A PATHWISE COMPARISON RESULT FOR PARALLEL QUEUES

PASCAL MOYAL

Abstract. We introduce the appropriate framework for pathwise comparison of multiple server queues under general stationary ergodic assumptions. In particular, we show in what sense it is better to have more servers for a system under FCFS (First Come, First Served) or equivalently, more queues in a system of parallel queues under the JSW (Join the Shortest Workload) allocation policy. These comparison results are based on the recursive representation of Kiefer and Wolfowitz, and on a non-mass conservative generalization of the Schur-Convex semi-ordering.

1. Introduction

Consider a queueing system of $S$ parallel queues: there are $S$ servers, and to each one of them is associated a particular line. Moreover the $S$ lines are independent of one another, so any incoming customer chooses a server upon arrival and stay in the corresponding queue until service.

The main allocation policy considered in this paper is Join the Shortest Workload (JSW): the workload, i.e. the quantity of service to be completed by each one of the $S$ servers, is known at any arrival time, and any arriving customer joins the queue having the shortest workload.

Another approach of queues with several servers assumes that the incoming customers are all put in the same queue during their wait. In the sequel, such models will be referred to as multiple server queues. It is intuitively clear that multiple server queues are more flexible than parallel queues, in that they offer to the customers the possibility of changing their service destination, to minimize their waiting time. In fact, parallel queues are nothing but a particular case of multiple server queues, having a service discipline which does not allow changes of queues. More precisely, It is well known that a parallel queue under the JSW policy is equivalent to a multiple server queue providing service in First Come, First Served (FCFS).

For multiple server queues GI/GI, Foss [5, 6, 7] show in various senses, the optimality of FCFS (First Come, First Served) among a very wide class of service disciplines (including those leading to parallel queues). This legitimates in se the use of the JSW policy, and this is why most of the literature on this topic focus on this allocation policy.

Therefore, the results in this paper concerning parallel queues under JSW can be interpreted as concerning multiple server queues under the optimal discipline FCFS. We adopt the first representation for technical purposes: under the most general assumption (stationary ergodic arriving process and service sequence), it leads to a simple representation by a stochastic recursive sequence, keeping track.

2000 Mathematics Subject Classification. Primary: 60F17, Secondary: 60K25 and 60B12.

Key words and phrases. Queueing systems, Parallel queues, Stochastic comparison.
upon arrival times, of all residual service times, in increasing order (we call it the service profile of the system - see Kiefer and Wolfowitz equation (2)).

Following this approach, Neveu [10] shows a stability condition for parallel queues under the JSW policy. This stability result is inherited from the monotonicity of the service profile sequence in some sense, and a minimal stationary profile is given by Loynes’s Theorem [8]. Foss [5] and Brandt [3] then introduce the concept of maximal stationary profile for JSW queues.

In this paper, we show that this representation provides the appropriate framework for a pathwise comparison of parallel queues. We show in Section 4 in what sense, a JSW system of $S$ queues performs better at equilibrium than one of $N$ queues for $N \leq S$. Equivalently, we show en route that for a system of fixed size $S$, JSW performs better than a policy consisting in sending the arriving customers to another (fixed) queue in the workload hierarchy. The key technical point relies on the monotonicity of the recursive representation (2), and on the introduction of a particular partial semi-ordering, generalizing the Schur-convex ordering to the non-mass conservative case. All the definitions, basic results and comparison lemmas are left to the Appendix.

2. Notation

In what follows, $\mathbb{R}$ denotes the real line and $\mathbb{R}^+$, the subset of non-negative numbers. Denote $\mathbb{N}$ (respectively $\mathbb{N}^+$, $\mathbb{Z}$), the subset of non-negative (resp. positive, relative) integers. For any two elements $p$ and $q$ in $\mathbb{N}$, $[p,q]$ denotes the finite family $\{p,p+1,...,q\}$. Let $S \in \mathbb{N}^*$. We denote for all $x,y \in \mathbb{R}^S$ and $\lambda \in \mathbb{R}$,

$$x = (x(1), x(2), ..., x(S)) ;$$
$$\lambda x = (\lambda x(1), ..., \lambda x(S)) ;$$
$$x + y = (x(1) + y(1), ..., x(S) + y(S)) ;$$
$$0 = (0, ..., 0) ;$$
$$e_i = (0, ..., 1_{i}, ..., 0) \text{ for all } i \in [1, S] ;$$
$$x^+ = (x(1)^+, x(2)^+, ..., x(S)^+) ;$$
$$\bar{x}, \text{ the fully ordered version of } x, \text{ i.e. } \bar{x}(1) \leq ... \leq \bar{x}(S) .$$

Denote then $(\mathbb{R}^+)^S$ the subset of fully ordered vectors, furnished with the euclidian norm.

3. A system of $S$ parallel queues

We consider a queueing system having $S$ servers ($S \in \mathbb{N}^*$) working in parallel, each of them providing service in First in, first out (FIFO). We assume that a global information is available to all entering customers, on the quantity of work to be completed by each of the $S$ servers (termed workload of the corresponding server). Unless explicitly mentioned, we assume that the discipline is JSW for Join the Shortest Workload, i.e. any customer choses to join the queue having the smallest workload. This allocation policy is equivalent to a multi-server queue in First in, First out (FIFO), as noticed in the introduction.
Palm representation of the system. The input of the system is represented by a marked stationary point process, for which \((\Omega, \mathcal{F}, \mathbf{P}, \theta)\) denotes the Palm probability space. We thereby assume that \(\mathbf{P}\)-a.s. a customer enters the system at time 0 (the latter is denoted by \(C_0\)), requesting a service time of duration \(\sigma\) (in an arbitrary time unit), and the following customer \(C_1\) enters the system at time \(\xi\). As a consequence of these basic assumptions, customer \(C_n\) (where \(n \in \mathbb{Z}\)) requests a service duration \(\sigma \circ \theta^n\), and the inter-arrival epoch between the arrivals of \(C_n\) and \(C_{n+1}\) equals \(\xi \circ \theta^n\). We finally assume that \(\theta\) is \(\mathbf{P}\)-stationary and ergodic, that \(\sigma\) and \(\xi\) are integrable and that \(\xi > 0\), \(\mathbf{P}\)-a.s.. The reader is referred to the classical monograph [2, ?] for a complete overview of the ergodic-theoretical representation of queueing systems.

If \(E\) is a partially ordered Polish space having a minimal point 0, \(X\) is a \(E\)-valued random variable (r.v. for short) and \(\Psi\) is a \(E\)-valued random mapping, the stochastic recursive sequence (SRS) descending from \(X\) and driven by \(\Psi\) is defined by
\[
\begin{cases}
X_n &= X; \\
X_{n+1} &= \Psi \circ \theta^n(X_n), \; n \in \mathbb{N}.
\end{cases}
\]

It is then routine to check that a time-stationary sequence having, on a reference probability space, the same distribution as \(\{X_n\}\) corresponds to a solution \(X\) defined on \(\Omega\) to the functional equation
\[
X \circ \theta = \Psi(X), \; \text{a.s.}
\]

As stated by Loynes’s Theorem ([8], see as well section 2.5.2 of [2]), in the case where \(\Psi\) is a.s. \(\prec\)-nondecreasing and continuous, a solution to (1) is given by the almost sure limit of the so-called \textit{Loynes's sequence} \(\{M_n\}\), defined by
\[
\begin{cases}
M_0 &= 0; \\
M_{n+1} &= \Psi \circ \theta^{-1}(M_n \circ \theta^{-1}), \; n \in \mathbb{N}.
\end{cases}
\]

The service profile. We follow the representation of Kiefer and Wolfowitz: all systems of \(S\) parallel queues will be represented upon the arrival of \(C_n\), by the \((\mathbb{R}^+)^S\)-valued r.v. \(V_n\), representing the residual workloads of all servers (i.e. for each server, the sum of the service times of all the customers in line, and of the residual service time of the customer in service, if any, counted in time unit). The workloads of the servers are sorted increasingly, i.e. \(V_n(1) \leq V_n(2) \leq ... \leq V_n(S)\). \(V_n\) will be called \textit{service profile} of the system.

Kiefer and Wolfowitz celebrated equation reads as follows: if the initial service profile \(V_0\) upon the arrival of \(C_0\) is fixed, then for all \(n\),
\[
V_{n+1} = G \circ \theta^n(V_n),
\]
where the random map \(G\) is defined as follows:
\[
G : \begin{cases}
(\mathbb{R}^+)^S &\rightarrow (\mathbb{R}^+)^S; \\
(u) &\mapsto [u + \sigma \mathbf{e}_1 - \xi, 1]
\end{cases}
\]
So from (1), a stationary service profile is a solution to the equation
\[
V \circ \theta = G(V) \; \text{a.s.}
\]
Clearly, \(G\) is a.s. \(\prec\)-non-decreasing and continuous, and this is how Loynes’s scheme is used, to construct the \(\prec\)-minimal solution \(V\) of (3). Moreover, all \(S\) coordinates of \(V\) are a.s. finite whenever \(\mathbf{E}[\sigma] < S \mathbf{E}[\xi]\) (see [10]).
In what follows, we will also be led to consider a stochastic recursion driven by the following mappings, for $P \in [1, S]$,

\[
\tilde{G}^P : \begin{cases} 
(R^+)^S \to (R^+)^S; \\
\{u \mapsto [u + \sigma e_P - \xi]^+}
\end{cases}
\]

which corresponds, as easily seen, to sending any incoming customer to the queue having the $P$-th least workload upon arrival (see below).

4. Comparison results

We now introduce our main results.

**Comparing two JSW systems.** We show hereafter, in which sense increasing the number of servers improves the performances of the system. Let $1 \leq N \leq S$ two integers, and consider two JSW systems having respectively $S$ and $N$ servers. We add to all the parameters involved, exponents $^S$ and $^N$ to emphasize the dependence on the number of servers.

**Theorem 1.** If $E[\sigma] < NE[\xi]$, then the minimal solutions of (3), $V^S$ and $V^N$ respectively for $S$ and $N$ servers, satisfy $P$-a.s.

\[
\forall \ell \in [1, N], \quad V^S(S - N + \ell) \leq V^N(\ell);
\]
\[
\sum_{i=1}^{S} V^S(i) \leq \sum_{i=1}^{N} V^N(i).
\]

**Proof.** Define by $\{M^S_n\}$ (resp. $\{M^N_n\}$), Loynes’s sequence for the service profile of the system of $S$ (resp. $N$) parallel queues, representing the service profile seen by customer $C_0$ upon arrival, when assuming that customer $C_n$ entered an empty system (and accordingly for $\{M^P_n\}$). Let us consider for all $n \in \mathbb{N}$, the $(\mathbb{R}^+)^S$-valued

---

**Figure 1.** The service profile
r.v. $\tilde{M}_n^S$ defined by

$$\tilde{M}_n^S = \left(0, \ldots, 0, M_n^S(1), M_n^S(2), \ldots, M_n^S(N) \right).$$

It is easily checked that for all $n$,

$$\tilde{M}_{n+1}^S = \tilde{G}^{S-N+1} \circ \theta^{-1} \left(\tilde{M}_n^S \circ \theta^{-1}\right),$$

where $\tilde{G}^{S-N+1}$ is defined by (4) for $P := S - N + 1$. Therefore, as $\{M_n^S\}$ and $\{\tilde{M}_n^S\}$ descend from the same initial value 0, applying item (i) of Lemma 5 yields by a simple induction that a.s. for all $n \in \mathbb{N}$, $M_n^S \prec \tilde{M}_n^S$, or in other words

$$\forall \ell \in [P, S], \ M_n^S(\ell) \leq \tilde{M}_n^S(\ell);$$

$M_n^S \prec \tilde{M}_n^S$,

where the partial orderings "$\prec\prec$" and "$\prec\prec\prec$" are introduced in Definition 2. This implies in turn that

$$\forall \ell \in [1, N], \ M_n^S(S - N + \ell) \leq M_n^S(\ell);$$

$$\sum_{i=1}^S M_n^S(i) \leq \sum_{i=1}^N M_n^S(i)$$

and (5) and (6) follow by taking the limit. $\square$

The latter result makes precise, in what sense the largest system performs better in steady state: it minimizes the total workload from (6) and as a straightforward consequence of (5), the offered waiting time:

$$V^S(1) \leq V^N(1), \ P - \ a.s..$$

**Remark 1.** Let us emphasize once again on the fact that these results can be adapted to multiple server queues operating in First Come, First Served: at equilibrium, (6) means that the server having the largest virtual workload among $S$ servers is almost surely less loaded than that of a system of $N$ servers, and all the same for the second largest, the third largest, etc... (5) means that the total workload of the system of $S$ servers is a.s. less than that of the system of $N$ servers, and (11) states that the system of $S$ servers offers almost surely a smaller waiting time than that of size $N$.

**Remark 2.** All these pathwise comparison results can be translated from the Palm representation that is adopted here, onto the primary time-stationary queue, using a classical representation argument à la Strassen. More precisely, it can be deduced from Theorem 7 that on a reference probability space $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{P})$ on which the primary stationary sequences of inter-arrival and service times are defined, the time-stationary distributions of the service profiles $\tilde{V}^S$ and $\tilde{V}^N$ satisfy

$$\tilde{V}^S(S - N + \ell) \leq_{st} \tilde{V}^N(\ell) \text{ for all } \ell \in [1, N];$$

$$\sum_{i=1}^S \tilde{V}^S(i) \leq_{st} \sum_{i=1}^N \tilde{V}^N(i),$$

where "$\leq_{st}$" denotes the strong (increasing) stochastic ordering associated to $\tilde{P}$. 
Comparing allocations. We can reason as above, to compare allocation policies for two systems of same size. So consider a system of $S$ parallel queues under two alternative allocation policies, JSW as above, and the following one: fix $P \leq S$, and suppose that any incoming customer is sent upon arrival, to the queue having the $P$-th smallest workload among the $S$ queues (regardless of whether there is an empty queue or not). Clearly, the corresponding service profile sequence $\{\tilde{V}_n; n \in \mathbb{N}\}$ obeys the following recurrence

$$\tilde{V}_{n+1} = \tilde{G}^p(V_n), \ n \in \mathbb{N}.$$ 

Therefore, as for the previous result we can easily show the following transient comparison of service profile sequences.

**Theorem 2.** Whenever $V_0 \prec_p \tilde{V}_0$, we have $V_n \prec_p \tilde{V}_n$ for all $n \in \mathbb{N}$, i.e.

$$V_n \prec_p \tilde{V}_n; \quad \forall \ell \geq P, V_n(\ell) \leq \tilde{V}_n(\ell),$$

where $\{V_n; n \in \mathbb{N}\}$ and $\{\tilde{V}_n; n \in \mathbb{N}\}$ denote the service profile sequences for JSW and the allocation to the $P$-th least workload, respectively.

Clearly, the latter transient comparison result can be extended to the equilibrium state, whenever the latter exists. From assertion (ii) of Lemma 5, the random map $\tilde{G}^p$ is a.s. non-decreasing and continuous on $(\mathbb{R}^+)^S$, and it is easily checked that under condition $E[\sigma] < (S - P + 1)E[\xi]$, there exists a minimal finite solution $\tilde{V}$, given by

$$\tilde{V} = \left(0, \ldots, 0, V^0\right)_{p-1 \text{ terms}},$$

where $V^0$ is the stationary service profile for a JSW system for $N = S - P + 1$ servers, introduced in the previous section. Then, Theorem 1 amounts to saying that $V \prec_p \tilde{V}$ a.s., which can be retrieved to the limit in Theorem 2.

**Appendix A. Partial (semi-) orderings on $(\mathbb{R}^+)^S$**

Let $S \geq 1$. The Schur-convex ordering "$\prec_c$" on $\mathbb{R}^S$ is defined as follows (see [1], [9] and the Chapter 4 of [2] for an exhaustive presentation).

**Definition 1.** For all $u, v \in \mathbb{R}^S$, we write $u \prec_c v$ whenever

$$\begin{cases} \sum_{i=1}^S u(i) = \sum_{i=1}^S v(i), \\
\sum_{i=1}^k u(i) \leq \sum_{i=1}^k v(i), \quad \text{for all } k \in [2, S].
\end{cases}$$

We also equip $(\mathbb{R}^+)^S$ with the following partial (semi-) orderings:

**Definition 2.** Let $u$ and $v \in (\mathbb{R}^+)^S$.

(i) We denote $u \prec v$ if

$$u(i) \leq v(i) \text{ for all } i \in [1, S].$$
(ii) We denote \( u \prec v \) if
\[
\sum_{i=k}^{S} u(i) \leq \sum_{i=k}^{S} v(i), \quad \text{for all } k \in [1, S].
\]

(ii) Let \( \mathcal{P} \in [1, S] \). We write \( u \prec_p v \) if
\[
\begin{align*}
&\{ u \prec v; \\
&u(\ell) \leq v(\ell) \text{ for all } \ell \in [\mathcal{P}, S].
\end{align*}
\]

The ordering "\( \prec \)" is nothing but the natural coordinatewise partial ordering on \((\mathbb{R}^+_0)^S\), whereas "\( \prec_* \)" and "\( \prec_p \)" are two partial semi-orderings, which are non-mass conservative declinations of the Schur-convex ordering "\( \prec_c \)".

We present hereafter several basic comparison results related to these orderings, useful for proving our claim. The proofs of Lemmas 1 and 2 below are straightforward.

**Lemma 1.** For all \( x, y \in \mathbb{R}^S \),
\[
(8) \quad u \prec v \iff -u \prec -v.
\]

**Lemma 2.** For all \( u, v \in (\mathbb{R}^+_0)^S \) and all real numbers \( x \leq y \),
\[
(9) \quad u \prec v \iff u + x.e_1 \prec v + y.e_1.
\]

For all \( u \in \mathbb{R}^S \) and any permutation \( \gamma \) of \([1, S]\), define
\[
u_{\gamma} = (u(\gamma(1)), ..., v(\gamma(S))).
\]

A mapping \( F : \mathbb{R}^S \to \mathcal{E} \) is then termed symmetric whenever \( F(u) = F(u_{\gamma}) \) for all \( u \in \mathbb{R}^S \) and all permutations \( \gamma \). We then have (see e.g. [2], Prop. 4.1.1 p.262 and Lemma 4.1.1 p.266) that

**Lemma 3.**
(i) For all \( u, v \in \mathbb{R}^S \) such that \( u \prec_c v \),
\[
(10) \quad u - v \prec_c u - v.
\]

(ii) For all \( u, v \in \mathbb{R}^S \),
\[
(11) \quad \bar{u} - \bar{v} \prec_c \bar{u} - \bar{v}.
\]

**Lemma 4.** Let \( u, v \in (\mathbb{R}^+_0)^S \) such that \( u \prec_* v \). Then,
(i) For all \( x \in \mathbb{R} \),
\[
[u - x.1]^+ \prec_* [v - x.1]^+.
\]

(ii) For all \( j \in [1, S] \) such that \( u(j) \leq v(j) \), for all \( y \in \mathbb{R}^+ \),
\[
u + y.e_j \prec_* v + y.e_j.
\]

**Proof.**
(i) The result is trivial if \( u(N) \leq x \). Else, for all \( k \in [1, S] \), for some \( \ell \geq k \),
\[
\sum_{i=k}^{S} (u(i) - x)^+ = \sum_{i=\ell}^{S} u(i) - x \leq \sum_{i=\ell}^{S} v(i) - x \leq \sum_{i=k}^{S} v(i) - x^+.
\]
(ii) For all $k > j$,
\[
\sum_{i=k}^{S} (u + y \cdot e_j)(i) = \left( \sum_{i=k}^{S} u(i) \right) \vee \left( \sum_{i=k+1}^{S} u(i) \right) \\
\leq \left( \sum_{i=k}^{S} v(i) \right) \vee \left( \sum_{i=k+1}^{S} v(i) \right) \\
= \sum_{i=k}^{S} (\overline{v + y \cdot e_j})(i),
\]
whereas for all $k \leq j$,
\[
\sum_{i=k}^{S} (u + y \cdot e_j)(i) = \sum_{i=k}^{S} u(i) + y \leq \sum_{i=k}^{S} v(i) + y = \sum_{i=k}^{S} (\overline{v + y \cdot e_j})(i).
\]

\[\square\]

**Lemma 5.** Let $P \in [1, S]$ and $u, v \in (\mathbb{R}^+)^S$ be such that $u \prec_p v$. Then, a.s.,

(i) $G(u) \prec_p \tilde{G}^p(v)$;

(ii) $\tilde{G}^p(u) \prec_p G^p(v)$.

**Proof.** (i) Set $u(S + 1) = \infty$ and $v(S + 1) = \infty$. We have for all $\ell \in [P, S]$, a.s.
\[
G(u)(\ell) = \left[ u(\ell) \lor ((u(1) + \sigma) \land u(\ell + 1)) - \xi \right]^+ \\
\leq \left[ v(\ell) \lor ((v(P) + \sigma) \land v(\ell + 1)) - \xi \right]^+ \\
= \tilde{G}^p(v)(\ell).
\]
This implies in particular that for all $k \geq P$,
\[
\sum_{i=k}^{S} G(u)(i) \leq \sum_{i=k}^{S} \tilde{G}^p(v)(i) \text{ a.s.}
\]
Consequently, it suffices to check that
\[
\sum_{i=k}^{S} G(u)(i) \leq \sum_{i=k}^{S} \tilde{G}^p(v)(i) \text{ for all } k < P.
\]

We have for all such $k$,
\[
\sum_{i=k}^{S} G(u)(i) = \sum_{i=k+1}^{S} [u(i) - \xi]^+ + [(u(1) + \sigma) \lor u(k) - \xi]^+;
\]
\[
\sum_{i=k}^{S} \tilde{G}^p(v)(i) = \sum_{i=k; i \neq P}^{S} [v(i) - \xi]^+ + [v(P) + \sigma - \xi]^+.
\]
If $u(k) \geq u(1) + \sigma$, then (12) equals
\[
\sum_{i=k}^{S} [u(i) - \xi]^+ \leq \sum_{i=k}^{S} [v(i) - \xi]^+ \leq \sum_{i=k}^{S} \tilde{G}^p(v)(i) \text{ a.s.,}
\]
where we used item (i) of Lemma 4 in the first inequality. Consequently only the case where \( u(k) < u(1) + \sigma \) remains to be treated. Then, (12) equals

\[
[u(1) + \sigma - \xi]^{+} + \sum_{i=k+1}^{S} [u(i) - \xi]^{+}.
\]

The vector \((u(1), u(k+1), ..., u(S))\) is fully ordered in \((\mathbb{R}^{+})^{S-k+1}\) and a.s.,

\[
\begin{pmatrix}
\xi - \sigma, & \xi, & \cdot \cdot \cdot, & \xi \\
1, & 2, & \cdot \cdot \cdot, & S-k+1
\end{pmatrix}
= \begin{pmatrix}
\xi, & \xi - \sigma, & \cdot \cdot \cdot, & \xi \\
1, & P-k+1, & \cdot \cdot \cdot, & S-k+1
\end{pmatrix}.
\]

So in view of (8) and (10),

\[
(u(1), u(k+1), ..., u(P), ..., u(S)) - (\xi - \sigma, \xi, \cdot \cdot \cdot, \xi)
\prec_{c} (u(1), u(k+1), ..., u(P), ..., u(S)) - \left( \xi, \cdot \cdot \cdot, \xi, \xi - \sigma, \xi, \cdot \cdot \cdot, \xi \right),
\]

where \( \prec_{c} \) is the Schur-convex semi-ordering on \( \mathbb{R}^{S-k+1} \). Therefore, from the symmetry and convexity of the map

\[
\begin{pmatrix}
(\mathbb{R}^{+})^{S-k+1} \rightarrow \mathbb{R} \\
u \mapsto \sum_{i=1}^{S-k+1} u^{+},
\end{pmatrix}
\]
in view of (9) the sum (12) satisfies a.s.

\[
[u(1) + \sigma - \xi]^{+} + \sum_{i=k+1}^{S} [u(i) - \xi]^{+}
\leq [u(1) - \xi]^{+} + \sum_{i=k+1; i \neq P}^{S} [u(i) - \xi]^{+} + [u(P) + \sigma - \xi]^{+}
\leq \sum_{i=1; i \neq P}^{S} [u(i) - \xi]^{+} + [u(P) + \sigma - \xi]^{+}.
\]

On the other hand, as \( u(P) \leq v(P) \), from the assertions (ii) and then (i) of Lemma 4 entail a.s.

\[
[u + \sigma \cdot e_{P} - \xi \cdot 1]^{+} \prec \Sigma_{i \neq P} [v + \sigma \cdot e_{P} - \xi \cdot 1]^{+}.
\]

In particular,

\[
\sum_{i=k; i \neq P}^{S} [u(i) - \xi]^{+} + [u(P) + \sigma - \xi]^{+} \leq \sum_{i=k; i \neq P}^{S} [v(i) - \xi]^{+} + [v(P) + \sigma - \xi]^{+},
\]

and we deduce (11) from (12), (13), (15) and (16).

(ii) For all \( \ell \in [P, S] \), a.s.

\[
\hat{G}^{\ell}(u)(\ell) = [u(\ell) \lor ((u(P) + \sigma) \land u(\ell + 1))] - \xi^{+}
\leq [v(\ell) \lor ((v(P) + \sigma) \land v(\ell + 1))] - \xi^{+}
= \hat{G}^{\ell}(v)(\ell).
\]
On another hand, as $u(P) \leq v(P)$, Lemma 4 entail a.s.

\[ \tilde{G}^p(u) \prec \widetilde{G}^p(v). \]

\[ \Box \]

Acknowledgment. The author would like to thank Sergey Foss for a very insightful discussion on this problem.

REFERENCES

[1] B.C. Arnold Majorization and the Lorenz Order: A Brief Introduction, Lecture Notes in Statistics, 43, Springer-Verlag, New York (1987).
[2] F. Baccelli and P. Brémaud. Elements of Queueing Theory (2nd ed.). Springer, 2002.
[3] A. Brandt. On stationary waiting times and limiting behavior of queues with many servers I: the general G/G/m/∞ case. Elektron. Inform. u. Kybernet. 21, 47–64, 1985.
[4] S. Foss. On the approximation of multichannel service systems. Siberian Math. Zh., 21(6), 132–140, 1980.
[5] S. Foss. Comparison of service disciplines in multichannel service systems. Siberian Math. Zh., 22(1), 190–197, 1981.
[6] S. Foss. Comparison of Service Disciplines in G/GI/m Queues. INRIA Research Report, 1097, 1989.
[7] S. Foss and N. Chernova. On optimality of the FCFS discipline in multiserver queueing systems and networks. UDC 519.2
[8] R.M. Loynes. The stability of queues with non-independent interarrivals and service times. Proceedings of the Cambridge Philosophical Society, 58, 497–520, 1962.
[9] A.W. Marshall and I. Olkin Inequalities: Theory of Majorization and Its Applications, Academic Press, New York (1979).
[10] J. Neveu. Construction de files d’attente stationnaires (in French). Lecture Notes in Control and Information Sciences, 60, Springer-Verlag, 31–41, 1983.