POLLING SYSTEMS WITH SYNCHRONIZATION CONSTRAINTS *

Khamisy Asad
Electrical Engineering
Technion — Israel Institute of Technology
Haifa, 32000 Israel

Eitan Altman
INRIA
Centre Sophia Antipolis
06565 Valbonne Cedex, France

Moshe Sidi
Electrical Engineering
Technion — Israel Institute of Technology
Haifa, 32000 Israel

November 28, 1995

Abstract

We introduce in this paper a new service discipline for polling systems that arises when there are precedence (or synchronization) constraints between the order that jobs in different queues should be served. We describe these constraints by a set of trees of depth of one or two; i.e. there are N stations which are "fathers" of (zero or more) synchronized stations ("children"). We assume that jobs that arrive to synchronized stations may need data obtained from jobs that arrived prior to them to their corresponding "father" station, in order to be processed.

Our work was motivated by the "globally gated" discipline introduced recently by Boxma, Levy and Yechiali [7] which exhibits a special case of synchronization.

We analyze the performance of the new service discipline and obtain expressions for the first two moments and LST of the waiting times in different stations, and expressions for the moments and LST of other quantities of interest such as cycle durations and generalized station times.

We obtain a "pseudo" conservation law for the synchronized gated discipline, and we find the optimal topology of the network that minimizes the weighted sum of the mean waiting times, obtained in the "pseudo" conservation law.

*This work was supported by grant number XXX of the XXX

Finally numerical examples are given for illustrating the dependence of the performance of the synchronized gated discipline to different parameters of the network.

1 INTRODUCTION

The term *Polling System* often refers in the literature to a system where a single server serves jobs from different queues (channels) in a pre-determined order such as cyclic order or polling table with generally nonzero switchover times (see e.g. Takagi [16]). The service disciplines most often studied are the exhaustive, gated and limited service regimes:

- Gated discipline: The server serves only those jobs which are waiting at the station when it is polled. The jobs that arrive during the service time are set aside to be served after the next polling instant to that station.
- Exhaustive discipline: The server serves all jobs in each queue until it is empty, before moving to the next queue.
- Limited service: At each queue, service is completed if the number of jobs that have been served equals to some threshold or, if the queue empties before the threshold was obtained.

We propose a new service discipline that arises when there are precedence (or synchronization) constraints between the order that jobs in different queues should be served. We describe these constraints by a set of trees of depth of one or two; i.e. there are N stations which are "fathers" of (zero or more) synchronized stations ("children"). We assume that jobs that arrive to synchronized stations may need data obtained from jobs that arrived prior to them to their corresponding "father" station, in order to be processed.

We suggest the following "synchronized" gated service discipline in order to meet these constraints: The server moves cyclically between the stations. When it arrives to a "father" station it serves only those jobs which are waiting at the station when it is polled, as in the standard gated discipline. In addition to that, it marks upon arrival to that "father" station, the jobs present in all its' children stations. When it arrives to a synchronized station (a "child") it serves only the jobs that were marked when the server last arrived to its father station. In both cases, jobs that are not served remain in the queues. Similarly to the standard gated or exhaustive schemes, the server keeps switching from one queue to the next even when there are no jobs present in the system.

Different models for Precedence Constraints have been developed and studied in the past years, see e.g. [1] [2] [3] [4] [13]. Our paper is the first to analyze several precedence constraints that may appear in polling systems. Both the standard gated discipline and the "globally gated" discipline introduced recently by Boxma, Levy and Yechiali [7] are special cases of our model. In the latter "globally gated" discipline, there is one selected station, such that at each polling instant to that station gates are closed in all the stations. In all the stations the jobs that are served are only those that were present at the polling instant of the selected station.

We show that the expected cycle duration in the synchronized gated discipline is equal to the one obtained in the standard gated service discipline.

Our main goal is to obtain expressions for the expected waiting times in different stations. An efficient method to do so, that proved to be useful for the standard gated discipline (see e.g. [8],[9]), is to first characterize the statistics of the "station times", (which are the times between the polling instants of two consecutive stations). This method relies on the fact that the vectors of the station times in the last cycle form a Markov chain. In our case we define a "station time" to be the time between the polling instants of two consecutive "father" stations. The fact that we may restrict to these quantities (instead of defining station times for all stations) follows from the fact that "gates" are closed only when visiting a "father" station. This allows us to obtain a reduced number of equations for the Laplace transforms and for the second moments of the station times, and to obtain explicit expressions for the first moments. Unlike the standard gated discipline, the vectors of the station times in the last cycle do not form a Markov chain. We overcome this problem by enlarging the state vector to include station times of the two last cycles.

We show that for any station, the expectation of the number of jobs present in it's queue when the station starts to be served is greater (for "child" station) or equal (for "father" station) to the expectation obtained for the standard gated service discipline.

We obtain a "pseudo" conservation law for the synchronized gated discipline, and we find the optimal topology of the network that minimizes the weighted sum of the mean waiting times, obtained in the "pseudo" conservation law.

We faced a non-trivial problem in the notation, when trying to enumerate the different stations in a way that grasps both the (cyclic) order at which they are served, as well as the order according to the precedence constraints. We present an enumeration that partially meets these goals and is yet simple and convenient for the subsequent analysis.

The structure of the paper is as follows. In section 2, we describe the model along with the assumptions and several definitions and notations used throughout the paper. In section 3, we describe the evolution equations of the system and we derive a recursion equality for the joint generating functions of the station times. In section 4, we derive the first moments of the station times in steady-state, and we obtain a set of $N(2N-1)$ equations from which the second moments of the station times in steady-state can be derived. These moments are then used in section 5 to obtain the average waiting times in the different stations of the network. In section 5, we provide a delay analysis of a polling system with synchronization constraints. In this section, we derive the Laplace transform and the expectation of the waiting times for jobs in "father" and in "child" stations, separately.In section 6, we derive a "pseudo" conservation law for a polling system with synchronization constraints and compare it with the standard gated discipline. In this section, we also derive an optimal topology with respect to the conservation law. In section 7, we give 3 numerical examples of polling systems with synchronization constraints for illustrating the dependence of the performance of the synchronized gated discipline to the different parameters of the network. In section 8, we summarize and discuss possible extensions and applications of our results.

2 The Model and Notation

The polling system has M stations, of which N are "fathers" and the rest are "children". We label some arbitrary "father" type station by 1. We then enumerate the rest of the fathers stations according to the (cyclic) order of service between them; they are thus given the numbers 2,...,N. Let $m(i)$ denote the number of "child" stations that are situated between (the father) station i and the subsequent father station. These stations are indexed by the set $\{i_j, j = 1, ..., m(i)\}$; the subscript j is given according to the order that they are served. Sometimes we shall use for convenience i_0 to denote the father station i. For each station $i_j, j \neq 0$ we denote by $s(i_j)$ the father of that station, and we set $s(i_0) = i$.

The queue capacity at each station is infinite. The arrival process at each station is independent and Poisson with arrival rate λ_{i_j}. The service times are all independent; their generating function in the i_jth station is denoted by $B^*_{i_j}(s)$, and first and second moments are denoted by $\bar{b}_{i_j}$ and $b^{(2)}_{i_j}$ respectively. The time it takes between the end of service in station i_j till the polling instant in the next station is called the i_jth *walking time*. Let d_{i_j} be a RV that stands for this time. We assume that the walking times are independent and their Laplace transform, first and second moments are denoted by $D^*_{i_j}(s)$, $\bar{d}_{i_j}$ and $d^{(2)}_{i_j}$ respectively; moreover, these times, the arrival times and the service durations are mutually independent.

Let $\rho_{i_j} := \lambda_{i_j} \, \bar{b}_{i_j}$ and denote, $\rho := \sum_{i=1}^{N} \sum_{j=0}^{m(i)} \rho_{i_j}$.

We make throughout the paper the standard stability assumption, i.e. $\rho < 1$.

In order to obtain the evolution equations for the system we look at the random times $\tau(k)$, which represent the kth polling instant to a father station. Thus a father station that is polled at $\tau(k)$ is also polled at $\tau(k+N)$, $\tau(k+2N)$ and so on. If station i is polled at time $\tau(k)$, we define by $\tau(k_j)$ the first time that station i_j is polled after time $\tau(k)$. With some abuse of notation, we shall refer to station k (and k_j) as the station that is visited at time $\tau(k)$ ($\tau(k_j)$ respectively). If $s(k_j) = i$ then $\hat{s}(k_j)$ will denote the largest time index $k' \le k$ such that station i is visited at time $\tau(k')$. $\tau(\hat{s}(k_j))$ is thus the largest polling time of station $s(k_j)$ before $\tau(k_j)$. With some abuse of notation, we shall use $m(k)$ to denote the number of "child" stations that are situated between (the father) station visited at $\tau(k)$ and the subsequent father station.

Generalizing the definitions for the standard gated discipline in [8], we define for any "father" type station k the *station time* θ_k as the (random) time it takes to serve that station as well as all the subsequent "child" stations k_j, plus the overall walking time between the kth station and the $k+1$st.

Let $d(k)$ be the RV that represents the total walking time between the father station k to the next father station. Its generating function is denoted by $D^*_k(s)$. For a given $i = 1, 2, ..., N$, note that $d(i+lN), l = 0, 1, 2, ...$, are i.i.d. RV. We denote by d_{i_f} (f stands for "father" station) a generic RV with that distribution. Its generating function is denoted by $D^*_{i_f}(s)$, and the first and second moments are denoted by $\bar{d}_{i_f}$ and $d^{(2)}_{i_f}$,respectively. Note that $D^*_{i_f}(s) = \prod_{j=0}^{m(i)} D^*_{i_j}(s)$.
Denote
$d := \sum_{i=1}^{N} \sum_{j=0}^{m(i)} d_{i_j}$, $\quad \overline{d} := E[d]$, $\quad d^{(2)} := E\left[d^2\right]$

The *state* of the system at the kth polling instance is taken to be the following vector of 2N-1 station times:

$$\vartheta_k = \{\theta_{k-(2N-2)}, \theta_{k-(2N-3)}, ..., \theta_{k-1}, \theta_k\} \tag{1}$$

Comparing this choice of state to the one in the standard gated discipline (see [8]), we note that whereas in the standard gated discipline only station times from the previous "cycle" are considered, we need here station times from the two previous cycles (the number of station times considered may be reduced, as explained latter in section 8). The need for those would become clear in the analysis to follow. Yet, we do not need to define and incorporate station times for all the M stations in order to obtain the dynamics of the system in the father stations.

Let T_l^k be a RV that stands for the time it takes to serve station k_l and denote $T(k) := \sum_{l=0}^{m(k)} T_l^k$.

3 The Generating Functions

3.1 The Evolution Equations

The system evolves according to the following dynamics:

$$\vartheta_{k+1} = \{\theta_{k-(2N-3)}, \theta_{k-(2N-4)}, ..., \theta_{k-1}, \theta_k, T(k+1) + d(k+1)\}$$

Note that T_j^k is the time it takes to serve the jobs that arrived to station k_j during the station times $\theta_{\hat{s}(k_j)-N}, \theta_{\hat{s}(k_j)-N+1}, ..., \theta_{\hat{s}(k_j)-1}$ This fact motivated the incorporation of station times from more than one cycle into the state of the system.

Note that the values of 2N-2 components of ϑ_k and of ϑ_{k+1} are common.

3.2 The Joint Generating Functions of the Station Times

For $\underline{x} := \{x_0, x_1, ..., x_{2N-2}\}$ we define

$$G_k^*(\underline{x}) := E \exp\left[-\sum_{j=0}^{2N-2} x_j \theta_{k-j}\right]$$

Define the following indicator functions:

$$I^k(j,n) = \begin{cases} 1, & k - \hat{s}(k_j) \le n \le k - \hat{s}(k_j) + N - 1 \\ 0, & \text{otherwise} \end{cases}$$

$$n = 0, 1, ..., 2N-2 \qquad j = 0, 1, ..., m(k)$$

Lemma 1 *The generating functions $G_k^*(\underline{x})$, $k = 1, 2, \ldots$ satisfy the following recursion equality:*

$$G_{k+1}^*(\underline{x}) = D_{k+1}^*(x_0) G_k^*(\underline{\alpha}) \tag{2}$$

where $\underline{\alpha} := \{\alpha_0, \alpha_1, \ldots, \alpha_{2N-2}\}$ is given by:

$$\alpha_n = x_{n+1} \cdot 1\{n \neq 2N-2\} + \sum_{j=0}^{m(k+1)} I^{k+1}(j,n) \lambda_{(k+1)_j} \left(1 - B_{(k+1)_j}^*(x_0)\right)$$

Proof:

$$G_{k+1}^*(\underline{x}) = E \exp\left[-\sum_{j=0}^{2N-2} x_j \theta_{k+1-j}\right] = E \exp\left[-\sum_{j=1}^{2N-2} x_j \theta_{k-j+1} - x_0\left(T(k+1) + d(k+1)\right)\right] =$$

$$E \exp\left[-x_0 d(k+1)\right] E\left\{E\left(\exp\left[-\sum_{j=1}^{2N-2} x_j \theta_{k-j+1} - x_0 T(k+1)\right] \middle| \vartheta_k\right)\right\} =$$

$$D_{k+1}^*(x_0) E\left\{\exp\left[-\sum_{j=1}^{2N-2} x_j \theta_{k-j+1}\right] E\left(\exp\left[-x_0 T(k+1)\right] \middle| \vartheta_k\right)\right\}$$

Since $T(k+1) := \sum_{j=0}^{m(k)} T_j^{k+1}$ and since the summands are independent given ϑ_k, we obtain:

$$G_{k+1}^*(\underline{x}) = D_{k+1}^*(x_0) E\left\{\exp\left[-\sum_{j=1}^{2N-2} x_j \theta_{k-j+1}\right] \prod_{j=0}^{m(k+1)} E\left(\exp\left[-x_0 T_j^{k+1}\right] \middle| \vartheta_k\right)\right\} \tag{3}$$

The last expression can be evaluated by direct integration (see c.f. [8] for the more simple case of standard exhaustive discipline). However, this cumbersome calculation can be avoided by recalling the well known solution of the following

Auxiliary Problem: Assume a Poisson flow of jobs with rate λ. Let T, be the (random) time that it takes to serve those jobs which arrived during the random time L whose Laplace transform is $L^*(s)$. The service times of different jobs are i.i.d. random variables with Laplace transform $B^*(s)$; these times, the arrival times and L are mutually independent. Then the Laplace Transform of the distribution of T is given by:

$$T^*(s) = E \exp\{-T\} = L^*\left(\lambda - \lambda B^*(s)\right)$$

and as special case,

$$E\left(\exp\{-sT\} \middle| L = l\right) = \exp\left\{-\left(\lambda - \lambda B^*(s)\right) l\right\}$$

Using this fact we obtain from (3):

$$G^*_{k+1}(\underline{x}) =$$

$$D^*_{k+1}(x_0)E\left\{\exp\left[-\sum_{j=1}^{2N-2} x_j\theta_{k-j+1}\right]\prod_{j=0}^{m(k+1)}\exp\left[-\lambda_{(k+1)_j}\left(1-B^*_{(k+1)_j}(x_0)\right)\sum_{l=1}^{N}\theta_{\hat{s}((k+1)_j)-l}\right]\right\} =$$

$$D^*_{k+1}(x_0)E\exp\left\{-\sum_{j=1}^{2N-2} x_j\theta_{k-j+1} - \sum_{j=0}^{m(k+1)}\left[\lambda_{(k+1)_j}\left(1-B^*_{(k+1)_j}(x_0)\right)\sum_{l=1}^{N}\theta_{\hat{s}((k+1)_j)-l}\right]\right\}$$

The theorem then follows where the α_n are the coefficients that multiply the appropriate θ's in the last expression. ■

Next, we obtain the the steady-state statistical behavior of the vectors of station times. Assume that at $\tau(1)$ station number 1 was polled. Under standard stability conditions ($\rho < 1$), for each i, $1 \leq i \leq N$, ϑ_{i+lN} converge in distribution to some random vector $\tilde{\vartheta}_i$ as $l \to \infty$; we denote its components by $\tilde{\theta}_{i-j}, j = 0, ..., 2N-2$ and the Laplace transform by $\tilde{G}_i(\underline{x})$, i.e.

$$\tilde{G}_i(\underline{x}) := E\exp\left[-\sum_{j=0}^{2N-2} x_j\tilde{\theta}_{i-j}\right]$$

Note that for each i, $1 \leq i \leq N$, $\tilde{\theta}_{i-N}$ has the same distribution as $\tilde{\theta}_i$.

In order to obtain recursive expressions for $\tilde{G}_i(\underline{x})$ we define the following indicator functions. Let $i = 1, 2, ..., N \quad n = 0, 1, ..., 2N-2 \quad j = 0, 1, ..., m(i)$.
If $s(i_j) \leq i$ then

$$\tilde{I}^i(j,n) = \begin{cases} 1, & i - s(i_j) \leq n \leq i - s(i_j) + N - 1 \\ 0, & \text{otherwise} \end{cases}$$

If $s(i_j) > i$ then

$$\tilde{I}^i(j,n) = \begin{cases} 1, & i + N - s(i_j) \leq n \leq i - s(i_j) + 2N - 1 \\ 0, & \text{otherwise} \end{cases}$$

Theorem 2 *The generating functions $\tilde{G}_i(\underline{x})$, $\quad i = 1,...,N$ can be computed by solving the following set of equations:*

$$\tilde{G}_{i+1}(\underline{x}) := D^*_{(i+1)_f}(x_0)\tilde{G}_i(\underline{\tilde{\alpha}}), \qquad i = 1,2,...,N \tag{4}$$

where $\underline{\tilde{\alpha}} := \{\tilde{\alpha}_0, \tilde{\alpha}_1, ..., \tilde{\alpha}_{2N-2}\}$ *is given by:*

$$\tilde{\alpha}_n = x_{n+1} \cdot 1\{n \neq 2N-2\} + \sum_{j=0}^{m(i+1)} \tilde{I}^{i+1}(j,n)\lambda_{(i+1)_j}\left(1 - B^*_{(i+1)_j}(x_0)\right) \tag{5}$$

Proof: Follows immediately from Lemma 1. ■

Remark: In (4) and (5) we understand by $i+1$ to be equal to 1 when $i = N$.

4 The Moments

4.1 The First Moments

Let C_{i_j} be the random variable which represents the i_jth cycle duration in steady-state regime, i.e. the time it takes between the polling of station i_j till the next time it is polled. Clearly $E[C] := E[C_{i_j}]$ does not depend on i or j. Let $c(k)$ be the random variable which represents the kth cycle duration. Let $\delta(k) := \sum_{n=0}^{N-1} d(kN+n)$, i.e. the total walking time during the kth cycle.

Proposition 3 *The expectation E[C] satisfies*

$$E[C] = \frac{\bar{d}}{1-\rho} \tag{6}$$

for cyclic polling systems with arbitrary service discipline, provided that
(i) the server does not idle (except for the walking times),
(ii) With probability one,

$$\lim_{k\to\infty} \frac{1}{k}\sum_{l=1}^{k} c(l) = E[C] \quad , \qquad \lim_{k\to\infty} \frac{1}{k}\sum_{l=1}^{k} \delta(l) = \bar{d}$$

Proof: The fraction of time during which the server is busy is equal to ρ (a.s.). We show below that the fraction of time during which the server is not serving is equal to $\bar{d}/E[C]$ from which we obtain

$$1-\rho = \frac{\bar{d}}{E[C]}$$

which establishes the proposition. It remains to show that the fraction of time during which the server is not serving is equal to $\bar{d}/E[C]$. This quantity is given by

$$\frac{\lim_{k\to\infty}\sum_{l=1}^{k}\delta(l)}{\lim_{k\to\infty}\sum_{l=1}^{k}c(l)} = \frac{\lim_{k\to\infty}\frac{1}{k}\sum_{l=1}^{k}\delta(l)}{\lim_{k\to\infty}\frac{1}{k}\sum_{l=1}^{k}c(l)} = \frac{\bar{d}}{E[C]}$$

■

The conditions (i) and (ii) of the proposition clearly hold in our case.

We are now ready to calculate the expectations of the station times.

$$E[\theta_k] = E\left\{E[\theta_k|\vartheta_{k-1}]\right\} = E[d(k)] + \sum_{j=0}^{m(k)}\rho_{k_j}\sum_{l=1}^{N}E\left[\theta_{\hat{s}(k_j)-l}\right]$$

which yields in the steady-state regime

$$E[\tilde{\theta}_i] = \bar{d}_{i_f} + \sum_{j=0}^{m(i)}\rho_{i_j}E[C] = \bar{d}_{i_f} + \bar{\rho}_i\frac{\bar{d}}{1-\rho} \quad ,i=1,2,...,N \tag{7}$$

Where in (7), we define

$$\bar{\rho}_i \;:=\; \sum_{j=0}^{m(i)}\rho_{i_j}. \tag{8}$$

Note that for each i, $i=1,2,...,N$, $E[\tilde{\theta}_{i-l}]$ is defined for $l=1,2,...,2N-2$, where

$$E[\tilde{\theta}_{i-l}] = \begin{cases} E[\tilde{\theta}_{i-l}]\ ; & i-l>0 \\ E[\tilde{\theta}_{N+i-l}]\ ; & -(N-1)\le i-l\le 0 \\ E[\tilde{\theta}_{2N+i-l}]\ ; & -(2N-3)\le i-l\le -N \end{cases} \tag{9}$$

4.2 The Second Moments

In what follows, we shall use the following notation: for two integers i and j, let

$$i \ominus j := \begin{cases} i-j, & i \geq j \\ i-j+2N, & \text{otherwise} \end{cases} \tag{10}$$

Define for $i = 1, ..., N, \quad j = 1, ..., 2N$, the following second central moments,

$$r_{ij} := \begin{cases} E\left[(\tilde{\theta}_i - E\tilde{\theta}_i)(\tilde{\theta}_j - E\tilde{\theta}_j)\right] \ ; & j \leq i \\ E\left[(\tilde{\theta}_i - E\tilde{\theta}_i)(\tilde{\theta}_{j-2N} - E\tilde{\theta}_j)\right] \ ; & j > i+1 \end{cases} \tag{11}$$

And define,

$$\rho_i(j) := \sum_{l=0}^{m(i)} \tilde{I}^i(l, i \ominus (j+1))\rho_{i_l} \tag{12}$$

We further define for $i = N+1, N+2, ..., 2N, \quad i-j \neq 2N-1$

$$r_{ij} := \begin{cases} r_{(i-N)(j+N)} \ ; & 1 \leq j \leq N \\ r_{(i-N)(j-N)} \ ; & N+1 \leq j \leq 2N \end{cases} \tag{13}$$

Note that, in the above definitions of $r_{ij}, \tilde{\theta}_j$ corresponds to a station time that *comes before* the station time $\tilde{\theta}_i$. Also, note that r_{ij} and r_{ji} need not be equal.

Theorem 4 *Using the definitions in (10)-(13), the second moments of the station times in steady-state can be calculated by solving the following $N(2N-1)$ equations:*

$$r_{ij} = \begin{cases} \sum_{m=i+1}^{2N} \rho_i(m) r_{jm} + \sum_{m=1}^{j-1} \rho_i(m) r_{jm} + \sum_{m=j}^{i-1} \rho_i(m) r_{mj}; & j < i \\ \sum_{m=i+1}^{j-1} \rho_i(m) r_{jm} + \sum_{m=j}^{2N} \rho_i(m) r_{mj} + \sum_{m=1}^{i-1} \rho_i(m) r_{mj}; & j > i+1 \end{cases} \tag{14}$$

$$r_{ii} \ = \ \mathrm{var}(d_{i_f}) + \sum_{n=0}^{2N-2} \left(\sum_{l=0}^{m(i)} \tilde{I}^i(l,n) \lambda_{i_l} b_{i_l}^{(2)} \right) E[\tilde{\theta}_{i-n-1}] +$$

$$\sum_{\substack{j=1 \\ j \neq i, i+1}}^{2N} \rho_i(j) r_{ij} + \rho_i(i+1) \sum_{\substack{j=1 \\ j \neq i}}^{2N} \rho_i(j) r_{j(i+1)} \tag{15}$$

$$i = 1, 2, ..., N, \qquad j = 1, 2, ..., 2N$$

Proof: To obtain the second moments, we differentiate twice (4) with respective to the different x_j's.

$$\begin{aligned}
\frac{\partial^2 \tilde{G}_i(\underline{x})}{\partial x_0^2} &= \frac{\partial^2}{\partial x_0^2}\left[D_i^*(x_0)\tilde{G}_{i-1}(\underline{\alpha})\right] \\
&= \frac{d^2 D_i^*(x_0)}{dx_0^2}\tilde{G}_{i-1}(\underline{\alpha}) + 2\frac{dD_i^*(x_0)}{dx_0}\sum_{n=0}^{2N-2}\frac{d\alpha_n}{dx_0}\frac{\partial \tilde{G}_{i-1}(\underline{\alpha})}{\partial \alpha_n} \\
&\quad + D_i^*(x_0)\sum_{n=0}^{2N-2}\frac{d^2\alpha_n}{dx_0^2}\frac{\partial \tilde{G}_{i-1}(\underline{\alpha})}{\partial \alpha_n} + D_i^*(x_0)\sum_{n=0}^{2N-2}\sum_{m=0}^{2N-2}\frac{d\alpha_n}{dx_0}\frac{d\alpha_m}{dx_0}\frac{\partial^2 \tilde{G}_{i-1}(\underline{\alpha})}{\partial \alpha_n \partial \alpha_m}
\end{aligned} \tag{16}$$

Denote by $\underline{0}$ the vector with $2N-1$ zeros and define $\rho_{i_j} := \lambda_{i_j} b_{i_j}$.
Let $i = 1, ..., N$, and $n, m = 0, 1, ..., 2N-2$. We have:

$$\left.\frac{dD_i^*(x_0)}{dx_0}\right|_{\underline{x}=\underline{0}} = -\overline{d}_{i_f}, \qquad \left.\frac{d^2 D_i^*(x_0)}{dx_0^2}\right|_{\underline{x}=\underline{0}} = d_{i_f}^{(2)},$$

$$\left.\frac{d\tilde{\alpha}_n}{dx_0}\right|_{\underline{x}=\underline{0}} = \sum_{j=0}^{m(i)} \tilde{I}^i(j,n)\rho_{i_j}, \qquad \left.\frac{d^2\tilde{\alpha}_n}{dx_0^2}\right|_{\underline{x}=\underline{0}} = -\sum_{j=0}^{m(i)} \tilde{I}^i(j,n)\lambda_{i_j} b_{i_j}^{(2)},$$

$$\left.\frac{\partial \tilde{G}_{i-1}(\underline{\tilde{\alpha}})}{\partial \tilde{\alpha}_n}\right|_{\underline{x}=\underline{0}} = -E\left[\tilde{\theta}_{i-n-1}\right], \qquad \left.\frac{\partial^2 \tilde{G}_{i-1}(\underline{\tilde{\alpha}})}{\partial \tilde{\alpha}_n \partial \tilde{\alpha}_m}\right|_{\underline{x}=\underline{0}} = -E\left[\tilde{\theta}_{i-n-1}\tilde{\theta}_{i-m-1}\right]$$

We thus obtain by computing (17) at $\underline{x} = \underline{0}$:

$$\begin{aligned}
E\left[\tilde{\theta}_i^2\right] &= d_{i_f}^{(2)} + \sum_{n=0}^{2N-2}\left\{\sum_{j=0}^{m(i)} \tilde{I}^i(j,n)\lambda_{i_j}\left(2\overline{d}_{i_f}\overline{b}_{i_j} + b_{i_j}^{(2)}\right)\right\} E\left[\tilde{\theta}_{i-n-1}\right] \\
&\quad + \sum_{n=0}^{2N-2}\sum_{m=0}^{2N-2}\left\{\sum_{l=0}^{m(i)}\sum_{j=0}^{m(i)} \tilde{I}^i(l,n)\tilde{I}^i(j,m)\rho_{i_l}\rho_{i_j}\right\} E\left[\tilde{\theta}_{i-n-1}\tilde{\theta}_{i-m-1}\right]
\end{aligned} \tag{17}$$

Note that

$$E\left[\tilde{\theta}_i\right] = \sum_{n=0}^{2N-2}\left\{\sum_{j=0}^{m(i)}\tilde{I}^i(j,n)\rho_{i_j}\right\}E\left[\tilde{\theta}_{i-n-1}\right]+\overline{d}_{i_f} \tag{18}$$

We thus obtain

$$\begin{aligned}\left(E\left[\tilde{\theta}_i\right]\right)^2 &= \sum_{n=0}^{2N-2}\sum_{m=0}^{2N-2}\left\{\sum_{l=0}^{m(i)}\sum_{j=0}^{m(i)}\tilde{I}^i(l,n)\tilde{I}^i(j,m)\rho_{i_l}\rho_{i_j}\right\}E\left[\tilde{\theta}_{i-n-1}\right]E\left[\tilde{\theta}_{i-m-1}\right]\\ &+2\overline{d}_{i_f}\sum_{n=0}^{2N-2}\left(\sum_{j=0}^{m(i)}\tilde{I}^i(j,n)\rho_{i_j}\right)E\left[\tilde{\theta}_{i-n-1}\right]+(\overline{d}_{i_f})^2\end{aligned} \tag{19}$$

Similarly, we obtain by computing $\dfrac{\partial^2\tilde{G}_i(\underline{x})}{\partial x_0\partial x_l}$ at $\underline{x}=\underline{0}$:

$$E\left[\tilde{\theta}_i\tilde{\theta}_{i-l}\right]=\overline{d}_{i_f}E\left[\tilde{\theta}_{i-l}\right]+\sum_{n=0}^{2N-2}\left(\sum_{j=0}^{m(i)}\tilde{I}^i(j,n)\rho_{i_j}\right)E\left[\tilde{\theta}_{i-n-1}\tilde{\theta}_{i-l}\right],\qquad l=1,...,2N-2 \tag{20}$$

Note that

$$E\left[\tilde{\theta}_i\right]E\left[\tilde{\theta}_{i-l}\right]=\left\{\sum_{n=0}^{2N-2}\left(\sum_{j=0}^{m(i)}\tilde{I}^i(j,n)\rho_{i_j}\right)E\left[\tilde{\theta}_{i-n-1}\right]+\overline{d}_{i_f}\right\}E\left[\tilde{\theta}_{i-l}\right] \tag{21}$$

From (17), (19) (20) and (21) we obtain (14) and

$$\begin{aligned}r_{ii} &= \mathrm{var}(d_{i_f})+\sum_{n=0}^{2N-2}\left(\sum_{l=0}^{m(i)}\tilde{I}^i(l,n)\lambda_{i_l}b_{i_l}^{(2)}\right)E[\tilde{\theta}_{i-n-1}]+\\ &\sum_{j=1}^{i-1}\rho_i(j)\left(\sum_{m=i+1}^{2N}\rho_i(m)r_{jm}+\sum_{m=1}^{j-1}\rho_i(m)r_{jm}+\sum_{m=j}^{i-1}\rho_i(m)r_{mj}\right)+\\ &\sum_{j=i+1}^{2N}\rho_i(j)\left(\sum_{m=i+1}^{j-1}\rho_i(m)r_{jm}+\sum_{m=j}^{2N}\rho_i(m)r_{mj}+\sum_{m=1}^{i-1}\rho_i(m)r_{mj}\right)\end{aligned} \tag{22}$$

Substituting (14) in (22) finally yields (15). ■

5 The Waiting Times

Standard techniques can be used to compute the Laplace transform and the expectation of the waiting times for jobs in "father" stations, using the previous expressions for the Laplace transform of the station times, (see e.g. [8] or [9]). For the synchronized stations the situation is more complex, and we obtain the distribution of the waiting times through a similar method to the one used in [16].

5.1 Waiting times in the fathers stations

Let W_i denote the waiting time at an arbitrary moment in steady-state regime in the "father" station i, $i = 1, 2, ..., N$, and let $W_i^*(s)$ denote its Laplace transform. Let $C_i^*(s)$ denote the Laplace transform of the ith cycle duration in steady-state C_i, $i = 1, ..., N$. We have:

$$\begin{aligned} C_i^*(s) \quad &:= \quad E\left[\exp\{-sC_i\}\right] = E\left[\exp\left(-s\sum_{l=1}^{N}\tilde{\theta}_{i-l}\right)\right] = \\ & \tilde{G}_i(\underline{x})\Big|_{\substack{x_0 = x_{N+1} = x_{N+2} = ...x_{2N-2} = 0 \\ x_1 = x_2 = ... = x_N = s}} \end{aligned} \tag{23}$$

The Laplace transform of the waiting time is then given by

$$W_i^*(s) = \frac{1}{E[C]} \frac{C_i^*\left[\lambda_i - \lambda_i B_i^*(s)\right] - C_i^*(s)}{s - \lambda_i + \lambda_i B_i^*(s)}$$

and the expectation of the waiting time is given by

$$E\left[W_i\right] = E[C]\frac{1+\rho_i}{2} + \frac{(1+\rho_i)\,\mathrm{var}\,(C_i)}{2E[C]}$$

where

$$\begin{aligned} \mathrm{var}\,(C_i) \quad &= \quad E\left[\left(\sum_{l=1}^{N}\left(\tilde{\theta}_{i-l} - E\tilde{\theta}_{i-l}\right)\right)^2\right] = \sum_{n=i-N}^{i-1}\ \sum_{m=i-N}^{i-1} E\left[\left(\tilde{\theta}_n - E\tilde{\theta}_n\right)\left(\tilde{\theta}_m - E\tilde{\theta}_m\right)\right] \\ &= \quad \sum_{n=1}^{i-1}\left(\sum_{m=i+N}^{2N} r_{nm} + \sum_{m=1}^{n-1} r_{nm} + \sum_{m=n}^{i-1} r_{mn}\right) + \sum_{n=i+N}^{2N}\left(\sum_{m=i+N}^{n-1} r_{nm} + \sum_{m=n}^{2N} r_{mn} + \sum_{m=1}^{i-1} r_{mn}\right) \end{aligned}$$

5.2 Waiting times in the synchronized stations

In this sub-section we compute the Laplace transform and the expectation of the waiting times for jobs in the synchronized stations, namely for the stations:

$$i_j \ , \quad i = 1, 2, ..., N \quad j = 0, 1, ..., m(i)$$

The formulas we obtain for these quantities hold also for jobs in the "father" stations.
In order to proceed, we define the following random variables:
L_{i_j} := the number of jobs that a job departing from station i_j sees at station i_j in steady-state regime.
$X_{i_j}^{i_j}$:= the number of jobs in station i_j at the polling instant to that station.
$X_{s(i_j)}^{i_j}$:= the number of jobs in station i_j at the polling instant to its father $s(i_j)$.
Let $\tilde{s}(i_l) := s(i_l) - N \cdot 1\{s(i_l) > i\}$.
We define the RV $d_{i \to j}$ to be the total walking time from station i to station i_j. Let $D^*_{i \to j}(s)$ denote its moment generating function. Then, $D^*_{i \to j}(s) = \prod_{n=0}^{j-1} D^*_{i_j}(s)$.

Define the moment generating function $Q_{i_j}(z) := E\left[z^{L_{i_j}}\right]$.

In the appendix, we compute $Q_{i_j}(z)$ and obtain

$$Q_{i_j}(z) = \frac{B^*_{i_j}(\lambda_{i_j} - \lambda_{i_j} z)}{E\left[X_{s(i_j)}^{i_j}\right]} \times \frac{E\left[\left(B^*_{i_j}(\lambda_{i_j} - \lambda_{i_j} z)\right)^{X_{s(i_j)}^{i_j}} \cdot z^{X_{i_j}^{i_j} - X_{s(i_j)}^{i_j}}\right] - E\left[z^{X_{i_j}^{i_j}}\right]}{B^*_{i_j}(\lambda_{i_j} - \lambda_{i_j} z) - z} \tag{24}$$

The waiting time distribution is related to $Q_{i_j}(z)$ through

$$Q_{i_j}(z) = W^*_{i_j}\left(\lambda_{i_j} - \lambda_{i_j} z\right) B^*_{i_j}\left(\lambda_{i_j} - \lambda_{i_j} z\right) \tag{25}$$

Note also that

$$E\left[X_{s(i_j)}^{i_j}\right] = E\left[E\left(X_{s(i_j)}^{i_j} \middle| \tilde{\vartheta}_{s(i_j)-1}\right)\right] = E\left[\lambda_{i_j} \sum_{l=\tilde{s}(i_j)-1}^{\tilde{s}(i_j)-N} \tilde{\theta}_l\right] = \lambda_{i_j} E\left[C_{s(i_j)}\right] = \lambda_{i_j} E[C] \tag{26}$$

Using (24)-(26), we obtain

$$W^*_{i_j}(s) = \frac{1}{E[C]} \times \frac{E\left[\left(B^*_{i_j}(\lambda_{i_j} - \lambda_{i_j} z)\right)^{X_{s(i_j)}^{i_j}} \cdot z^{X_{i_j}^{i_j} - X_{s(i_j)}^{i_j}}\right]_{z=1-s/\lambda_{i_j}} - E\left[z^{X_{i_j}^{i_j}}\right]_{z=1-s/\lambda_{i_j}}}{s - \lambda_{i_j} + \lambda_{i_j} B^*_{i_j}(s)} \tag{27}$$

In what follows, we calculate the terms which appear in the numerator of (27).

- $$E\left[\left(B^*_{i_j}(\lambda_{i_j}-\lambda_{i_j}z)\right)^{X^{i_j}_{s(i_j)}}\cdot z^{X^{i_j}_{i_j}-X^{i_j}_{s(i_j)}}\right]=E\left(E\left[\left(B^*_{i_j}(\lambda_{i_j}-\lambda_{i_j}z)\right)^{X^{i_j}_{s(i_j)}}\cdot z^{X^{i_j}_{i_j}-X^{i_j}_{s(i_j)}}\middle|\tilde{\vartheta}_{i-1}\right]\right)=$$

$$D^*_{i\to j}\left(\lambda_{i_j}-\lambda_{i_j}z\right)E\quad\left(\exp\left\{-\lambda_{i_j}(1-z)\sum_{l=\tilde{s}(i_j)}^{i-1}\tilde{\theta}_l-\sum_{l=0}^{j-1}\left[\lambda_{i_l}\left(1-B^*_{i_l}(\lambda_{i_j}-\lambda_{i_j}z)\right)\sum_{n=1}^{N}\tilde{\theta}_{\tilde{s}(i_l)-n}\right]\right\}\right.$$
$$\left.\times\exp\left\{-\lambda_{i_j}\left(1-B^*_{i_j}(\lambda_{i_j}-\lambda_{i_j}z)\right)\sum_{l=\tilde{s}(i_j)-N}^{\tilde{s}(i_j)-1}\tilde{\theta}_l\right\}\right)$$

The last equality is obtained using the auxiliary problem in the end of this Section. We thus obtain:

$$E\left[\left(B^*_{i_j}(\lambda_{i_j}-\lambda_{i_j}z)\right)^{X^{i_j}_{s(i_j)}}\cdot z^{X^{i_j}_{i_j}-X^{i_j}_{s(i_j)}}\right]=D^*_{i\to j}\left(\lambda_{i_j}-\lambda_{i_j}z\right)\tilde{G}^*_{i-1}(\underline{\beta})$$

Where $\underline{\beta}=\{\beta_0,\beta_1,...,\beta_{2N-2}\}$ is given by

$$\begin{aligned}\beta_n \quad = \quad & 1\left\{0\le n\le i-\tilde{s}(i_j)-1\right\}\cdot\lambda_{i_j}(1-z)+\sum_{l=0}^{j-1}\lambda_{i_l}\left(1-B^*_{i_l}(\lambda_{i_j}-\lambda_{i_j}z)\right)\tilde{I}^i(l,n)+\\ & 1\left\{i-\tilde{s}(i_j)\le n\le i-\tilde{s}(i_j)+N-1\right\}\cdot\lambda_{i_j}\left(1-B^*_{i_j}(\lambda_{i_j}-\lambda_{i_j}z)\right)\end{aligned}\tag{28}$$

$$n=0,1,...,2N-2$$

- $$E\left[z^{X^{i_j}_{i_j}}\right]=$$

$$D^*_{i\to j}\left(\lambda_{i_j}-\lambda_{i_j}z\right)E\exp\left\{-\lambda_{i_j}(1-z)\sum_{l=\tilde{s}(i_j)-N}^{i-1}\tilde{\theta}_l-\sum_{l=0}^{j-1}\left[\lambda_{i_l}\left(1-B^*_{i_l}(\lambda_{i_j}-\lambda_{i_j}z)\right)\sum_{n=1}^{N}\tilde{\theta}_{\tilde{s}(i_l)-n}\right]\right\}$$

From which we obtain:

$$E\left[z^{X^{i_j}_{i_j}}\right]=D^*_{i\to j}\left(\lambda_{i_j}-\lambda_{i_j}z\right)\tilde{G}^*_{i-1}(\underline{\gamma})$$

Where $\underline{\gamma}=\{\gamma_0,\gamma_1,...,\gamma_{2N-2}\}$ is given by

$$\gamma_n=1\left\{0\le n\le i-\tilde{s}(i_j)+N-1\right\}\cdot\lambda_{i_j}(1-z)+\sum_{l=0}^{j-1}\lambda_{i_l}\left(1-B^*_{i_l}(\lambda_{i_j}-\lambda_{i_j}z)\right)\tilde{I}^i(l,n)$$

$$n = 0, 1, ..., 2N - 2$$

Next, we calculate the expected waiting time in a synchronized station.

$$E[Wi_j] = -\frac{d}{ds} W^*_{i_j}(s)\Big|_{s=0} =$$

$$\frac{(1+\rho_{i_j})\mathrm{var}\left(X^{i_j}_{s(i_j)}\right)}{2\lambda_{i_j} E\left[X^{i_j}_{s(i_j)}\right]} + \frac{E\left[X^{i_j}_{s(i_j)}\left(X^{i_j}_{i_j} - X^{i_j}_{s(i_j)}\right)\right] - E\left[X^{i_j}_{s(i_j)}\right] E\left[X^{i_j}_{i_j} - X^{i_j}_{s(i_j)}\right]}{\lambda_{i_j} E\left[X^{i_j}_{s(i_j)}\right]} +$$

$$\frac{E\left[X^{i_j}_{i_j} - X^{i_j}_{s(i_j)}\right]}{\lambda_{i_j}} + \frac{(1+\rho_{i_j}) E\left[X^{i_j}_{s(i_j)}\right]}{2\lambda_{i_j}} - \frac{(1+\rho_{i_j})}{2\lambda_{i_j}}$$

Where $E\left[X^{i_j}_{s(i_j)}\right]$ is given in (26) and,

- $E\left[X^{i_j}_{i_j} - X^{i_j}_{s(i_j)}\right] = E\left[E\left(X^{i_j}_{i_j} - X^{i_j}_{s(i_j)}\Big|\, \tilde{\vartheta}_{i-1}\right)\right] =$

$$E\left[\lambda_{i_j} \sum_{l=\tilde{s}(i_j)}^{i-1} \tilde{\theta}_l + \lambda_{i_j}\left(E\left[d_{i\to j}\right] + \sum_{l=0}^{j-1} \rho_{i_l} \sum_{n=1}^{N} \tilde{\theta}_{\tilde{s}(i_l)-n}\right)\right]$$

Note that, $E\left[\sum_{n=1}^{N} \tilde{\theta}_{\tilde{s}(i_l)-n}\right] = E\left[C_{s(i_j)}\right] = E[C]$. Then, we have

$$E\left[X^{i_j}_{i_j} - X^{i_j}_{s(i_j)}\right] = \lambda_{i_j}\left[\sum_{l=\tilde{s}(i_j)}^{i-1} E[\tilde{\theta}_l] + \sum_{l=0}^{j-1}\left(\bar{d}_{i_l} + \rho_{i_l} E[C]\right)\right] \tag{29}$$

Where $E[\tilde{\theta}_l]$ is given by (7) and (9).

- $\mathrm{var}\left(X^{i_j}_{s(i_j)}\right) = \lambda^2_{i_j}\mathrm{var}[C_{s(i_j)}] + \lambda_{i_j} E[C]$

(see e.g. [10] pp. 186-187 Eq. 5.58, 5.61).

- $E\left[X^{i_j}_{s(i_j)}\left(X^{i_j}_{i_j} - X^{i_j}_{s(i_j)}\right)\right] - E\left[X^{i_j}_{s(i_j)}\right] E\left[X^{i_j}_{i_j} - X^{i_j}_{s(i_j)}\right] =$

$$= \lambda^2_{i_j} E\left[\left(\sum_{n=1}^{N}\left(\tilde{\theta}_{\tilde{s}(i_j)-n} - E\left[\tilde{\theta}_{\tilde{s}(i_j)-n}\right]\right)\right) \times \left(\sum_{m=\tilde{s}(i_j)}^{i-1}\left(\tilde{\theta}_m - E[\tilde{\theta}_m]\right)\right)\right] + \tag{30}$$

$$\lambda^2_{i_j} \sum_{l=0}^{j-1} \rho_{i_l} E\left[\left(\sum_{n=1}^{N}\left(\tilde{\theta}_{\tilde{s}(i_l)-n} - E\left[\tilde{\theta}_{\tilde{s}(i_l)-n}\right]\right)\right) \times \left(\sum_{m=1}^{N}\left(\tilde{\theta}_{\tilde{s}(i_j)-m} - E[\tilde{\theta}_{\tilde{s}(i_j)-m}]\right)\right)\right] \tag{31}$$

(30) is given by

$$(30) = \begin{cases} \lambda_{i_j}^2 \sum_{n=s(i_j)}^{i-1} \left(\sum_{m=s(i_j)+N}^{2N} r_{nm} + \sum_{m=1}^{s(i_j)-1} r_{nm} \right) \; ; & s(i_j) \le i \\ \\ \lambda_{i_j}^2 \sum_{n=s(i_j)}^{s(i_j)+N-1} \left(\sum_{m=s(i_j)+N}^{2N} r_{mn} + \sum_{m=1}^{i-1} r_{mn} \right) \; ; & s(i_j) > i \end{cases} \tag{32}$$

To compute (31), we have to compute

$$E\left[\left(\sum_{n=1}^{N} \left(\tilde{\theta}_{\tilde{s}(i_l)-n} - E\left[\tilde{\theta}_{\tilde{s}(i_l)-n}\right]\right)\right) \times \left(\sum_{m=1}^{N} \left(\tilde{\theta}_{\tilde{s}(i_j)-m} - E[\tilde{\theta}_{\tilde{s}(i_j)-m}]\right)\right)\right] \tag{33}$$

for, i_l, $\quad i = 1, 2, ..., N \qquad l, j = 1, 2, ..., m(i)$

For $s(i_l) \le s(i_j) \le i$ (and similarly for $s(i_j) \le s(i_l) \le i$) we have

$$\begin{aligned} (33) = \quad & \sum_{n=s(i_l)}^{s(i_j)-1} \left(\sum_{m=s(i_l)+N}^{2N} r_{nm} + \sum_{m=1}^{s(i_l)-1} r_{nm} \right) + \\ & \sum_{n=1}^{s(i_l)-1} \left(\sum_{m=s(i_l)+N}^{2N} r_{nm} + \sum_{m=1}^{n-1} r_{nm} + \sum_{m=n}^{s(i_l)-1} r_{mn} \right) + \\ & \sum_{n=s(i_j)+N}^{2N} \left(\sum_{m=s(i_l)+N}^{n-1} r_{nm} + \sum_{m=n}^{2N} r_{mn} + \sum_{m=1}^{s(i_l)-1} r_{mn} \right) \end{aligned}$$

For the case that $s(i_j), s(i_l) \ge i$ and $s(i_l) \le s(i_j)$ (and similarly for $s(i_l) \ge s(i_j)$) we get

$$(33) = \sum_{n=s(i_l)+N}^{s(i_j)+N-1} \sum_{m=s(i_l)}^{s(i_l)+N-1} r_{nm} + \sum_{n=s(i_j)}^{s(i_l)+N-1} \left(\sum_{m=s(i_l)}^{n-1} r_{nm} + \sum_{m=n}^{s(i_l)+N-1} r_{mn} \right)$$

Finally, for the case that $s(i_j) > i$ and $s(i_l) < i$ (and similarly for the case that $s(i_l) > i$ and $s(i_j) < i$) we get

$$\begin{aligned} (33) = \quad & \sum_{n=s(i_j)}^{s(i_l)+N-1} \left(\sum_{m=s(i_l)+N}^{2N} r_{mn} + \sum_{m=1}^{s(i_l)-1} r_{mn} \right) + \\ & \sum_{n=s(i_l)+N}^{s(i_j)+N-1} \left(\sum_{m=s(i_l)+N}^{n-1} r_{nm} + \sum_{m=n}^{2N} r_{mn} + \sum_{m=1}^{s(i_l)-1} r_{mn} \right) \end{aligned}$$

Auxiliary Problem:

We calculate the distribution of the number of jobs that arrive to station i_j during T_{i_l} – the time that station i_l is served. We denote this number by $Y_{i_j,l}$. The jobs that are served at station i_l are those that arrived to that station during $C_{s(i_l)}$; we denote them by n. The moment generating function of $Y_{i_j,l}$ satisfies:

$$Y^*_{i_j,l}(z) = T^*_{i_l}(\lambda_{i_j} - \lambda_{i_j} z) = E\exp\left\{-\lambda_{i_j}(1-z)T_{i_l}\right\} = E\left[E\exp\left\{-\lambda_{i_j}(1-z)T_{i_l}\middle| n\right\}\right]$$

$$= E\left[\left(B^*_{i_l}(\lambda_{i_j} - \lambda_{i_j} z)\right)^n\right] = E\left\{E\left[\left(B^*_{i_l}(\lambda_{i_j} - \lambda_{i_j} z)\right)^n \middle| \tilde{\vartheta}_{i-1}\right]\right\}$$

$$= E\exp\left\{-\lambda_{i_l}\left(1 - B^*_{i_l}(\lambda_{i_j} - \lambda_{i_j} z)\right)\sum_{m=1}^{N}\tilde{\theta}_{\tilde{s}(i_l)-m}\right\}$$

6 A Pseudo Conservation Law

6.1 Computing the Conservation Law

In Boxma and Groenendijk [5] a "pseudo" conservation law has been derived for polling systems with mixed service strategies. The derivation of this law is based on a work decomposition theorem. Recently, Boxma [6] extended this theorem to much more general single server systems with non-serving intervals. Here, we adapt *Theorem 1* of Boxma and Groenendijk [5] to our network with the synchronized gated discipline.

The derivation of "pseudo" conservation law for our network proceeds in the same way as in Boxma and Groenendijk [5] and we may obtain an equivalent equation to equation (3.10) in [5] with the appropriate modifications in notation,

$$\begin{aligned}\sum_{i=1}^{N}\sum_{j=0}^{m(i)}\rho_{i_j}E\left[W_{i_j}\right] \quad = \quad & \rho\frac{\sum_{i=1}^{N}\sum_{j=0}^{m(i)}\lambda_{i_j}b^{(2)}_{i_j}}{2(1-\rho)} + \rho\frac{d^{(2)}}{2\overline{d}} \\ & + \frac{\overline{d}}{2(1-\rho)}\left[\rho^2 - \sum_{i=1}^{N}\sum_{j=0}^{m(i)}\rho^2_{i_j}\right] + \sum_{i=1}^{N}\sum_{j=0}^{m(i)}E\left[M^{(1)}_{i_j}\right] \qquad (34)\end{aligned}$$

Where $M_{i_j}^{(1)}, \quad i=1,2,...,N, \quad j=0,1,..,m(i)$ is defined as in [5], and is equal to the amount of work in station i_j at a departure epoch of the server from station i_j .

The meaning of the terms in the right-hand side of (34) is given in Boxma and Groenendijk [5] just after equation (3.10). It turns out that $E\left[M_{i_j}^{(1)}\right]$ is the only term in the right-hand side of (34) which depends on the service strategy at station i_j, as expected. $\sum_{i=1}^{N}\sum_{j=0}^{m(i)} E\left[M_{i_j}^{(1)}\right]$ represents the mean amount of work that arrived at the stations during the last service periods of these stations, but that was not handled by the server at those service periods. $E\left[M_{i_j}^{(1)}\right]$ *depends on the service strategy at station* i_j . We now turn to the determination of $E\left[M_{i_j}^{(1)}\right]$ in the case of synchronized gated discipline.

For $i=1,2,...,N$ we have the following,

$$E\left[M_{i}^{(1)}\right] = \rho_i^2 E[C] = \rho_i^2 \frac{\overline{d}}{1-\rho} \tag{35}$$

For the synchronized stations, $\quad i=1,2,...,N, \quad j=1,2,..,m(i)$ we have the following,

$$E\left[M_{i_j}^{(1)}\right] = \rho_{i_j}^2 E[C] + E\left[X_{i_j}^{i_j} - X_{s(i_j)}^{i_j}\right] b_{i_j} = \rho_{i_j}^2 \frac{\overline{d}}{1-\rho} + E\left[X_{i_j}^{i_j} - X_{s(i_j)}^{i_j}\right] b_{i_j} \tag{36}$$

Using (7) and (9) in (29), we obtain

$$E\left[X_{i_j}^{i_j} - X_{s(i_j)}^{i_j}\right] = \lambda_{i_j}\left[\sum_{l=\tilde{s}(i_j)}^{i-1}(\bar{\rho}_l E[C] + \bar{d}_{l_f}) + \sum_{l=0}^{j-1}(\rho_{i_l} E[C] + \bar{d}_{i_l})\right] \tag{37}$$

where in (37), we used the following difinitions,

$$\bar{\rho}_{l-N} := \bar{\rho}_l \quad and \quad \bar{d}_{(l-N)_f} := \bar{d}_{l_f}, \quad l=1,2,...,N$$

Substitution of (35), (36) and (37) in (34), we obtain a "pseudo" conservation law for our network,

$$\sum_{i=1}^{N}\sum_{j=0}^{m(i)}\rho_{i_j}E\left[W_{i_j}\right] = \rho\frac{\sum_{i=1}^{N}\sum_{j=0}^{m(i)}\lambda_{i_j}b_{i_j}^{(2)}}{2(1-\rho)}+\rho\frac{d^{(2)}}{2\overline{d}}+\frac{\overline{d}}{2(1-\rho)}\left[\rho^2+\sum_{i=1}^{N}\sum_{j=0}^{m(i)}\rho_{i_j}^2\right]$$
$$+\sum_{i=1}^{N}\sum_{j=0}^{m(i)}\rho_{i_j}\left[\sum_{l=\tilde{s}(i_j)}^{i-1}\left(\frac{\overline{d}}{1-\rho}\overline{\rho}_l+\overline{d}_{l_f}\right)+\sum_{l=0}^{j-1}\left(\frac{\overline{d}}{1-\rho}\rho_{i_l}+\overline{d}_{i_n}\right)\right] \quad (38)$$

The last term in the right-hand side of (38) reflects the effect of the presence of synchronization. The weighted sum of the mean waiting times *increases* in comparison with the standard gated discipline. Without synchronization this term is dropped, and we got the "pseudo" conservation law for the standard gated discipline.

6.2 Optimizing the topology

Assume that we are free to choose the topology of the polling system; more precisely, we assume that we may change the placement of the different stations on the cyclic path of the server. We assume that the expectation of the walking time is equal for all stations.

Our goal is to minimize the weighted sum of the mean waiting times at the left-hand side of equation (38).

We use the "pseudo" conservation law from the last Section to obtain the optimal topology.

Theorem 5 *An optimal topology for the order of different stations is obtained by placing for each father station all it's children between him and the subsequent father station in the cycle. In that case, the weighted sum of the mean waiting times is not sensitive to the order between father stations nor to the internal order between the child stations of any given father station.*

Proof:

We first note that for any topology, the weighted sum of the mean waiting times may be reduced if for each father station, all it's children are positioned between him and the subsequent father station in the cycle. In that case, we obtain from equation (38)

$$\sum_{i=1}^{N}\sum_{j=0}^{m(i)} \rho_{i_j} E\left[W_{i_j}\right] = \rho\frac{\sum_{i=1}^{N}\sum_{j=0}^{m(i)}\lambda_{i_j} b_{i_j}^{(2)}}{2(1-\rho)} + \rho\frac{d^{(2)}}{2\overline{d}} + \frac{\overline{d}}{2(1-\rho)}\left[\rho^2 + \sum_{i=1}^{N}\sum_{j=0}^{m(i)} \rho_{i_j}^2\right]$$
$$+\sum_{i=1}^{N}\sum_{j=0}^{m(i)} \rho_{i_j}\left[\sum_{n=0}^{j-1}\overline{d}_{i_n} + \frac{\overline{d}}{1-\rho}\sum_{n=0}^{j-1}\rho_{i_n}\right] \qquad (39)$$

Under the assumption that $\bar{d}_{i_n}$ does not depend on i_n we see that the weighted sum of the mean waiting times is independent of the order of the father stations; Moreover we see that the optimization problem decomposes to N independent optimization problems; for each $1 \leq i \leq N$ we wish to minimize

$$\sum_{j=0}^{m(i)} \rho_{i_j}\left[\sum_{n=0}^{j-1}\overline{d}_{i_n} + \frac{\overline{d}}{1-\rho}\sum_{n=0}^{j-1}\rho_{i_n}\right]$$

However, these expressions are invariant under any change between the order of child stations.
■

7 Numerical Examples

To obtain the expected waiting times at the different stations one must calculate the second moments of the station times. This involves solving $N(2N-1)$ equations given in (14) and (15). It is convenient to solve these equations by successive substitutions. We have used the *Gauss-Seidel iterative procedure* (see, e.g. [11],p.111 for formal description of the Gauss-Seidel procedure) for the solution of r_{ij} and a two dimensional array which contains $2N^2$ floating point numbers to store r_{ij} through successive iterations. In the *Gauss-Seidel iterative procedure* r_{ij} of the nth iteration is computed from the most recent r_{ij}'s available; these may include some r_{ij}'s which where lastly computed in the $n-1st$ iteration, and some that have already been computed in the nth iteration. There are another iterative schemes to calculate the second moments of the station times, for numerical comparisons of the iterative schemes in polling systems see [12].

In fact, there are approximately $4N^3$ multiplications and $4N^3$ additions per iteration. Humblet [9] reports that this procedure does indeed converge for a similar set of equations (for the standard gated discipline).

We write a program which carry out the calculations of the expected waiting times as computed in the paper. Its input is an *arbitrary* network topology with the parameters, $\overline{b}_{i_j}, b_{i_j}^{(2)}, \lambda_{i_j}, \overline{d}_{i_j}$ and $d_{i_j}^{(2)}$, for each station i_j in the network and its output is the expected waiting times in the different stations. These waiting times are then substituted in the left-hand side of the conservation law (38) and compared with the right-hand side of the conservation law.

In what follows, several numerical examples are demonstrated. In each of them, the expected waiting times in the different stations of the network are calculated and agree with the conservation law. Ten iterations were sufficient to obtain the second moments with accuracy of about $10^{-6}\%$. In all the networks plotted in the following figures, a small circle indicates a "child" station and a small box indicates a "father" station.

1. Example I The purpose of this example is to demonstrate the solution of a general polling system with several precedence constraints, and to compare its performance with the standard gated discipline and the "optimal topology" version (see section 6).

A schematic figure of network 1 appears in Fig. 1. It consists of 10 stations, 5 of them are "father" stations and the others are "child" stations. An arrow is drawn from a "father" station to its child. The network is symmetric with deterministic walking times. The following parameters are considered:

$$\overline{b}_{i_j} = 0.25, \quad b_{i_j}^{(2)} = 0.125, \quad \overline{d}_{i_j} = 1, \quad d_{i_j}^{(2)} = 1$$

The expected waiting times in the different stations of the network and the average waiting time in the network ($\overline{W}_{net1}$) are plotted in Fig. 2 as functions of the total system load, ρ. All "father" stations have approximately the same expected waiting time , and it is the smallest among the stations of the network. For "child" station, the expected waiting time increases as the distance (number of stations in the direction of the server movement) between its "father" and it increases.

Schematic figures of networks 2 and 3 appears in Fig. 1. They are symmetric networks with the same parameters as in network 1. The average waiting times in networks 1,2 and 3 ($\overline{W}_{net1}, \overline{W}_{net2}\ and\ \overline{W}_{net3}$), are plotted in Fig. 3. Network 2 is the "optimal topology" version of network 1, as described in section 6. The average waiting time in network 2 is lower than the average waiting time in network 1 as proved in section 6. Network 3 is the "gated" version of network 1, where all the stations are "father" stations. The average waiting time in network 3 is the smallest among the three networks, as noticed in section 6.

2. Example II In this example, we show the effect of the distance between a "child" station and its "father" station on the expected waiting times in the "father" stations and the "child" stations, and also on the average waiting time in the network. To show this, we use 3 networks as shown in Fig. 4 (the networks 4-6). All the networks are symmetric networks with the following parameters:

$$\overline{b}_{i_j} = 0.25, \quad b_{i_j}^{(2)} = 1, \quad \overline{d}_{i_j} = 0.01, \quad d_{i_j}^{(2)} = 0.001$$

In Fig. 5, the expected waiting times in a "father" and "child" stations and the average waiting times in network(i), $i = 4, 5, 6$, $(\overline{W}_{father}(neti), \overline{W}_{child}(neti) \ and \ \overline{W}_{neti})$, are plotted as functions of the total network load, ρ. The expected waiting time in a "father" station (it is the same for all "father" stations) *decreases* as the distance between "father" and "child" stations increases. For "child" station opposite behavior is observed. The average waiting time in the network remains approximately unchanged as the distance increases, this can be explained by the fact that $\overline{d}$ and $\overline{d}^{(2)}$ are small, and in that case the "pseudo" conservation law becomes conservation law. We use the following parameters to illustrate the behavior of the network for larger $\overline{d}$ and $\overline{d}^{(2)}$,

$$\overline{b}_{i_j} = 0.25, \quad b_{i_j}^{(2)} = 0.125, \quad \overline{d}_{i_j} = 1, \quad d_{i_j}^{(2)} = 1$$

The expected waiting times in a "father" and "child" stations are plotted in Fig. 6 as functions of the total network load ρ, for the networks 4,5 and 6. The expected waiting time for a "father" station is approximately the same for the three networks, and for a "child" station it increases as the distance increases. In Fig. 7 we plot the average waiting time in the network, for the networks 4,5 and 6. It can be seen from Fig. 7 that the average waiting time in the network increases noticeably as the distance increases.

3. Example III In this example we demonstrate the effect of the number of synchronized stations in the network on the expected waiting times in the "father" and "child" stations and also on the average waiting time in the network. As extreme cases of our analysis we have the following two service discipline: On the one extreme, there is the standard gated scheme (without synchronized stations at all) and on the other extreme there is the globally gated scheme, in which all the stations except one (the "father" station) are synchronized stations.

In [7], it was reported that for symmetric, globally gated network in which $d = 0$, stations positioned at the "first half circle" prefer globally gated regime, (in the sense that lower expected waiting times are achieved) while the others prefer the standard gated scheme. In order to investigate the behavior of polling systems with synchronization constraints for small values of d, we use 6 networks as demonstrated in Fig. 8 (the networks 7-12). The station numbers that appear in brackets indicate the station numbers used in the following figures (Fig. 9-12). In each of the networks there are 6 stations and the only station with any children is station 1. Network 7 is the standard gated version and network 12 is the globally gated version. The other networks are versions that lie between the globally gated and the standard gated schemes. For all the networks, we use the following parameters:

$$\overline{b}_{i_j} = 0.25, \quad b_{i_j}^{(2)} = 1, \quad d_{i_j}^{(2)} = \overline{d}_{i_j}^2, \quad \rho = 0.75$$

That is $var(d_{i_j}) = 0$.

In Fig. 9,10,11 and 12, the expected waiting times in stations 1,3,5 (the stations enclosed in brackets in Fig. 8) and in the network, respectively, are plotted as functions of the average total walking time in the network, $\overline{d}$. For each graph in Fig. 9-12, we print the number of the network in which the graph was obtained near the graph, (in brackets).

From Fig. 9, the expected waiting time in station 1 decreases when moving from network 7 to network 12, for all values of $\overline{d}$. That is, station 1 prefer more synchronized stations in the network. From Fig. 10, station 3, which is positioned in the "first half circle", prefers network 8 over the other networks for values of $\overline{d}$ greater than 0.18, and prefers the globally gated regime over the other networks for values of $\overline{d}$ smaller than 0.18, and prefers the globally gated regime over the gated scheme for values of $\overline{d}$ smaller than 0.32. From Fig. 11, station 5, which is positioned in the "second half circle", doesn't prefer the globally gated regime for any value of $\overline{d}$, and prefers network 10 (the last network before it becomes synchronized) for all values of $\overline{d}$. From Fig. 12, the average waiting time in the network *increases* when moving from network 7 to network 12, for all values of $\overline{d} > 0$, that is, the average waiting time in the network increases as the number of synchronized stations increases. Note that, for $d \to 0$, the "pseudo" conservation law becomes a conservation law and the average waiting time in the network doesn't depend on the service strategy, and it is the same for all the networks.

8 Concluding Remarks

We present in this Section some further applications and possible extensions of our results, and we compare the performance of our model with the corresponding standard gated discipline.

1. comparison of the performance. By the corresponding standard gated discipline model we mean the system which is obtained by using the same number and order of stations and the same statistics of arrivals, service times of customers and walking times, but replacing the synchronized gating by the standard gating discipline. We therefore keep the notation $i = 1, ..., N$ and $i_j,\ j = 1, ..., m(i)$ also when referring to the standard gated discipline without assuming any more that stations i_j are synchronized to other stations.

As we saw in Section 4.1 (equation 6), the expectation of the cycle duration remains the same. Another quantity that remains unchanged is $E[X_i^i]$:the expected number of jobs in a father station $i,\ i = 1, ..., N$ seen at the polling instants to that station. This can be seen in equation (26) (which holds also for $j = 0$). However, the synchronization mechanism causes an increase in these quantities in all synchronized stations. $E[X_{s(i_j)}^{i_j}]$ as computed in the synchronized regime turns to be equal to $E[X_{i_j}^{i_j}]$ in the corresponding standard gated regime. However, $E[X_{i_j}^{i_j}]$ as computed in the synchronized regime, is strictly larger for any $j = 1, ..., m(i)$ than the same quantity in the corresponding standard gated regime. The difference is equal to the right-hand side of equation (29). Finally, as we saw in Section 6 (equation 38), the weighted sum of the mean waiting times increases when adding synchronization; The difference is given in the last term of equation (38).

In [7], the individual waiting times for the symmetric gated scheme and for the globally gated scheme were compared. In the case of polling system with general synchronization constraints, the comparison is impossible, because no explicit expression exist for the expected waiting times, instead they are expressed as functions of the second moments which are given implicitly by a set of $N(2N-1)$ equations, and poses a different structure from the equivalent set of equations in the standard gated scheme.

We carried out the calculation of the waiting times distributions and moments through an extension of the "station time" approach. The alternative "buffer occupancy" approach (as can be found in [16]) i.e. the direct calculation of the distributions of the quantities X_k^l (where k and l stand for arbitrary stations), was not suitable for our synchronized model; the problem of using that method directly is that it becomes extremely complex to appropriately represent information needed from more than one cycle back in a way that yields recursive equations for calculating the moment generating functions.

2. The exhaustive synchronized discipline Assume the same partition between father and child stations as in the gated synchronized model, and assume that the father stations are served by an exhaustive discipline. We require that the precedence constraints are to hold between the father and its child stations, i.e. jobs in any child station can be processed only if all jobs that arrived previously to the father station have been served. In order to meet these precedence constraints we may use the following synchronization discipline: at the *end* of the service of every father station, gates are closed in all its child stations. When the server arrives to a child station it serves only the jobs that arrived there before the gate was closed. The same methodology that we developed in this paper can be applied for analyzing this new discipline.

3. Nested Precedence Constraints In parallel computations we often encounter precedence constraints which are much more complex than the ones we analyzed; these are described either by trees of arbitrary depth or by more general precedence graphs (allowing cyclic dependence between jobs) (see e.g. [2], [3] and references therein). We could consider a nested precedence order between stations for creating gates, e.g. when station 1 is polled it closes a gate in station 4, and when station 4 is polled it closes a gate in station 7. However, it can easily be seen that such a discipline fails to meet with the original nested precedence constraints. For example, a problem arises in station 7, where jobs that arrived after the polling instant to station 1 will be served before the jobs that arrived to station 4 after the polling instant to station 1. The same kind of problem arises when using variants of the exhaustive synchronized discipline. Therefore, in order to fulfill nested precedence constraints one should devise more complex service discipline.

4. Priority Polling Systems Consider a single server, cyclic system with N stations, and non-zero switch-over times between the stations. Customers from $m(i)+1$ different classes, $i_j, j = 0, 1, ..., m(i)$, arrive at each station $i, i = 1, 2, ..., N$ according to independent poison processes with intensity λ_{i_j} for class i_j jobs. Assume that service at each station i is given according to pre-assigned priorities, where the priority of class $i_j, j = 0, 1, ..., m(i)$ jobs increases as the index j decreases. Assume that the service strategy is the standard gated strategy, where at each polling epoch to station i a gate is closed at that station and jobs are serviced according to their priorities. Assume also that the server takes vacation each time he finished service of priority class at station i. The polling system with priorities and vacations between the service of priority classes, we have just described turns to be a special case of the synchronized polling system described in this paper, where we have N father stations and for each father station all of its children are situated between him and the following father station as described in the model description.

5. Reducing The Complexity Actually, the evolution of the system can be described in terms of $N + Dist$ station times in place of $2N - 1$ station times described in section 2 (equation 1). Where, $Dist$ equals to the largest distance (number of father stations in the direction of the server movement), between a father station and one of its child stations. Then, $0 \leq Dist \leq N - 1$. $Dist = 0$ when all fathers have their children situated between them and the subsequent father stations , and $Dist = N - 1$ when for some child station i_j, its father is the station $i + 1$. This fact may be used to reduce the complexity of the equations for the second moments. Note that for the case without synchronization, $Dist = 0$ and the state is described by N station times as in the standard gated discipline. We avoid the use of $Dist$ to simplify the notations.

9 Appendix: obtaining equation (21)

Consider the regeneration points as the time instants at which the father station $s(i_j)$ is visited and all stations are empty. A regeneration interval is the period between two consecutive regeneration points. Denote a cycle as the period between two consecutive polling instants to station $s(i_j)$. The nth cycle is denoted by C_n. Let $M :=$ the number of cycles in a regeneration interval.

Define the following random variables:

$\sigma^{i_j}(m, n)$ (resp. $\sigma^{s(i_j)}(m, n)$) is the service completion instant for the mth job served at station i_j ($s(i_j)$ respectively), in the nth cycle

$L_{i_j}(m, n)$ is the number of jobs that a job departing from station i_j sees at station i_j at time $\sigma^{i_j}(m, n)$.

$X_{i_j}^{i_j}(n)$ (resp. $X_{s(i_j)}^{i_j}(n)$) is the number of jobs in station i_j at the polling instant to that station, (to its father station $s(i_j)$ respectively) in the nth cycle.

When adding a hat to the above random variables (and to others) we shall mean that n refers to the nth cycle in a typical regeneration interval.

Due to the well known PASTA phenomenon (Poisson Arrivals See Time Averages, [14]) and the fact that only one customer may leave at a time, $Q_{i_j}(z)$ (defined in section 5.2) also stands for the the moment generating function of the number of jobs at station i_j in steady state regime at an *arbitrary time.*

Following Takagi ([16] pp. 77-79, 109) we have

$$E\left[Q_{i_j}(z)\right] = E\left[z^{L_{i_j}}\right] = \frac{E\sum_{n=1}^{M}\left[\sum_{m=1}^{\hat{X}^{i_j}_{s(i_j)}(n)} z^{\hat{L}_{i_j}(m,n)}\right]}{\lambda_{i_j} E\sum_{n=1}^{M}\hat{C}_n}$$

Where L_{i_j} was defined in section 5.2.

Since we assume that the system is stable, it follows that

$$\lim_{t\to\infty}\frac{1}{t}\sum_{n=1}^{t} C_n = EC \qquad w.p.1$$

On the other hand, from the theory of Renewal Reward Processes (e.g. [15] p. 279) we obtain

$$\lim_{t\to\infty}\frac{1}{t}\sum_{n=1}^{t} C_n = \frac{E\sum_{n=1}^{M}\hat{C}_n}{EM} \qquad w.p.1$$

Combining these equations we obtain

$$EC = \frac{E\sum_{n=1}^{M}\hat{C}_n}{EM}$$

(in some references such as in [16] Wald's Theorem is used improperly to obtain the same result).

By using similar arguments we obtain

$$E\left[\sum_{m=1}^{X^{i_j}_{s(i_j)}} z^{L_{i_j}(m)}\right] = \frac{E\sum_{n=1}^{M}\left[\sum_{m=1}^{\hat{X}^{i_j}_{s(i_j)}(n)} z^{\hat{L}_{i_j}(m,n)}\right]}{EM}$$

from which it follows that

$$E\left[Q_{i_j}(z)\right] = \frac{E\left[\sum_{m=1}^{X^{i_j}_{s(i_j)}} z^{L_{i_j}(m)}\right]}{\lambda_{i_j} EC}$$

The rest of the calculation follows as in [16].

References

[1] Baccelli F. and Z. Liu, "On the Execution of Parallel Programs on Multiprocessor Systems – A queueing Theory Approach", preprint.

[2] Baccelli F. and Z. Liu, "On the Stability Condition of a Precedence-Based Queueing Discipline", *Adv. Appl. Prob.* **21** pp. 883-898, 1989.

[3] Baccelli F. and A. Makowski, "Queueing Models for Systems with Synchronization Constraints", *Proceeding of the IEEE*, Vol. 77 No. 1, pp. 138-161, January 1989.

[4] Bayer N., *A Hardware-Synchronized/Scheduled Multiprocessor Model*, M.Sc. thesis, Technion – Israel Institute of Technology, January 1989.

[5] Boxma O. J. and Groenendijk, W. P., "Pseudo-conservation laws in cyclic-service systems", *Journal of Applied Prob.* **24**, 949-964, 1987.

[6] Boxma O. J., "Workloads and Waiting Times in Single-Server Systems with Multiple Customer Classes", Report BS-R8906, Center for Math. and Computer Sciences, Amsterdam, March 1989. To appear in *Queueing Systems*.

[7] Boxma O. J., H. Levy and U. Yechiali, "Cyclic Reservation Schemes for Efficient Operation of Multiple-Queue Single-Server Systems", submitted to Annals of OR.

[8] Ferguson M. J. and Y. J. Aminetzah, "Exact Results for Nonsymmetric Token Ring Systems", *IEEE Trans. on Communications* **33** No. 3, March 1985.

[9] Humblet P., "Source coding for communication concentrators", Electron. Syst. Lab., MIT, Cambridge, ESL-R-798, 1978.

[10] Kleinrock L., *Queueing Systems*, Vol. 1, John Wiley, 1975.

[11] Kushner H., *Introduction to Stochastic Control* Holt, Rinehart and Winston, New York, 1971.

[12] Levy H., "Delay Computation and Dynamic Behavior of Non-symmetric Polling Systems", *Performance Evaluation 10 (1989) 35-51*, North-Holland.

[13] Marsan A. M., G. Chiola and G. Conte, "Performance of Task Synchronization in Computer Systems", *Proc. 1st Int. Conf. on Computers and Applications*, pp. 9-16, June 1984.

[14] Ronald W. Wolff, "Poisson Arrivals See Time Averages", *Operrations Research*, Vol. 30, No. 2, March-April 1982.

[15] Ross M. S., *Introduction to Probability Models*, Academic Press, 3rd edition.

[16] Takagi H., *Analysis of Polling Systems*, The MIT Press, 1986.