TECHNICAL RESEARCH REPORT

Stochastic Comparison Results for Non-Blocking Switches with Output Queueing

by Y.B. Kim, A.M. Makowski

CSHCN T.R. 96-6
(ISR T.R. 96-24)

CENTER FOR SATELLITE AND HYBRID COMMUNICATION NETWORKS

Stochastic Comparison Results for Non–Blocking Switches with Output Queueing

Young B. Kim †   Armand M. Makowski†‡
dosa@isr.umd.edu   armand@isr.umd.edu
(301) 405-2960   (301) 405-6844

Abstract

We propose a systematic approach to quantify the impact of nonuniform traffic on the performance of non–blocking switches with output queueing. We do so in the context of a simple queueing model where cells arrive to input ports according to independent Bernoulli processes, and are switched to an output port under a random routing mechanism. We give conditions on pairs of input rate vectors and switching matrices which ensure various stochastic comparisons for performance measures of interest. These conditions are formulated in terms of the majorization ordering while the comparison results are expressed in the strong and convex increasing orderings.

Key words: Stochastic majorization, Stochastic convexity, Bernoulli routing, Crossbar switches

1 Introduction

Space–division packet switching has been recognized as a key component in the ongoing evolution towards future high–performance communication networks [1, 5]. This is due to the high capacity, viz., in the range 10–100Gps, that space–division packet switching can achieve through the use of a highly parallel switching fabric with simple per packet processing distributed among many high–speed VLSI circuits.

†Electrical Engineering Department and Institute for Systems Research, University of Maryland, College Park, MD 20742. The work of these authors was supported partially through NSF Engineering Research Centers Program Grant NSFD CDR-88-03012 and through NASA Grant NAGW277S.

‡Corresponding author.

In non–blocking space–division packet switches, it is always possible to establish a connection between any pair of idle input and output ports. However, output contention arises when more than one cell at different input ports demand to be routed to the same output. As the contending cells cannot be placed on the output port at the same time, buffering has to be provided in order to store the cell(s) which cannot be served. Several buffering strategies have been reported in the literature [4, 20], with proposed solutions depending on a variety of factors such as the speed of input and output lines relative to the cell transfer time across the switching fabric, and implementation complexity.

Noteworthy among proposed buffering strategies is *output queueing* which we adopt in this paper. Consider a non–blocking crossbar switch with K input and L output ports. The switch operates in a synchronous mode with time divided into consecutive slots of equal duration. At the beginning of a time slot, new cells arrive into the system; the destination of a cell is immediately declared upon arrival. The switching fabric operates at K times the speed of the input and output lines, and each output port is equipped with an infinite capacity buffer, thereafter referred to as its output buffer. Under the output queueing strategy, all cells which arrive during a time slot and which are destined for a given output port, are transported across the switch during that single time slot, and put into the output buffer. This is indeed possible under the assumption made on the speed of the switching fabric. However, during any time slot at most one cell in each output buffer can be transmitted on the corresponding output line,

The simplest model of this synchronous crossbar switch with output queueing is that of a collection of L discrete–time queues, one for each output port, operating in parallel and fed by K independent Bernoulli processes under a random routing assignment. The arrival process at the k^{th} input port, $k = 1, \ldots, K$, is a Bernoulli process with parameter λ_k, $0 < \lambda_k < 1$. The output addressing scheme is described by a stochastic matrix $\boldsymbol{R} \equiv (r_{k\ell})$, called the switching or routing matrix, with the following implementation in each time slot: A cell that arrives at the k^{th} input port at the beginning of a time slot is destined for the ℓ^{th} output port with probability $r_{k\ell}$, $k = 1, \ldots, K, \ell = 1, \ldots, L$; this assignment is carried out independently over time across input ports, and independently of the arrival streams which are assumed mutually independent.

The performance analysis for this model is typically carried out under the *uni-*

form traffic and routing assumptions, which are specified by

$$\lambda_1 = \cdots = \lambda_K \equiv \lambda \tag{1.1}$$

and

$$r_{k\ell} = \frac{1}{L} \equiv u_{k\ell}, \quad k = 1, \ldots, K; \; \ell = 1, \ldots, L. \tag{1.2}$$

A distinct advantage of assuming (1.1)–(1.2) is the fact that the input rate vector and switching matrix being symmetric, it suffices to analyze a single queue in order to obtain information concerning most performance measures of interest.

In reality, however, traffic offered to the switch is most likely to be nonuniform, and it is not clear how this will affect its performance. As a case in point, with $K = L$, if cells arriving at the k^{th} input port are always routed to the k^{th} output port, $k = 1, 2, \ldots$, there is no output contention and the best possible performance is achieved. This is in sharp contrast with the worst case scenario where *all* incoming cells are destined to the *same* output port, thereby creating severe congestion at the corresponding output buffer. Various attempts have been made to understand the range of possibilities that result from nonuniform traffic patterns. These efforts have been recently reported in the numerical studies [9, 10, 11, 18, 22], and have focused on packet switches with output queueing as well as with input queueing (and combination thereof). As nonuniform traffic refers to any traffic pattern different from (1.1)–(1.2), the number of possible nonuniform traffic patterns is simply huge due to the large number of parameters involved, and this precludes a *systematic* exploration of all cases. In fact, most analyses under nonuniform traffic have considered only very specific traffic patterns, e.g., bi–group traffic [9, 11, 18, 22], hot–spot traffic [14, 22] and point–to–point traffic [21, 22].

Given this state of affairs, in the context of the simple queueing model introduced earlier, we seek to understand in a more *systematic* manner the behavior of the output queueing switch as a function of the input rate vector $\boldsymbol{\lambda} \equiv (\lambda_1, \ldots, \lambda_K)$ and of the switching matrix $\boldsymbol{R}$. Specifically, we focus on finding conditions on pairs $(\boldsymbol{\lambda}, \boldsymbol{R})$ and $(\boldsymbol{\lambda}', \boldsymbol{R}')$ of input rate vectors and switching matrices which ensure various stochastic comparisons for the corresponding performance measures. Switch performance is quantified by output queueing delays and buffer sizes, and we distinguish performance measures associated with output ports, e.g., the queue size at the ℓ^{th} output buffer and the delay incurred by a cell leaving through the ℓ^{th} output port, from measures which are associated with input ports, e.g., the delay incurred by a cell that enters the switch by the k^{th} input port.

We formulate the conditions on the pairs $(\boldsymbol{\lambda}, \boldsymbol{R})$ and $(\boldsymbol{\lambda}', \boldsymbol{R}')$ in terms of the (weak) majorization ordering. The comparison results are expressed in the strong and convex increasing orderings for distributions, and not merely in terms of the first moments of the performance measures. The results are summarized in Section 5, where only the steady–state version is presented; however it should be clear from the discussion given in Sections 7–9 that transient versions hold as well. The results are derived through the combined use of recent ideas from the theory of stochastic convexity, and of techniques from the theory of stochastic orderings. In the process we establish several majorization properties for sums of independent Bernoulli rvs; some of these results given in Section 6 appear to be new.

In this paper we establish only one–dimensional results, i.e., results pertaining to a particular queue or port. However, these results can already be used to obtain bounds on system performance. In particular, as we show in [7], under certain load constraints, we can identify the best and worst scenarios. We refer the reader to the companion paper [6, 8] for a collection of multi–dimensional comparison results which yield traffic and switch configurations for optimal load balancing.

The paper is organized as follows: The model of interest is described in details in Section 2. Delay measures are introduced in Section 3, and the statistical equilibrium for the system is discussed in Section 4. The main results are presented Section 5, and their proofs can be found in Sections 7–9. In Section 6 we have isolated intermediary results on sums of Bernoulli random variables which are of independent interest. Several proofs have been relegated to two technical appendices.

A few words on the notation in use: Throughout K and L denote given positive integers. The k^{th} component of any element $\boldsymbol{x}$ in $\mathbb{R}^K$ is denoted either by x^k or by x_k, $k = 1, \ldots, K$, so that $\boldsymbol{x} \equiv (x^1, \ldots, x^K)$ or $(x_1, \ldots, x_K)$. A similar convention is used for random variables (rvs). For any vector $\boldsymbol{x} = (x_1, \ldots, x_K)$ in $\mathbb{R}^K$, let $x_{(1)} \leq x_{(2)} \leq \cdots \leq x_{(K)}$ denote the components of $\boldsymbol{x}$ arranged in increasing order. For vectors $\boldsymbol{x}$ and $\boldsymbol{y}$ in $\mathbb{R}^K$, we say that $\boldsymbol{x}$ is *majorized* by $\boldsymbol{y}$, and write $\boldsymbol{x} \prec \boldsymbol{y}$, whenever the conditions

$$\sum_{i=1}^{k} x_{(i)} \geq \sum_{i=1}^{k} y_{(i)}, \quad k = 1, 2, \ldots, K \tag{1.3}$$

and

$$\sum_{i=1}^{K} x_i = \sum_{i=1}^{K} y_i, \tag{1.4}$$

hold. If conditions (1.3) all hold without (1.4), then we say that $\boldsymbol{x}$ is *weakly su-*

permajorized by $\boldsymbol{y}$, and write $\boldsymbol{x} \prec^w \boldsymbol{y}$. Additional information regarding (weak) majorization can be found in [13].

The notation $\leq_{st}$ (resp. $\leq_{icx}$) stands for the the strong stochastic (resp. convex increasing) ordering on the collection of distributions [15, 19]. Finally two $\mathbb{R}$–valued rvs X and Y are said to be *equal in law* if they have the same distribution, a situation we denote by $X =_{st} Y$.

2 The Model

All rvs are defined on some probability triple $(\Omega, \mathcal{F}, \mathbf{P})$, and let $\mathbf{E}$ denote the corresponding expectation operator. With K input ports and L output ports, the queueing model of interest is parameterized by a vector of rates $\boldsymbol{\lambda}$ (in $[0,1]^L$) and by probability vectors $\boldsymbol{r}_k = (r_{k1}, \ldots, r_{kL})$ (in $\mathcal{S}_L \equiv \{\boldsymbol{r} = (r_1, \ldots, r_L) \in [0,1]^L : \sum_{\ell=1}^L r_\ell = 1\}$), $k = 1, \ldots, K$. We organize these K vectors into the $K \times L$ routing matrix $\mathbf{R}$ given by

$$\boldsymbol{R} = \begin{bmatrix} \boldsymbol{r}_1 \\ \vdots \\ \boldsymbol{r}_K \end{bmatrix} = \begin{bmatrix} r_{11} \ldots r_{1L} \\ \vdots \\ r_{K1} \ldots r_{KL} \end{bmatrix}.$$

With each set of such vectors, we associate $\{0,1\}$–valued rvs $\{A_{t+1}^k(\lambda_k),\ t = 0,1,\ldots\}$ and $\{1,\ldots,L\}$–valued rvs $\{\nu_t^k(\boldsymbol{r}_k),\ t = 0,1,\ldots\}$, $k = 1,\ldots,K$. During the discussion we make the following assumptions: (i) For each $k = 1,\ldots,K$, the rvs $\{A_{t+1}^k(\lambda_k),\ t = 0,1,\ldots\}$ are *i.i.d.* rvs with

$$\mathbf{P}\left[A_{t+1}^k(\lambda_k) = 1\right] = 1 - \mathbf{P}\left[A_{t+1}^k = 0\right] = \lambda_k$$

for all $t = 0,1,\ldots$; (ii) For each $k = 1,\ldots,K$, the rvs $\{\nu_t^k(\boldsymbol{r}_k),\ t = 0,1,\ldots\}$ are *i.i.d.* rvs with

$$\mathbf{P}\left[\nu_t^k(\boldsymbol{r}_k) = \ell\right] = r_{k\ell}, \quad \ell = 1,\ldots,L$$

for all $t = 0,1,\ldots$; and (iii) The $2K$ collections of rvs $\{A_{t+1}^k(\lambda_k),\ t = 0,1,\ldots\}$ and $\{\nu_t^k(\boldsymbol{r}_k),\ t = 0,1,\ldots\}$, $k = 1,\ldots,K$, are *mutually independent.*

These quantities have a ready interpretation in the context of the output queueing system described earlier: At the beginning of time slot $[t, t+1)$, new cells arrive into the system, with $A_{t+1}^k(\lambda_k)$ cell arriving at the k^{th} input port, $k = 1,\ldots,K$. The destination of a cell arriving at the k^{th} input port is encoded in the rv $\nu_t^k(\boldsymbol{r}_k)$, and is immediately declared upon arrival. All cells which arrive during a time slot and which are destined for a given output port, are transported across the switch during that single time slot, and put into the output buffer in *random* order. With the notation

$$\boldsymbol{\xi}_{t+1}^{\ell}(\boldsymbol{\lambda}, \boldsymbol{R}) \equiv \sum_{k=1}^{K} \mathbf{1}[\nu_t^k(\boldsymbol{r}_k) = \ell] A_{t+1}^k(\lambda_k), \quad \ell = 1, \ldots, L, \quad t = 0, 1, \ldots$$

we see that a batch of $\boldsymbol{\xi}_{t+1}^{\ell}(\boldsymbol{\lambda}, \boldsymbol{R})$ cells are destined for the ℓ^{th} output port during time slot $[t, t+1)$.

During any time slot at most one cell can be transmitted, or equivalently, served. Let $Q_t^{\ell}(\boldsymbol{\lambda}, \boldsymbol{R})$ denote the number of cells present at the beginning of time slot $[t, t+1)$ in the ℓ^{th} output buffer, $\ell = 1, \ldots, L$. If we assume the system to be initially empty at time $t = 0$, then the queue size processes evolve according to the recursions

$$\begin{aligned} Q_0^{\ell}(\boldsymbol{\lambda}, \boldsymbol{R}) &= 0; \\ Q_{t+1}^{\ell}(\boldsymbol{\lambda}, \boldsymbol{R}) &= \left[Q_t^{\ell}(\boldsymbol{\lambda}, \boldsymbol{R}) - 1\right]^+ + \xi_{t+1}^{\ell}(\boldsymbol{\lambda}, \boldsymbol{R}), \ \ell = 1, \ldots, L, \ t = 0, 1, \ldots. \end{aligned} \tag{2.1}$$

In deriving (2.1) we made the following operational assumption: If the ℓ^{th} output queue were empty at the beginning of a time slot, no cell arriving at that output queue during that time slot is eligible for transmission during the time slot. Instead of this "gated" transmission strategy, we could also consider a "cut–through" strategy according to which, if the ℓ^{th} output queue were empty at the beginning of a time slot, cells arriving at that output queue during that time slot are eligible for transmission during the time slot. In that case, the dynamics (2.1) have to be replaced by

$$Q_{t+1}^{\ell}(\boldsymbol{\lambda}, \boldsymbol{R}) = \left[Q_t^{\ell}(\boldsymbol{\lambda}, \boldsymbol{R}) - 1 + \xi_{t+1}^{\ell}(\boldsymbol{\lambda}, \boldsymbol{R})\right]^+, \quad \ell = 1, \ldots, L, \ t = 0, 1, \ldots.$$

The results derived here hold under either strategy, but for the sake of definiteness, we carry out the discussion only in the context of the gated strategy with queue dynamics (2.1).

3 Delay Measures

For each $\ell = 1, \ldots, L$, we denote by $D_n^{\ell}(\boldsymbol{\lambda}, \boldsymbol{R})$ the delay of the n^{th} cell, $n = 1, 2, \ldots$, to arrive at the ℓ^{th} output port, i.e., $D_n^{\ell}(\boldsymbol{\lambda}, \boldsymbol{R})$ represents the time that elapses between the arrival of the n^{th} cell at the ℓ^{th} output port and the end of its transmission. At each of the output queues, we assume that batches are processed in order of arrival, i.e., all cells in the m^{th} batch are served before the cells in the $(m+1)^{rst}$

batch, $m = 1, 2, \ldots$, but the order of service within a given batch is random. As a result, the delay process of the n^{th} cell can be decomposed into two successive stages: First, all the cells which have arrived in earlier time slots (and which must belong to different batches) are serviced. Then, the cells belonging to the same batch as the n^{th} cell are processed in random order. We can thus write

$$D_n^\ell(\boldsymbol{\lambda}, \boldsymbol{R}) = W_n^\ell(\boldsymbol{\lambda}, \boldsymbol{R}) + B_n^\ell(\boldsymbol{\lambda}, \boldsymbol{R}), \quad n = 1, 2, \ldots \tag{3.1}$$

where the rv $W_n^\ell(\boldsymbol{\lambda}, \boldsymbol{R})$ counts the number of slots required for transmitting all the cells in the batches which have arrived before that containing the n^{th} cell, and the rv $B_n^\ell(\boldsymbol{\lambda}, \boldsymbol{R})$ denotes the number of slots that the n^{th} cell needs to wait before it is served, once the batch to which it belongs starts being served. We can also interpretate $B_n^\ell(\boldsymbol{\lambda}, \boldsymbol{R})$ as the position of the n^{th} cell in its batch.

We also consider performance measures which are associated with the input ports: Fix $k = 1, \ldots, K$. We denote by $T_n^k(\boldsymbol{\lambda}, \boldsymbol{R})$ the delay of the n^{th} cell, $n = 1, 2, \ldots$, to arrive at the k^{th} input port, i.e., $T_n^k(\boldsymbol{\lambda}, \boldsymbol{R})$ represents the time that elapses between the arrival of the n^{th} cell at the k^{th} input port and the end of its transmission. This performance measure is closely related to the following notion of virtual delay: For each $t = 0, 1, \ldots$, let $H_t^k(\boldsymbol{\lambda}, \boldsymbol{R})$ denote the delay of a virtual cell to arrive at the k^{th} input port at the beginning of the slot $[t, t+1)$, i.e., $H_t^k(\boldsymbol{\lambda}, \boldsymbol{R})$ represents the time that elapses between the arrival of a fictitious cell at the k^{th} input port at time t and the end of its transmission. We see that $T_n^k(\boldsymbol{\lambda}, \boldsymbol{R})$ coincides with the virtual delay $H_t^k(\boldsymbol{\lambda}, \boldsymbol{R})$ when t is the arrival time of the n^{th} cell to arrive at the k^{th} input port.

We can compute $H_t^k(\boldsymbol{\lambda}, \boldsymbol{R})$ as follows: If $\nu_t^k(\boldsymbol{r}_k) = \ell$, $\ell = 1, \ldots, L$, then this fictitious cell is routed to the ℓ^{th} output port, together with cells which may have arrived at the other input ports during slot $[t, t+1)$ and which are also routed to the ℓ^{th} output port. There are $N_{t+1}^{k,\ell}(\boldsymbol{\lambda}, \boldsymbol{R})$ such cells, with

$$N_{t+1}^{k,\ell}(\boldsymbol{\lambda}, \boldsymbol{R}) \equiv \sum_{j \neq k: 1, \ldots, K} \mathbf{1}[\nu_t^j(\boldsymbol{r}_j) = \ell] A_{t+1}^j(\lambda_j), \quad \ell = 1, \ldots, L, \quad t = 0, 1, \ldots,$$

and this batch of $N_{t+1}^{k,\ell}(\boldsymbol{\lambda}, \boldsymbol{R}) + 1$ cells therefore arrive at the ℓ^{th} output queue during the time slot $[t, t+1)$, where $Q_t^\ell(\boldsymbol{\lambda}, \boldsymbol{R})$ cells are already awaiting transmission. Consequently, we have

$$H_t^k(\boldsymbol{\lambda}, \boldsymbol{R}) = \sum_{\ell=1}^{L} \mathbf{1}[\nu_t^k(\boldsymbol{r}_k) = \ell] \left(Q_t^\ell(\boldsymbol{\lambda}, \boldsymbol{R}) + J_t^{k,\ell}(\boldsymbol{\lambda}, \boldsymbol{R}) \right), \quad t = 0, 1, \ldots \tag{3.2}$$

where for each $\ell = 1, \ldots, L$, $J_t^{k,\ell}(\boldsymbol{\lambda}, \boldsymbol{R})$ denotes the random position of the fictitious cell in the batch of size $N_{t+1}^{k,\ell}(\boldsymbol{\lambda}, \boldsymbol{R}) + 1$. It is plain that for $n = 0, 1, \ldots, K-1$,

$$\mathbf{P}[J_t^{k,\ell}(\boldsymbol{\lambda}, \boldsymbol{R}) = i+1 | N_{t+1}^{k,\ell}(\boldsymbol{\lambda}, \boldsymbol{R}) = n] = \frac{1}{n+1}, \quad i = 0, 1, \ldots, n$$

so that

$$\mathbf{P}[J_t^{k,\ell}(\boldsymbol{\lambda}, \boldsymbol{R}) = i+1] = \sum_{n=i}^{K-1} \frac{1}{n+1} \mathbf{P}[N_{t+1}^{k,\ell}(\boldsymbol{\lambda}, \boldsymbol{R}) = n], \quad i = 0, 1, \ldots, K-1.$$

Moreover, the rvs $\boldsymbol{Q}_t(\boldsymbol{\lambda}, \boldsymbol{R}) \equiv (Q_t^1(\boldsymbol{\lambda}, \boldsymbol{R}), \ldots, Q_t^L(\boldsymbol{\lambda}, \boldsymbol{R}))$, $\nu_t^k(\boldsymbol{r}_k)$ and $\left(N_{t+1}^{k,1}(\boldsymbol{\lambda}, \boldsymbol{R}), J_t^{k,1}(\boldsymbol{\lambda}, \boldsymbol{R}), \ldots, N_{t+1}^{k,L}(\boldsymbol{\lambda}, \boldsymbol{R}), J_t^{k,L}(\boldsymbol{\lambda}, \boldsymbol{R})\right)$ are mutually independent under the unforced operational assumptions.

4 The Steady–State Regime

As some of the results below are concerned with performance measures for the system in statistical equilibrium, we now discuss the existence of such a steady–state regime in some details. To set the notation, for any sequence of $\mathbb{R}^d$–valued rvs $\{X_t,\ t = 0, 1, \ldots\}$, we denote its weak limit by X (as t goes to ∞) whenever it exists and write $X_t \Longrightarrow_t X$ to denote this weak convergence [2]. We call X the stationary version of the sequence $\{X_t,\ t = 0, 1, \ldots\}$.

The recursions (2.1) are very similar to the Lindley recursion for single server queues, and by arguments similar to those used in that context, we can show the following facts: Define

$$\rho_\ell(\boldsymbol{\lambda}, \boldsymbol{R}) \equiv \sum_{k=1}^{K} \lambda_k r_{k\ell}, \quad \ell = 1, \ldots, L \tag{4.1}$$

as the offered load to the ℓ^{th} output buffer. Whenever the conditions $\rho_\ell(\boldsymbol{\lambda}, \boldsymbol{R}) < 1$, $\ell = 1, \ldots, L$, are satisfied simultaneously, there exists an $\mathbb{N}^L$–valued rv $\boldsymbol{Q}(\boldsymbol{\lambda}, \boldsymbol{R}) \equiv (Q^1(\boldsymbol{\lambda}, \boldsymbol{R}), \ldots, Q^L(\boldsymbol{\lambda}, \boldsymbol{R}))$ such that $\boldsymbol{Q}_t(\boldsymbol{\lambda}, \boldsymbol{R}) \Longrightarrow_t \boldsymbol{Q}(\boldsymbol{\lambda}, \boldsymbol{R})$. In such circumstances, the system is termed *stable* and $\boldsymbol{Q}(\boldsymbol{\lambda}, \boldsymbol{R})$ is called the steady–state queue size vector or the queue size in statistical equilibrium.

If for some $\ell = 1, \ldots, L$, we only have $\rho_\ell(\boldsymbol{\lambda}, \boldsymbol{R}) < 1$, then the one–dimensional convergence $Q_t^\ell(\boldsymbol{\lambda}, \boldsymbol{R}) \Longrightarrow_t Q^\ell(\boldsymbol{\lambda}, \boldsymbol{R})$ still takes place, in which case the ℓ^{th} output queue is said to be stable.

We now turn to delay measures. Fix $\ell = 1, \ldots, L$, and assume the stability condition $\rho_\ell(\boldsymbol{\lambda}, \boldsymbol{R}) < 1$. For each $n = 1, 2, \ldots$, with t_n denoting the arrival epoch of the batch containing the n^{th} cell, we have the relation $W_n^\ell(\boldsymbol{\lambda}, \boldsymbol{R}) = Q_{t_n}^\ell(\boldsymbol{\lambda}, \boldsymbol{R})$. Because the arrival of batches to the ℓ^{th} output port is governed by the Bernoulli sequence $\{\mathbf{1}[\xi_{t+1}^\ell(\boldsymbol{\lambda}, \boldsymbol{R}) > 0],\ t = 0, 1, \ldots\}$, we get $W_n^\ell(\boldsymbol{\lambda}, \boldsymbol{R}) \Longrightarrow_n Q^\ell(\boldsymbol{\lambda}, \boldsymbol{R})$ upon invoking the property that Bernoulli arrivals see time average (BASTA) [12]. Hence, $W_n^\ell(\boldsymbol{\lambda}, \boldsymbol{R}) \Longrightarrow_n W^\ell(\boldsymbol{\lambda}, \boldsymbol{R})$ with $W^\ell(\boldsymbol{\lambda}, \boldsymbol{R}) =_{st} Q^\ell(\boldsymbol{\lambda}, \boldsymbol{R})$.

For each $n = 1, 2, \ldots$, let $G_n^\ell(\boldsymbol{\lambda}, \boldsymbol{R})$ denote the size of the batch that contains the n^{th} cell to arrive at the ℓ^{th} output port. Interpreting $B_n^\ell(\boldsymbol{\lambda}, \boldsymbol{R})$ as the (random) position of the n^{th} cell within this batch, we readily see that

$$\mathbf{P}[B_n^\ell(\boldsymbol{\lambda}, \boldsymbol{R}) = i] = \sum_{j=i}^{K} \frac{1}{j} \mathbf{P}[G_n^\ell(\boldsymbol{\lambda}, \boldsymbol{R}) = j], \quad i = 1, \ldots, K. \tag{4.2}$$

Because batch sizes are i.i.d. rvs all distributed according to the rv $\xi_1^\ell(\boldsymbol{\lambda}, \boldsymbol{R})$, it is well known [3] that $G_n^\ell(\boldsymbol{\lambda}, \boldsymbol{R}) \Longrightarrow_n G^\ell(\boldsymbol{\lambda}, \boldsymbol{R})$, where the rv $G^\ell(\boldsymbol{\lambda}, \boldsymbol{R})$ is distributed according to

$$\mathbf{P}[G^\ell(\boldsymbol{\lambda}, \boldsymbol{R}) = i] = \frac{1}{\mathbf{E}[\xi_1^\ell(\boldsymbol{\lambda}, \boldsymbol{R})]} i \mathbf{P}[\xi_1^\ell(\boldsymbol{\lambda}, \boldsymbol{R}) = i], \quad i = 1, \ldots, K. \tag{4.3}$$

Therefore, $B_n^\ell(\boldsymbol{\lambda}, \boldsymbol{R}) \Longrightarrow_n B^\ell(\boldsymbol{\lambda}, \boldsymbol{R})$ with

$$\begin{aligned} \mathbf{P}[B^\ell(\boldsymbol{\lambda}, \boldsymbol{R}) = i] &= \frac{1}{\mathbf{E}[\xi_1^\ell(\boldsymbol{\lambda}, \boldsymbol{R})]} \sum_{j=i}^{K} \frac{1}{j} \cdot j \mathbf{P}[\xi_1^\ell(\boldsymbol{\lambda}, \boldsymbol{R}) = j] \\ &= \frac{1}{\mathbf{E}[\xi_1^\ell(\boldsymbol{\lambda}, \boldsymbol{R})]} \mathbf{P}[\xi_1^\ell(\boldsymbol{\lambda}, \boldsymbol{R}) \geq i], \quad i = 1, \ldots, K. \end{aligned} \tag{4.4}$$

In other words, the rv $B^\ell(\boldsymbol{\lambda}, \boldsymbol{R})$ is the forward recurrence time associated with $\xi_1^\ell(\boldsymbol{\lambda}, \boldsymbol{R})$.

Because for each $n = 1, 2, \ldots$, the rvs $W_n^\ell(\boldsymbol{\lambda}, \boldsymbol{R})$ and $B_n^\ell(\boldsymbol{\lambda}, \boldsymbol{R})$ are independent, we obtain from (3.1) that $D_n^\ell(\boldsymbol{\lambda}, \boldsymbol{R}) \Longrightarrow_n D^\ell(\boldsymbol{\lambda}, \boldsymbol{R})$ for some rv $D^\ell(\boldsymbol{\lambda}, \boldsymbol{R})$ given by $D^\ell(\boldsymbol{\lambda}, \boldsymbol{R}) =_{st} Q^\ell(\boldsymbol{\lambda}, \boldsymbol{R}) + B^\ell(\boldsymbol{\lambda}, \boldsymbol{R})$ with $W^\ell(\boldsymbol{\lambda}, \boldsymbol{R})$ and $B^\ell(\boldsymbol{\lambda}, \boldsymbol{R})$ independent rvs.

In view of the independence mentioned at the end of Section 3, whenever $\boldsymbol{Q}_t(\boldsymbol{\lambda}, \boldsymbol{R}) \Longrightarrow_t \boldsymbol{Q}(\boldsymbol{\lambda}, \boldsymbol{R})$, we conclude from (3.2) that there exists an $\mathbb{N}^L$–valued rv $\boldsymbol{H}(\boldsymbol{\lambda}, \boldsymbol{R}) = (H^1(\boldsymbol{\lambda}, \boldsymbol{R}), \ldots, H^L(\boldsymbol{\lambda}, \boldsymbol{R}))$ such that $\boldsymbol{H}_t(\boldsymbol{\lambda}, \boldsymbol{R}) \Longrightarrow_t \boldsymbol{H}(\boldsymbol{\lambda}, \boldsymbol{R})$ and

$$H^k(\boldsymbol{\lambda}, \boldsymbol{R}) =_{st} \sum_{\ell=1}^{L} \mathbf{1}[\nu_0^k(\boldsymbol{r}_k) = \ell] \left(Q^\ell(\boldsymbol{\lambda}, \boldsymbol{R}) + J_0^{k,\ell}(\boldsymbol{\lambda}, \boldsymbol{R}) \right), \quad k = 1, \ldots, K,$$

where the rvs $\boldsymbol{Q}(\boldsymbol{\lambda}, \boldsymbol{R})$, $\nu_0^k(\boldsymbol{r}_k)$ and $(J_0^{k,1}(\boldsymbol{\lambda}, \boldsymbol{R}), \ldots, J_0^{k,L}(\boldsymbol{\lambda}, \boldsymbol{R}))$ are mutually independent. Using the BASTA property, this time with respect to the arrival process $\{A_{t+1}^k(\lambda_k),\ t = 0, 1, \ldots\}$, we find $T_n^k(\boldsymbol{\lambda}, \boldsymbol{R}) \Longrightarrow_n T^k(\boldsymbol{\lambda}, \boldsymbol{R})$ with

$$T^k(\boldsymbol{\lambda}, \boldsymbol{R}) =_{st} H^k(\boldsymbol{\lambda}, \boldsymbol{R}), \quad k = 1, \ldots, K. \tag{4.5}$$

5 The Main Results

We now present the main stochastic comparison results that describe how changes in arrival rates and routing probabilities affect the various performance measures. To simplify the presentation, for each rate vector $\boldsymbol{\lambda}$ and routing matrix $\boldsymbol{R}$, we write

$$\boldsymbol{\gamma}_\ell(\boldsymbol{\lambda}, \boldsymbol{R}) \equiv (\lambda_1 r_{1\ell}, \ldots, \lambda_K r_{K\ell}), \quad \ell = 1, \ldots, L.$$

We begin with results concerning performance measures that are associated with a single output destination; proofs are available in Section 7.

Theorem 5.1 *Assume that for some $\ell = 1, \ldots, L$, the comparison*

$$\boldsymbol{\gamma}_\ell(\boldsymbol{\lambda}, \boldsymbol{R}) \prec^w \boldsymbol{\gamma}_\ell(\boldsymbol{\lambda}', \boldsymbol{R}') \tag{5.1}$$

holds. Then, we have

$$Q_t^\ell(\boldsymbol{\lambda}', \boldsymbol{R}') \leq_{icx} Q_t^\ell(\boldsymbol{\lambda}, \boldsymbol{R}), \quad t = 0, 1, \ldots. \tag{5.2}$$

If in addition $\rho_\ell(\boldsymbol{\lambda}, \boldsymbol{R}) < 1$, then in statistical equilibrium we have

$$Q^\ell(\boldsymbol{\lambda}', \boldsymbol{R}') \quad \leq_{icx} \quad Q^\ell(\boldsymbol{\lambda}, \boldsymbol{R}), \tag{5.3}$$

$$B^\ell(\boldsymbol{\lambda}', \boldsymbol{R}') \quad \leq_{st} \quad B^\ell(\boldsymbol{\lambda}, \boldsymbol{R}) \tag{5.4}$$

and

$$D^\ell(\boldsymbol{\lambda}', \boldsymbol{R}') \quad \leq_{icx} \quad D^\ell(\boldsymbol{\lambda}, \boldsymbol{R}). \tag{5.5}$$

Under (5.1), the stability condition $\rho_\ell(\boldsymbol{\lambda}, \boldsymbol{R}) < 1$ implies $\rho_\ell(\boldsymbol{\lambda}', \boldsymbol{R}') < 1$, so that the ℓ^{th} output queue is stable in both systems and the comparisons (5.3)–(5.5) are indeed meaningful.

We next turn to results concerning the delay measures associated with input ports. Throughout, for any element $\boldsymbol{x}$ in $\mathbb{R}^K$ we write $\boldsymbol{x}^{(k)}$ to denote the vector in $\mathbb{R}^{K-1}$ obtained from $\boldsymbol{x}$ by removing its k^{th} component, $k = 1, 2, \ldots, K$. The first set of results is presented in Theorem 5.2 and discussed in Section 8.

Theorem 5.2 *Fix $k = 1, \ldots, K$. Assume that the comparisons*

$$\boldsymbol{\gamma}_\ell(\boldsymbol{\lambda}, \boldsymbol{R}) \prec^w \boldsymbol{\gamma}_\ell(\boldsymbol{\lambda}', \boldsymbol{R}') \quad \text{and} \quad \boldsymbol{\gamma}_\ell(\boldsymbol{\lambda}, \boldsymbol{R})^{(k)} \prec^w \boldsymbol{\gamma}_\ell(\boldsymbol{\lambda}', \boldsymbol{R}')^{(k)}, \quad \ell = 1, \ldots, L$$

simultaneously hold, and that $\boldsymbol{r}_k = \boldsymbol{r}'_k$. If $\rho_\ell(\boldsymbol{\lambda}, \boldsymbol{R}) < 1$, $\ell = 1, \ldots, L$, then we have

$$T^k(\boldsymbol{\lambda}', \boldsymbol{R}') \leq_{icx} T^k(\boldsymbol{\lambda}, \boldsymbol{R}). \tag{5.6}$$

The conclusion (5.6) will simultaneously hold for *all* $k = 1, \ldots, K$ provided the conditions $\boldsymbol{R} = \boldsymbol{R}'$, and $\boldsymbol{\gamma}_\ell(\boldsymbol{\lambda}, \boldsymbol{R})^{(k)} \prec^w \boldsymbol{\gamma}_\ell(\boldsymbol{\lambda}', \boldsymbol{R})^{(k)}$, $k = 1, \ldots, K$, $\ell = 1, \ldots, L$, simultaneously hold, in which case the conditions $\boldsymbol{\gamma}_\ell(\boldsymbol{\lambda}, \boldsymbol{R}) \prec^w \boldsymbol{\gamma}_\ell(\boldsymbol{\lambda}', \boldsymbol{R})$, $\ell = 1, \ldots, L$, are now automatically implied [13, B.2, p. 109].

To formulate the second set of results concerning delay measures associated with input ports, we need to place restrictions on the switching matrices: The addressing scheme is said to be *input independent* if its switching matrix $\boldsymbol{R}$ has all its row identical, say $\boldsymbol{r}_k = \boldsymbol{r}$, $k = 1, \ldots, K$, for some vector $\boldsymbol{r}$ in $\mathcal{S}_L$. Bi-group and hot-spot traffic patterns are instances of input independent addressing schemes. Under this constraint, we explore how the routing vector $\boldsymbol{r}$ affects the delay performance measure (3.2), as the input rate vector $\boldsymbol{\lambda}$ remains fixed. The dependency on the pair $(\boldsymbol{\lambda}, \boldsymbol{R})$ will be abbreviated to read $(\boldsymbol{\lambda}, \boldsymbol{r})$, where $\boldsymbol{r}$ is the common row of $\boldsymbol{R}$. The main result along these lines is contained in Theorem 5.3 below, and its proof discussed in Section 9.

Theorem 5.3 *Fix $k = 1, \ldots, K$, and consider two input independent switching matrices $\boldsymbol{R}$ and $\boldsymbol{R}'$ with common rows $\boldsymbol{r}$ and $\boldsymbol{r}'$, respectively. If $\rho_\ell(\boldsymbol{\lambda}, \boldsymbol{R}') < 1$, $\ell = 1, \ldots, L$, then we also have*

$$T^k(\boldsymbol{\lambda}, \boldsymbol{r}) \leq_{icx} T^k(\boldsymbol{\lambda}, \boldsymbol{r}') \tag{5.7}$$

provided $\boldsymbol{r} \prec \boldsymbol{r}'$.

From Theorem 5.3, we immediately conclude that for input delay measures, the uniform addressing scheme $\boldsymbol{U} \equiv (u_{k\ell})$ (given by 1.2) is best amongst all input independent schemes.

Furthermore, in the comparison of Theorem 5.1, if the total load (4.1) to the ℓ^{th} output queue is constrained to some given value, then condition (5.1) is equivalent to

$$\boldsymbol{\gamma}_\ell(\boldsymbol{\lambda}, \boldsymbol{R}) \prec \boldsymbol{\gamma}_\ell(\boldsymbol{\lambda}', \boldsymbol{R}'). \tag{5.8}$$

Theorem 5.1 thus suggests a way to obtain lower and upper bounds on the queue size metrics (among other things) by seeking the "extremizers" in the conditions (5.8) under certain load constraints. This leads to the generic optimization problems discussed by the authors in [7]. There we also identify the worst and best cases, and under some special circumstances, show that uniform addressing and uniform traffic patterns exhibit optimality properties.

6 On Sums of Bernoulli Random Variables

To prepare for the proof of Theorem 5.1 given in the next section, we begin with several comparison results for sums of independent Bernoulli rvs; some of these facts are well known while others appear to be new. For p in $[0,1]$, let $X(p)$ denote a $\{0,1\}$–valued rv with $\mathbf{P}[X(p)=1]=p$. Moreover for $\boldsymbol{p}$ in $[0,1]^K$, we define the rv $S_K(\boldsymbol{p})$ as the sum

$$S_K(\boldsymbol{p}) \equiv \sum_{i=1}^{K} X_i(p_i)$$

where the rvs $X_1(p_1),\ldots,X_K(p_K)$ are assumed mutually independent. For any mapping $\varphi : I\!N \to \mathbb{R}$, we also define the mapping $\Phi_K : [0,1]^K \to \mathbb{R}$ by

$$\Phi_K(\boldsymbol{p}) \equiv \mathbf{E}[\varphi(S_K(\boldsymbol{p}))], \quad \boldsymbol{p} \in [0,1]^K. \tag{6.1}$$

Lemma 6.1 *For any mapping $\varphi : I\!N \to \mathbb{R}$, the mapping $\Phi_K : [0,1]^K \to \mathbb{R}$ given by (6.1) is*

1. *Schur–concave if φ is integer–convex;*

2. *increasing if φ is increasing.*

Claim 1 is established in [13, F.1, p. 360], and Claim 2 follows by an easy coupling argument. Lemma 6.1 easily translates into the following comparison results for sums of independent Bernoulli rvs.

Lemma 6.2 *Let $\boldsymbol{p}$ and $\boldsymbol{q}$ be vectors in $[0,1]^K$. Then the following statements hold:*

1. *If $\boldsymbol{p} \prec \boldsymbol{q}$, then $S_K(\boldsymbol{q}) \leq_{cx} S_K(\boldsymbol{p})$;*

2. *If $\boldsymbol{p} \prec^w \boldsymbol{q}$, then $S_K(\boldsymbol{q}) \leq_{icx} S_K(\boldsymbol{p})$.*

Proof. (Claim 1) For any integer–convex mapping $\varphi : I\!N \to \mathbb{R}$, the mapping Φ_K given by (6.1) is Schur–concave by Claim 1 of Lemma 6.1. The condition $\boldsymbol{p} \prec \boldsymbol{q}$ thus implies $\Phi_K(\boldsymbol{q}) \leq \Phi_K(\boldsymbol{p})$, and the conclusion $S_K(\boldsymbol{q}) \leq_{cx} S_K(\boldsymbol{p})$ follows from the definition of the ordering $\leq_{cx}$.

(Claim 2) As it is well known [13, 5.A.9, p. 123], the condition $\boldsymbol{p} \prec^w \boldsymbol{q}$ is equivalent to the existence of a vector $\boldsymbol{r}$ (a priori in $\mathbb{R}^K$) such that $\boldsymbol{r} \leq \boldsymbol{p}$ and $\boldsymbol{r} \prec \boldsymbol{q}$. The constraint $\boldsymbol{r} \leq \boldsymbol{p}$ is equivalent to $r_k \leq p_k$, $k = 1, \ldots, K$, whence $r_k \leq 1$, $k = 1, \ldots, K$, because $\boldsymbol{p}$ belongs to $[0,1]^K$. From $\boldsymbol{r} \prec \boldsymbol{q}$, we get $\min r_k \geq \min q_k \geq 0$, so that $r_k \geq 0$, $k = 1, \ldots, K$. Therefore, $\boldsymbol{r}$ is an element in $[0,1]^K$.

Consider now a mapping $\varphi : I\!N \to \mathbb{R}$ which is integer–convex and increasing. Upon invoking Lemma 6.1, we get from Claim 2 that $\Phi_K(\boldsymbol{r}) \leq \Phi_K(\boldsymbol{p})$ because $\boldsymbol{r} \leq \boldsymbol{p}$, and from Claim 1 that $\Phi_K(\boldsymbol{q}) \leq \Phi_K(\boldsymbol{r})$ because $\boldsymbol{r} \prec \boldsymbol{q}$. Hence, $\Phi_K(\boldsymbol{q}) \leq \Phi_K(\boldsymbol{p})$ and the comparison $S_K(\boldsymbol{q}) \leq_{icx} S_K(\boldsymbol{p})$ follows as in the first part of the proof. ■

Taking our cue from (4.4), with each non–zero vector $\boldsymbol{p}$ in $[0,1]^K$, we associate an $I\!N$–valued rv $B_K(\boldsymbol{p})$ with probability distribution given by

$$\mathbf{P}[B_K(\boldsymbol{p}) = i] \equiv \frac{1}{\mathbf{E}[S_K(\boldsymbol{p})]} \mathbf{P}[S_K(\boldsymbol{p}) \geq i], \quad i = 1, \ldots, K.$$

The rv $B_K(\boldsymbol{p})$ is known as the forward recurrence time associated with $S_K(\boldsymbol{p})$. For any mapping $\varphi : I\!N \to \mathbb{R}$, straightforward calculations show that

$$\begin{aligned} \mathbf{E}[\varphi(B_K(\boldsymbol{p}))] &= \frac{1}{\mathbf{E}[S_K(\boldsymbol{p})]} \sum_{i=1}^{K} \varphi(i) \sum_{j=i}^{K} \mathbf{P}[S_K(\boldsymbol{p}) = j] \\ &= \frac{1}{\mathbf{E}[S_K(\boldsymbol{p})]} \sum_{j=1}^{K} \mathbf{P}[S_K(\boldsymbol{p}) = j] \left(\sum_{i=1}^{j} \varphi(i) \right) \\ &= \frac{1}{\mathbf{E}[S_K(\boldsymbol{p})]} \mathbf{E}[\widehat{\varphi}(S_K(\boldsymbol{p}))], \quad \boldsymbol{p} \in [0,1]^K \end{aligned} \tag{6.2}$$

where the mapping $\widehat{\varphi} : I\!N \to \mathbb{R}$ is defined by

$$\widehat{\varphi}(0) \equiv 0, \quad \widehat{\varphi}(j) \equiv \sum_{i=1}^{j} \varphi(i), \quad j = 1, 2, \ldots. \tag{6.3}$$

Proposition 6.1 *Let $\boldsymbol{p}$ and $\boldsymbol{q}$ be non–zero vectors in $[0,1]^K$. If $\boldsymbol{p} \prec \boldsymbol{q}$, then*

$$B_K(\boldsymbol{q}) \leq_{st} B_K(\boldsymbol{p}). \tag{6.4}$$

Proof. We need to show that

$$\mathbf{E}[\varphi(B_K(\boldsymbol{q}))] \leq \mathbf{E}[\varphi(B_K(\boldsymbol{p}))] \tag{6.5}$$

for any increasing mapping $\varphi : I\!N \rightarrow \mathbb{R}$. By Claim 1 of Lemma 6.2, the condition $\boldsymbol{p} \prec \boldsymbol{q}$ implies $S_K(\boldsymbol{q}) \leq_{cx} S_K(\boldsymbol{p})$, whence

$$\mathbf{E}[\widehat{\varphi}(S_K(\boldsymbol{q}))] \leq \mathbf{E}[\widehat{\varphi}(S_K(\boldsymbol{p}))] \tag{6.6}$$

for any increasing mapping φ because the mapping $\widehat{\varphi} : I\!N \rightarrow \mathbb{R}$ is then integer–convex. We obtain (6.5) via (6.2) upon combining (6.6) with the equality $\mathbf{E}[S_K(\boldsymbol{p})] = \mathbf{E}[S_K(\boldsymbol{q})]$ derived from the condition $\boldsymbol{p} \prec \boldsymbol{q}$. ■

Under the condition $\boldsymbol{p} \prec \boldsymbol{q}$, the validity of (6.6) is an immediate consequence of Lemma 6.2 once we note the equality of the means. It is then natural to wonder whether the conclusion (6.4) still holds under the weaker condition $\boldsymbol{p} \prec^w \boldsymbol{q}$. In order to answer this question in the affirmative, we need the following result.

Proposition 6.2 *Let* $\boldsymbol{p}$ *and* $\boldsymbol{q}$ *be non–zero vectors in* $[0,1]^K$. *If* $\boldsymbol{p} \leq \boldsymbol{q}$, *then the comparison* $B_K(\boldsymbol{q}) \leq_{st} B_K(\boldsymbol{p})$ *also holds.*

To the best of the authors' knowledge, Proposition 6.2 appears to be new; its proof is given in Appendix A.

Proposition 6.3 *Let* $\boldsymbol{p}$ *and* $\boldsymbol{q}$ *be non–zero vectors in* $[0,1]^K$. *If* $\boldsymbol{p} \prec^w \boldsymbol{q}$, *then the comparison* $B_K(\boldsymbol{q}) \leq_{st} B_K(\boldsymbol{p})$ *still holds.*

Proof. As in the proof of Claim 2 of Lemma 6.2, the condition $\boldsymbol{p} \prec^w \boldsymbol{q}$ is equivalent to the existence of a vector $\boldsymbol{r}$ (in $[0,1]^K$) such that $\boldsymbol{r} \leq \boldsymbol{p}$ and $\boldsymbol{r} \prec \boldsymbol{q}$. The desired conclusion is now immediate once we note that by Proposition 6.1, we already have $B_K(\boldsymbol{q}) \leq_{st} B_K(\boldsymbol{r})$, and that $B_K(\boldsymbol{r}) \leq_{st} B_K(\boldsymbol{p})$ holds by Proposition 6.2. ■

7 A Proof of Theorem 5.1

We begin by noting that under (5.1) the comparison

$$\xi^{\ell}_{t+1}(\boldsymbol{\lambda}', \boldsymbol{R}') \leq_{icx} \xi^{\ell}_{t+1}(\boldsymbol{\lambda}, \boldsymbol{R}), \quad t = 0, 1, \ldots \tag{7.1}$$

is a simple rephrasing of Claim 2 of Lemma 6.2. Therefore the validity of (5.2) can be established by a straightforward induction argument as is done for the Lindley recursion [15, Theorem 8.6.2, p. 274]: The basis step follows by assumption because $Q_0^\ell(\boldsymbol{\lambda}', \boldsymbol{R}') = Q_0^\ell(\boldsymbol{\lambda}, \boldsymbol{R}) = 0$. Next, we assume that $Q_t^\ell(\boldsymbol{\lambda}', \boldsymbol{R}') \leq_{icx} Q_t^\ell(\boldsymbol{\lambda}, \boldsymbol{R})$ for some $t = 0, 1, \ldots$. Obviously,

$$[Q_t^\ell(\boldsymbol{\lambda}', \boldsymbol{R}') - 1]^+ \leq_{icx} [Q_t^\ell(\boldsymbol{\lambda}, \boldsymbol{R}) - 1]^+ \tag{7.2}$$

because $\leq_{icx}$ propagates under convex increasing transformations. The rvs $\xi_{t+1}^\ell(\boldsymbol{\lambda}', \boldsymbol{R}')$ and $Q_t^\ell(\boldsymbol{\lambda}', \boldsymbol{R}')$ (resp. $\xi_{t+1}^\ell(\boldsymbol{\lambda}, \boldsymbol{R})$ and $Q_t^\ell(\boldsymbol{\lambda}, \boldsymbol{R})$) being independent, we conclude from (7.1) and (7.2) that the comparison $Q_{t+1}^\ell(\boldsymbol{\lambda}', \boldsymbol{R}') \leq_{icx} Q_{t+1}^\ell(\boldsymbol{\lambda}, \boldsymbol{R})$ holds because $\leq_{icx}$ is preserved under convolution. This completes the induction step.

Under (5.1), the stability condition $\rho_\ell(\boldsymbol{\lambda}, \boldsymbol{R}) < 1$ implies $\rho_\ell(\boldsymbol{\lambda}', \boldsymbol{R}') < 1$, so that the ℓ^{th} output queue is stable in both cases. It is simple matter to show (say by transform techniques) that the steady–state queue size rvs $Q^\ell(\boldsymbol{\lambda}, \boldsymbol{R})$ and $Q^\ell(\boldsymbol{\lambda}', \boldsymbol{R}')$ both have all their moments finite. On the other hand we also note that $Q_t^\ell(\boldsymbol{\lambda}, \boldsymbol{R}) \leq_{st} Q_{t+1}^\ell(\boldsymbol{\lambda}, \boldsymbol{R}) \leq_{st} Q^\ell(\boldsymbol{\lambda}, \boldsymbol{R})$ for all $t = 0, 1, \ldots$; this monotonicity result follows by an easy induction argument [19, Theorem 2.2.8, p. 48] which is omitted for the sake of brevity. Combining these remarks, we readily conclude that the rvs $\{Q_t^\ell(\boldsymbol{\lambda}, \boldsymbol{R}),\ t = 0, 1, \ldots\}$ are uniformly integrable, whence

$$\lim_{t \to \infty} \mathbf{E}\left[Q_t^\ell(\boldsymbol{\lambda}, \boldsymbol{R})\right] = \mathbf{E}\left[Q^\ell(\boldsymbol{\lambda}, \boldsymbol{R})\right],$$

and Proposition 1.3.2 of [19, p. 10] can now be applied on (5.2) to yield the conclusion (5.3).

The comparison (5.4) is a restatement of Proposition 6.3; in particular $B^\ell(\boldsymbol{\lambda}', \boldsymbol{R}') \leq_{icx} B^\ell(\boldsymbol{\lambda}, \boldsymbol{R})$ by virtue of the fact that the ordering $\leq_{st}$ is stronger than $\leq_{icx}$. Finally, the comparison (5.5) follows from (3.1) (in statistical equilibrium) upon combining this last remark with the independence of the rvs.

8 A Proof of Theorem 5.2

The steps leading to (5.6) can be traced back to the following remark which readily follows from (3.2): For any mapping $\varphi : \mathbb{R} \to \mathbb{R}$, the expression

$$\mathbf{E}[\varphi(H_t^k(\boldsymbol{\lambda}, \boldsymbol{R}))] = \sum_{\ell=1}^{L} r_{k\ell} \mathbf{E}\left[\varphi\left(Q_t^\ell(\boldsymbol{\lambda}, \boldsymbol{R}) + J_t^{k,\ell}(\boldsymbol{\lambda}, \boldsymbol{R})\right)\right], \quad t = 0, 1, \ldots \tag{8.1}$$

holds for each $k = 1, \ldots, K$. This fact suggests the need for the following intermediary lemma.

Lemma 8.1 *Fix $k = 1, \ldots, K$, and $\ell = 1, \ldots, L$. If the condition*

$$\boldsymbol{\gamma}_\ell(\boldsymbol{\lambda}, \boldsymbol{R})^{(k)} \prec^w \boldsymbol{\gamma}_\ell(\boldsymbol{\lambda}', \boldsymbol{R}')^{(k)} \tag{8.2}$$

holds, then

$$J_t^{k\ell}(\boldsymbol{\lambda}', \boldsymbol{R}') \leq_{icx} J_t^{k\ell}(\boldsymbol{\lambda}, \boldsymbol{R}), \quad t = 0, 1, \ldots. \tag{8.3}$$

Proof. By Claim 2 of Lemma 6.2, the comparison

$$N_{t+1}^{k\ell}(\boldsymbol{\lambda}', \boldsymbol{R}') \leq_{icx} N_{t+1}^{k\ell}(\boldsymbol{\lambda}, \boldsymbol{R}), \quad t = 0, 1, \ldots \tag{8.4}$$

holds under (8.2).

Next, for any mapping $\varphi : \mathbb{R} \to \mathbb{R}$, straightforward calculations show that

$$\begin{aligned} \mathbf{E}[\varphi(J_t^{k,\ell}(\boldsymbol{\lambda}, \boldsymbol{R}))] &= \sum_{i=0}^{K-1} \varphi(i+1) \sum_{n=i}^{K-1} \frac{1}{n+1} \mathbf{P}[N_{t+1}^{k,\ell}(\boldsymbol{\lambda}, \boldsymbol{R}) = n] \\ &= \sum_{n=0}^{K-1} \mathbf{P}[N_{t+1}^{k,\ell}(\boldsymbol{\lambda}, \boldsymbol{R}) = n] \frac{1}{n+1} \sum_{i=0}^{n} \varphi(i+1) \\ &= \mathbf{E}[\varphi_{av}(N_{t+1}^{k\ell}(\boldsymbol{\lambda}, \boldsymbol{R}))], \quad t = 0, 1, \ldots \end{aligned} \tag{8.5}$$

where we have set

$$\varphi_{av}(n) \equiv \frac{1}{n+1} \sum_{i=0}^{n} \varphi(i+1), \quad n = 0, 1, \ldots. \tag{8.6}$$

In Appendix B, we show that the mapping $\varphi_{av} : I\!N \to \mathbb{R}$ is integer–increasing convex whenever the mapping $\varphi : \mathbb{R} \to \mathbb{R}$ is increasing convex. Therefore, for every increasing convex mapping $\varphi : \mathbb{R} \to \mathbb{R}$, the inequality (8.4) yields

$$\mathbf{E}[\varphi_{av}(N_{t+1}^{k,\ell}(\boldsymbol{\lambda}', \boldsymbol{R}'))] \leq \mathbf{E}[\varphi_{av}(N_{t+1}^{k,\ell}(\boldsymbol{\lambda}, \boldsymbol{R}))], \quad t = 0, 1, \ldots$$

and the conclusion (8.3) immediately follows via (8.5). ■

The next result can be interpreted as a transient version of Theorem 5.2.

Proposition 8.1 *Fix $k = 1, \ldots, K$ and assume conditions (5.1) and (8.2) to hold for each $\ell = 1, \ldots, L$. If $\boldsymbol{r}_k = \boldsymbol{r}'_k$, then*

$$H_t^k(\boldsymbol{\lambda}', \boldsymbol{R}') \leq_{icx} H_t^k(\boldsymbol{\lambda}, \boldsymbol{R}), \quad t = 0, 1, \ldots. \tag{8.7}$$

Proof. Fix $\ell = 1, \ldots, L$ and $t = 0, 1, \ldots$. Combining Theorem 5.1 and Lemma 8.1 we immediately get under the enforced independence that

$$Q_t^\ell(\boldsymbol{\lambda}', \boldsymbol{R}') + J_t^{k,\ell}(\boldsymbol{\lambda}', \boldsymbol{R}') \leq_{icx} Q_t^\ell(\boldsymbol{\lambda}, \boldsymbol{R}) + J_t^{k,\ell}(\boldsymbol{\lambda}, \boldsymbol{R})$$

as we recall that $\leq_{icx}$ is closed under convolution. Therefore, for each $\ell = 1, \ldots, L$, we have

$$\mathbf{E}\left[\varphi(Q_t^\ell(\boldsymbol{\lambda}', \boldsymbol{R}') + J_t^{k,\ell}(\boldsymbol{\lambda}', \boldsymbol{R}'))\right] \leq \mathbf{E}\left[\varphi(Q_t^\ell(\boldsymbol{\lambda}, \boldsymbol{R}) + J_t^{k,\ell}(\boldsymbol{\lambda}, \boldsymbol{R}))\right] \tag{8.8}$$

for every increasing convex mapping $\varphi : \mathbb{R} \to \mathbb{R}$. In that case, with $\boldsymbol{r}_k = \boldsymbol{r}'_k$, we combine (8.1) and (8.8) to get

$$\mathbf{E}\left[\varphi(H_t^k(\boldsymbol{\lambda}', \boldsymbol{R}'))\right] \leq \mathbf{E}\left[\varphi(H_t^k(\boldsymbol{\lambda}, \boldsymbol{R}))\right]$$

and the conclusion (8.7) is obtained. ■

We are now in position to complete the proof of Theorem 5.2: Assume the system to be stable, i.e., $\rho_\ell(\boldsymbol{\lambda}, \boldsymbol{R}) < 1$, $\ell = 1, \ldots, L$. It was already pointed out in Section 7 that for each $\ell = 1, \ldots, L$, the rvs $\{Q_t^\ell(\boldsymbol{\lambda}, \boldsymbol{R}),\ t = 0, 1, \ldots\}$ are uniformly integrable. On the other hand, for each $k = 1, 2, \ldots, K$, the rvs $\{J_t^{k,\ell}(\boldsymbol{\lambda}, \boldsymbol{R}),\ t = 0, 1, \ldots\}$ all have bounded support $\{1, \ldots, K\}$, whence are uniformly integrable, and it is plain from (3.2) that the rvs $\{H_t^k(\boldsymbol{\lambda}, \boldsymbol{R}),\ t = 0, 1, \ldots\}$ are also uniformly integrable. Applying Proposition 1.3.2 of [19, p. 10] to the transient comparison (8.7), we get

$$H^k(\boldsymbol{\lambda}', \boldsymbol{R}') \leq_{icx} H^k(\boldsymbol{\lambda}, \boldsymbol{R}), \quad t = 0, 1, \ldots$$

in statistical equilibrium, and the conclusion (5.6) is now immediate from (4.5).

9 A Proof of Theorem 5.3

The proof of Theorem 5.3 relies on several notions of stochastic convexity which have recently received a great deal of attention [16, 17]: With Θ denoting a convex subset of $\mathbb{R}$, we say that the collection $\{X(\theta),\ \theta \in \Theta\}$ of $\mathbb{R}$–valued rvs is

1. stochastically increasing and convex – in short *SICX* – if for any increasing and convex function $\varphi : \mathbb{R} \to \mathbb{R}$, the mapping $\theta \to \mathbf{E}[\varphi(X(\theta))]$ is increasing and convex on Θ (whenever defined);

2. stochastically increasing and convex in sample path sense – in short *SICX(sp)* – if for any four values θ_i, $i = 1,2,3,4$, in Θ, satisfying $\theta_1 \leq \theta_2 \leq \theta_3 \leq \theta_4$ and $\theta_1 + \theta_4 = \theta_2 + \theta_3$, there exist four rvs $\widehat{X}_i$, $i = 1,2,3,4$, defined on a common probability space such that $\widehat{X}_i =_{st} X(\theta_i)$, $i = 1,2,3,4$, and the four rvs satisfy the inequalities

$$\widehat{X}_2 + \widehat{X}_3 \leq \widehat{X}_1 + \widehat{X}_4 \quad \text{and} \quad \widehat{X}_j \leq \widehat{X}_4, \quad j = 1,2,3.$$

The reader is referred to [16, 17] for proofs and additional details concerning these notions of stochastic convexity. For our purpose here, the single most important fact relates to the stochastic convexity of Bernoulli rvs, a property first pointed out in [16, Example 4.4, p. 438]:

Lemma 9.1 *For $0 \leq p \leq 1$, let $X(p)$ denote a $\{0,1\}$–valued Bernoulli rv with $\mathbf{P}[X(p) = 1] = p$. Then $\{X(p),\ p \in [0,1]\}$ is SICX(sp).*

Next, consider an arrival vector $\boldsymbol{\lambda}$ and a switching matrix $\boldsymbol{R}$ which is input independent, with common row vector $\boldsymbol{r}$. For each $k = 1, \ldots, K$, and $\ell = 1, \ldots, L$, in the notation of Lemma 9.1, we have $\mathbf{1}[\nu_t^k(\boldsymbol{r}_k) = \ell] =_{st} X(r_\ell)$ so that the distribution of the rvs $\xi_{t+1}^{\ell}(\boldsymbol{\lambda}, \boldsymbol{r})$ and $N_{t+1}^{k,\ell}(\boldsymbol{\lambda}, \boldsymbol{r})$ are fully determined by the vectors $r_\ell \boldsymbol{\lambda}$ and $r_\ell \boldsymbol{\lambda}^{(k)}$, respectively. We account for this dependency by modifying the notation to $\xi_{t+1}^{\ell}(r_\ell \boldsymbol{\lambda})$ and $N_{t+1}^{k,\ell}(r_\ell \boldsymbol{\lambda}^{(k)})$, respectively; similar modifications are made for derived quantities.

Lemma 9.2 *Fix $k = 1, \ldots, K$, $\ell = 1, \ldots, L$ and $t = 0, 1, \ldots$. For every input rate vector $\boldsymbol{\lambda}$, the following statements hold:*

1. *The collection of rvs $\{J_t^{k\ell}(r_\ell \boldsymbol{\lambda}^{(k)}),\ r_\ell \in [0,1]\}$ is SICX;*

2. *The collection of rvs $\{Q_t^{\ell}(r_\ell \boldsymbol{\lambda}),\ r_\ell \in [0,1]\}$ is SICX.*

Proof. (Claim 1) By Lemma 9.1, the collection of Bernoulli rvs $\{\mathbf{1}[\nu_t^j(\boldsymbol{r}_j) = \ell],\ r_\ell \in [0,1]\}$ is SICX(sp), and so is the collection of rvs $\{A_{t+1}^j(\lambda_j)\mathbf{1}[\nu_t^j(\boldsymbol{r}_j) = \ell],\ r_\ell \in [0,1]\}$, $j = 1, \ldots, K$. Because the SICX(sp) property is stable under convolution [16, Theorem 3.10, p. 436], the collection of rvs $\{N_t^{k\ell}(r_\ell \boldsymbol{\lambda}^{(k)}),\ r_\ell \in [0,1]\}$ is SICX(sp), thus also SICX [16, Theorem 3.6, p. 435]. The desired conclusion readily follows from this last fact, the relation (8.5) and Lemma B.1.

(Claim 2) By the argument given in the proof of Claim 1, the collection of rvs $\{\xi^\ell_{t+1}(r_\ell \boldsymbol{\lambda}),\ r_\ell \in [0,1]\}$ is SICX(sp). Because this property is preserved under convex increasing transformations [16, Proposition 3.5, p.434], an easy induction argument using the recursion (2.1) shows that the collection of rvs $\{Q^\ell_t(r_\ell \boldsymbol{\lambda}),\ r_\ell \in [0,1]\}$ is SICX(sp), thus SICX [16, Theorem 3.6, p. 435]. ■

Proposition 9.1 *Fix $k = 1, \ldots, K$, and consider two input independent switching matrices $\boldsymbol{R}$ and $\boldsymbol{R}'$ with common rows $\boldsymbol{r}$ and $\boldsymbol{r}'$, respectively. If $\boldsymbol{r} \prec \boldsymbol{r}'$, then*

$$H^k_t(\boldsymbol{\lambda}, \boldsymbol{r}) \leq_{icx} H^k_t(\boldsymbol{\lambda}, \boldsymbol{r}'), \quad \boldsymbol{\lambda} \in [0,1]^K, \quad t = 0, 1, \ldots. \tag{9.1}$$

Proof. Fix the input rate vector $\boldsymbol{\lambda}$ and $t = 0, 1, \ldots$. For any mapping $\varphi : \mathbb{R} \to \mathbb{R}$, we can use (8.1) to write

$$\mathbf{E}\left[\varphi(H^k_t(\boldsymbol{\lambda}, \boldsymbol{r}))\right] = \sum_{\ell=1}^{L} r_\ell \Phi^k_t(r_\ell \boldsymbol{\lambda}), \quad \boldsymbol{r} \in \mathcal{S}_L$$

where for each $\ell = 1, \ldots, L$, we have set

$$\Phi^k_t(r_\ell \boldsymbol{\lambda}) \equiv \mathbf{E}\left[\varphi\left(Q^\ell_t(\boldsymbol{\lambda}, \boldsymbol{r}) + J^{k,\ell}_t(\boldsymbol{\lambda}, \boldsymbol{r})\right)\right], \quad r_\ell \in [0,1];$$

these expectations are indeed independent of ℓ.

Under the enforced independence assumptions, we see by Lemma 9.2 that the collection of rvs $\{Q^\ell_t(r_\ell \boldsymbol{\lambda}) + J^{k,\ell}_t(r_\ell \boldsymbol{\lambda}^{(k)}),\ r_\ell \in [0,1]\}$ is also SICX [17, Theorem 5.3, p. 521]. Therefore, for any increasing and convex mapping $\varphi : \mathbb{R} \to \mathbb{R}$, the mappings $r_\ell \to \Phi^k_t(r_\ell \boldsymbol{\lambda})$, $\ell = 1, \ldots, L$, are increasing and convex, and the mappings $r_\ell \to r_\ell \Phi^k_t(r_\ell \boldsymbol{\lambda})$, $\ell = 1, \ldots, L$, are therefore convex on the interval $[0,1]$. By a well–known result of Schur [13, C.1., p. 64], the mapping $\boldsymbol{r} \to \mathbf{E}\left[\varphi(H^k_t(\boldsymbol{\lambda}, \boldsymbol{r}))\right]$ is thus Schur–convex on $\mathcal{S}_L$, whence

$$\mathbf{E}\left[\varphi(H^k_t(\boldsymbol{\lambda}, \boldsymbol{r}))\right] \leq \mathbf{E}\left[\varphi(H^k_t(\boldsymbol{\lambda}, \boldsymbol{r}'))\right]$$

whenever $\boldsymbol{r} \prec \boldsymbol{r}'$, and the conclusion (9.1) follows from the definition of $\leq_{icx}$. ■

The final step to establish (5.7) from (9.1) is simply as in the proof of Theorem 5.2, and is therefore omitted.

A Appendix

Let $\boldsymbol{p}^K = (p_1, \ldots, p_K)$ denote an arbitrary element of $[0,1]^K$. For any function $\varphi : I\!N \to \mathbb{R}$, we define

$$\Phi^\star_K(\boldsymbol{p}^K) \equiv \mathbf{E}\left[\varphi(B_K(\boldsymbol{p}^K))\right] = \frac{\widehat{\Phi}_K(\boldsymbol{p}^K)}{\mathbf{E}[S_K(\boldsymbol{p}^K)]}, \quad \boldsymbol{p}^K \in [0,1]^K$$

where with the notation (6.3), we have set

$$\widehat{\Phi}_K(\boldsymbol{p}^K) \equiv \mathbf{E}\left[\widehat{\varphi}(S_K(\boldsymbol{p}^K))\right], \quad \boldsymbol{p}^K \in [0,1]^K.$$

The main result is contained in Lemma A.1.

Lemma A.1 *Consider an increasing mapping $\varphi : I\!N \to \mathbb{R}$. For each $K = 1, 2, \ldots,$ the mapping $\Phi^\star_K : [0,1]^K \to \mathbb{R}$ is increasing, or equivalently, $B_K(\boldsymbol{p}^K) \leq_{st} B_K(\boldsymbol{q}^K)$ whenever $\boldsymbol{p}^K \leq \boldsymbol{q}^K$ in $[0,1]^K$.*

Proof. For each $K = 0, 1, \ldots,$ we view any element $\boldsymbol{p}^{K+1} = (p_1, \ldots, p_K, p_{K+1})$ of $[0,1]^{K+1}$ as the concatenation of the vector $\boldsymbol{p}^K = (p_1, \ldots, p_K,)$ (in $[0,1]^K$) with the scalar p_{K+1} (in $[0,1]$). With this notation, we find that

$$\begin{aligned}
\widehat{\Phi}_{K+1}(\boldsymbol{p}^{K+1}) &= \mathbf{E}[\widehat{\varphi}\left(S_K(\boldsymbol{p}^K) + X_{K+1}(p_{K+1})\right)] \\
&= p_{K+1}\mathbf{E}[\widehat{\varphi}(S_K(\boldsymbol{p}^K) + 1)] + (1 - p_{K+1})\mathbf{E}[\widehat{\varphi}(S_K(\boldsymbol{p}^K))] \\
&= p_{K+1}\mathbf{E}[\widehat{\varphi}(S_K(\boldsymbol{p}^K) + 1) - \widehat{\varphi}(S_K(\boldsymbol{p}^K))] + \mathbf{E}[\widehat{\varphi}(S_K(\boldsymbol{p}^K))] \\
&= p_{K+1}\mathbf{E}[\varphi(S_K(\boldsymbol{p}^K) + 1)] + \mathbf{E}[\widehat{\varphi}(S_K(\boldsymbol{p}^K))], \qquad \text{(A.1)}
\end{aligned}$$

and

$$\mathbf{E}[S_{K+1}(\boldsymbol{p}^{K+1})] = \mathbf{E}[S_K(\boldsymbol{p}^K)] + p_{K+1}. \qquad \text{(A.2)}$$

Therefore,

$$\Phi^\star_{K+1}(\boldsymbol{p}^{K+1}) = \frac{p_{K+1}\mathbf{E}[\varphi(S_K(\boldsymbol{p}^K) + 1)] + \mathbf{E}[\widehat{\varphi}(S_K(\boldsymbol{p}^K))]}{p_{K+1} + \mathbf{E}[S_K(\boldsymbol{p}^K)]}. \qquad \text{(A.3)}$$

To show that the mapping $\Phi^\star_{K+1} : [0,1]^{K+1} \to \mathbb{R}$ is increasing, it suffices to show that $\Phi^\star_{K+1}$ is increasing in p_{K+1} (with $\boldsymbol{p}^K$ fixed). Differentiating (A.3) with respect to p_{K+1}, we find

$$\frac{\partial}{\partial p_{K+1}}\Phi^\star_{K+1}(\boldsymbol{p}^{K+1}) = \frac{\mathbf{E}[S_K(\boldsymbol{p}^K)]\mathbf{E}[\varphi(S_K(\boldsymbol{p}^K) + 1)] - \mathbf{E}[\widehat{\varphi}(S_K(\boldsymbol{p}^K))]}{(p_{K+1} + \mathbf{E}[S_K(\boldsymbol{p}^K)])^2}$$

and the desired conclusion now follows if we can show

$$\mathbf{E}[S_K(\boldsymbol{p}^K)]\mathbf{E}[\varphi(S_K(\boldsymbol{p}^K)+1)] - \mathbf{E}[\widehat{\varphi}(S_K(\boldsymbol{p}^K))] \geq 0, \quad \boldsymbol{p}^K \in [0,1]^K. \tag{A.4}$$

We shall prove that this is indeed the case by induction on K.

- **The basis step**: When $K = 1$, we see that

$$\begin{aligned}
&\mathbf{E}[S_1(p_1)]\mathbf{E}[\varphi(S_1(p_1)+1)] - \mathbf{E}[\widehat{\varphi}(S_1(p_1))] \\
&\quad = \quad p_1(\varphi(1)(1-p_1) + \varphi(2)p_1) - p_1\varphi(1) \\
&\quad = \quad (\varphi(2) - \varphi(1))p_1^2 \geq 0
\end{aligned}$$

because the mapping $\varphi : I\!N \to \mathbb{R}$ is assumed increasing.

- **The induction step**: Next, suppose that (A.4) holds for some $K = 1, 2, \ldots$. We then observe from (A.1)–(A.2) that

$$\begin{aligned}
& \mathbf{E}[S_{K+1}(\boldsymbol{p}^{K+1})]\mathbf{E}[\varphi(S_{K+1}(\boldsymbol{p}^{K+1})+1)] - \mathbf{E}[\widehat{\varphi}(S_{K+1}(\boldsymbol{p}^{K+1}))] \\
= \quad & (p_{K+1} + \mathbf{E}[S_K(\boldsymbol{p}^K)])\Big(p_{K+1}\mathbf{E}[\varphi(S_K(\boldsymbol{p}^K)+2)] \\
+ \quad & (1-p_{K+1})\mathbf{E}[\varphi(S_K(\boldsymbol{p}^K)+1)]\Big) - \Big(p_{K+1}\mathbf{E}[\varphi(S_K(\boldsymbol{p}^K)+1)] + \mathbf{E}[\widehat{\varphi}(S_K(\boldsymbol{p}^K))]\Big) \\
= \quad & p_{K+1}\Big(p_{K+1} + \mathbf{E}[S_K(\boldsymbol{p}^K)]\Big)\Big(\mathbf{E}[\varphi(S_K(\boldsymbol{p}^K)+2)] - \mathbf{E}[\varphi(S_K(\boldsymbol{p}^K)+1)]\Big) \\
+ \quad & \mathbf{E}[S_K(\boldsymbol{p}^K)]\mathbf{E}[\varphi(S_K(\boldsymbol{p}^K)+1)] - \mathbf{E}[\widehat{\varphi}(S_K(\boldsymbol{p}^K))] \\
\geq \quad & 0
\end{aligned}$$

because $\mathbf{E}[\varphi(S_K(\boldsymbol{p}^K)+2)] \geq \mathbf{E}[\varphi(S_K(\boldsymbol{p}^K)+1)]$ by the monotonicity of φ, and upon using the induction hypothesis. This completes the proof of the induction step. ■

B Appendix

With any mapping $\varphi : I\!N \to \mathbb{R}$, we associate the "*averaged*" mapping $\varphi_{av} : I\!N \to \mathbb{R}$ introduced in (8.6). The following result, which is used in the proof of Proposition 8.1, shows how several properties of φ are inherited by φ_{av}.

Proposition B.1 *For any mapping $\varphi : I\!N \to \mathbb{R}$, the mapping $\varphi_{av} : I\!N \to \mathbb{R}$ defined by (8.6) is*

1. *integer–increasing if φ is integer–increasing;*

2. *integer–convex if φ is integer–convex.*

Proof. (Claim 1) For $n = 0, 1, \ldots$, we have

$$\varphi_{av}(n+1) - \varphi_{av}(n) = \frac{1}{n+2}\left[\varphi(n+2) - \frac{1}{n+1}\sum_{i=0}^{n}\varphi(i+1)\right] \geq 0 \qquad \text{(B.1)}$$

because φ is integer–increasing.

(Claim 2) For $n = 0, 1, \ldots$, using (B.1), we can write

$$\begin{aligned}
& \varphi_{av}(n+2) - 2\varphi_{av}(n+1) + \varphi_{av}(n) \\
= \; & \frac{1}{n+3}\left[\varphi(n+3) - \frac{1}{n+2}\sum_{i=0}^{n+1}\varphi(i+1)\right] \\
& -\frac{1}{n+2}\left[\varphi(n+2) - \frac{1}{n+1}\sum_{i=0}^{n}\varphi(i+1)\right] \\
= \; & \frac{1}{n+3}\varphi(n+3) - \frac{1}{n+2}\varphi(n+2) \\
& - \frac{1}{n+2}\left[\frac{1}{n+3}\sum_{i=0}^{n+1}\varphi(i+1) - \frac{1}{n+1}\sum_{i=0}^{n}\varphi(i+1)\right] \\
= \; & \frac{1}{n+3}\varphi(n+3) - \frac{1}{n+2}\varphi(n+2) \\
& - \frac{1}{n+2}\left[\frac{1}{n+3}\varphi(n+2) - \frac{2}{(n+1)(n+3)}\sum_{i=0}^{n}\varphi(i+1)\right] \\
= \; & \frac{1}{n+3}\left[\varphi(n+3) - \frac{n+4}{n+2}\varphi(n+2) + \frac{2}{(n+1)(n+2)}\sum_{i=0}^{n}\varphi(i+1)\right]. \qquad \text{(B.2)}
\end{aligned}$$

Setting

$$\widetilde{\varphi}_{av}(n) \equiv \varphi(n+3) - \frac{n+4}{n+2}\varphi(n+2) + \frac{2}{(n+1)(n+2)}\sum_{i=0}^{n}\varphi(i+1), \quad n = 0, 1, \ldots,$$

we see from (B.2) that the integer–convexity of φ_{av} is equivalent to

$$\widetilde{\varphi}_{av}(n) \geq 0, \quad n = 0, 1, \ldots. \qquad \text{(B.3)}$$

We shall prove this claim by induction on n.

• **The basis step**: For $n = 0$, we have $\widetilde{\varphi}_{av}(0) = \varphi(3) - 2\varphi(2) + \varphi(1) \geq 0$ because φ is integer–convex.

• **The induction step**: Suppose that (B.3) holds for some $n = 0, 1, \ldots$. Because $\varphi(n+4) \geq 2\varphi(n+3) - \varphi(n+2)$ by the integer–convexity of φ, we observe that

$$
\begin{aligned}
&\widetilde{\varphi}_{av}(n+1) \\
&= \varphi(n+4) - \frac{n+5}{n+3}\varphi(n+3) + \frac{2}{(n+2)(n+3)} \sum_{i=0}^{n+1} \varphi(i+1) \\
&\geq 2\varphi(n+3) - \varphi(n+2) - \frac{n+5}{n+3}\varphi(n+3) + \frac{2}{(n+2)(n+3)} \sum_{i=0}^{n+1} \varphi(i+1) \\
&= \frac{n+1}{n+3}\varphi(n+3) - \frac{(n+1)(n+4)}{(n+2)(n+3)}\varphi(n+2) + \frac{2}{(n+2)(n+3)} \sum_{i=0}^{n} \varphi(i+1) \\
&= \frac{n+1}{n+3}\widetilde{\varphi}_{av}(n) \geq 0
\end{aligned}
$$

upon using the induction hypothesis. This completes the proof of the induction step. ■

References

[1] H. Ahmadi and W. E. Denzel, "A survey of modern high–performance switching techniques," *IEEE Journal on Selected Areas in Communications* **JSAC–7** (1989), pp. 1091–1103.

[2] P. Billingsley, *Probability and Measure*, (Second Edition), J. Wiley & Sons, New York (NY), 1986.

[3] P. J. Burke, "Delays in single–server queues with batch input," *Operations Research* **23** (1975), pp. 830–833.

[4] M. Hluchyj and M. J. Karol, "Queueing in high–performance packet switching," *IEEE Journal on Selected Areas in Communications* **JSAC–6** (1987), pp. 1587–1597.

[5] M. J. Karol, M. Hluchyj, and S. P. Morgan, "Input versus output queueing on a space–division packet switch," *IEEE Transactions on Communications* **COM–35** (1988), pp. 1347–1356.

[6] Y.–B. Kim, *On the Performance of Crossbar Switches: Stochastic Comparison and Large Size Asymptotics*, Electrical Engineering Department, University of Maryland, College Park (MD). Expected May 1996.

[7] Y.–B. Kim and A.M. Makowski, "Simple optimization problems via majorization ordering," *IEEE Transactions on Automatic Control*, submitted 1996.

[8] Y.–B. Kim and A.M. Makowski, "Load balancing for non–blocking switches with output queueing," in preparation.

[9] M. J. Lee and S.–Q. Li, "A study of traffic imbalances in a fast packet switch," *Proceedings of the 1989 Infocom Conference* (1989), pp. 538–547.

[10] M. J. Lee and S.–Q. Li, "Performance of a nonblocking space–division packet switch in a time variant nonuniform traffic environment," *IEEE Transactions on Communications* **COM–39** (1991), pp. 1515–1524.

[11] S.–Q. Li, "Nonuniform traffic analysis on a nonblocking space–division packet switch," *IEEE Transactions on Communications* **COM–38** (1990), pp. 1085–1096.

[12] A.M. Makowski, B. Melamed and W. Whitt, "On averages seen by arrivals in discrete time," Proceedings of the 28th IEEE Conference on Decision and Control, Tampa (FL), December 1989, pp. 1084–1086.

[13] A.W. Marshall and I. Olkin, *Inequalities: Theory of Majorization and Its Applications*, Academic Press, New York (NY), 1979.

[14] G. F. Pfister and V. A. Norton, "'Hot spot' contention and combining in multistage interconnection networks," *IEEE Transactions on Computers* **C–34** (1985), pp. 943–948.

[15] S. Ross, *Stochastic Processes*, J. Wiley & Sons, New York (NY), 1984.

[16] M. Shaked and J.G. Shanthikumar, "Stochastic convexity and its applications," *Advances in Applied Probability* **20** (1988), pp. 427–446.

[17] M. Shaked and J.G. Shanthikumar, "Parametric stochastic convexity and concavity of stochastic processes," *Annals of the Institute of Statistical Mathematics* **42** (1990), pp. 509–531.

[18] J. S.–C. Chen and T.E. Stern, "Throughput analysis, optimal buffer allocation, and traffic imbalance study of a generic nonblocking packet switch," *IEEE Journal on Selected Areas in Communications* **JSAC–9** (1991), pp. 439–449.

[19] D. Stoyan, *Comparison Methods for Queues and Other Stochastic Models*, English Translation (D.J. Daley, Editor), J. Wiley & Sons, New York (NY), 1984.

[20] F. A. Tobagi, "Fast packet switch architectures for broadband integrated services digital networks," *Proceedings of IEEE* **8** (1990), pp. 133–167.

[21] L. T. Wu, "Mixing traffic in a buffered banyan network," *Proceedings of ACM 9th Data Communication Symposium* (1985), pp. 134–139.

[22] H. Yoon, M.T. Liu, K.Y. Lee and Y.M. Kim, "The Knockout switch under nonuniform traffic," *IEEE Transactions on Communications* **COM–43** (1995), pp. 2149–2156.