On Heavy-Traffic Optimal Scaling of *c*-Weighted MaxWeight Scheduling in Input-Queued Switches

Yingdong Lu, Siva Theja Maguluri, Mark S. Squillante, Tonghoon Suk

Abstract—**We consider the asymptotically optimal control of input-queued switches under a cost-weighted variant of MaxWeight scheduling, for which we establish theoretical properties that include showing the algorithm exhibits optimal heavy-traffic queue-length scaling. Our results are expected to be of theoretical interest more broadly than input-queued switches.**

I. INTRODUCTION

Data centers form the backbone of today's big data revolution. The interchange of data within a data center is facilitated by huge $n \times n$ input-queued switches (IQSs) [9]. Hence, understanding scheduling problems in IQSs is essential for real-world data center networks [8]. MaxWeight scheduling, first introduced for wireless networks [11] and then for IQSs [7], is well-known for being throughput optimal. However, the issue of delay-optimal scheduling for switches is less clear. MaxWeight scheduling has been shown to be asymptotically optimal in heavy traffic for an objective function of the summation of the squares of queue lengths (QLs) under complete resource pooling [10]. MaxWeight scheduling has also been shown to have optimal scaling in heavy traffic for an objective function of the summation of QLs under all ports saturated [6], which was then extended to the case of incompletely saturated ports [5]. Otherwise, the question of delay-optimal scheduling in IQSs remains open for general objective functions.

We seek to gain fundamental insights on the delay-optimal properties of a generalized MaxWeight scheduling policy in $n \times n$ IQSs in which a linear cost function of QL (delay) is associated with each queue. Specifically, we extend the results in [6] to include per-queue costs and prove that a cost-weighted generalization of MaxWeight scheduling has optimal scaling in heavy traffic for an objective function consisting of a linear function of the steady-state average QLs. Our results shed light on the delay optimality of MaxWeight scheduling and its variants more generally, including extensions to more general objective functions. In addition, our results are expected to be of theoretical interest beyond IQS and related models as implied by our extension of the drift method, first introduced in [2] and together with its subsequent developments. This paper extends an earlier version [3] to include a tighter universal lower bound (l.b.) on the average weighted QL and an explicit expression for the weighted sum of QLs in heavy traffic in general $n \times n$ IQSs.

Y. Lu, M.S. Squillante and T. Suk are with Mathematical Sciences, IBM Research, {`yingdong, mss, tonghoon.suk`}`@us.ibm.com`. S.T. Maguluri is with School of Industrial and Systems Engineering, Georgia Institute of Technology, `siva.theja@gatech.edu`; this research was supported in part by the NSF grant CCF 1850439.

§II presents our mathematical model and formulation, and §III presents our analysis of cost-weighted MaxWeight scheduling, followed by conclusions and some proofs.

II. MODEL AND FORMULATION

Consider an IQS with n input ports and n output ports. Each input port has a queue associated with every output port that stores packets waiting to be transmitted to the output port. Let $(i,j) \in \mathcal{I}:=\{(i,j) : i,j \in [n]\}$, $[n]:=\{1,\ldots,n\}$, index the queue associated with the ith input port and the jth output port. Let $c_{ij} > 0$ denote the cost associated with queue (i,j) and define $\mathbf{c} := (c_{ij}) \in \mathbb{R}^{n^2}$. Further define a new inner product on $\mathbb{R}^{n^2}$ with respect to (w.r.t.) the vector $\mathbf{c}$ as follows

$$\langle x, y \rangle_{\mathbf{c}} := \sum_{ij} c_{ij} x_{ij} y_{ij}. \tag{1}$$

Hence, the corresponding norm of a vector $\mathbf{x} \in \mathbb{R}^{n^2}$ is given by $||\mathbf{x}||_{\mathbf{c}}^2 = \sum_{ij} c_{ij} x_{ij}^2$.

Packets arrive at queue (i,j) from a stochastic process. Time is slotted and denoted by $t \in \mathbb{Z}_+:=\{0,1,\ldots\}$. At each time t, a scheduling policy selects a set of queues from which to simultaneously transmit packets under the constraints: (1) At most one packet can be transmitted from an input port; (2) At most one packet can be transmitted to an output port. We refer to a *schedule* as a subset of queues that satisfies these constraints.

A schedule is formally described by an n^2-dimensional binary vector $\mathbf{s} = (s_{ij})_{(i,j)\in\mathcal{I}}$ such that $s_{ij} = 1$ if queue (i,j) is in the schedule, and $s_{ij} = 0$ otherwise. Let $\mathcal{P}$ denote the set of all maximal schedules, i.e.,

$$\mathcal{P} = \left\{ \left[\mathbf{s} \in \{0,1\}^{n^2}\right] \ : \ \begin{array}{l} \sum_{j\in[n]} s_{ij} = 1, \ \forall i \in [n] \\ \sum_{i\in[n]} s_{ij} = 1, \ \forall j \in [n] \end{array} \right\},$$

and $\mathbf{S}(t) \in \mathcal{P}$ the schedule for period t under the **c**-weighted MaxWeight scheduling algorithm defined below. Let $Q_{ij}(t) \in \mathbb{Z}_+$ denote the length of the infinite capacity queue (i,j) at time t under this MaxWeight policy and $A_{ij}(t) \in \mathbb{Z}_+$ the number of arrivals to queue (i,j) during $[t, t+1)$. The queueing dynamics then can be expressed as

$$Q_{ij}(t+1) = Q_{ij}(t) + A_{ij}(t) - S_{ij}(t) + U_{ij}(t), \tag{2}$$

where $U_{ij}(t) \in \mathbb{Z}_+$ denotes the unused service for queue (i,j) at time t. Any selected schedule is always a maximal schedule in $\mathcal{P}$, which results in an unused service at those queues with no packets to serve. We assume that $\{A_{ij}(t) : t \in \mathbb{Z}_+, (i,j) \in \mathcal{I}\}$ are independent random variables (r.v.s) and that, for each fixed $(i,j) \in \mathcal{I}$, $\{A_{ij}(t) : t \in \mathbb{Z}_+\}$ are identically distributed with $\mathbb{E}[A_{ij}(t)] =$

λ_{ij}. Define $\mathbf{Q}(t):=(Q_{ij}(t))_{(i,j)\in\mathcal{I}}$, $\mathbf{A}(t):=(A_{ij}(t))_{(i,j)\in\mathcal{I}}$, $\mathbf{S}(t):=(S_{ij}(t))_{(i,j)\in\mathcal{I}}$ and $\mathbf{U}(t):=(U_{ij}(t))_{(i,j)\in\mathcal{I}}$.

Consider the above IQS model under the **c**-weighted MaxWeight scheduling Algorithm 1.

Algorithm 1 c-Weighted MaxWeight Scheduling

Let $\mathbf{c} \in \mathbb{R}^{n^2}$ be a given positive weight (cost) vector, i.e., $c_{ij} > 0$, $\forall i, j$. Then, in every time slot t under the **c**-weighted MaxWeight algorithm, each queue is assigned a weight $c_{ij}Q_{ij}(t)$ and a schedule with the maximum weight is chosen, namely

$$\mathbf{S}(t) = \arg\max_{\mathbf{s}\in\mathcal{P}} \sum_{ij} c_{ij}Q_{ij}(t)s_{ij} = \arg\max_{\mathbf{s}\in\mathcal{P}} \langle \mathbf{Q}(t), \mathbf{s}\rangle_{\mathbf{c}}.$$

Ties are broken uniformly at random.

The objective function consists of minimizing a weighted summation of expected delay cost in steady state, based on which we establish delay-optimal properties of the **c**-weighted MaxWeight scheduling algorithm. Given the relationship between delays and QLs via Little's Law, we henceforth focus on cost as a function of the QLs. The objective function can then be expressed as $\min_{\pi\in\mathcal{M}} \mathbb{E}\left[\sum_{(i,j)\in\mathcal{I}} c_{ij}\overline{Q}_{ij}^{\pi}\right]$, where $\mathcal{M}$ denotes the set of all stationary policies. In this average cost case we also know that, as long as the arrival rate vector is within the capacity region [6], there exists an expected average cost optimal stationary policy; see, e.g., [1]. Note that [6] considers the specific case of $c_{ij} = 1, \forall i, j$.

III. HEAVY TRAFFIC ANALYSIS

We study the switch system when the arrival rate vector $\boldsymbol{\lambda}$ approaches a point on the boundary of the capacity region such that all the ports are saturated. In other words, we consider the arrival rate vector approaching the face $\mathcal{F}$ of the capacity region where

$$\mathcal{F} = \left\{ \boldsymbol{\lambda} \in \mathbb{R}_{+}^{n^2} : \langle \boldsymbol{\lambda}, \mathbf{e}_{\mathbf{c}}^{(i)}\rangle_{\mathbf{c}} = 1, \langle \boldsymbol{\lambda}, \tilde{\mathbf{e}}_{\mathbf{c}}^{(j)}\rangle_{\mathbf{c}} = 1, \forall i, j \in [n] \right\},$$

and where $\mathbf{e}_{\mathbf{c}}^{(i)} = \{\mathbf{x} \in \mathbb{R}^{n^2}, x_{ij} = \frac{1}{c_{ij}}, x_{i'j} = 0, \forall i' \neq i\}$ and $\tilde{\mathbf{e}}_{\mathbf{c}}^{(j)} = \{\mathbf{x} \in \mathbb{R}^{n^2}, x_{ij} = \frac{1}{c_{ij}}, x_{ij'} = 0, \forall j' \neq j\}$.

We will obtain an exact expression for the heavy traffic scaled weighted sum of QLs under the **c**-weighted MaxWeight algorithm in heavy traffic, along similar lines as [5] but with the dot product redefined in (1) and related technical differences. To obtain the desired result for heavy traffic performance under the **c**-weighted MaxWeight algorithm, we first provide a universal l.b. on the average weighted QL. We then establish that the QL vector concentrates close to a lower dimensional cone in the heavy traffic limit, which is called state space collapse. Finally, we exploit this state space collapse result to obtain an exact expression for the heavy traffic scaled weighted sum of QLs in heavy traffic. The proofs of the main results in III-B and III-C follow a similar logical approach to that in [6], though with derivations based on new methods

required to address important technical differences and details for the **c**-*weighted* MaxWeight algorithm due to the modified dot product, norms and projections w.r.t. **c**.

Throughout, we consider a *base family of switch systems* having arrival processes $\mathbf{A}^{(\epsilon)}(t)$ parameterized by $0 < \epsilon < 1$ such that the mean arrival rate vector is given by $\boldsymbol{\lambda}^{(\epsilon)} = \mathbb{E}[\mathbf{A}^{(\epsilon)}(t)] = (1-\epsilon)\boldsymbol{\nu}$ for some $\boldsymbol{\nu}$ in the relative interior of $\mathcal{F}$ with $\nu_{\min}:=\min_{ij}\nu_{ij} > 0$, and the arrival variance vector is given by $\mathrm{Var}(\mathbf{A}^{(\epsilon)}) = (\boldsymbol{\sigma}^{(\epsilon)})^2 < \infty$.

A. Universal Lower Bound

Consider a priority queueing system $\tilde{\mathbf{P}}^{(\ell)}$ under a fixed priority ordering $\mathbf{p}^{(\ell)} \in \mathcal{L}$ among all $L = n!$ schedules in the set $\mathcal{P}$, indexed by ℓ, where $\mathcal{L}$ is the set of all possible priority orderings of the L schedules. Let $\tilde{Q}_{\ell,l}(t)$, $l = 1,\ldots,L$, denote the QL process of the lth highest priority class in the system $\tilde{\mathbf{P}}^{(\ell)}$ under ordering $\mathbf{p}^{(\ell)}$. Let $\mathbf{m}^{(\ell)}(l)$ be the set of queues (i, j) of the switch contained within the lth priority class in $\mathbf{p}^{(\ell)}$, and $\tilde{A}_{\ell,l}(t)$ the composite arrival r.v. from all $A_{ij\in\mathbf{m}^{(\ell)}(l)}(t)$ that leads to the smallest QL among the queues (i, j) in the set $\mathbf{m}^{(\ell)}(l)$. Then, for the system $\tilde{\mathbf{P}}^{(\ell)}$, we can write an expression for the QL process of the highest priority class 1 as

$$[\tilde{Q}_{\ell,1}(t+1)]^2 - [\tilde{Q}_{\ell,1}(t)]^2$$
$$=[\tilde{Q}_{\ell,1}(t) + \tilde{A}_{\ell,1}(t+1) - 1 + \tilde{V}_{\ell,1}(t+1)]^2 - [\tilde{Q}_{\ell,1}(t)]^2$$
$$=[\tilde{Q}_{\ell,1}(t) + \tilde{A}_{\ell,1}(t+1) - 1]^2 - \tilde{V}_{\ell,1}^2(t+1) - \tilde{Q}_{\ell,1}^2(t)$$
$$=[\tilde{A}_{\ell,1}(t+1) - 1]^2 + 2\tilde{Q}_{\ell,1}(t)[\tilde{A}_{\ell,1}(t+1) - 1] - \tilde{V}_{\ell,1}^2(t+1)$$

where $\tilde{V}_{\ell,u}(t)$ denotes the time spent serving all lower priority classes $v > u$ and idling. From the relationship

$$[\tilde{Q}_{\ell,1}(t) + \tilde{A}_{\ell,1}(t+1) - 1 + \tilde{V}_{\ell,1}(t+1)]\tilde{V}_{\ell,1}(t+1) = 0,$$

we therefore have

$$[\tilde{Q}_{\ell,1}(t) + \tilde{A}_{\ell,1}(t+1) - 1]\tilde{V}_{\ell,1}(t+1) = -\tilde{V}_{\ell,1}^2(t+1).$$

Similarly, for the next highest priority class 2, we obtain

$$[\tilde{Q}_{\ell,2}(t+1)]^2 - [\tilde{Q}_{\ell,2}(t)]^2 = [\tilde{A}_{\ell,2}(t+1) - \tilde{V}_{\ell,1}(t+1)]^2$$
$$+ 2\tilde{Q}_{\ell,2}(t)[\tilde{A}_{\ell,1}(t+1) - \tilde{V}_{\ell,1}(t+1)] - \tilde{V}_{\ell,2}^2(t+1),$$

thus rendering in stationarity

$$\mathbb{E}[(1 - \tilde{A}_{\ell,1})^2] - 2\mathbb{E}[\tilde{Q}_{\ell,1}(1 - \tilde{A}_{\ell,1})] - \mathbb{E}[\tilde{V}_{\ell,1}^2] = 0.$$

Hence, $\tilde{Q}_{\ell,1}$ will be finite, and more specifically

$$\mathbb{E}[\tilde{Q}_{\ell,1}] = \frac{\mathbb{E}[(1 - \tilde{A}_{\ell,1})^2] - \mathbb{E}[\tilde{V}_{\ell,1}^2]}{1 - \mathbb{E}[\tilde{A}_{\ell,1}]} \leq \frac{\mathbb{E}[(1 - \tilde{A}_{\ell,1})^2]}{1 - \mathbb{E}[\tilde{A}_{\ell,1}]},$$

which then yields for $\tilde{Q}_{\ell,2}$ in stationarity

$$2\mathbb{E}[\tilde{Q}_{\ell,2}]\mathbb{E}[\tilde{V}_{\ell,1} - \tilde{A}_{\ell,2}] \geq \mathbb{E}[\tilde{A}_{\ell,2}^2] - 2\mathbb{E}[\tilde{A}_{\ell,2}]\mathbb{E}[\tilde{V}_{\ell,1}] - \mathbb{E}[\tilde{V}_{\ell,2}^2].$$

Continuing in this manner, we have in general for class l

$$\mathbb{E}[\tilde{A}_{\ell,l} - \tilde{V}_{\ell,l-1}^2] - 2\epsilon\mathbb{E}[\tilde{Q}_{\ell,l}] - \mathbb{E}[\tilde{V}_{\ell,l}^2] = 0, \quad \forall l = 2,\ldots,L.$$

Upon expanding the first term, we obtain

$$\mathbb{E}[\tilde{Q}_{\ell,l}] \geq \frac{\mathbb{E}[\tilde{A}_{\ell,l}^2] - 2\mathbb{E}[\tilde{A}_{\ell,l}]\mathbb{E}[\tilde{V}_{\ell,l-1}] - \mathbb{E}[\tilde{V}_{\ell,l}^2]}{2\epsilon}, \quad (3)$$

which, since we know $\mathbb{E}[\tilde{V}_{\ell,l}^2] = O(\epsilon)$, renders

$$\liminf_{\epsilon \downarrow 0} \epsilon \mathbb{E}[\tilde{Q}_{\ell,l}] \geq \mathbb{E}[\tilde{A}_{\ell,l}^2] - 2\mathbb{E}[\tilde{A}_{\ell,l}]\mathbb{E}[\tilde{V}_{\ell,l-1}]. \quad (4)$$

Let $\hat{\mathcal{Q}}_{\ell,l}^{(\epsilon)}$ and $\tilde{\mathcal{Q}}_{\ell,l}$ be the RHS of (3) and (4), respectively. Define $\hat{\mathcal{Q}}_{\ell,ij}^{(\epsilon)} := \min_{l:ij\in\mathbf{m}^{(\ell)}(l)} \hat{\mathcal{Q}}_{\ell,l}^{(\epsilon)}$ and $\hat{\mathcal{Q}}_{\ell,ij} := \min_{l:ij\in\mathbf{m}^{(\ell)}(l)} \tilde{\mathcal{Q}}_{\ell,l}$. We then have the desired universal l.b.

Proposition III.1. *Consider the base family of switches and fix a scheduling policy under which the system is stable for any $0 < \epsilon < 1$. Suppose the QL process $\mathbf{Q}^{(\epsilon)}(t)$ converges in distribution to a steady state random vector $\overline{\mathbf{Q}}^{(\epsilon)}$, and assume $(\boldsymbol{\sigma}^{(\epsilon)})^2 \to \boldsymbol{\sigma}^2$. Define*

$$\hat{\mathcal{Q}}_*^{(\epsilon)} := \min_{\mathbf{p}^{(\ell)}\in\mathcal{L}} \sum_{ij} c_{ij}\hat{\mathcal{Q}}_{\ell,ij}^{(\epsilon)}, \quad \hat{\mathcal{Q}}_* := \min_{\mathbf{p}^{(\ell)}\in\mathcal{L}} \sum_{ij} c_{ij}\hat{\mathcal{Q}}_{\ell,ij}.$$

Then, for each of these switch systems, the average weighted QL is lower bounded by $\mathbb{E}[\sum_{i,j} c_{ij}\overline{Q}_{ij}^{(\epsilon)}] \geq \hat{\mathcal{Q}}_*^{(\epsilon)}$, and, in the heavy-traffic limit as $\epsilon \downarrow 0$, we have

$$\liminf_{\epsilon \downarrow 0} \epsilon \mathbb{E}\left[\sum_{i,j} c_{ij}\overline{Q}_{ij}^{(\epsilon)}\right] \geq \hat{\mathcal{Q}}_*. \quad (5)$$

Proof. The overall average QL $\sum_l \mathbb{E}[\tilde{Q}_{\ell,l}]$ for each L-class priority queueing system $\tilde{\mathbf{P}}^{(\ell)}$ under ordering $\mathbf{p}^{(\ell)}$, $\forall\mathbf{p}^{(\ell)} \in \mathcal{L}$, forms the vertices of the performance region polytope in which must lie the overall average QL of any scheduling policy in the L-class queueing system. Since, by construction, the lth queue under any $\mathbf{p}^{(\ell)} \in \mathcal{L}$ can be scheduled whenever at least one queue (i,j) in $\mathbf{m}^{(\ell)}(l)$ has a packet, this polytope together with $\hat{\mathcal{Q}}_{\ell,ij}^{(\epsilon)}$ and $\hat{\mathcal{Q}}_{\ell,ij}$ provide a l.b. on the overall average QL of any scheduling policy in the original switch system. It follows that the average weighted QL under any scheduling policy in the switch is lower bounded by $\hat{\mathcal{Q}}_*^{(\epsilon)}$, with the corresponding heavy-traffic limit lower bounded by $\hat{\mathcal{Q}}_*$. □

Remark III.1. *The l.b. (5) – based on multiclass priority queues – improves upon the looser bound of $c_{\min}(\|\boldsymbol{\sigma}\|^2/2)$ – based on a single-class queue – established in [3].*

B. State Space Collapse

Since [6] considers $c_{ij} = 1, \forall i, j$, the state space collapse in our general case is to a different cone. To establish this state space collapse result, we first define the cone $\mathcal{K}_{\mathbf{c}}$ to be the cone spanned by the vectors $\mathbf{e}_{\mathbf{c}}^{(i)}$ and $\tilde{\mathbf{e}}_{\mathbf{c}}^{(j)}$, namely

$$\mathcal{K}_{\mathbf{c}} := \left\{ \mathbf{x} \in \mathbb{R}^{n^2} : x_{ij} = \frac{w_i + \tilde{w}_j}{c_{ij}}, \quad w_i, \tilde{w}_j \in \mathbb{R}_+ \right\}.$$

For any $\mathbf{x} \in \mathbb{R}^{n^2}$, define $\mathbf{x}_{\|\mathcal{K}_{\mathbf{c}}} := \arg\min_{\mathbf{y}\in\mathcal{K}_{\mathbf{c}}} \|\mathbf{x} - \mathbf{y}\|_{\mathbf{c}}$ to be the projection of $\mathbf{x}$ onto the cone $\mathcal{K}_{\mathbf{c}}$. The error after projection is denoted by $\mathbf{x}_{\perp\mathcal{K}_{\mathbf{c}}} = \mathbf{x} - \mathbf{x}_{\|\mathcal{K}_{\mathbf{c}}}$. To simplify the notation throughout the paper, we will write $\mathbf{x}_{\|_{\mathbf{c}}}$ to mean $\mathbf{x}_{\|\mathcal{K}_{\mathbf{c}}}$ and write $\mathbf{x}_{\perp_{\mathbf{c}}}$ to mean $\mathbf{x}_{\perp\mathcal{K}_{\mathbf{c}}}$. Let $\mathcal{S}_{\mathbf{c}}$ denote the space spanned by the cone $\mathcal{K}_{\mathbf{c}}$, or more formally

$$\mathcal{S}_{\mathbf{c}} = \left\{ \mathbf{x} \in \mathbb{R}^{n^2} : x_{ij} = \frac{w_i + \tilde{w}_j}{c_{ij}}, \quad w_i, \tilde{w}_j \in \mathbb{R} \right\}.$$

The projection of $\mathbf{x} \in \mathbb{R}^{n^2}$ onto the space $\mathcal{S}_{\mathbf{c}}$ is denoted by $\mathbf{x}_{\|\mathcal{S}_{\mathbf{c}}}$, with the error after projection denoted by $\mathbf{x}_{\perp\mathcal{S}_{\mathbf{c}}}$.

Now, consider the base family of switch systems under the $\mathbf{c}$-weighted MaxWeight scheduling algorithm with the maximum possible arrivals in any queue denoted by $A_{\max}$. Let the variance of the arrival process be such that $\|\boldsymbol{\sigma}^{(\epsilon)}\|^2 \leq \tilde{\sigma}^2$ for some $\tilde{\sigma}^2$ that is not dependent on ϵ. Let $\overline{\mathbf{Q}}^{(\epsilon)}$ denote the steady state random vector of the QL process for each switch system parameterized by ϵ. We then have the following proposition.

Proposition III.2. *For each system above with $0 < \epsilon \leq \nu'_{\min}$, the steady state QL vector satisfies*

$$\mathbb{E}\left[\|\overline{\mathbf{Q}}_{\perp_{\mathbf{c}}}^{(\epsilon)}\|^r\right] \leq (M_r)^r, \qquad \forall r \in \{1, 2, \ldots\},$$

where $\nu'_{\min}$ and M_r are functions of $r, \tilde{\sigma}, \boldsymbol{\nu}, A_{\max}, \nu_{\min}$ but independent of ϵ.

Proof. Omitting superscript (ϵ) to simplify the notation and clarify the presentation, our general approach consists of defining a Lyapunov function $W_{\perp_{\mathbf{c}}}(\overline{\mathbf{Q}}) := \|\overline{\mathbf{Q}}_{\perp_{\mathbf{c}}}\|_{\mathbf{c}}$ and its drift $\Delta W_{\perp_{\mathbf{c}}}(\overline{\mathbf{Q}}) := \left(W_{\perp_{\mathbf{c}}}(\mathbf{Q}(t+1)) - W_{\perp_{\mathbf{c}}}(\mathbf{Q}(t))\right)\mathbb{I}_{\{\mathbf{Q}(t)=\overline{\mathbf{Q}}\}}$, for all $\overline{\mathbf{Q}} \in \mathbb{R}^{n^2}$. Then, from Lemma A.1 in Appendix A, there exist positive numbers η, κ and D that depend on $\tilde{\sigma}, \boldsymbol{\nu}, A_{\max}$ and $\nu_{\min}$, but not on ϵ, such that

$$\mathbb{E}[\Delta W_{\perp_{\mathbf{c}}}(\overline{\mathbf{Q}})|\mathbf{Q}(t) = \overline{\mathbf{Q}}] \leq -\eta, \quad \forall\overline{\mathbf{Q}}, W_{\perp_{\mathbf{c}}}(\overline{\mathbf{Q}}) \geq \kappa,$$
$$\mathbb{P}[|\Delta W_{\perp_{\mathbf{c}}}(\overline{\mathbf{Q}})| \leq D] = 1, \qquad \forall\overline{\mathbf{Q}},$$

from which we derive, by Lemma 3 in [6],

$$\mathbb{E}\left[\|\overline{\mathbf{Q}}_{\perp_{\mathbf{c}}}^{(\epsilon)}\|^r\right] \leq (2\kappa)^4 + r\left(\frac{D+\eta}{\eta}\right)^r (4D)^r$$
$$\leq (2\kappa)^r + \sqrt{r}e\left(4D\frac{r}{e}\frac{D+\eta}{\eta}\right)^r$$
$$\leq 2\left(\max\left\{2\kappa, (\sqrt{r}e)^{1/r}4D\frac{r}{e}\frac{D+\eta}{\eta}\right\}\right)^r$$
$$= (M_r)^r = \left(2^{1/r}\max\left\{2\kappa, (\sqrt{r}e)^{1/r}4D\frac{r}{e}\frac{D+\eta}{\eta}\right\}\right)^r;$$

this is a function of $r, \tilde{\sigma}, \boldsymbol{\nu}, A_{\max}$ and $\nu_{\min}$, but independent of ϵ, hence completing the proof. □

Remark III.2. *The special case of $\mathbf{c} = \mathbf{1}$ renders the standard MaxWeight algorithm and our results coincide with the state space collapse in [6]. More generally, the capacity region and maximal face $\mathcal{F}$ are not dependent on the choice of the weight vector $\mathbf{c}$. However, for any positive weight vector, the state space collapses into the normal cone of the face $\mathcal{F}$ w.r.t. the dot product defined by the weight vector $\mathbf{c}$. This cone depends upon the choice of $\mathbf{c}$, and thus the choice of the weight vector "tilts" the cone of collapse.*

C. Weighted Sum of Queue Lengths in Heavy Traffic

We next exploit the above state space collapse result to obtain an exact expression for the heavy traffic scaled weighted sum of QLs in heavy traffic. Our main results are provided in the following theorem, with the next section

providing a general matrix solution approach to calculate the corresponding limit and obtain an explicit expression for this heavy traffic limit.

Theorem III.1. *Consider the base family of switches under the $\mathbf{c}$-weighted MaxWeight algorithm as in Proposition III.2. Then, in the heavy traffic limit as $\epsilon \downarrow 0$, we have*

$$\lim_{\epsilon \to 0} \epsilon \mathbb{E}\left[\sum_{ij} c_{ij} \overline{Q}_{ij}^{(\epsilon)}\right] = \frac{n}{2}\left\langle \boldsymbol{\sigma}^2, \boldsymbol{\zeta}\right\rangle_{\mathbf{c}}, \qquad (6)$$

where $\boldsymbol{\sigma}^2 = (\sigma_{ij}^2)_{ij}$, and the vector $\boldsymbol{\zeta}$ is defined by $\zeta_{ij} := \|(\mathbf{e}_{ij})_{\|\mathcal{S}_\mathbf{c}}\|_\mathbf{c}^2$ and the matrix $\mathbf{e}_{ij}$ by 1 in position (i,j) and 0 elsewhere.

Proof. We again omit the superscript (ϵ) to simplify the notation and clarify the presentation. Let $\mathbf{A}$ denote the arrival vector in steady state, which is distributed identically to the random vector $\mathbf{A}(t)$ for any t. Further let $\mathbf{S}(\overline{\mathbf{Q}})$ and $\mathbf{U}(\overline{\mathbf{Q}})$ denote the steady state schedule and unused service vector, respectively, both of which depend on the QL vector in steady state $\overline{\mathbf{Q}}$. Recalling the queueing dynamics in (2), define $\overline{\mathbf{Q}}^+ := \overline{\mathbf{Q}} + \mathbf{A} - \mathbf{S}(\overline{\mathbf{Q}}) + \mathbf{U}(\overline{\mathbf{Q}})$ to be the QL vector at time $(t+1)$, given the QL vector at time t is $\overline{\mathbf{Q}}$. Clearly, $\overline{\mathbf{Q}}^+$ and $\overline{\mathbf{Q}}$ have the same distribution.

The proof proceeds by setting the drift of the Lyapunov function $V(\mathbf{Q}) = \|\mathbf{Q}_{\|\mathcal{S}_\mathbf{c}}\|_\mathbf{c}^2$ to 0 in steady state, from which we obtain

$$\begin{aligned} 0 =& \mathbb{E}[V(\overline{\mathbf{Q}}^+) - V(\overline{\mathbf{Q}})] \\ =& \mathbb{E}[\|(\mathbf{A} - \mathbf{S}(\mathbf{Q}))_{\|\mathcal{S}_\mathbf{c}}\|_\mathbf{c}^2 + 2\langle \mathbf{Q}_{\|\mathcal{S}_\mathbf{c}}, (\mathbf{A} - \mathbf{S}(\mathbf{Q}))_{\|\mathcal{S}_\mathbf{c}}\rangle_\mathbf{c} \\ & - \|\mathbf{U}_{\|\mathcal{S}_\mathbf{c}}(\mathbf{Q})\|_\mathbf{c}^2 + 2\langle \mathbf{Q}_{\|\mathcal{S}_\mathbf{c}}^+, \mathbf{U}_{\|\mathcal{S}_\mathbf{c}}(\mathbf{Q})\rangle_\mathbf{c}]. \end{aligned}$$

This yields an equation of the form

$$\begin{aligned} 2\mathbb{E}\left[\langle \mathbf{Q}_{\|\mathcal{S}_\mathbf{c}}, (\mathbf{S}(\mathbf{Q}) - \mathbf{A})_{\|\mathcal{S}_\mathbf{c}}\rangle_\mathbf{c}\right] =& \mathbb{E}\left[\|(\mathbf{A} - \mathbf{S}(\mathbf{Q}))_{\|\mathcal{S}_\mathbf{c}}\|_\mathbf{c}^2\right] \\ & - \mathbb{E}\left[\|\mathbf{U}_{\|\mathcal{S}_\mathbf{c}}(\mathbf{Q})\|_\mathbf{c}^2\right] + 2\mathbb{E}\left[\langle \mathbf{Q}_{\|\mathcal{S}_\mathbf{c}}^+, \mathbf{U}_{\|\mathcal{S}_\mathbf{c}}(\mathbf{Q})\rangle_\mathbf{c}\right]. \end{aligned}$$

The desired result then follows from Lemmas A.3 and A.4 in Appendix B, matching the LHS and RHS of (6). $\square$

D. Explicit Expression for Heavy Traffic Limit

Given the important differences in the cone for our general case of $c_{ij} > 0$ in comparison to the cone in [6], we now seek to obtain an explicit expression for the RHS of (6). More specifically, we want to calculate ζ_{ij} for each (i,j). To start, let us consider the following n^2-dimensional vectors, given in matrix form for any $i, j \in [n-1]$,

$$B_{ij} = \begin{pmatrix} E_{ij} & -E_i \\ -E_j^T & 1 \end{pmatrix}$$

where E_{ij} is an $(n-1) \times (n-1)$ matrix with the (i,j)th element 1 and all other elements 0, E_i is an $(n-1)$-vector with the ith element 1 and all other elements 0, and superscript T denotes the transpose operator. The B_{ij} are linearly independent, and it can be readily verified that the $\mathbf{c}$-inner product of B_{ij} with any of $\mathbf{e}_\mathbf{c}^{(1)}, \ldots, \mathbf{e}_\mathbf{c}^{(n)}, \tilde{\mathbf{e}}_\mathbf{c}^{(1)}, \ldots, \tilde{\mathbf{e}}_\mathbf{c}^{(n-1)}$ is 0. Recalling $\zeta_{ij} = \|(\mathbf{e}_{ij})_{\|\mathcal{S}_\mathbf{c}}\|_\mathbf{c}^2$, our approach consists of calculating the projection of $\mathbf{e}_{ij}$ to the $\perp_\mathbf{c}$-space, which

yields its projection to the $\|_\mathbf{c}$-space. Additional details are provided in [4], including an alternative approach to establish an explicit expression for the RHS of (6).

Clearly $\langle \mathbf{e}_{ij}, B_{k\ell}\rangle_\mathbf{c} = c_{ij}$ only when $i = k$, $j = \ell$, and is 0 otherwise. Without loss of generality, let us consider $\mathbf{e}_{11}$. We seek to find the vector $(\mathbf{e}_{11})_{\perp \mathcal{S}_\mathbf{c}}$ that is perpendicular to all B_{ij} except B_{11}, as well as to $\mathbf{e}_\mathbf{c}^{(i)}$ and $\tilde{\mathbf{e}}_\mathbf{c}^{(j)}$. Let δ_{ij} denote the Kronecker delta. From the special structure of B_{ij}, we can see that $((\mathbf{e}_{11})_{\perp \mathcal{S}_\mathbf{c}})_{ij} = (c_{ij}\delta_{1i}\delta_{1j} - x_i(1 - \delta_{in}) - y_j(1 - \delta_{jn}) - z[(1 - \delta_{in})(1 - \delta_{jn}) - \delta_{in}\delta_{jn}])c_{ij}^{-1}$ for some $x_1, x_2, \ldots, x_{n-1}, y_1, y_2, \ldots, y_{n-1}, z$ satisfying a system of $2n - 1$ linear equations $G(x_1, x_2, \ldots, x_{n-1}, y_1, y_2, \ldots, y_{n-1}, z)^T = \mathbf{e}_{11}^{(2n-1)}$, where $\mathbf{e}_{11}^{(2n-1)}$ is a $(2n-1)$-dimensional vector with the first and $(n+1)$st elements being 1, and all others being 0, and $G = \begin{pmatrix} D_1 & C & H_1 \\ 0 & H_3^T & -c_{nn}^{-1} \\ C^T & D_2 & H_2 \end{pmatrix}$ is a $(2n-1) \times (2n-1)$ nonsingular matrix representing the coefficients of the system of equations. Here $C = [c_{ij}^{-1}]_{i,j=1,\ldots,n-1}$, $D_1 = \text{diag}[\sum_{j=1}^n c_{ij}^{-1}]_{i=1,\ldots,n-1}$ and $D_2 = \text{diag}[\sum_{i=1}^n c_{ij}^{-1}]_{j=1,\ldots,n-1}$ are $(n-1) \times (n-1)$ matrices, and $H_1 = [\sum_{j=1}^{n-1} c_{ij}^{-1}]_{i=1,\ldots,n-1}^T$, $H_2 = [\sum_{i=1}^{n-1} c_{ij}^{-1}]_{j=1,\ldots,n-1}^T$ and $H_3 = [c_{nj}^{-1}]_{j=1,\ldots,n-1}^T$ are $(n-1)$-dimensional column vectors. Hence, $(x_1, x_2, \ldots, x_{n-1}, y_1, y_2, \ldots, y_{n-1}, z)^T = G^{-1}\mathbf{e}_{11}^{(2n-1)}$. Meanwhile, we know $\|(\mathbf{e}_{11})_{\|\mathcal{S}_\mathbf{c}}\|_\mathbf{c}^2 = \langle (\mathbf{e}_{11})_{\|\mathcal{S}_\mathbf{c}}, (\mathbf{e}_{11})_{\|\mathcal{S}_\mathbf{c}}\rangle_\mathbf{c} = \langle \mathbf{e}_{11} - (\mathbf{e}_{11})_{\perp \mathcal{S}_\mathbf{c}}, \mathbf{e}_{11} - (\mathbf{e}_{11})_{\perp \mathcal{S}_\mathbf{c}}\rangle_\mathbf{c} = \langle \mathbf{e}_{11}, \mathbf{e}_{11}\rangle_\mathbf{c} - 2\langle \mathbf{e}_{11}, (\mathbf{e}_{11})_{\perp \mathcal{S}_\mathbf{c}}\rangle_\mathbf{c} + \langle (\mathbf{e}_{11})_{\perp \mathcal{S}_\mathbf{c}}, (\mathbf{e}_{11})_{\perp \mathcal{S}_\mathbf{c}}\rangle_\mathbf{c} = \langle \mathbf{e}_{11}, \mathbf{e}_{11}\rangle_\mathbf{c} - \langle \mathbf{e}_{11}, (\mathbf{e}_{11})_{\perp \mathcal{S}_\mathbf{c}}\rangle_\mathbf{c}$, where the last equality is due to the fact that $\langle \mathbf{e}_{11}, (\mathbf{e}_{11})_{\perp \mathcal{S}_\mathbf{c}}\rangle_\mathbf{c} = \langle (\mathbf{e}_{11})_{\perp \mathcal{S}_\mathbf{c}}, (\mathbf{e}_{11})_{\perp \mathcal{S}_\mathbf{c}}\rangle_\mathbf{c}$. This together with the above yields $\|(\mathbf{e}_{11})_{\|\mathcal{S}_\mathbf{c}}\|_\mathbf{c}^2 = z + x_1 + y_1$.

The same approach can be taken for all the other $\mathbf{e}_{ij}$. Specifically, we need to identify the vector $(\mathbf{e}_{ij})_{\perp \mathcal{S}_\mathbf{c}}$ that satisfies $\langle (\mathbf{e}_{ij})_{\perp \mathcal{S}_\mathbf{c}}, B_{ij}\rangle_\mathbf{c} = c_{ij}$ and $\langle (\mathbf{e}_{ij})_{\perp \mathcal{S}_\mathbf{c}}, B_{k\ell}\rangle_\mathbf{c} = 0$ for all other (k, ℓ). Here, $(\mathbf{e}_{ij})_{\perp \mathcal{S}_\mathbf{c}}$ can have a similar representation as above with the position of 1 changed from $(1,1)$ to (i,j). The variables x, y, z will satisfy a system of equations $G(x_1^{ij}, x_2^{ij}, \ldots, x_{n-1}^{ij}, y_1^{ij}, y_2^{ij}, \ldots, y_{n-1}^{ij}, z^{ij})^T = \mathbf{e}_{ij}^{(2n-1)}$ where G is the same matrix defined above and where $\mathbf{e}_{ij}^{(2n-1)}$ is a $(2n-1)$-dimensional vector with the ith and $(n+j)$th elements being 1, and all others being 0 when $j \leq n-1$, and with only the ith element being 1 and all others being 0 when $j = n$. Hence, following the above approach, we have in general $\|(\mathbf{e}_{ij})_{\|\mathcal{S}_\mathbf{c}}\|_\mathbf{c}^2 = z^{ij} + x_1^{ij} + y_1^{ij}$ with $(x_1^{ij}, \ldots, x_{n-1}^{ij}, y_1^{ij}, \ldots, y_{n-1}^{ij}, z^{ij})^T$ given by $G^{-1}\mathbf{e}_{ij}^{(2n-1)}$.

Remark III.3. *For the $n = 2$ case, the explicit expression above can be used to recover the heavy-traffic limit of $(1/2)\sum_{ij}\sigma_{ij}^2 c_{ij}(1 - [c_{ij}^2 / \sum_{i'j'} c_{i'j'}^2])$ in [3] for RHS of (6).*

IV. Conclusions

In this paper we considered the optimal control of $n \times n$ IQSs under the $\mathbf{c}$-weighted MaxWeight algorithm, with the goal of gaining fundamental insights on the delay-optimal properties of this cost-weighted variant of MaxWeight scheduling in real-world IQSs. We established

theoretical properties that include showing the **c**-weighted MaxWeight algorithm exhibits optimal scaling in heavy traffic under an objective function consisting of a general linear function of the steady-state average QLs. Our results shed light on the delay optimality of variants of MaxWeight scheduling and are expected to be of theoretical interest more broadly than IQSs.

Appendix

A. Proof of State Space Collapse

To simplify the notation, we use $\mathbb{E}_{\mathbf{Q}}[\cdot]$ to denote $\mathbb{E}[\cdot \mid \mathbf{Q}(t) = \mathbf{Q}]$ throughout this section.

Lemma A.1. *For Lyapunov function drift* $\Delta W_{\perp_{\mathbf{c}}}(\mathbf{Q}) := \big(W_{\perp_{\mathbf{c}}}(\mathbf{Q}(t+1)) - W_{\perp_{\mathbf{c}}}(\mathbf{Q}(t))\big)\mathbb{I}_{\{\mathbf{Q}(t)=\mathbf{Q}\}}$, *we have*

$$\mathbb{P}[|\Delta W_{\perp_{\mathbf{c}}}(\mathbf{Q})| \leq D] = 1, \quad \forall \mathbf{Q}, \tag{7}$$
$$\mathbb{E}_{\mathbf{Q}}[\Delta W_{\perp_{\mathbf{c}}}(\mathbf{Q})] \leq -\eta, \quad \forall \mathbf{Q}, W_{\perp_{\mathbf{c}}}(\mathbf{Q}) \geq \kappa, \tag{8}$$

for some positive numbers η, κ *and* D *that depend on* $\widetilde{\sigma}$, $\boldsymbol{\nu}$, $A_{\max}$ *and* $\nu_{\min}$, *but not on* ϵ.

Proof. First of all, (7) follows from

$$
\begin{aligned}
|\Delta W_{\perp_{\mathbf{c}}}(\mathbf{Q})| &\leq \left| \|\mathbf{Q}_{\perp_{\mathbf{c}}}(t+1)\|_{\mathbf{c}} - \|\mathbf{Q}_{\perp_{\mathbf{c}}}(t)\|_{\mathbf{c}} \right| \\
&\leq \|\mathbf{Q}(t+1) - \mathbf{Q}(t)\|_{\mathbf{c}} \\
&= \sqrt{\sum_{ij} c_{ij}(Q_{ij}(t+1) - Q_{ij}(t))^2} \\
&\leq \sqrt{\sum_{ij} c_{ij} A_{ij}^2} \leq n\sqrt{c_{\max}}\, A_{\max},
\end{aligned}
$$

with $D = n\sqrt{c_{\max}}\, A_{\max}$. To prove (8) we start with a version of Lemma 4 in [6], which can be shown to hold more generally for the new dot product by appropriately adapting the arguments in the proof of Lemma 7 in [2]. $\square$

Lemma A.2. *For all* $\mathbf{Q} \in \mathbb{R}^{n^2}$, *we have*

$$\Delta W_{\perp_{\mathbf{c}}}(\mathbf{Q}) \leq \frac{1}{2\|\mathbf{Q}_{\perp_{\mathbf{c}}}\|_{\mathbf{c}}}(\Delta V(\mathbf{Q}) - \Delta V_{\|_{\mathbf{c}}}(\mathbf{Q})), \tag{9}$$

where $V(\mathbf{Q}) := \|\mathbf{Q}\|_{\mathbf{c}}^2, V_{\|_{\mathbf{c}}}(\mathbf{Q}) := \|\mathbf{Q}_{\|_{\mathbf{c}}}\|_{\mathbf{c}}^2$ *and*

$$\Delta V(\mathbf{Q}) := \big(V(\mathbf{Q}(t+1)) - V(\mathbf{Q}(t))\big)\mathbb{I}_{\{\mathbf{Q}(t)=\mathbf{Q}\}}$$
$$\Delta V_{\|_{\mathbf{c}}}(\mathbf{Q}) := \big(V_{\|_{\mathbf{c}}}(\mathbf{Q}(t+1)) - V_{\|_{\mathbf{c}}}(\mathbf{Q}(t))\big)\mathbb{I}_{\{\mathbf{Q}(t)=\mathbf{Q}\}}.$$

Proof. Let us separately consider the two quantities $\Delta V(\mathbf{Q})$ and $\Delta V_{\|_{\mathbf{c}}}(\mathbf{Q})$, recalling the queueing dynamics in (2). For the first quantity, we obtain

$$
\begin{aligned}
&\mathbb{E}_{\mathbf{Q}}[\Delta V(\mathbf{Q})] \\
&= \mathbb{E}_{\mathbf{Q}}[\|\mathbf{Q}(t) + \mathbf{A}(t) - \mathbf{S}(t)\|_{\mathbf{c}}^2 - \|\mathbf{U}(t)\|_{\mathbf{c}}^2 - \|\mathbf{Q}(t)\|_{\mathbf{c}}^2] \\
&\leq \mathbb{E}_{\mathbf{Q}}[\|\mathbf{A}(t) - \mathbf{S}(t)\|_{\mathbf{c}}^2 + 2\langle \mathbf{Q}(t), \mathbf{A}(t) - \mathbf{S}(t)\rangle_{\mathbf{c}}] \\
&= \mathbb{E}_{\mathbf{Q}}\left[\sum_{ij} c_{ij} A_{ij}^2(t) + c_{ij}S_{ij}(t) - 2c_{ij}A_{ij}(t)S_{ij}(t) \right] \\
&\quad + 2\langle \mathbf{Q}, \boldsymbol{\lambda} - \mathbb{E}_{\mathbf{Q}}[\mathbf{S}(t)]\rangle_{\mathbf{c}} \\
&\leq \sum_{ij} c_{ij}(\lambda_{ij} + \sigma_{ij}^2) + \sum_{ij} c_{ij}S_{ij}(t) - 2\epsilon\langle \mathbf{Q}, \boldsymbol{\nu}\rangle_{\mathbf{c}} \\
&\quad + 2\min\langle \mathbf{Q}, \boldsymbol{\nu} - \mathbf{r}\rangle_{\mathbf{c}},
\end{aligned}
$$

where we exploit the facts that $\langle \mathbf{Q}(t+1), \mathbf{U}(t)\rangle_{\mathbf{c}} = 0$ and that arrivals are independent of the QL and service processes in each time slot, together with our definition of the **c**-weighted MaxWeight algorithm. The selection of $\mathbf{r}$ will be $\boldsymbol{\nu} + \frac{\nu_{\min}^c}{\|\mathbf{Q}_{\perp_{\mathbf{c}}}\|_{\mathbf{c}}}\mathbf{Q}_{\perp_{\mathbf{c}}}$, where $\boldsymbol{\nu}$ is an arrival rate vector that resides on the boundary of the capacity region with all input and output ports saturated and where $\nu_{\min}^c := \min \frac{\nu_{ij}}{c_{ij}}$. This selection of $\mathbf{r}$ guarantees that it is within the capacity region, which is readily verified by first observing $\nu_{ij} + \frac{\nu_{\min}^c}{\|\mathbf{Q}_{\perp_{\mathbf{c}}}\|_{\mathbf{c}}}\mathbf{Q}_{\perp_{\mathbf{c}},ij} \geq \nu_{ij} - \nu_{min} \geq 0$ and then observing $\langle \boldsymbol{\nu} + \frac{\nu_{\min}^c}{\|\mathbf{Q}_{\perp_{\mathbf{c}}}\|_{\mathbf{c}}}\mathbf{Q}_{\perp_{\mathbf{c}}}, \mathbf{e}^i\rangle_{\mathbf{c}} \leq 1$ and $\langle \boldsymbol{\nu} + \frac{\nu_{\min}^c}{\|\mathbf{Q}_{\perp_{\mathbf{c}}}\|_{\mathbf{c}}}\mathbf{Q}_{\perp_{\mathbf{c}}}, \tilde{\mathbf{e}}^j\rangle_{\mathbf{c}} \leq 1$. We therefore have

$$
\begin{aligned}
\mathbb{E}_{\mathbf{Q}}[\Delta V(\mathbf{Q})] \leq &\sum_{ij} c_{ij}(\lambda_{ij} + \sigma_{ij}^2) + nc_{\max} \\
&- 2\epsilon\langle \mathbf{Q}, \boldsymbol{\nu}\rangle_{\mathbf{c}} - 2\nu_{\min}^c\|\mathbf{Q}_{\perp_{\mathbf{c}}}\|_{\mathbf{c}},
\end{aligned}
$$

taking advantage of the fact that $\langle \mathbf{Q}_{\|_{\mathbf{c}}}, \mathbf{Q}_{\perp_{\mathbf{c}}}\rangle_{\mathbf{c}} = 0$. Turning to the second quantity, we obtain

$$
\begin{aligned}
\mathbb{E}_{\mathbf{Q}}[\Delta V_{\|_{\mathbf{c}}}] =& \mathbb{E}_{\mathbf{Q}}[\|\mathbf{Q}_{\|_{\mathbf{c}}}(t+1) - \mathbf{Q}_{\|_{\mathbf{c}}}(t)\|_{\mathbf{c}}^2] \\
&+ 2\mathbb{E}_{\mathbf{Q}}[\langle \mathbf{Q}_{\|_{\mathbf{c}}}(t), \mathbf{Q}_{\|_{\mathbf{c}}}(t+1) - \mathbf{Q}_{\|_{\mathbf{c}}}(t)\rangle_{\mathbf{c}}] \\
\geq & 2\mathbb{E}_{\mathbf{Q}}[\langle \mathbf{Q}_{\|_{\mathbf{c}}}(t), \mathbf{Q}_{\|_{\mathbf{c}}}(t+1) - \mathbf{Q}_{\|_{\mathbf{c}}}(t)\rangle_{\mathbf{c}}] \\
\geq & 2\mathbb{E}_{\mathbf{Q}}[\langle \mathbf{Q}_{\|_{\mathbf{c}}}(t), \mathbf{A}(t) - \mathbf{S}(t) + \mathbf{U}(t)\rangle_{\mathbf{c}}] \\
\geq & 2\langle \mathbf{Q}_{\|_{\mathbf{c}}}(t), \boldsymbol{\lambda}\rangle_{\mathbf{c}} - 2\mathbb{E}_{\mathbf{Q}}[\langle \mathbf{Q}_{\|_{\mathbf{c}}}(t), \mathbf{S}(t)\rangle_{\mathbf{c}}] \\
= & -2\epsilon\langle \mathbf{Q}_{\|_{\mathbf{c}}}(t), \boldsymbol{\nu}\rangle_{\mathbf{c}} - 2\mathbb{E}_{\mathbf{Q}}[\langle \mathbf{Q}_{\|_{\mathbf{c}}}(t), \mathbf{S}(t) - \boldsymbol{\nu}\rangle_{\mathbf{c}}] \\
= & -2\epsilon\langle \mathbf{Q}_{\|_{\mathbf{c}}}(t), \boldsymbol{\nu}\rangle_{\mathbf{c}},
\end{aligned}
$$

where we again take advantage of the above facts together with $\langle \mathbf{Q}_{\|_{\mathbf{c}}}(t), \mathbf{Q}_{\perp_{\mathbf{c}}}(t+1)\rangle_{\mathbf{c}} \leq 0$, both $\mathbf{Q}_{\|_{\mathbf{c}}}$ and $\mathbf{U}(t)$ being nonnegative componentwise, and properties related to the cone $\mathcal{K}_{\mathbf{c}}$ and its spanned space $\mathcal{S}_{\mathbf{c}}$.

Upon substituting the above expressions for both quantities into (9), we have

$$
\begin{aligned}
\mathbb{E}_{\mathbf{Q}}[\Delta W_{\perp_{\mathbf{c}}}(\mathbf{Q})] \leq & \frac{1}{2\|\mathbf{Q}_{\perp_{\mathbf{c}}}\|}\bigg[\sum_{ij} c_{ij}(\lambda_{ij} + \sigma_{ij}^2) \\
& + nc_{\max} - 2\epsilon\langle \mathbf{Q}, \boldsymbol{\nu}\rangle_{\mathbf{c}} - 2\nu_{\min}^c\|\mathbf{Q}_{\perp_{\mathbf{c}}}\|_{\mathbf{c}} + 2\epsilon\langle \mathbf{Q}_{\|_{\mathbf{c}}}(t), \boldsymbol{\nu}\rangle_{\mathbf{c}} \bigg] \\
\leq & \frac{\sum_{ij} c_{ij}(\lambda_{ij} + \sigma_{ij}^2) + nc_{\max}}{\|\mathbf{Q}_{\perp_{\mathbf{c}}}\|} - \nu_{\min}^c - \frac{\epsilon}{\|\mathbf{Q}_{\perp_{\mathbf{c}}}\|}\langle \mathbf{Q}_{\perp_{\mathbf{c}}}(t), \boldsymbol{\nu}\rangle_{\mathbf{c}}.
\end{aligned}
$$

Given $\epsilon < \nu_{\min}^c/(2\|\boldsymbol{\nu}\|_{\mathbf{c}})$, then on the set of $W_{\perp_{\mathbf{c}}}(\mathbf{Q}) \geq 4(\sum_{ij} c_{ij}(\lambda_{ij} + \sigma_{ij}^2) + nc_{\max})/\nu_{\min}^c$, we obtain

$$
\begin{aligned}
&\mathbb{E}_{\mathbf{Q}}[\Delta W_{\perp_{\mathbf{c}}}(\mathbf{Q})] \\
&\leq \frac{1}{2\|\mathbf{Q}_{\perp_{\mathbf{c}}}\|}(\sum_{ij} c_{ij}(\lambda_{ij} + \sigma_{ij}^2) + nc_{\max} - 2\epsilon\langle \mathbf{Q}, \boldsymbol{\nu}\rangle_{\mathbf{c}} \\
&\quad - 2\nu_{\min}^c\|\mathbf{Q}_{\perp_{\mathbf{c}}}\|_{\mathbf{c}} + 2\epsilon\langle \mathbf{Q}_{\|_{\mathbf{c}}}(t), \boldsymbol{\nu}\rangle_{\mathbf{c}}) \\
&\leq \frac{\sum_{ij} c_{ij}(\lambda_{ij} + \sigma_{ij}^2) + nc_{\max}}{2\|\mathbf{Q}_{\perp_{\mathbf{c}}}\|} - \nu_{\min}^c - \epsilon\|\boldsymbol{\nu}\|_{\mathbf{c}} \\
&\leq \frac{\sum_{ij} c_{ij}(\lambda_{ij} + \sigma_{ij}^2) + nc_{\max}}{2\|\mathbf{Q}_{\perp_{\mathbf{c}}}\|} - \frac{\nu_{\min}^c}{2} \leq -\frac{\nu_{\min}^c}{4}.
\end{aligned}
$$

Hence, (8) holds with $\eta = -\nu_{\min}^c/4$. $\square$

B. Proof of Theorem III.1

Lemma A.3. *In the limit as $\epsilon \downarrow 0$, we have*

$$n\mathbb{E}\left[\left\langle \mathbf{Q}_{||\mathcal{S}_{\mathbf{c}}}, (\mathbf{S}(\mathbf{Q}) - \mathbf{A})_{||\mathcal{S}_{\mathbf{c}}}\right\rangle_{\mathbf{c}}\right] = \lim_{\epsilon \to 0} \epsilon\mathbb{E}\left[\sum_{ij} c_{ij}\overline{Q}_{ij}^{(\epsilon)}\right].$$

Proof. The LHS of above can be written as

$$2\mathbb{E}\left[\left\langle \mathbf{Q}_{||\mathcal{S}_{\mathbf{c}}}, (\mathbf{S}(\mathbf{Q}) - \mathbf{A})_{||\mathcal{S}_{\mathbf{c}}}\right\rangle_{\mathbf{c}}\right]$$

$$= 2\epsilon\mathbb{E}\left[\left\langle \mathbf{Q}_{||\mathcal{S}_{\mathbf{c}}}, \boldsymbol{\nu}\right\rangle_{\mathbf{c}}\right] + 2\mathbb{E}\left[\left\langle \mathbf{Q}_{||\mathcal{S}_{\mathbf{c}}}, \mathbf{S}(\mathbf{Q}) - \boldsymbol{\nu}\right\rangle_{\mathbf{c}}\right]$$

$$= \frac{2}{n}\epsilon\mathbb{E}\left[\left\langle \mathbf{Q}_{||\mathcal{S}_{\mathbf{c}}}, \mathbf{1}\right\rangle_{\mathbf{c}}\right] + 2\epsilon\mathbb{E}\left[\left\langle \mathbf{Q}_{||\mathcal{S}_{\mathbf{c}}}, \boldsymbol{\nu} - \frac{1}{n}\mathbf{1}\right\rangle_{\mathbf{c}}\right]$$

$$= \frac{2}{n}\epsilon\mathbb{E}\left[\left\langle \mathbf{Q}, \mathbf{1}\right\rangle_{\mathbf{c}}\right] - \frac{2}{n}\epsilon\mathbb{E}\left[\left\langle \mathbf{Q}_{\perp\mathcal{S}_{\mathbf{c}}}, \mathbf{1}\right\rangle_{\mathbf{c}}\right],$$

where the second equality follows from the fact that $\mathbf{S}(\overline{\mathbf{Q}}), \boldsymbol{\nu} \in \mathcal{F}$, and therefore $\mathbf{S}(\overline{\mathbf{Q}}) - \boldsymbol{\nu}$ is orthogonal to the space spanned by the normal vectors of $\mathcal{F}$, i.e., to the space $\mathcal{S}_{\mathbf{c}}$; and the next to last equality follows from the fact that $\boldsymbol{\nu}, \mathbf{1}/n \in \mathcal{F}$. Since the second term of the last equation goes to 0 as $\epsilon \downarrow 0$ by the state space collapse from Proposition III.2, we have

$$\lim_{\epsilon \downarrow 0} \mathbb{E}\left[\left\langle \mathbf{Q}_{||\mathcal{S}_{\mathbf{c}}}, (\mathbf{S}(\mathbf{Q}) - \overline{\mathbf{A}})_{||\mathcal{S}_{\mathbf{c}}}\right\rangle_{\mathbf{c}}\right] = \lim_{\epsilon \downarrow 0} \frac{\epsilon}{n}\mathbb{E}\left[\sum_{ij} c_{ij}\overline{Q}_{ij}\right],$$

thus yielding the LHS of (6) in Theorem III.1. $\qquad\square$

Lemma A.4. *In the limit as $\epsilon \downarrow 0$, we have*

$$\mathbb{E}\left[\|(\mathbf{A} - \mathbf{S}(\mathbf{Q}))_{||\mathcal{S}_{\mathbf{c}}}\|_{\mathbf{c}}^2\right] - \mathbb{E}\left[\|\mathbf{U}_{||\mathcal{S}_{\mathbf{c}}}(\mathbf{Q})\|_{\mathbf{c}}^2\right]$$

$$+ 2\mathbb{E}\left[\left\langle \mathbf{Q}_{||\mathcal{S}_{\mathbf{c}}}^+, \mathbf{U}_{||\mathcal{S}_{\mathbf{c}}}(\mathbf{Q})\right\rangle_{\mathbf{c}}\right] = \frac{n}{2}\left\langle \boldsymbol{\sigma}^2, \boldsymbol{\zeta}\right\rangle_{\mathbf{c}}. \quad (10)$$

Proof. First of all, from the equation

$$0 = \mathbb{E}\left[\sum_{i,j} Q_{ij}(t+1) - \sum_{ij} Q(t)|\mathbf{Q}(t) = \overline{\mathbf{Q}}\right]$$

$$= \mathbb{E}\left[\sum_{i,j} A_{ij} - \sum_{i,j} S_{i,j}(\overline{\mathbf{Q}}) - \sum_{i,j} U_{i,j}\right],$$

we can conclude that $\mathbb{E}[\sum_{i,j} U_{ij}(\mathbf{Q})] = n\epsilon$, which implies that the second term on the LHS of (10) converges to 0 as $\epsilon \downarrow 0$:

$$\mathbb{E}\left[\|\mathbf{U}_{||\mathcal{S}_{\mathbf{c}}}(\mathbf{Q})\|_{\mathbf{c}}^2\right] \leq \mathbb{E}\left[\sum_{i,j} c_{ij}U_{ij}(\overline{\mathbf{Q}})^2\right] = \mathbb{E}\left[\sum_{i,j} c_{ij}U_{ij}(\overline{\mathbf{Q}})\right]$$

$$\leq c_{\max}n\epsilon \to 0, \quad \text{as } \epsilon \downarrow 0.$$

For the third term on the LHS of (10), we have

$$2\mathbb{E}\left[\left\langle \mathbf{Q}_{||\mathcal{S}_{\mathbf{c}}}^+, \mathbf{U}_{||\mathcal{S}_{\mathbf{c}}}(\mathbf{Q})\right\rangle_{\mathbf{c}}\right] = 2\mathbb{E}\left[\left\langle \mathbf{Q}^+, \mathbf{U}_{||\mathcal{S}_{\mathbf{c}}}(\mathbf{Q})\right\rangle_{\mathbf{c}}\right]$$

$$- 2\mathbb{E}\left[\left\langle \mathbf{Q}_{\perp\mathcal{S}_{\mathbf{c}}}^+, \mathbf{U}_{||\mathcal{S}_{\mathbf{c}}}(\mathbf{Q})\right\rangle_{\mathbf{c}}\right]$$

$$= -2\mathbb{E}\left[\left\langle \mathbf{Q}_{\perp\mathcal{S}_{\mathbf{c}}}^+, \mathbf{U}_{||\mathcal{S}_{\mathbf{c}}}(\mathbf{Q})\right\rangle_{\mathbf{c}}\right],$$

because $\overline{Q}_{ij}^+ = 0$ if $U_{ij}(\overline{Q}) = 1$. Furthermore,

$$\left|\mathbb{E}\left[\left\langle \mathbf{Q}_{||\mathcal{S}_{\mathbf{c}}}^+, \mathbf{U}_{||\mathcal{S}_{\mathbf{c}}}(\mathbf{Q})\right\rangle_{\mathbf{c}}\right]\right| \leq \sqrt{\mathbb{E}\left[\|\mathbf{Q}_{\perp\mathcal{S}_{\mathbf{c}}}^+\|^2\right]\mathbb{E}\left[\|\mathbf{U}_{||\mathcal{S}_{\mathbf{c}}}(\mathbf{Q})\|^2\right]}$$

$$\leq M_2\sqrt{\mathbb{E}\left[\|\mathbf{U}_{||\mathcal{S}_{\mathbf{c}}}(\mathbf{Q})\|^2\right]} \leq M_2\sqrt{2n\epsilon},$$

where the first inequality is just Cauchy-Schwartz, M_2 is the constant in Proposition III.2, and the last inequality is due to $\mathbb{E}[\sum_{i,j} U_{ij}(\mathbf{Q})] = n\epsilon$. This then implies that the third term also converges to 0 as $\epsilon \downarrow 0$.

Finally, turning to investigate the first term, let $f_1, f_2, \ldots, f_{2n-1}$ be an orthonormal base for space $\mathcal{S}$. Then, from basic properties of the space, there exist $v_{\ell,i}$ and $\tilde{v}_{\ell,j}$ such that $f_{\ell ij} = \frac{v_{\ell i} + \tilde{v}_{\ell j}}{c_{ij}}$. Thus, we can derive

$$\mathbb{E}[\|(\mathbf{A} - \mathbf{S}(\mathbf{Q}))_{||\mathcal{S}_{\mathbf{c}}}\|^2] = \sum_{\ell=1}^{2n-1} \mathbb{E}\left[\langle \mathbf{A} - \mathbf{S}(\mathbf{Q}), f_\ell\rangle_{\mathbf{c}}^2\right]$$

$$= \sum_{\ell=1}^{2n-1} \text{Var}\left[\sum_i v_{\ell i}\sum_j A_{ij} + \sum_j \tilde{v}_{\ell j}\sum_i A_{ij}\right]$$

$$= \sum_{\ell=1}^{2n-1}\left[\sum_i v_{\ell i}^2\sum_j \sigma_{ij}^2 + \sum_j \tilde{v}_{\ell j}^2\sum_j \sigma_{ij}^2 + 2\sum_{ij} v_{\ell i}\tilde{v}_{\ell j}\sigma_{ij}^2\right]$$

$$= \sum_{ij} c_{ij}\sigma_{ij}\sum_{\ell=1}^{2n-1}\left(\frac{v_{\ell i} + \tilde{v}_{\ell j}}{c_{ij}}\right)^2 c_{ij} = \sum_{ij} c_{ij}\sigma_{ij}\sum_{\ell=1}^{2n-1}\langle f_\ell, e_{ij}\rangle_{\mathbf{c}}^2$$

$$= \sum_{ij} c_{ij}\sigma_{ij}\|(e_{ij})_{||\mathcal{S}_{\mathbf{c}}}\|^2 = \left\langle \boldsymbol{\sigma}^2, \boldsymbol{\zeta}\right\rangle_{\mathbf{c}},$$

which establishes the desired result. $\qquad\square$

REFERENCES

[1] R. Cavazos-Cadena, L.I. Sennott. Comparing recent assumptions for the existence of average optimal stationary policies. *Op. Res. Letters*, 11: 33–37, 1992.

[2] A. Eryilmaz, R. Srikant. Asymptotically tight steady-state queue length bounds implied by drift conditions. *QUESTA*, 72:311–359, 2012.

[3] Y. Lu, S. T. Maguluri, M. S. Squillante, T. Suk. Optimal Dynamic Control for Input-Queued Switches in Heavy Traffic. *Proc. ACC*, 2018.

[4] Y. Lu, S. T. Maguluri, M. S. Squillante, T. Suk. Optimal Dynamic Control for Input-Queued Switches in Heavy Traffic with Improved Bounds. arXiv:1907.02115v2, 2020.

[5] S. T. Maguluri, S. K. Burle, R. Srikant. Optimal heavy-traffic queue length scaling in an incompletely saturated switch. *QUESTA*, 88:279–309 , 2018.

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

[7] N. McKeown, V. Anantharam, J. Walrand. Achieving 100% throughput in an input-queued switch. *Proc. INFOCOM*, 1996.

[8] J. Perry, A. Ousterhout, H. Balakrishnan, D. Shah, H. Fugal. Fastpass: A centralized zero-queue datacenter network. *ACM SIGCOMM CCR*, 44(4): 307–318, 2014.

[9] A. Singh et al. Jupiter rising: A decade of clos topologies and centralized control in Google's datacenter network, *ACM SIGCOMM CCR*, 45(4): 183–197, 2015.

[10] A. L. Stolyar. Maxweight scheduling in a generalized switch: State space collapse and workload minimization in heavy traffic. *Ann. Appl. Probab.*, 14(1):1–53, 02 2004.

[11] L. Tassiulas, A. Ephremides. Stability properties of constrained queueing systems and scheduling policies for maximum throughput in multihop radio networks. *IEEE TAC*, 37:1936–1948, 1992.