with $\|\boldsymbol{e}^{(r,k)}\|^2$'s by a constant A_1, and (c) follows from the definition of $\boldsymbol{\rho}$ in (17). We can derive a lower bound on $\Delta\|\boldsymbol{n}_{\shortparallel}\|^2$ in a similar way:

$$
\Delta\|\boldsymbol{n}_{\shortparallel}\|^2
$$

$$
\overset{(a)}{=} \sum_r \Bigg(\sum_{k \in [K_r]} \lambda_r \beta_{r,k} \Big(\|(\boldsymbol{n} + \boldsymbol{e}^{(r,k)})_{\shortparallel}\|^2 - \|\boldsymbol{n}_{\shortparallel}\|^2 \Big)
+ \sum_{\substack{k_1, k_2 \in [K_r] \\ k_1 \neq k_2}} n_{r,k_1} x_r (S_r)_{k_1,k_2} \cdot \Big(\|(\boldsymbol{n} - \boldsymbol{e}^{(r,k_1)} + \boldsymbol{e}^{(r,k_2)})_{\shortparallel}\|^2 - \|\boldsymbol{n}_{\shortparallel}\|^2 \Big)
+ \sum_{k \in [K_r]} n_{r,k} x_r \sum_{k'} (-S_r)_{k,k'} \Big(\|(\boldsymbol{n} - \boldsymbol{e}^{(r,k)})_{\shortparallel}\|^2 - \|\boldsymbol{n}_{\shortparallel}\|^2 \Big) \Bigg)
$$

$$
\overset{(b)}{\geq} \sum_r \Bigg(\sum_{k \in [K_r]} \lambda_r \beta_{r,k} \Big(2\langle \boldsymbol{n}_{\shortparallel}, \boldsymbol{e}^{(r,k)} \rangle \Big)
+ \sum_{\substack{k_1, k_2 \in [K_r] \\ k_1 \neq k_2}} n_{r,k_1} x_r (S_r)_{k_1,k_2} \Big(2\langle \boldsymbol{n}_{\shortparallel}, -\boldsymbol{e}^{(r,k_1)} + \boldsymbol{e}^{(r,k_2)} \rangle \Big)
+ \sum_{k \in [K_r]} n_{r,k} x_r \sum_{k'} (-S_r)_{k,k'} \Big(2\langle \boldsymbol{n}_{\shortparallel}, -\boldsymbol{e}^{(r,k)} \rangle \Big) \Bigg)
$$

$$
\overset{(c)}{=} 2\langle \boldsymbol{n}_{\shortparallel}, (-S^T)(\boldsymbol{\rho} - \boldsymbol{n}\boldsymbol{x}) \rangle
$$

$$
= 2\langle \boldsymbol{n}_{\shortparallel}, (-S^T)(\boldsymbol{\rho}^{(0)} - \boldsymbol{n}\boldsymbol{x}) \rangle - 2\epsilon \langle \boldsymbol{n}_{\shortparallel}, (-S^T)\boldsymbol{\rho}^{(0)} \rangle, \quad (44)
$$

where (a) still follows from the transitions rates of the flow count process, (b) follows from that $\langle \boldsymbol{n}_{\shortparallel}, \boldsymbol{n}_\perp \rangle = 0$ and $\langle \boldsymbol{n}_{\shortparallel}, (\boldsymbol{n} + \boldsymbol{e}^{(r,k)})_\perp \rangle \leq 0$, $\langle \boldsymbol{n}_{\shortparallel}, (\boldsymbol{n} - \boldsymbol{e}^{(r,k_1)} + \boldsymbol{e}^{(r,k_2)})_\perp \rangle \leq 0$, $\langle \boldsymbol{n}_{\shortparallel}, (\boldsymbol{n} - \boldsymbol{e}^{(r,k)})_\perp \rangle \leq 0$ since perpendicular components are in the polar cone of the cone $\mathcal{K}$, and (c) still follows from the definition of $\boldsymbol{\rho}$. Combining the above bounds (43) and (44) we have

$$
\Delta\|\boldsymbol{n}_\perp\|
\leq \frac{1}{\|\boldsymbol{n}_\perp\|} \langle \boldsymbol{n} - \boldsymbol{n}_{\shortparallel}, (-S^T)(\boldsymbol{\rho}^{(0)} - \boldsymbol{n}\boldsymbol{x}) \rangle
- \epsilon \frac{\langle \boldsymbol{n}_\perp, (-S^T)\boldsymbol{\rho}^{(0)} \rangle}{\|\boldsymbol{n}_\perp\|} + \frac{A_1}{2\|\boldsymbol{n}_\perp\|}
$$

$$
\leq \frac{1}{\|\boldsymbol{n}_\perp\|} \langle \boldsymbol{n} - \boldsymbol{n}_{\shortparallel}, (-S^T)(\boldsymbol{\rho}^{(0)} - \boldsymbol{n}\boldsymbol{x}) \rangle
+ \epsilon \|(-S^T)\boldsymbol{\rho}^{(0)}\| + \frac{A_1}{2\|\boldsymbol{n}_\perp\|},
$$

which completes the proof of the claim. □

5.2 Proof of Lemma 5 (Tail Bound)

Before presenting the proof, let us first define some parameters needed. We still fix an $\epsilon > 0$ and omit the superscript $^{(\epsilon)}$ for conciseness. Recall that $q_{\boldsymbol{n}\boldsymbol{n}'}$ is the transition rate from state $\boldsymbol{n}$ to $\boldsymbol{n}'$ of the flow count process $(\boldsymbol{N}(t)\colon t \geq 0)$, where a state is a vector stacking together the flow counts on every route in every phase. Let

$$
\overline{q} = \sup_{\boldsymbol{n}} (-q_{\boldsymbol{n}\boldsymbol{n}}),
$$

$$
\nu_0 = \sup_{\boldsymbol{n}, \boldsymbol{n}'\colon q_{\boldsymbol{n}\boldsymbol{n}'} > 0} \big| \|\boldsymbol{n}'\| - \|\boldsymbol{n}\| \big|, \; \eta_0 = \sup_{\boldsymbol{n}} \sum_{\boldsymbol{n}'\colon \|\boldsymbol{n}\| < \|\boldsymbol{n}'\|} q_{\boldsymbol{n}\boldsymbol{n}'},
$$

$$
\nu_1 = \sup_{\boldsymbol{n}, \boldsymbol{n}'\colon q_{\boldsymbol{n}\boldsymbol{n}'} > 0} \big| \|\boldsymbol{n}'_\perp\| - \|\boldsymbol{n}_\perp\| \big|, \; \eta_1 = \sup_{\boldsymbol{n}} \sum_{\boldsymbol{n}'\colon \|\boldsymbol{n}_\perp\| < \|\boldsymbol{n}'_\perp\|} q_{\boldsymbol{n}\boldsymbol{n}'},
$$

$$
\alpha = \frac{\eta_1 \nu_1}{\eta_1 \nu_1 + \sqrt{\epsilon}}.
$$

It can be verified that $\overline{q} < +\infty$, $\nu_0 < +\infty$ and $\nu_1 < +\infty$. Note that by this definition of α, the constant a in (39) of Lemma 5 equals to $\eta_1 \nu_1$.

We also need the following lemma to bound the distribution of $\sum_{r,k} \overline{N}_{r,k}$. The proof of this lemma is given in our technical report [23] due to space limit. In the proof, we analyze the drift $\Delta\|\boldsymbol{n}\|$, and then apply a continuous-time version of the exponential-type tail bound in [2]. Note that the definition of β in (46) below corresponds to $b = \eta_0 \nu_0 / (\kappa_{\min} A_2 / \mu_{\max})$ for the constant b in (39) of Lemma 5.

Lemma 6. *For any nonnegative $\epsilon \leq \epsilon_{\max}$, the distribution of $\sum_{r,k} \overline{N}_{r,k}$ has the following exponential tail bound: for any nonnegative integer j,*

$$
\mathbb{P}\Bigg(\sum_{r,k} \overline{N}_{r,k} > \frac{\mu_{\max} A_1 A_3}{2\kappa_{\min} A_2 \epsilon} + 2\nu_0 A_3 j \Bigg) \leq \beta^{j+1}, \quad (45)
$$

where $\mu_{\max}, A_2, A_3$ are positive constants, and

$$
\beta = \frac{\eta_0 \nu_0}{\eta_0 \nu_0 + \epsilon \kappa_{\min} A_2 / \mu_{\max}} < 1. \quad (46)
$$

Now we are ready to prove Lemma 5, where we derive the tail bound for $\|\overline{\boldsymbol{N}}_\perp\|$.

Proof of Lemma 5. Note that $\mathbb{E}[\|\overline{\boldsymbol{N}}_\perp\|] < +\infty$ since $\mathbb{E}[\sum_{r,k} \overline{N}_{r,k}] < +\infty$ by the proof of Lemma 6. Let A denote $\frac{A_1}{2\xi_1\sqrt{\epsilon}}$. Fix a $c \geq A - \nu_1$. Let $\hat{V}(\boldsymbol{n}) = \max\{c, \|\boldsymbol{n}_\perp\|\}$. Let π denote the distribution of $\overline{\boldsymbol{N}}$. Then similar to the proof of the exponential-type bound in [2], since $\mathbb{E}[\hat{V}(\overline{\boldsymbol{N}})] < +\infty$ and $\overline{q} < +\infty$,

$$
0 = \sum_{\boldsymbol{n}} \pi(\boldsymbol{n}) \sum_{\boldsymbol{n}'\colon \boldsymbol{n}' \neq \boldsymbol{n}} Q_{\boldsymbol{n}\boldsymbol{n}'} (\hat{V}(\boldsymbol{n}') - \hat{V}(\boldsymbol{n}))
$$

$$
= \sum_{\boldsymbol{n}\colon \|\boldsymbol{n}_\perp\| \leq c - \nu_1} \pi(\boldsymbol{n}) \sum_{\boldsymbol{n}'\colon \boldsymbol{n}' \neq \boldsymbol{n}} Q_{\boldsymbol{n}\boldsymbol{n}'} (\hat{V}(\boldsymbol{n}') - \hat{V}(\boldsymbol{n})) \quad (47)
$$

$$
+ \sum_{\boldsymbol{n}\colon c - \nu_1 < \|\boldsymbol{n}_\perp\| \leq c + \nu_1} \pi(\boldsymbol{n}) \sum_{\boldsymbol{n}'\colon \boldsymbol{n}' \neq \boldsymbol{n}} Q_{\boldsymbol{n}\boldsymbol{n}'} (\hat{V}(\boldsymbol{n}') - \hat{V}(\boldsymbol{n})) \quad (48)
$$

$$+\sum_{\boldsymbol{n}:\,\|\boldsymbol{n}_\perp\|>c+\nu_1}\pi(\boldsymbol{n})\sum_{\boldsymbol{n}':\,\boldsymbol{n}'\neq\boldsymbol{n}}Q_{\boldsymbol{n}\boldsymbol{n}'}(\hat V(\boldsymbol{n}')-\hat V(\boldsymbol{n})). \tag{49}$$

(i) The first summand (47) is 0 since when $\|\boldsymbol{n}_\perp\|\le c-\nu_1$, $\hat V(\boldsymbol{n}')=\hat V(\boldsymbol{n})=c$ for $\boldsymbol{n}'$ with $Q_{\boldsymbol{n}\boldsymbol{n}'}>0$.

(ii) Consider the second summand (48). We can check that for any two states $\boldsymbol{n}$ and $\boldsymbol{n}'$, either

$$0\le\hat V(\boldsymbol{n}')-\hat V(\boldsymbol{n})\le\|\boldsymbol{n}'_\perp\|-\|\boldsymbol{n}_\perp\|,$$

or

$$\|\boldsymbol{n}'_\perp\|-\|\boldsymbol{n}_\perp\|\le\hat V(\boldsymbol{n}')-\hat V(\boldsymbol{n})\le 0,$$

regardless of the relation between c and $\|\boldsymbol{n}'_\perp\|$, $\|\boldsymbol{n}_\perp\|$. Then,

$$\begin{aligned}
&\sum_{\boldsymbol{n}':\,\boldsymbol{n}'\neq\boldsymbol{n}}Q_{\boldsymbol{n}\boldsymbol{n}'}(\hat V(\boldsymbol{n}')-\hat V(\boldsymbol{n}))\\
&=\sum_{\boldsymbol{n}':\,\hat V(\boldsymbol{n}')>\hat V(\boldsymbol{n})}Q_{\boldsymbol{n}\boldsymbol{n}'}(\hat V(\boldsymbol{n}')-\hat V(\boldsymbol{n}))\\
&\quad+\sum_{\boldsymbol{n}':\,\hat V(\boldsymbol{n}')\le\hat V(\boldsymbol{n})}Q_{\boldsymbol{n}\boldsymbol{n}'}(\hat V(\boldsymbol{n}')-\hat V(\boldsymbol{n}))\\
&\le\sum_{\boldsymbol{n}':\,\hat V(\boldsymbol{n}')>\hat V(\boldsymbol{n})}Q_{\boldsymbol{n}\boldsymbol{n}'}(\hat V(\boldsymbol{n}')-\hat V(\boldsymbol{n}))\\
&\le\sum_{\boldsymbol{n}':\,\|\boldsymbol{n}'_\perp\|>\|\boldsymbol{n}_\perp\|}Q_{\boldsymbol{n}\boldsymbol{n}'}\nu_1\\
&\le\eta_1\nu_1.
\end{aligned}$$

Thus the second summand satisfies

$$\begin{aligned}
&\sum_{\boldsymbol{n}:\,c-\nu_1<\|\boldsymbol{n}_\perp\|\le c+\nu_1}\pi(\boldsymbol{n})\sum_{\boldsymbol{n}':\,\boldsymbol{n}'\neq\boldsymbol{n}}Q_{\boldsymbol{n}\boldsymbol{n}'}(\hat V(\boldsymbol{n}')-\hat V(\boldsymbol{n}))\\
&\le\eta_1\nu_1\Big(\mathbb{P}(\|\overline{\boldsymbol{N}}_\perp\|>c-\nu_1)-\mathbb{P}(\|\overline{\boldsymbol{N}}_\perp\|>c+\nu_1)\Big).
\end{aligned}$$

(iii) Consider the third summand (49). When $\|\boldsymbol{n}_\perp\|>c+\nu_1$, $\hat V(\boldsymbol{n})=\|\boldsymbol{n}_\perp\|$ and $\hat V(\boldsymbol{n}')=\|\boldsymbol{n}'_\perp\|$ for $\boldsymbol{n}'$ with $Q_{\boldsymbol{n}\boldsymbol{n}'}>0$. Therefore,

$$\begin{aligned}
&\sum_{\boldsymbol{n}:\,\|\boldsymbol{n}_\perp\|>c+\nu_1}\pi(\boldsymbol{n})\sum_{\boldsymbol{n}':\,\boldsymbol{n}'\neq\boldsymbol{n}}Q_{\boldsymbol{n}\boldsymbol{n}'}(\hat V(\boldsymbol{n}')-\hat V(\boldsymbol{n}))\\
&=\sum_{\boldsymbol{n}:\,\|\boldsymbol{n}_\perp\|>c+\nu_1}\pi(\boldsymbol{n})\Delta\|\boldsymbol{n}_\perp\|\\
&=\sum_{\substack{\boldsymbol{n}:\,\|\boldsymbol{n}_\perp\|>c+\nu_1\\ \frac{\|\boldsymbol{n}_\perp\|}{\sum_{r,k}n_{r,k}}\ge\frac{\xi_2\sqrt{\epsilon}}{\kappa_{\min}C_{\min}}}}\pi(\boldsymbol{n})\Delta\|\boldsymbol{n}_\perp\|+\sum_{\substack{\boldsymbol{n}:\,\|\boldsymbol{n}_\perp\|>c+\nu_1\\ \frac{\|\boldsymbol{n}_\perp\|}{\sum_{r,k}n_{r,k}}<\frac{\xi_2\sqrt{\epsilon}}{\kappa_{\min}C_{\min}}}}\pi(\boldsymbol{n})\Delta\|\boldsymbol{n}_\perp\|\\
&\overset{(a)}{\le}-\sqrt{\epsilon}\mathbb{P}\left(\|\overline{\boldsymbol{N}}_\perp\|>c+\nu_1,\frac{\|\overline{\boldsymbol{N}}_\perp\|}{\sum_{r,k}\overline{N}_{r,k}}\ge\frac{\xi_2\sqrt{\epsilon}}{\kappa_{\min}C_{\min}}\right)\\
&\quad+(\xi_1+1)\sqrt{\epsilon}\mathbb{P}\left(\|\overline{\boldsymbol{N}}_\perp\|>c+\nu_1,\frac{\|\overline{\boldsymbol{N}}_\perp\|}{\sum_{r,k}\overline{N}_{r,k}}<\frac{\xi_2\sqrt{\epsilon}}{\kappa_{\min}C_{\min}}\right)\\
&\overset{(b)}{=}-\sqrt{\epsilon}\mathbb{P}(\|\overline{\boldsymbol{N}}_\perp\|>c+\nu_1)\\
&\quad+\xi_2\sqrt{\epsilon}\mathbb{P}\left(\|\overline{\boldsymbol{N}}_\perp\|>c+\nu_1,\frac{\|\overline{\boldsymbol{N}}_\perp\|}{\sum_{r,k}\overline{N}_{r,k}}<\frac{\xi_2\sqrt{\epsilon}}{\kappa_{\min}C_{\min}}\right)\\
&\le-\sqrt{\epsilon}\mathbb{P}(\|\overline{\boldsymbol{N}}_\perp\|>c+\nu_1)\\
&\quad+\xi_2\sqrt{\epsilon}\mathbb{P}\left(\sum_{r,k}\overline{N}_{r,k}>\frac{(c+\nu_1)\kappa_{\min}C_{\min}}{\xi_2\sqrt{\epsilon}}\right).
\end{aligned}$$

The inequality (a) follows from the drift bounds given in Lemma 4, and (b) follows from the choice of ξ_1 and ξ_2 in the proof of Lemma 4.

Combining the three summands we have

$$\begin{aligned}
\mathbb{P}(\|\overline{\boldsymbol{N}}_\perp\|>c+\nu_1)&\le\frac{\eta_1\nu_1}{\eta_1\nu_1+\sqrt{\epsilon}}\mathbb{P}(\|\overline{\boldsymbol{N}}_\perp\|>c-\nu_1)\\
&\quad+\frac{\xi_2\sqrt{\epsilon}}{\eta_1\nu_1+\sqrt{\epsilon}}\mathbb{P}\left(\sum_{r,k}\overline{N}_{r,k}>\frac{(c+\nu_1)\kappa_{\min}C_{\min}}{\xi_2\sqrt{\epsilon}}\right).
\end{aligned}$$

Recall that we let α denote $\frac{\eta_1\nu_1}{\eta_1\nu_1+\sqrt{\epsilon}}$. Let $c=A+(2j-1)\nu_1$ for a nonnegative integer j. Then

$$\begin{aligned}
\mathbb{P}(\|\overline{\boldsymbol{N}}_\perp\|>A+2\nu_1 j)&\le\alpha\mathbb{P}(\|\overline{\boldsymbol{N}}_\perp\|>A+2\nu_1(j-1))\\
&\quad+\xi_2(1-\alpha)\mathbb{P}\left(\sum_{r,k}\overline{N}_{r,k}>\frac{(A+2\nu_1 j)\kappa_{\min}C_{\min}}{\xi_2\sqrt{\epsilon}}\right).
\end{aligned}$$

Now we use Lemma 6 to bound the last probability above. Recall that we have chosen ξ_1 and ξ_2 in (42) such that $\xi_2-\xi_1=2$. We can further require that

$$\xi_1\xi_2=\frac{\kappa_{\min}^2C_{\min}A_2}{\mu_{\max}A_3}. \tag{50}$$

It can be verified that such constants ξ_1 and ξ_2 are well-defined. Also recall that $A=\frac{A_1}{2\xi_1\sqrt{\epsilon}}$. Define a constant $\theta=\frac{\nu_1\kappa_{\min}C_{\min}}{\xi_2\nu_0A_3}$. Then

$$\begin{aligned}
&\mathbb{P}\left(\sum_{r,k}\overline{N}_{r,k}>\frac{(A+2\nu_1 j)\kappa_{\min}C_{\min}}{\xi_2\sqrt{\epsilon}}\right)\\
&=\mathbb{P}\left(\sum_{r,k}\overline{N}_{r,k}>\frac{A_1\kappa_{\min}C_{\min}}{2\xi_1\xi_2\epsilon}+\frac{2\nu_1\kappa_{\min}C_{\min}}{\xi_2\sqrt{\epsilon}}j\right)\\
&\le\mathbb{P}\left(\sum_{r,k}\overline{N}_{r,k}>\frac{\mu_{\max}A_1A_3}{2\kappa_{\min}A_2\epsilon}+2\nu_0A_3\left\lfloor\frac{j\theta}{\sqrt{\epsilon}}\right\rfloor\right)\\
&\le\beta^{\lfloor j\theta/\sqrt{\epsilon}\rfloor+1}\\
&\le\left(\beta^{\theta/\sqrt{\epsilon}}\right)^j.
\end{aligned}$$

Therefore,

$$\begin{aligned}
&\mathbb{P}(\|\overline{\boldsymbol{N}}_\perp\|>A+2\nu_1 j)\\
&\le\alpha\mathbb{P}(\|\overline{\boldsymbol{N}}_\perp\|>A+2\nu_1(j-1))+\xi_2(1-\alpha)\left(\beta^{\theta/\sqrt{\epsilon}}\right)^j.
\end{aligned}$$

Using this inequality for $k-1,k-2,\cdots$ yields

$$\begin{aligned}
&\mathbb{P}\left(\|\overline{\boldsymbol{N}}_\perp\|>\frac{A_1}{2\xi_1\sqrt{\epsilon}}+2\nu_1 j\right)\\
&\le\alpha^{j+1}+\xi_2(1-\alpha)\sum_{i=0}^{j}\alpha^i\left(\beta^{\theta/\sqrt{\epsilon}}\right)^{j-i}.
\end{aligned}$$

This completes the proof of Lemma 5. □

6. BACKLOG BOUND

In this section, we prove the backlog bound in Theorem 2, by deriving upper and lower bounds that are asymptotically tight. We obtain these bounds by setting the steady-state drift of the Lyapunov function $V(\boldsymbol{n})=\|\boldsymbol{n}^s_{\shortparallel}\|$ to 0, where $\boldsymbol{n}^s_{\shortparallel}$ is the projection of the state $\boldsymbol{n}$ onto the *subspace* where the cone $\mathcal{K}$ lies in, i.e., the subspace $\mathcal{S}$ spanned by $\boldsymbol{b}^{(\ell)}$'s. Note that in this section, we often consider the projection onto the *subspace* instead of the projection onto the *cone*. We use the superscript ^{s} to indicate when the projection is onto the subspace. Then $\boldsymbol{n}_{\shortparallel}^s$ can be written as

$$\boldsymbol{n}_{\shortparallel}^s = \sum_l \alpha_\ell^s \boldsymbol{b}^{(\ell)}, \tag{51}$$

where the coefficients α_ℓ^s's can be negative. The projection onto the subspace is a linear operator, i.e., $(\boldsymbol{y}+\boldsymbol{z})_{\shortparallel}^s = \boldsymbol{y}_{\shortparallel}^s + \boldsymbol{z}_{\shortparallel}^s$ for any $\boldsymbol{y}, \boldsymbol{z} \in \mathbb{R}^K$.

For a state $\boldsymbol{n}$, recall that p_ℓ is the Lagrange multiplier for the capacity constraint of link ℓ. A key step in deriving the backlog bounds is to show that the p_ℓ's are close to the α_ℓ^s's in heavy traffic. We know that when $\boldsymbol{n} = \boldsymbol{n}_{\shortparallel}^s$, i.e., $\|\boldsymbol{n}_\perp\| = \|\boldsymbol{n}_\perp^s\| = 0$, the Lagrangian multipliers p_ℓ's are equal to the coefficients α_ℓ^s's of the projection. Then intuitively, when $\|\boldsymbol{n}_\perp\|$ is small, the rate allocation based on $\boldsymbol{n}$ should not be far away from the rate allocation based on $\boldsymbol{n}_{\shortparallel}^s$, and thus the p_ℓ's should not be far away from the α_ℓ's. Then we can use the state-space collapse result to bound the difference $|\alpha_\ell^s - p_\ell|$ in heavy traffic. Specifically, the following lemma bounds the difference $|\alpha_\ell^s - p_\ell|$ using $\|\boldsymbol{n}_\perp\|$, where notice that $\boldsymbol{n}_\perp$ is the projection onto the cone. The proof of this lemma is given in Section 6.2.

LEMMA 7. *There exists a constant $B_3 > 0$ such that for any state $\boldsymbol{n}$ and any link ℓ,*

$$|\alpha_\ell^s - p_\ell| \le B_3 \|\boldsymbol{n}_\perp\|^{1/2} \left(\sum_{r,k} n_{r,k} \right)^{1/2}. \tag{52}$$

6.1 Proof of Theorem 2

PROOF. For ease of notation, we fix an ϵ and omit the superscript $^{(\epsilon)}$ when it is clear from context. We obtain the backlog bounds by analyzing the drift of $\|\boldsymbol{n}_{\shortparallel}^s\|^2$. Similar to (44)(a),

$$\begin{aligned}
&\Delta\|\boldsymbol{n}_{\shortparallel}^s\|^2 \\
&= \sum_r \Bigg(\sum_{k\in[K_r]} \lambda_r \beta_{r,k} \Big(2\langle \boldsymbol{n}_{\shortparallel}^s, (\boldsymbol{e}^{(r,k)})_{\shortparallel}^s \rangle + \|(\boldsymbol{e}^{(r,k)})_{\shortparallel}^s\|^2 \Big) \\
&\quad + \sum_{\substack{k_1,k_2\in[K_r]\\ k_1\neq k_2}} n_{r,k_1} x_r (S_r)_{k_1,k_2} \Big(2\langle \boldsymbol{n}_{\shortparallel}^s, -(\boldsymbol{e}^{(r,k_1)})_{\shortparallel}^s + (\boldsymbol{e}^{(r,k_2)})_{\shortparallel}^s \rangle \\
&\qquad\qquad + \| -(\boldsymbol{e}^{(r,k_1)})_{\shortparallel}^s + (\boldsymbol{e}^{(r,k_2)})_{\shortparallel}^s \|^2 \Big) \\
&\quad + \sum_{k\in[K_r]} n_{r,k} x_r \Big(-\sum_{k'} (S_r)_{k,k'} \Big) \\
&\qquad\qquad \cdot \Big(-2\langle \boldsymbol{n}_{\shortparallel}^s, (\boldsymbol{e}^{(r,k)})_{\shortparallel}^s \rangle + \|(\boldsymbol{e}^{(r,k)})_{\shortparallel}^s\|^2 \Big) \Bigg) \\
&\overset{(a)}{=} 2\langle \boldsymbol{n}_{\shortparallel}^s, (-S^T)(\boldsymbol{\rho} - \boldsymbol{n}\boldsymbol{x}) \rangle + B_1 \\
&= -2\epsilon \langle \boldsymbol{n}_{\shortparallel}^s, (-S^T)\boldsymbol{\rho}^{(0)} \rangle + 2\langle \boldsymbol{n}_{\shortparallel}^s, (-S^T)(\boldsymbol{\rho}^{(0)} - \boldsymbol{n}\boldsymbol{x}) \rangle + B_1,
\end{aligned}$$

where in (a) we have used the fact that $\langle \boldsymbol{n}_{\shortparallel}^s, (\boldsymbol{e}^{(r,k)})_{\shortparallel}^s \rangle = \langle \boldsymbol{n}_{\shortparallel}^s, \boldsymbol{e}^{(r,k)} \rangle$ for any r,k since $\langle \boldsymbol{n}_{\shortparallel}^s, (\boldsymbol{e}^{(r,k)})_\perp^s \rangle = 0$, and

$$\begin{aligned}
B_1 = \sum_r \Bigg(&\sum_{k\in[K_r]} \lambda_r \beta_{r,k} \|(\boldsymbol{e}^{(r,k)})_{\shortparallel}^s\|^2 \\
&+ \sum_{\substack{k_1,k_2\in[K_r]\\ k_1\neq k_2}} n_{r,k_1} x_r (S_r)_{k_1,k_2} \| -(\boldsymbol{e}^{(r,k_1)})_{\shortparallel}^s + (\boldsymbol{e}^{(r,k_2)})_{\shortparallel}^s \|^2 \\
&+ \sum_{k\in[K_r]} n_{r,k} x_r \Big(-\sum_{k'} (S_r)_{k,k'} \Big) \|(\boldsymbol{e}^{(r,k)})_{\shortparallel}^s\|^2 \Bigg).
\end{aligned}$$

When the system is in steady state, we have $\mathbb{E}[\Delta\|\overline{\boldsymbol{N}}_{\shortparallel}^s\|^2] = 0$. Therefore,

$$\epsilon\mathbb{E}[\langle \overline{\boldsymbol{N}}_{\shortparallel}^s, (-S^T)\boldsymbol{\rho}^{(0)} \rangle] = \mathbb{E}[\langle \overline{\boldsymbol{N}}_{\shortparallel}^s, (-S^T)(\boldsymbol{\rho}^{(0)} - \boldsymbol{n}\boldsymbol{x}) \rangle] + \mathbb{E}[B_1]/2. \tag{53}$$

We analyze the three terms in (53) term by term.

(i) We first consider the term $\epsilon\mathbb{E}[\langle \overline{\boldsymbol{N}}_{\shortparallel}^s, (-S^T)\boldsymbol{\rho}^{(0)} \rangle]$ and show that it is close to $\epsilon\mathbb{E}[\sum_{r,k} \overline{N}_{r,k}]$. For any state $\boldsymbol{n}$, since $\boldsymbol{n}_{\shortparallel}^s \in \mathcal{S}$, it can be written as

$$\boldsymbol{n}_{\shortparallel}^s = \sum_{l\in\mathcal{L}} \alpha_\ell^s \boldsymbol{b}^{(\ell)}.$$

Thus

$$\langle \boldsymbol{n}_{\shortparallel}^s, (-S^T)\boldsymbol{\rho}^{(0)} \rangle = \sum_\ell \alpha_\ell^s \langle \boldsymbol{b}^{(\ell)}, (-S^T)\boldsymbol{\rho}^{(0)} \rangle \overset{(a)}{=} \sum_\ell \alpha_\ell^s C_\ell, \tag{54}$$

where (a) follows from arguments similar to those in the proof of condition (C1) in (30)–(34). We also know that $\sum_{r,k} n_{r,k} = \sum_\ell p_\ell C_\ell$. Let $C_{\max} = \max_\ell C_\ell$. Then

$$\begin{aligned}
\left| \sum_{r,k} n_{r,k} - \langle \boldsymbol{n}_{\shortparallel}^s, (-S^T)\boldsymbol{\rho}^{(0)} \rangle \right| &\le \sum_\ell |\alpha_\ell^s - p_\ell| C_\ell \\
&\le C_{\max} B_3 \|\boldsymbol{n}_\perp\|^{1/2} \left(\sum_{r,k} n_{r,k} \right)^{1/2},
\end{aligned}$$

where the second inequality follows from Lemma 7. Therefore,

$$\begin{aligned}
&\epsilon\mathbb{E}\left[\left| \sum_{r,k} \overline{N}_{r,k} - \langle \overline{\boldsymbol{N}}_{\shortparallel}^s, (-S^T)\boldsymbol{\rho}^{(0)} \rangle \right| \right] \\
&\le \epsilon C_{\max} B_3 \mathbb{E}\left[\|\overline{\boldsymbol{N}}_\perp\|^{1/2} \left(\sum_{r,k} \overline{N}_{r,k} \right)^{1/2} \right] \\
&\overset{(a)}{\le} \epsilon C_{\max} B_3 \mathbb{E}[\|\overline{\boldsymbol{N}}_\perp\|]^{1/2} \left(\mathbb{E}\left[\sum_{r,k} \overline{N}_{r,k} \right] \right)^{1/2} \\
&\overset{(b)}{=} O(\epsilon^{1/4}),
\end{aligned}$$

where (a) follows from Cauchy-Schwarz inequality, and (b) follows from the state-space collapse result in Theorem 1 and the bound on $\mathbb{E}[\sum_{r,k} \overline{N}_{r,k}]$ indicated by Lemma 6.

(ii) Next, we bound the term $\mathbb{E}[\langle \overline{\boldsymbol{N}}_{\shortparallel}^s, (-S^T)(\boldsymbol{\rho}^{(0)} - \boldsymbol{n}\boldsymbol{x}) \rangle]$. Again, since $\boldsymbol{n}_{\shortparallel}^s \in \mathcal{S}$ and recall the condition (C1) for the inner product, we have

$$\begin{aligned}
\langle \boldsymbol{n}_{\shortparallel}^s, (-S^T)(\boldsymbol{\rho}^{(0)} - \boldsymbol{n}\boldsymbol{x}) \rangle &= \sum_\ell \alpha_\ell^s \langle \boldsymbol{b}^{(\ell)}, (-S^T)(\boldsymbol{\rho}^{(0)} - \boldsymbol{n}\boldsymbol{x}) \rangle \\
&= \sum_\ell \alpha_\ell^s U_\ell.
\end{aligned}$$

By Lemma 7 and Hölder's inequality:

$$\mathbb{E}[|\alpha_\ell^s U_\ell|] = \mathbb{E}[|(\alpha_\ell^s - p_\ell)U_\ell|]$$
$$\leq B_3\left(\mathbb{E}\left[\|\overline{\boldsymbol{N}}_\perp\|^{\frac{\tau_1}{2}}\left(\sum_{r,k}\overline{N}_{r,k}\right)^{\frac{\tau_1}{2}}\right]\right)^{\frac{1}{\tau_1}}\left(\mathbb{E}[U_\ell^{\tau_2}]\right)^{\frac{1}{\tau_2}},$$

where we pick τ_1 and τ_2 such that τ_1 is an even integer with $\tau_1 > 4$ and $\frac{1}{\tau_1} + \frac{1}{\tau_2} = 1$. Using Cauchy-Schwarz inequality we have

$$\left(\mathbb{E}\left[\|\overline{\boldsymbol{N}}_\perp\|^{\frac{\tau_1}{2}}\left(\sum_{r,k}\overline{N}_{r,k}\right)^{\frac{\tau_1}{2}}\right]\right)^{\frac{1}{\tau_1}}$$
$$\leq (\mathbb{E}[\|\overline{\boldsymbol{N}}_\perp\|^{\tau_1}])^{\frac{1}{2\tau_1}}\left(\mathbb{E}\left[\left(\sum_{r,k}\overline{N}_r\right)^{\tau_1}\right]\right)^{\frac{1}{2\tau_1}}$$
$$= O(\epsilon^{-\frac{3}{4}}),$$

where again the last equality follows from the state-space collapse result in Theorem 1 and the bound on $\mathbb{E}[\sum_{r,k}\overline{N}_{r,k}]$ indicated by Lemma 6. Next we bound $\mathbb{E}[U_\ell^{\tau_2}]$. We can prove that $\mathbb{E}[U_\ell] = \epsilon C_\ell$ by considering the Lyapunov function $w_\ell(\boldsymbol{n}) = \langle \boldsymbol{b}^{(\ell)}, \boldsymbol{n}\rangle$. Its drift is

$$\Delta w_\ell(\boldsymbol{n}) = -\epsilon\langle \boldsymbol{b}^{(\ell)}, (-S^T)(\boldsymbol{\rho}^{(0)}\rangle + \langle \boldsymbol{b}^{(\ell)}, (-S^T)\boldsymbol{\rho}^{(0)} - \boldsymbol{n}\boldsymbol{x})$$
$$= -\epsilon C_\ell + U_\ell.$$

Since in the steady state $\mathbb{E}[\Delta w_\ell(\overline{\boldsymbol{N}})] = 0$, we have $\mathbb{E}[U_\ell] = \epsilon C_\ell$. Since $0 \leq U_\ell \leq C_\ell$, there holds

$$\left(\mathbb{E}\left[U_\ell^{\tau_2}\right]\right)^{\frac{1}{\tau_2}} \leq \left(\mathbb{E}\left[U_\ell \cdot C_\ell^{\tau_2 - 1}\right]\right)^{\frac{1}{\tau_2}} = \epsilon^{\frac{1}{\tau_2}} C_\ell.$$

Combining these bounds we have $\mathbb{E}[|\alpha_\ell^s U_\ell|] = O(\epsilon^{\frac{1}{4} - \frac{1}{\tau_1}})$, and thus

$$\mathbb{E}[|\langle \overline{\boldsymbol{N}}_{\shortparallel}^s, (-S^T)(\boldsymbol{\rho}^{(0)} - \boldsymbol{n}\boldsymbol{x})|] = \mathbb{E}\left[\left|\sum_\ell \alpha_\ell^s U_\ell\right|\right] = O(\epsilon^{\frac{1}{4} - \frac{1}{\tau_1}}).$$

(iii) Lastly, we claim that the last term $\mathbb{E}[B_1]/2 = (1-\epsilon)L$. The proof is given in our technical report [23] due to space limit.

Combining (i), (ii) and (iii) for the terms in (53) gives

$$\mathbb{E}\left[\sum_{r,k}\overline{N}_{r,k}\right] = \frac{L}{\epsilon} + o\left(\frac{1}{\epsilon}\right),$$

which completes the proof. □

6.2 Proof of Lemma 7

Proof. We first bound the distance between the instant rate allocation and the load in the following claim.

Claim 2. *There exists a constant $B_2 > 0$ such that for any state $\boldsymbol{n}$,*

$$\sum_\ell p_\ell \|\widehat{\boldsymbol{b}}^{(\ell)} - \boldsymbol{b}^{(\ell)}\|^2 \leq B_2\|\boldsymbol{n}_\perp\|.$$

Proof of the Claim. Consider the term $\langle \boldsymbol{n}, (-S^T)(\boldsymbol{\rho}^{(0)} - \boldsymbol{n}\boldsymbol{x})\rangle$. By the proof of Lemma 4, we know that

$$\langle \boldsymbol{n}, (-S^T)(\boldsymbol{\rho}^{(0)} - \boldsymbol{n}\boldsymbol{x})\rangle \leq -\kappa_{\min}\sum_\ell p_\ell\|\widehat{\boldsymbol{b}}^{(\ell)} - \boldsymbol{b}^{(\ell)}\|^2. \quad (55)$$

On the other hand, by the duality principle for minimum norm problems [12],

$$\|\boldsymbol{n}_\perp\| = \sup_{\boldsymbol{y}\in\mathcal{K}^\circ:\ \|\boldsymbol{y}\|\leq 1}\langle \boldsymbol{n}, \boldsymbol{y}\rangle,$$

where $\mathcal{K}^\circ$ is the polar cone of the cone $\mathcal{K}$. It is easy to see that $\|S^T(\boldsymbol{\rho}^{(0)} - \boldsymbol{n}\boldsymbol{x})\|$ is well-defined. Let $\boldsymbol{y} = S^T(\boldsymbol{\rho}^{(0)} - \boldsymbol{n}\boldsymbol{x})/\|S^T(\boldsymbol{\rho}^{(0)} - \boldsymbol{n}\boldsymbol{x})\|$. We can verify that $\boldsymbol{y} \in \mathcal{K}^\circ$ since $\langle \boldsymbol{b}^{(\ell)}, S^T(\boldsymbol{\rho}^{(0)} - \boldsymbol{n}\boldsymbol{x})\rangle = -U_\ell \leq 0$ for all ℓ. Thus

$$-\langle \boldsymbol{n}, (-S^T)(\boldsymbol{\rho}^{(0)} - \boldsymbol{n}\boldsymbol{x})\rangle = \|S^T(\boldsymbol{\rho}^{(0)} - \boldsymbol{n}\boldsymbol{x})\|\langle \boldsymbol{n}, \boldsymbol{y}\rangle$$
$$\leq \|S^T(\boldsymbol{\rho}^{(0)} - \boldsymbol{n}\boldsymbol{x})\|\|\boldsymbol{n}_\perp\|. \quad (56)$$

Combining (55) and (56) gives

$$\sum_\ell p_\ell\|\widehat{\boldsymbol{b}}^{(\ell)} - \boldsymbol{b}^{(\ell)}\|^2 \leq \frac{\|S^T(\boldsymbol{\rho}^{(0)} - \boldsymbol{n}\boldsymbol{x})\|}{\kappa_{\min}}\|\boldsymbol{n}_\perp\|.$$

Since each entry of the rate allocation $\boldsymbol{n}\boldsymbol{x}$ can be bounded using a constant independent of $\boldsymbol{n}$ and ϵ, there exists a constant $B_2 > 0$ such that $\frac{\|S^T(\boldsymbol{\rho}^{(0)} - \boldsymbol{n}\boldsymbol{x})\|}{\kappa_{\min}} \leq B_2$. Therefore,

$$\sum_\ell p_\ell\|\widehat{\boldsymbol{b}}^{(\ell)} - \boldsymbol{b}^{(\ell)}\|^2 \leq B_2\|\boldsymbol{n}_\perp\|,$$

which completes the proof of the claim. □

Next we bound $|\alpha_\ell^s - p_\ell|$ using this claim. We first write each $(n_\perp^s)_{r,k} = n_{r,k} - (n_{\shortparallel}^s)_{r,k}$ in the following form

$$(n_\perp^s)_{r,k} = n_{r,k}x_r\sum_{l:l\in r}p_\ell - \rho_{r,k}^{(0)}\sum_{l:l\in r}\alpha_\ell^s$$
$$= \rho_{r,k}^{(0)}\sum_{l:l\in r}(p_\ell - \alpha_\ell^s) + (n_{r,k}x_r - \rho_{r,k}^{(0)})\sum_{l:l\in r}p_\ell.$$

Then for each (r,k),

$$\left|\sum_{l:l\in r}(\alpha_\ell^s - p_\ell)\right| \leq \frac{1}{\rho_{r,k}^{(0)}}|n_{r,k}x_r - \rho_{r,k}^{(0)}|\sum_{l:l\in r}p_\ell + \left|\frac{(n_\perp^s)_{r,k}}{\rho_{r,k}^{(0)}}\right| \quad (57)$$

By the claim above,

$$\sum_r\sum_{k:k\in[K_r]}(n_{r,k}x_r - \rho_{r,k}^{(0)})^2\sum_{l:l\in r}p_\ell = \sum_\ell p_\ell\|\widehat{\boldsymbol{b}}^{(\ell)} - \boldsymbol{b}^{(\ell)}\|^2$$
$$\leq B_2\|\boldsymbol{n}_\perp\|.$$

Since each summand on the left hand side is nonnegative, we have that for each (r,k),

$$(n_{r,k}x_r - \rho_{r,k}^{(0)})^2\sum_{l:l\in r}p_\ell \leq B_2\|\boldsymbol{n}_\perp\|.$$

Inserting this to (57) we get

$$\left|\sum_{l:l\in r}(\alpha_\ell^s - p_\ell)\right|$$
$$\leq \frac{\sqrt{B_2}}{\rho_{r,k}^{(0)}}\|\boldsymbol{n}_\perp\|^{1/2}\left(\sum_{l:l\in r}p_\ell\right)^{1/2} + \left|\frac{(n_\perp^s)_{r,k}}{\rho_{r,k}^{(0)}}\right|$$
$$\overset{(a)}{\leq} \frac{\sqrt{B_2}}{\rho_{r,k}^{(0)}\sqrt{C_{\min}}}\|\boldsymbol{n}_\perp\|^{1/2}\left(\sum_{r,k}n_{r,k}\right)^{1/2} + \left|\frac{(n_\perp^s)_{r,k}}{\rho_{r,k}^{(0)}}\right|,$$

where (a) follows from that $\sum_{l:l\in r} p_\ell \leq \sum_\ell p_\ell$ and $\sum_\ell p_\ell C_\ell = \sum_{r,k} n_{r,k}$. Since $\|\boldsymbol{n}_\perp^s\| \leq \|\boldsymbol{n}_\perp\| \leq \|\boldsymbol{n}\|$, (a) indicates that there exists a constant B_4 such that

$$\left|\sum_{l:l\in r}(\alpha_\ell^s - p_\ell)\right| \leq B_4\|\boldsymbol{n}_\perp\|^{1/2}\left(\sum_{r,k} n_{r,k}\right)^{1/2}.$$

Next we bound $|\alpha_\ell^s - p_\ell|$ for each link ℓ. Recall that H is the routing matrix defined as $H = (h_{lr})_{l\in\mathcal{L},r\in\mathcal{R}}$ with $h_{lr} = 1$ if $l \in r$, and $h_{lr} = 0$ otherwise. Since we assume that H has full row rank, HH^T is invertible. Let $\overline{h}^{(l)T}$ be the l^{th} row of $(HH^T)^{-1}H$. Then $\overline{h}^{(l)T}H^T = \overline{e}^{(\ell)}$, where $\overline{e}^{(\ell)}$ is a $L \times 1$ vector with the l^{th} entry being 1 and other entries being 0. Thus

$$\alpha_\ell^s - p_\ell = \overline{h}^{(l)T}H^T(\boldsymbol{\alpha}^s - \boldsymbol{p}) = \sum_r \overline{h}_r^{(\ell)} \sum_{l':l'\in r}(\alpha_{l'}^s - p_{l'}).$$

Therefore, we have the following bound

$$\begin{aligned}|\alpha_\ell^s - p_\ell| &\leq \sum_r \left|\overline{h}_r^{(\ell)}\right| \left|\sum_{l':l'\in r}(\alpha_{l'}^s - p_{l'})\right| \\ &\leq \left(\sum_r \left|\overline{h}_r^{(\ell)}\right|\right) B_4\|\boldsymbol{n}_\perp\|^{1/2}\left(\sum_{r,k} n_{r,k}\right)^{1/2}.\end{aligned}$$

Note that $\overline{h}_r^{(\ell)}$'s are constants independent of ϵ, i.e., $B_3 = (\sum_r |\overline{h}_r^{(\ell)}|)B_4$ is a constant. Thus,

$$|\alpha_\ell^s - p_\ell| \leq B_3\|\boldsymbol{n}_\perp\|^{1/2}\left(\sum_{r,k} n_{r,k}\right)^{1/2},$$

which completes the proof of Lemma 7.

7. CONCLUSIONS AND FUTURE WORK

In this paper, we studied a bandwidth sharing network under the proportionally fair bandwidth allocation policy for a general, dense class of file-size distributions. We obtained asymptotically tight bounds on the expected number of files in steady state in the heavy-traffic regime. These bounds show that the mean delay of file transfers under proportionally fair policy in heavy traffic does not depend on file-size distributions beyond the mean file-sizes, which gives delay insensitivity of proportionally fair policy in heavy traffic. Our results indicate that the backlog bound given by diffusion approximation is valid in steady state, thus complementing the diffusion approximation result of Vlasiou et al. [20].

With these results, some interesting extensions deserve further exploration. Our state-space collapse result is for steady states of the systems, so it gives a possible direction for proving the interchange of limits for the diffusion approximation in [20], which still remains an open problem. We are also interested in bounds on higher moments of the backlog, which may also be obtained using the drift-based framework, since Eryilmaz and Srikant [7] derived bounds on higher moments of the backlog for join-the-shortest-queue and MaxWeight in their settings.

8. ACKNOWLEDGMENT

This work was supported in part by NSF Grants ECCS-1609202, ECCS-1739344, ECCS-1739189, ECCS-1609370, CMMI-1562276, CIF-1409106, the U.S. Army Research Office (ARO Grant No. W911NF-16-1-0259), and the U.S. Office of Naval Research (ONR Grant No. N00014-15-1-2169).

References

[1] S. Asmussen. *Applied Probability and Queues*. Springer-Verlag New York, 2003.

[2] D. Bertsimas, D. Gamarnik, and J. N. Tsitsiklis. Performance of multiclass markovian queueing networks via piecewise linear lyapunov functions. *Ann. Appl. Probab.*, 11(4):1384–1428, 11 2001.

[3] T. Bonald and L. Massoulié. Impact of fairness on internet performance. In *Proc. ACM SIGMETRICS Int. Conf. Measurement and Modeling of Computer Systems*, pages 82–91, Cambridge, MA, 2001.

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

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

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

[7] A. Eryilmaz and R. Srikant. Asymptotically tight steady-state queue length bounds implied by drift conditions. *Queueing Syst.*, 72(3-4):311–359, Dec. 2012.

[8] B. Hajek. Hitting-time and occupation-time bounds implied by drift analysis with applications. *Adv. Appl. Probab.*, 14(3):502–525, 1982.

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

[10] F. Kelly. Charging and rate control for elastic traffic. *Eur. T. Telecommun.*, 8(1):33–37, 1997.

[11] Y. Lu, S. T. Maguluri, M. S. Squillante, and T. Suk. On optimal weighted-delay scheduling in input-queued switches. *arXiv:1704.02302 [math.OC]*, Apr. 2017.

[12] D. G. Luenberger. *Optimization by Vector Space Methods*. John Wiley & Sons, 1969.

[13] S. T. Maguluri and R. Srikant. Heavy traffic queue length behavior in a switch under the maxweight algorithm. *Stoch. Syst.*, 6(1):211–250, 2016.

[14] S. T. Maguluri, R. Srikant, and L. Ying. Heavy traffic optimal resource allocation algorithms for cloud computing clusters. *Perform. Eval.*, 81:20–39, 2014.

[15] S. T. Maguluri, S. K. Burle, and R. Srikant. Optimal heavy-traffic queue length scaling in an incompletely saturated switch. In *Proc. ACM SIGMETRICS Int. Conf. Measurement and Modeling of Computer Systems*, pages 13–24, Antibes Juan-les-Pins, France, 2016.

[16] L. Massoulié and J. W. Roberts. Bandwidth sharing and admission control for elastic traffic. *Telecommun. Syst.*, 15(1):185–201, 2000.

[17] F. Paganini, A. Tang, A. Ferragut, and L. L. H. Andrew. Network stability under alpha fair bandwidth allocation with general file size distribution. *IEEE Trans. Autom. Control*, 57(3):579–591, Mar. 2012.

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

[19] E. D. Sontag. *Mathematical Control Theory: Deterministic Finite Dimensional Systems*. Springer-Verlag, 1990.

[20] M. Vlasiou, J. Zhang, and B. Zwart. Insensitivity of proportional fairness in critically loaded bandwidth sharing networks. *arXiv:1411.4841v2 [math.PR]*, 2014.

[21] C.-H. Wang, S. T. Maguluri, and T. Javidi. Heavy traffic queue length behavior in switches with reconfiguration delay. In *Proc. IEEE Int. Conf. Computer Communications (INFOCOM)*, Atlanta, GA, May 2017.

[22] W. Wang, K. Zhu, L. Ying, J. Tan, and L. Zhang. MapTask scheduling in MapReduce with data locality: Throughput and heavy-traffic optimality. *IEEE/ACM Trans. Netw.*, 24:190–203, Feb. 2016.

[23] W. Wang, S. T. Maguluri, R. Srikant, and L. Ying. Heavy-traffic delay insensitivity in connection-level models of data transfer with proportionally fair bandwidth sharing. Technical report, 2017.

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

[25] Q. Xie and Y. Lu. Priority algorithm for near-data scheduling: Throughput and heavy-traffic optimality. In *Proc. IEEE Int. Conf. Computer Communications (INFOCOM)*, pages 963–972, Hong Kong, China, Apr. 2015.

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

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

APPENDIX

A. PROOFS OF LEMMAS 1, 2 AND 3

Proof of Lemma 1. It suffices to prove that for each route r, M_r is well-defined and positive definite since M is block-diagonal.

We first prove that M_r is well-defined, i.e., the integral below is (entry-wise) finite:

$$M_r = \frac{1}{\lambda_r^{(0)}} \int_0^{+\infty} \frac{\exp(S_r\sigma)\mathbf{1}\mathbf{1}^T \exp(S_r^T\sigma)}{\boldsymbol{\beta}_r(-S_r)^{-1}\exp(S_r\sigma)\mathbf{1}} d\sigma.$$

Let $\boldsymbol{P}(\sigma) = \exp(S_r\sigma)\mathbf{1}$ and $\chi = (-S_r)^{-1}$. Then from the properties of phase-type distributions we know that $\chi_{k_1k_2}$ is the expected time spent in phase k_2 given that phase k_1 is the initial state. Let $\overline{\boldsymbol{\chi}} = (\boldsymbol{\beta}_r\chi)^T$. Then $\overline{\chi}_k$ is the expected time spent in phase k for the initial distribution $\boldsymbol{\beta}_r$. Therefore, $\overline{\chi}_k > 0$ for all $k \in [K_r]$. With the above notation, the (k_1, k_2)th entry of M_r can be written as

$$\begin{aligned}(M_r)_{k_1,k_2} &= \frac{1}{\lambda_r^{(0)}} \int_0^{+\infty} \frac{P_{k_1}(\sigma)P_{k_2}(\sigma)}{\sum_k \overline{\chi}_k P_k(\sigma)} d\sigma \\ &\le \frac{1}{\lambda_r^{(0)}} \int_0^{+\infty} \frac{P_{k_1}(\sigma)P_{k_2}(\sigma)}{\overline{\chi}_{k_2} P_{k_2}(\sigma)} d\sigma \\ &= \frac{1}{\lambda_r^{(0)}\overline{\chi}_{k_2}} \int_0^{+\infty} P_{k_1}(\sigma) d\sigma.\end{aligned}$$

By our assumptions, S_r is an upper triangular matrix with all the main diagonal entries being negative. So S_r is invertible and $\lim_{\sigma\to+\infty}\exp(S_r^T\sigma)$ is an all-zero matrix. Therefore, $P_{k_1}(\sigma)$ is integrable and thus M_r is well-defined.

Next we prove that M_r is positive definite. Let $G(u)$ denote the complementary cumulative distribution function (CCDF) of the file size distribution on route r. Then

$$G(u) = \boldsymbol{\beta}_r \exp(S_r u)\mathbf{1}, \quad \int_0^{+\infty} G(u)du = \frac{1}{\overline{\mu}_r}.$$

The denominator inside the integral of M_r can be written as

$$\boldsymbol{\beta}_r(-S_r)^{-1}\exp(S_r\sigma)\mathbf{1} = \frac{1}{\overline{\mu}_r} - \int_0^{\sigma} G(u)du,$$

which is positive for all $\sigma \ge 0$. Therefore, it is obvious that M_r is positive semi-definite. Further, M_r is positive definite if and only if there exists no $\boldsymbol{y} \neq \mathbf{0}$ such that

$$\boldsymbol{y}^T \exp(S_r\sigma)\mathbf{1} = \mathbf{0}, \text{ for all } \sigma \ge 0, \tag{58}$$

where $\mathbf{0}$ is an all-zero vector with dimension $K_r \times 1$. If we view the pair $(S_r, \mathbf{1})$ as the (A, B) matrix of a control system, (58) is equivalent to the controllability of the system (A, B) [19]. By the Popov-Belevitch-Hautus (PBH) test (also referred to as Hautus Lemma) in control theory [19], this is equivalent to that $\text{rank}[\lambda I - S_r, \mathbf{1}] = K_r$ for each eigenvalue λ of the matrix S_r, which is further equivalent to that S_r^T has no eigenvector $\boldsymbol{v}$ such that $\boldsymbol{v}^T\mathbf{1} = 0$. Now let us look at the eigenvectors of S_r^T. With the mixture Erlang assumption, recall that S_r^T has a block-diagonal structure given in (8) with rate $\mu_r^{(b)}$ for each block b, where $\mu_r^{(b)}$'s are positive and distinct. Then the $-\mu_r^{(b)}$'s are the eigenvalues of S_r^T. Let $\boldsymbol{v}$ be an eigenvector associated with the eigenvalue $-\mu_r^{(b)}$. Then $\boldsymbol{v}$ satisfies that

$$\begin{bmatrix} (S_r^{(1)})^T + \mu_r^{(b)}I & 0 & \cdots & 0 \\ 0 & (S_r^{(2)})^T + \mu_r^{(b)}I & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & (S_r^{(B_r)})^T + \mu_r^{(b)}I \end{bmatrix} \boldsymbol{v}$$
$$= \mathbf{0}.$$

Since $(S_r^{(b')})^T + \mu_r^{(b)} I$ is full rank for all $b' \neq b$ and

$$(S_r^{(b)})^T + \mu_r^{(b)} I = \begin{bmatrix} 0 & 0 & \cdots & 0 \\ 2\mu_r^{(b)} & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 2\mu_r^{(b)} & 0 \end{bmatrix},$$

the eigenvector $\boldsymbol{v}$ has only one nonzero entry. Then $\boldsymbol{v}^T\mathbf{1} \neq 0$. This completes the proof that M_r is positive definite. □

Proof of Lemma 2. By the definitions of $\boldsymbol{\rho}_r^{(0)}$ and M_r in (17) and (27), respectively,

$$\begin{aligned}
&(\boldsymbol{\rho}_r^{(0)})^T M_r(-S_r)^T \\
&= \boldsymbol{\beta}_r(-S_r)^{-1} \int_0^{+\infty} \frac{\exp(S_r\sigma)\mathbf{1}\mathbf{1}^T \exp(S_r^T\sigma)(-S_r)^T}{\boldsymbol{\beta}_r(-S_r)^{-1}\exp(S_r\sigma)\mathbf{1}} d\sigma \\
&= \int_0^{+\infty} \mathbf{1}^T \exp(S_r^T\sigma)(-S_r)^T d\sigma \\
&= -\mathbf{1}^T \exp(S_r^T\sigma)\Big|_0^{+\infty} \\
&= \mathbf{1}^T.
\end{aligned}$$

Here we have used that fact that $\lim_{\sigma\to+\infty}\exp(S_r^T\sigma)$ is an all-zero matrix since S_r is an upper triangular matrix and its main diagonal entries are all negative. This completes the proof. □

Proof of Lemma 3. We first derive another representation of $M_r(-S_r^T) + (-S_r)M_r$. Let

$$M_r(t) = \frac{1}{\lambda_r^{(0)}} \int_0^t \frac{\exp(S_r\sigma)\mathbf{1}\mathbf{1}^T\exp(S_r^T\sigma)}{\boldsymbol{\beta}_r(-S_r)^{-1}\exp(S_r\sigma)\mathbf{1}} d\sigma.$$

Then

$$M_r(-S_r^T) + (-S_r)M_r = \lim_{t\to+\infty}\Big(M_r(t)(-S_r^T) + (-S_r)M_r(t)\Big).$$

We can verify that

$$\begin{aligned}
&M_r(t)(-S_r^T) + (-S_r)M_r(t) \\
&= -\frac{1}{\lambda_r^{(0)}} \frac{\exp(S_r t)\mathbf{1}\mathbf{1}^T\exp(S_r^T t)}{\boldsymbol{\beta}_r(-S_r)^{-1}\exp(S_r t)\mathbf{1}} + \frac{1}{\overline{\rho}_r}\mathbf{1}\mathbf{1}^T \\
&\quad + \frac{1}{\lambda_r^{(0)}} \int_0^t \frac{\exp(S_r\sigma)\mathbf{1}\mathbf{1}^T\exp(S_r^T\sigma)}{\boldsymbol{\beta}_r(-S_r)^{-1}\exp(S_r\sigma)\mathbf{1}} \\
&\qquad \cdot \frac{\boldsymbol{\beta}_r\exp(S_r\sigma)\mathbf{1}}{\boldsymbol{\beta}_r(-S_r)^{-1}\exp(S_r\sigma)\mathbf{1}} d\sigma. \qquad (59)
\end{aligned}$$

We have proved that M_r is well-defined, so

$$\lim_{t\to0}\left(-\frac{1}{\lambda_r^{(0)}} \frac{\exp(S_r t)\mathbf{1}\mathbf{1}^T\exp(S_r^T t)}{\boldsymbol{\beta}_r(-S_r)^{-1}\exp(S_r t)\mathbf{1}}\right) = (0)_{K_r\times K_r},$$

where $(0)_{K_r\times K_r}$ is the all-zero $K_r \times K_r$ matrix.

Now it suffices to prove that there exists a constant $\kappa_r > 0$ such that for any $\sigma \geq 0$,

$$\frac{\boldsymbol{\beta}_r\exp(S_r\sigma)\mathbf{1}}{\boldsymbol{\beta}_r(-S_r)^{-1}\exp(S_r\sigma)\mathbf{1}} \geq 2\kappa_r, \qquad (60)$$

since combining this with (59) implies that for any $\boldsymbol{y} \in \mathbb{R}^{K_r}$,

$$\begin{aligned}
&\boldsymbol{y}^T\left(\frac{1}{2}M_r(-S_r^T) + \frac{1}{2}(-S_r)M_r - \kappa_r M_r\right)\boldsymbol{y} \\
&= \lim_{t\to+\infty} \boldsymbol{y}^T\left(\frac{1}{2}M_r(t)(-S_r^T) + \frac{1}{2}(-S_r)M_r(t) - \kappa_r M_r(t)\right)\boldsymbol{y} \\
&\geq \lim_{t\to+\infty} \frac{1}{2\lambda_r^{(0)}} \int_0^t \frac{\boldsymbol{y}^T\exp(S_r\sigma)\mathbf{1}\mathbf{1}^T\exp(S_r^T\sigma)\boldsymbol{y}}{\boldsymbol{\beta}_r(-S_r)^{-1}\exp(S_r\sigma)\mathbf{1}} \cdot (2\kappa_r)d\sigma \\
&\quad - \kappa_r\boldsymbol{y}^T M_r\boldsymbol{y} + \frac{1}{2\overline{\rho}_r}(\boldsymbol{y}^T\mathbf{1})^2 \\
&= \frac{1}{2\overline{\rho}_r}(\boldsymbol{y}^T\mathbf{1})^2 \\
&\geq 0.
\end{aligned}$$

Let $g(u)$ and $G(u)$ denote the probability density function (PDF) and the complementary cumulative distribution function (CCDF) of the file size distribution on route r, respectively. Then

$$G(\sigma) = \int_\sigma^{+\infty} g(u)du = \boldsymbol{\beta}_r\exp(S_r\sigma)\mathbf{1},$$

and

$$\int_\sigma^{+\infty} G(u)du = \boldsymbol{\beta}_r(-S_r)^{-1}\exp(S_r\sigma)\mathbf{1}.$$

Thus (60) is equivalent to that there exists a constant $\kappa_r > 0$ such that for any $u \geq 0$, $\frac{g(u)}{G(u)} \geq 2\kappa_r$, i.e., the hazard function is lower bounded by $2\kappa_r$. Since S_r is the sub-transition rate matrix associated with a mixture of Erlang distributions, the eigenvalues of S_r are the rates of the phases, denoted by $\mu_{r,k}$'s with $k \in [K_r]$, which are all positive. Consider the Jordan canonical form $S_r = \Phi J \Phi^{-1}$. Then $\exp(S_r u) = \Phi\exp(Ju)\Phi^{-1}$, where the (i,j)th entry of $\exp(Ju)$ is either $e^{-\mu_{r,i}u}u^{j-1}/(j-1)!$ or 0. So $G(u)$ can be written as

$$G(u) = \sum_{i,j\in[K_r]} c_{ij}e^{-\mu_{r,i}u}\frac{u^{j-1}}{(j-1)!}$$

for some constants c_{ij}, and thus

$$g(u) = -G'(u) = \sum_{i,j\in[K_r]} c_{ij}\mu_{r,i}e^{-\mu_{r,i}u}\frac{u^{j-1}}{(j-1)!}.$$

Therefore, $\lim_{u\to+\infty}\frac{g(u)}{G(u)} \geq \min_{k\in[K_r]}\mu_{r,k} > 0$. It can be verified that $g(u) > 0$ and $G(u) > 0$ for any $u \geq 0$. Thus there exists a constant $\kappa_r > 0$ such that for any $u \geq 0$, $\frac{g(u)}{G(u)} \geq 2\kappa_r$. □