Polling on a space with general arrival and service time distribution

Eitan ALTMAN, Serguei FOSS

INRIA B.P.93, 2004 Route des Lucioles
06902 Sophia-Antipolis Cedex, France

Abstract

We consider a polling system on a space with general independent arrival, service and walking times. We present necessary and sufficient conditions for stability, based on drift analysis of both the workload and the residual interarrival times within some embedded times called cycles.

Keywords: Polling systems, stability condition, general arrival, service and walking times.

1 Introduction

We consider a polling system on a space; a single server follows some random path that starts at some point e_0 and returns to that point after having scanned the whole space (in general, within one path, some parts of the space may be scanned several times). We consider some general independent arrival process of customers. The location of the arrivals are i.i.d. If the probability distribution of the location of arrivals have some mass at some points then queueing may occur there. The server can serve one customer at a time. In addition to the service time, the server also requires some random time (that may depend on its trajectory) to move from one customer to another.

Such systems have been widely used as models in wide areas of applications in telecommunication systems in general and for Local Area Networks in particular, computer systems, reader-head's movement in a computer's hard disk, flexible manufacturing systems, road traffic control, repair and maintenance. For overviews on polling systems, see [23, 24]. Growing attention was given in recent years to establish rigorously sufficient and necessary stability conditions for such systems.

Several different approaches were used to obtain ergodicity as a measure of stability, see [4, 6, 7, 12, 14, 21, 27] for the stability of "discrete" polling systems (consisting of

N queues). Kroese and Schmidt [16, 17], studied the stability of polling on a graph (i.e. "continuous" polling models), and Altman and Levy [5] studied the stability of non cyclic polling in two and higher dimensional spaces. In [4, 7] Altman et al. further present sufficient and necessary conditions for stronger notions of stability, namely the geometric ergodicity and the geometric rate of convergence of moments of the process of queues' length, (embedded at polling instants) and of other performance measures. All the above references assumed Poisson arrivals and general independent walking times and service times. Recently, Sherma [22] obtained ergodicity results for i.i.d. arrivals, using Cramer type lower and upper bounds on the inter-arrival distribution. These imply, in particular, the existence of all moments of the interarrival times. Moreover, sufficient conditions for Central Limit Theorems and the Law of Iterated Logarithm were obtained in [22] there (see also [7]).

The contribution of this paper is to consider polling with general independent arrival process, with only first moment conditions on the service, inter-arrival and walking times. This further relaxes previous results that we obtained in [1] that still required some second moment conditions.

We obtain sufficient (and necessary) conditions for stability, by which we mean the positive recurrence of some compact sets for the workload process. We show that if the rate ρ at which work arrives to the system (defined later) is smaller than one then the system is stable. Under some additional standard mixing conditions, this further implies ergodicity.

An alternative sample-path approach for studying the stability of polling systems is given by Altman et al. [2, 3]. Different type of conditions are imposed on a single sample path, which yields several notions of stability under suitable assumptions.

The structure of the paper is as follows. In Section 2 we present the model. We define a "cycle", and present a drift analysis of the behavior of the workload and the residual interarrival times (Section 3) within a single cycle. Results on positive recurrence and on ergodicity are finally presented in Section 4.

2 model

Let $(\mathcal{E}, d)$ be a Polish (complete separable) space with metric d. Assume $\mathcal{E}$ to be compact and that there exists a continuous map from $[0,1]$ onto $(\mathcal{E}, d)$ and a continuous map from $(\mathcal{E}, d)$ onto $[0,1]$ (here $[0,1]$ is considered as a space with the natural Borel topology). Customers arrive to points ("nodes", "stations") that are located in $(\mathcal{E}, d)$. There is a single server that is polling points from this space. Below we give a precise description for a routing (polling) mechanism of a server, for arrival (input) stream and for service policies the server uses.

Routing mechanism: Consider a set $\mathcal{G}'$ of continuous functions $g : \mathcal{R}_+ \to \mathcal{E}$ such that

- (i) $g(0) = e_0$ (where e_0 is a fixed element from $\mathcal{E}$);
- (ii) for any $g \in \mathcal{G}'$, there exists a finite positive number t_g such that
  - (ii,a) $g(t) = g(0) = e_0$ for all $t \geq t_g$;
  - (ii,b) the set $\{g(t); \quad 0 \leq t \leq t_g\}$ coinsides with $\mathcal{E}$.

Introduce a distance (metric) $\hat{d}$ on the functional space $\mathcal{G}'$: $\hat{d}(g_1, g_2) = \sup_{t \geq 0} d(g_1(t), g_2(t))$. It is not difficult to see that $(\mathcal{G}', \hat{d})$ is a Polish space, too. Introduce a subset $\mathcal{G} \subset \mathcal{G}'$ of all functions satisfying

- (ii,c) for any $0 \leq t < t_g$, $\inf\{u > t : \quad g(u) \neq g(t)\} = t$.

Let now $(\Omega, \mathcal{F}, \mathbf{P})$ be a probability space and $\hat{\Phi} : (\Omega, \mathcal{F}, \mathbf{P}) \to (\mathcal{G}', \hat{d})$ be a random element, $\hat{\Phi} \equiv \{\hat{\Phi}(t); \quad t \geq 0\} \equiv \{\hat{\Phi}(t, \omega); \quad t \geq 0\}$. Assume in addition that $\mathbf{P}(\hat{\Phi} \in \mathcal{G}) = 1$. Put $D = t_{\hat{\Phi}}; \quad \Phi = \{\hat{\Phi}(t); \quad 0 \leq t \leq D\}$. It is clear that D is a real-valued RV (random variable) and that $\mathbf{P}(D > 0) = 1$.

Consider some examples.

Example 1. Elevator-type routing, $\mathcal{E} = [0, 1]$. The distribution $\mathbf{P}$ is concentrated on the subspace of all functions g such that $g(0) == e_0 = 0$; g is increasing on some interval $(0, u_g)$, where $g(u_g) = 1$; is decreasing on (u_g, t_g) and $g(t) = 0$ for $t \geq t_g$.

Example 2. Reflected Brownian motion, $\mathcal{E} = [0, 1]$. There are given some fixed drift δ and variance v; the distribution $\mathbf{P}$ is concentrated on the subspace of all functions g such that $g(0) = 0$; $g(t), t \in [0, t_g)$ is a realisation of a Brownian motion with drift δ and variance v, reflected at 1. t_g may be taken as the first instant for which $g([0, t_g)) = [0, 1]$, and $g(t_g) = 0$.

Example 3. $\mathcal{E}$ is a unit circle parametrize by the angle $[0, 2\pi]$. We may now consider the movement of the server (as a function of the angle) as in the examples above. (For the case of Brownian motion, the reflection at 2π may be omitted.) We may also consider the simple movement in one direction: the distribution $\mathbf{P}$ is concentrated on the subspace of all functions g such that $g(0) = 0$ with g is non-decreasing on some interval $[0, t_g)$. (Note that points 0 and 2π are the same, so condition (ii,a) holds.)

Example 4. $\mathcal{E}$ is a connected directed graph, e_0 is one of its nodes, $\mathcal{G}$ is a family of functions which scan all the points on all links (arcs). Again, we may now consider the movement of the server as a function of a parameter $\theta \in [0, 1]$ as in the Examples above.

Introduce now a sequence $\{(\Phi_n, D_n)\}_{n\geq 1}$ of independent copies of (Φ, D). We consider a polling system on $(\mathcal{E}, d)$, with a single server that starts (at time instant $t=0$) from point e_0 and follows paths $\Phi_1, \Phi_2, \ldots$ consequently. The server spends some time traveling and some time serving customers that she finds on her way; if, during time $t > 0$, she spends time $a(t)$ traveling and time $b(t)$ serving $(a(t)+b(t)=t)$, and if $a(t) = \sum_1^{n-1} D_i + c(t)$ for some $n \geq 0$; $0 < c(t) \leq D_n$ (where $\sum_1^0 := 0$), then at moment t she is at point $\Phi_n(c(t)) \in \mathcal{E}$.

We call a "cycle" n the period of time when the server follows path Φ_n, $n \geq 1$. Cycle 1 starts at $t=0$. (ii, b) means that during one cycle the server scans (a.s.) all the points from $\mathcal{E}$ and (ii, c) - that she does not stay at a point unless she serves a customer there.

Input process: Customers arrive to the system to arbitrary locations on the space. Arriving customers are routed to some location according to a routing sequence. If the distribution of this location has mass at some points, then queueing phenomena may occur there. The first customer arrives at moment $\chi(0) > 0$, has a service time σ^1 and is located at point α^1. For any $n \geq 1$, the customer $(n+1)$ arrives at time instant $\chi(0) + \sum_1^n \tau^i$, has a service time σ^{n+1} and a location α^{n+1}.

The processes of inter-arrival times $\{\tau^n\}$ and of service times and of locations of customers $\{(\sigma^n, \alpha^n)\}$ are defined on the same probability space $(\Omega, \mathcal{F}, \mathbf{P})$ and are assumed to be independent (here $\alpha^n : (\Omega, \mathcal{F}, \mathbf{P}) \to (\mathcal{E}, d)$). In addition, they are jointly independent of the routing mechanism (Φ_n, D_n) and each of them forms an i.i.d. sequence.

Remark 2.1. The results of this paper may be formulated for a model with a more general input process of a regenerative type, and the scheme of the proofs is the same since the assumptions still allow for a Markovian analysis. More precisely, consider some regenerative times $\overline{T}_1, \overline{T}_2, \ldots$. Define $\tau^i = \overline{T}_{i+1} - \overline{T}_i$. The location of arrivals and the workload they require in each regenerative period is given by a measure M defined on $\mathcal{E} \times [0, \tau^i)$. For any Borel sets $B_1 \subset \mathcal{E}$, $B_2 \subset \mathbb{R}$, $M(B_1 \times B_2)$ is the amount of workload that arrives to points in B_2 at time within B_1. This construction includes inputs with random batch arrivals, and under fairly general conditions, it allows to describe K independent i.i.d. arrival streams into K nodes [13].

Description of the global state space for the system: There is given an initial "state" $X(0)$ of the system as a 5-dimensional "vector"

$$X(0) = \{S(0), \Psi(0), \vec{V}(0), \chi(0), (D(0), \underline{\Phi}^0)\} \quad \text{where}$$

- (i) $S(0) = e_0$;
- (ii) $\Psi(0)$ is an initial "configuration" of work, i.e. an initial set of points (which is a subset of $\mathcal{E}$) at which customers are present, and we assume $\Psi(0)$ to contain (a.s.) a

finite number of points. Put $\Psi(0) = \{y_i\}_{i=1}^{N(0)}$, where $N(0) = ||\Psi(0)||$ is the number of points in $\Psi(0)$. Then,

- (iii) $\vec{V}(0) := \{\vec{V}_{y_i}(0)\}_{i=1}^{N(0)}$ is the workload vector, where, for any i, $\vec{V}_{y_i}(0)$ is a sequence $(n_{y_i}(0); x_{1,y_i}(0), \dots, x_{n_{y_i},y_i}(0))$ and $n_{y_i}(0)$ is a number of customers at point y_i which are ordered, and x_{j,y_i} are their service times. Further,

- (iv) $\chi(0)$ was defined before as a first arrival epoch;

- (v) $D(0) = D_1$, $\underline{\Phi}^0 = \Phi_1$.

Similarly, for any $t > 0$, the current state $X(t)$ is a 5-dimensional "vector"

$$\{S(t), \Psi(t), \vec{V}(t), \chi(t), (D(t), \underline{\Phi}^t)\},$$

where

- (i)-(iii) the first 3 coordinates have a natural meaning;

- (iv) $\chi(t)$ is a residual interarrival time at time instant t;

- (v) $D(t)$ is the residual length of a current "walking cycle", and $\underline{\Phi}^t$ is its residual path which can be defined by the following procedure: assume the server to finish her (n-1)st walking cycles up to time t (where $n = n(t) \geq 1$) and denote $\mathcal{T}_n$ the moment when she finished the last one of them (and started the nth cycle), and by $\xi = \xi(t)$ the time she spends with walking on time interval $(\mathcal{T}_n, t)$ (note that, by definition, $\xi < D_n$). Then $\underline{\Phi}^t = \{\underline{\Phi}^t(u); 0 \leq u \leq D_n - \xi\}$ where $\underline{\Phi}^t(u) = \Phi_n(u + \xi)$.

We assume the process $X(t)$ to be right-continuous, so, "at time t" means "at time $t+$". Denote by $\mathcal{F}_t := \sigma\{X(s); s \leq t\}$, $t \in \mathcal{R}_+$ a right-continuous filtration describing the internal history of the process.

Finally, we define the 3-dimensional "vector" $Y(t) = \{\Psi(t), \vec{V}(t), \chi(t)\}$, and the right-continuous filtration $\mathcal{F}_t^Y := \sigma\{Y(s); s \leq t\}$, $t \in \mathcal{R}_+$.

Service policies: We assume that at any point e where the queueing phenomena does not occur (i.e. $\mathbf{P}(\alpha^1 = e) = 0$) we have to serve a customer if there is any.

The number of customers that are served at point $e \in \mathcal{E}$, $\mathbf{P}(\alpha^1 = e) > 0$, at each visit is determined by some polling discipline (such as the gated, exhaustive etc.) which may be random. For the sake of simplicity of explanation, we assume this policy to depend on the current state only (the natural generalizations when the policy depends on the history lead to the same type of results but they involve technical difficulties; the ideas of the proof do not change. In particular, one may easily consider policies that are function of the configuration in the beginning of the cycle).

Let $\tilde{\mathcal{E}}$ be a set of points $e \in \mathcal{E}$ such that $\mathbf{P}(\alpha^1 = e) > 0$; $\mathcal{V}$ be a state space for vectors $\vec{V}(t)$ (equiped by a natural topology) and $\mathcal{N}$ be a set of positive integers. Note that $\tilde{\mathcal{E}}$ has to be countable (finite or infinite).

Let $h : \tilde{\mathcal{E}} \times \mathcal{V} \to \mathcal{N} \bigcup \{\infty\}$ be a measurable function (we call it a "decision function"), such that $h_e(\vec{v}) \geq 1$ if $\vec{v}_e \neq \emptyset$. The "decision" $\{h_e(v) = k\}$ means that if we arrive to point e and $\vec{v}$ is such that $\vec{v}_e = (n; x_1, \ldots, x_n)$, $n \geq 1$ (at point e, we find n customers with service times $x_1, \ldots, x_n$) then we have to serve (in FIFO order) customers till the moment when first either we serve k customers or the queue becomes empty. In general, k may be greater than n - it means that we can serve new arriving customers, too; in particular, $\{h_e(\vec{v}) = \infty\}$ means "exhaustive" service - the service till the moment when the queue becomes empty.

Example 5: Time-limited policy. There is given a fixed $T^* > 0$. Put $h_e(v) = n$, if $\sum_1^n x_i \leq T^*$, and $h_e(v) = \min\{k \geq 1 : \sum_1^k x_i \geq T^*\}$.

Remark 2.2. When the server has a "Brownian motion" type of movement, then a station is visited infinitely many times during a cycle, and if we leave a non-empty "node" ("station") we typically return to it immediately after serving it. This implies in practice that policies behave like exhaustive ones. If we wish to consider other type of policies, we should either enlarge the state space (e.g. in order to allow a station to be served only once in a cycle), or allow the service policies to depend on the past history too. In that case, similar results can be obtained by appropriately reformulating the model.

Remark 2.3. The assumptions made above imply that $X(t)$ is a Markov process. In particular, it means the following. Assume $\mathcal{F}_0$ to be given. Fix any k and $(\tau^i, \alpha^i, \sigma^i)$ for $1 \leq i \leq k$. Then the behavior of the underlined process $X(t)$ becomes completely deterministic on time interval $[0, \chi(0) + \sum_1^k \tau^i)$. Therefore, if we denote by T the duration of the first cycle, then the function

$$I\{T \in [\chi(0) + \sum_1^{k-1} \tau^i, \chi(0) + \sum_1^k \tau^i)\}$$

is a deterministic function of the parameters $X(0)$ and $\tau^1, \ldots, \tau^k$ and (α^i, σ^i) for $1 \leq i \leq k$. The latter means that if we denote by η_1 the number of exogenous arrivals in the first cycle, then η_1 is a "stopping time" in the sense that, for any k, for any event $A \in \mathcal{F}_0$, for any event B from the σ-algebra generated by $\{\tau^i, \alpha^i, \sigma^i\}_{i=1}^k$, the event $\{\eta_1 = k\} \cap A \cap B$ does not depend on the σ-algebra generated by $\{\tau^i, \alpha^i, \sigma^i\}_{i>k}$. The same "stopping time" properties are true for any $n \geq 1$ - for the moment $\mathcal{T}_n$ (the time at which cycle n begins), and for the total number $\eta_1 + ... + \eta_{n-1}$ of arrivals on $(0, \mathcal{T}_n)$. Similarly, denote by $\mathcal{F}_{\mathcal{T}_n}$

the σ-algebra generated by events $\{\mathcal{T}_n \le t\} \cap A$, $A \in \mathcal{F}$. Then $\mathcal{T}_{n+1}$ and η_n have similar "stopping time" properties w.r. to $\mathcal{F}_{\mathcal{T}_n}$.

3 Drift analysis in a single cycle

We first focus on drift analysis within a single (first) cycle that starts at $t = 0$. We are interested to upper bound in some sense the difference between the work present at the beginning of the cycle and the end of the cycle. Introduce the following notation:

- $V_1 \equiv V(0) \equiv ||\vec{V}(0)|| = \sum_{i=1}^{N} \sum_{j=1}^{n_i} x_{j,i}$, where $N := N(0)$,$n_i := n_{y_i}(0)$, $x_{j,i} := x_{j,y_i}(0)$.
- $\chi := \chi(0)$
- We assume that $0 < E\tau^i < \infty$, $0 \le \mathbf{E}\sigma^i < \infty$ and put $\rho := E\sigma^1/E\sigma^1$.
- $D := D(0) \equiv D_1$ We assume $0 < E[D] < \infty$;
- $\eta := \eta_1$
- $\mathcal{F}^l$:= the σ-algebra generated by $\{\{\tau^i\}_{i>l}, \{\sigma^i\}_{i>l}, \{\alpha^i\}_{i>l}\}$. As we mentioned before, for all $l = 0, 1, 2, ...,$
$$\{\eta \le l\} \text{ is independent of } \mathcal{F}^l. \qquad (1)$$
- G:= past interarrival time, i.e. the time between the last arrival in the cycle and the end of the cycle, in case that at least one arrival occurs during the cycle, otherwise, it denotes the whole cycle time. It is a nonnegative RV satisfying
(1) $\forall l = 0, 1, ...$ and for all sets $A \in \mathcal{F}$, the event $\{G \in A\} \cap \{\eta = l\}$ is independent of $\mathcal{F}^l$, (2) $G \le \tau^\eta$ a.s. on the event $\{\eta \ge 1\}$.
- $\tau^0 := \chi$, (so that $G \le \tau^\eta$ a.s.). Thus, $T = G + \sum_{i=0}^{\eta-1} \tau^i$ (where, $\sum_{i=0}^{-1} \tau^i := 0$).
- $\chi' := \tau^\eta - G$ which represents the residual interarrival time $\chi(T)$ at time T (i.e. the time between T and the next arrival).
- $\Gamma = \sum_{j=1}^{\eta} \sigma^j$ is the total work arriving during the cycle.
- V^{out}:= the amount of work leaving the system during the cycle. It satisfies $V^{out} + D = T$.
- V_2:= the amount of work at time T. Thus $V_2 = V_1 + \Gamma - V^{out}$.

Lemma 3.1 *Assume $E[\eta|\mathcal{F}_0] < \infty$ a.s. The following holds a.s.:*

$$E\left(\sum_{i=1}^{\eta} \tau^i|\mathcal{F}_0\right) = E(\eta|\mathcal{F}_0)E\sigma^1$$

$$E(\Gamma|\mathcal{F}_0) = E(\eta|\mathcal{F}_0)E\sigma^1 = \rho E(T + \chi' - \chi|\mathcal{F}_0) \tag{2}$$

Proof: Follows from Wald's identity, (1), and since $T + \chi' - \chi = \sum_{i=1}^{\eta} \tau^i$. ■

Lemma 3.2 *Assume $\rho < 1$. Then we have a.s.*

$$E[\eta|\mathcal{F}_0] < \infty \tag{3}$$

$$E(T|\mathcal{F}_0) < \infty. \tag{4}$$

Proof: Consider the following polling policy $\overline{u}$. The server waits till the first arrival occurs in the cycle, then spends D walking time. It then serves all the the customers in the system, as well as all new arriving customers, until the system empties, incurring zero walking time. Denote by $\eta^{\overline{u}}$ the number of arrivals till the service of the system is completed. Clearly, the order in which customers are served does not influence the time it takes till the system empties, nor $\eta^{\overline{u}}$. So, without loss of generality, we shall assume that customers are served according to the FIFO discipline. It then easily follows that

$$\eta \leq \eta^{\overline{u}}, \qquad a.s. \tag{5}$$

Let $S_i = \sum_{k=1}^{i} \tau^k - \sigma^k$. Then $\eta^{\overline{u}} = \min\{k : S_k > V_1 + D\}$. It follows from renewal arguments that if $\rho < 1$ then $E[\eta^{\overline{u}}|\mathcal{F}_0]$ is finite a.s., which implies (3). (4) then follows from Lemma 3.1. ■

Lemma 3.3 *Assume $\rho < 1$. Then (i)*

$$E(V_2|\mathcal{F}_0) - V_1 = \rho D + (\rho - 1)E(V^{out}|\mathcal{F}_0) + \rho E(\chi' - \chi|\mathcal{F}_0) \qquad a.s. \tag{6}$$

(ii) For any $\epsilon > 0$, there exists some constant L such that a.s.

$$E[\chi' - \chi|\mathcal{F}_0] \leq -D + \epsilon E[V^{out}|\mathcal{F}_0] + (1+\epsilon)(D+L)1\{\chi \leq (1+\epsilon)(D+L)/\epsilon\}.$$

Proof: (i) follows since $V_2 - V_1 = \Gamma - V^{out}$ and $T = V^{out} + D$, and by (2). Next we establish (ii). We first note that if χ is large then we have some negative drift in the following sense:

$$E[\chi' - \chi|\mathcal{F}_0] = -(D + V^{out}) \tag{7}$$

on the set $\{V^{out}+D<\chi\}$. Note that the event $\{V^{out}+D<\chi\}$, and the RV $V^{out}1\{V^{out}+D<\chi\}$ are measurable on $\mathcal{F}_0$. To illustrate that, consider for example the exhaustive discipline ($h\equiv\infty$). Then $\{V^{out}+D<\chi\}=\{V_1+D<\chi\}$ and $V^{out}1\{V^{out}+D<\chi\}=V_1\{V_1+D<\chi\}$.

Next, assume that $V^{out}+D\geq\chi$. We make use of the fact that the expected residual interarrival time $\chi(t)$ at time t is of order $o(t)$ if at time zero, the residual time is zero [9, 10], i.e. for any $a>0$ there exist some constant b such that $\mathbf{E}\chi(t)$ is smaller than $at+b$ at any time t. Choose some arbitrary constant a and the corresponding constant b. Choose an arbitrary constant L. Define

$$\zeta(t):=\min\left\{k:\sum_{j=1}^{k}\tau^j>t\right\},\qquad A(t):=\sum_{j=1}^{\zeta(t)}\tau^j.$$

Note that (see Remark 2.3) we consider only polling policies that do not anticipate the future, i.e. such that $\mathcal{F}_t$ is independent of $\mathcal{F}^{\zeta(t-\chi)}$ for all t. Let $z_i=A(iL)-iL$ be the residual interarrival time at time $T+\chi+iL$, $i=0,1,\ldots$. Then $z_0=0$ and

$$E[z_{i+1}|\mathcal{F}_{\chi+iL}]\leq(z_i-L)I\{z_i>L\}+(aL+b)I\{z_i\leq L\}\leq z_i+b+aL.$$

(Note that z_i is measurable with respect to $\mathcal{F}_{\chi+iL}$ by the definition of the state $X(t)$ and the definition of $\mathcal{F}$.) Denote $Q:=\min\{i:\chi+iL>T\}=\min\{i:\chi+iL>V^{out}+D\}$. (Note that $Q\geq1$ a.s. on event $\{V^{out}+D\geq\chi\}$). Since we assume that the polling policies do not anticipate the future, Q is a stopping time with respect to $\mathcal{F}_{\chi+iL}$, $i=1,2\ldots$, i.e. $\{Q\leq i\}\in\mathcal{F}_{\chi+iL}$. It follows that $Z_i:=z_i-i(b+aL)$ is a $\mathcal{F}_{\chi+iL}$ supermartingale. Moreover, we observe that

- $T\geq\chi+(Q-1)L$, so $E[Q|\mathcal{F}_\chi]$ is a.s. finite, and $P(Q<\infty|\mathcal{F}_\chi)=1$;
- Let $Q^{\overline{u}}$ be defined as Q, when using policy $\overline{u}$. Then if $i\leq Q$, then

  $$Z_i\geq-i(b+aL)\geq-Q(b+aL)\geq-Q^{\overline{u}}(b+aL).$$

  Hence $Z_{\min(i,Q)}\geq-Q^{\overline{u}}(aL+b)$ for all integers i, $E[Q^{\overline{u}}|\mathcal{F}_\chi]<\infty$ (this again follows from Lemma 3.2), and $Q^{\overline{u}}\geq0$.

Thus, we may apply the sampling Theorem for supermartingale ([15] p. 266), to conclude that $E[Z_Q|\mathcal{F}_\chi]\leq Z_0$. Hence,

$$\begin{aligned}E[z_Q|\mathcal{F}_\chi]&\leq(aL+b)E[Q|\mathcal{F}_\chi]\leq(aL+b)E[(T-\chi)/L+1|\mathcal{F}_\chi]\\&=\left(a+\frac{b}{L}\right)(E[V^{out}|\mathcal{F}_\chi]+D+L-\chi)\end{aligned}$$

We note that on the set $V^{out} + D \geq \chi$ we have $\chi' \leq z_Q + L$, since, if $\chi' > L$ then $\chi' = LQ - T + z_Q \leq L + z_Q$. Consequently, on the set $V^{out} + D \geq \chi$ we have

$$E[\chi'|\mathcal{F}_\chi] \leq \left(a + \frac{b}{L}\right)(E[V^{out}|\mathcal{F}_0] + D + L - \chi) + L. \tag{8}$$

Choose any $a < \epsilon$, and L such that $a + b/L < \epsilon$. By combining (7) and (8), we obtain

$$\begin{aligned}
E[\chi' - \chi|\mathcal{F}_0] &\leq \epsilon E[V^{out}|\mathcal{F}_0] - E\left[1\left\{\chi > V^{out} + D\right\}\left(V^{out} + D\right)\Big|\mathcal{F}_0\right] \\
&\quad + \epsilon E\left[1\left\{\chi \leq V^{out} + D\right\}\left(-\chi + D + L + \frac{L}{\epsilon}\right)\Big|\mathcal{F}_0\right] \\
&= \epsilon E[V^{out}|\mathcal{F}_0] - E\left[1\left\{\chi > V^{out} + D\right\}\left(V^{out} + D + \epsilon(D + L) + L\right)\Big|\mathcal{F}_0\right] \\
&\quad - \epsilon E\left[1\left\{\chi \leq V^{out} + D\right\}\chi\Big|\mathcal{F}_0\right] + \epsilon(D + L) + L \\
&\leq \epsilon E[V^{out}|\mathcal{F}_0] - (1 + \epsilon)(D + L)P\left(\chi > V^{out} + D\Big|\mathcal{F}_0\right) \\
&\quad - (D + L)(1 + \epsilon)P\left(\epsilon\chi > (D + L)(1 + \epsilon), \chi \leq V^{out} + D\Big|\mathcal{F}_0\right) + (1 + \epsilon)(D + L) - D \\
&= \epsilon E[V^{out}|\mathcal{F}_0] + (1 + \epsilon)(D + L)P\left\{\chi \leq (D + L)(1 + \epsilon)/\epsilon, \chi \leq V^{out} + D\Big|\mathcal{F}_0\right\} - D \\
&\leq \epsilon E[V^{out}|\mathcal{F}_0] + (1 + \epsilon)(D + L)I\left\{\chi \leq (D + L)(1 + \epsilon)/\epsilon\right\} - D
\end{aligned}$$

which establishes (ii). ■

Corollary 3.4 *Assume $\rho < 1$. For any $\epsilon > 0$, there exist $M, C_2, \epsilon_2 > 0$ s.t.*

$$E[(V_2 + \chi') - (V_1 + \chi)|\mathcal{F}_0^Y] \leq -\epsilon_2 + (\rho + (1 + \rho)\epsilon - 1)E[V^{out}|\mathcal{F}_0^Y] + C_2 I\{\chi \leq M\},\ P - a.s.$$

Proof: Combining (i) and (ii) from Lemma 3.3, we obtain for any M

$$\begin{aligned}
&E[(V_2 + \chi') - (V_1 + \chi)|\mathcal{F}_0] \\
&\quad \leq -D + (\rho + (1 + \rho)\epsilon - 1)E[V^{out}|\mathcal{F}_0] \\
&\qquad + (1 + \rho)(1 + \epsilon)(D + L)I\{\chi \leq (1 + \epsilon)(D + L)/\epsilon\} \\
&\quad \leq -D + (\rho + (1 + \rho)\epsilon - 1)E[V^{out}|\mathcal{F}_0] + (1 + \rho)(1 + \epsilon)(D + L)I\{\chi \leq M\} \\
&\qquad + (1 + \rho)(1 + \epsilon)(D + L)I\{M < (1 + \epsilon)(D + L)/\epsilon\}.
\end{aligned}$$

Choose some M such that

$$0 < ED - (1 + \rho)(1 + \epsilon)E\left[(D + L)I\left\{\epsilon M < (1 + \epsilon)(D + L)\right\}\right] =: \epsilon_2.$$

Then, the Corollary follows by setting $C_2 := (1 + \rho)(1 + \epsilon)(ED + L)$. ■

4 Positive recurrence.

In this section we consider quantities that correspond to the sequence of consecutive cycles. We add the $i = 1, 2, ...$ as subscripts of quantities defined in Section 3 in order to denote a quantity at the beginning of cycle i. We consider the discrete-time processes $X_i = X(\mathcal{T}_i)$ and $Y_i = Y(\mathcal{T}_i)$, which are both time-homogeneous Markov chains.

Denote $U_1 = \{\tau^j, \sigma^j, \alpha^j\}_{j=1}^{\infty}$. Let θ^n be the shift operator, such that $(\tau^{n+j}, \sigma^{n+j}, \alpha^{n+j}) = (\tau^j, \sigma^j, \alpha^j) \circ \theta^n$ a.s. Put $\varphi_i = \eta(\mathcal{T}_i) \equiv \eta_1 + \ldots + \eta_{i-1}$ and define recursively

$$U_{i+1} = U_i \circ \theta^{\varphi_i} \equiv \{\tau_{i+1}^j, \sigma_{i+1}^j, \alpha_{i+1}^j\}_{j=1}^{\infty}. \tag{9}$$

Remark 2.3 imply that $\mathcal{F}_{\mathcal{T}_i}$ and $\mathcal{F}^{\varphi_i}$ are independent; in particular, for each i, $U_i =_d U_1$.

Theorem 4.1 *Assume that $\rho < 1$, and that there exists a nonnegative function H, such that $H(V) \to \infty$ as $V_1 \to \infty$, and $E\{V_i^{out}|X_i\} \geq H(V_i)$ a.s. for all i; Then one can choose constants $\tilde{L}$ and $\tilde{M}$, such that*
(i) for each initial condition X_1, $\tilde{\lambda} \equiv \min\{i \geq 1 : V_i \leq \tilde{L}; \chi_i \leq \tilde{M}\} < \infty$ a.s.;
(ii) $\sup E\{\tilde{\lambda}|X_1\} < \infty$, where we take a supremum on all non-random initial conditions X_1, such that $V_2 \leq \tilde{L}$ and $\chi_1 \leq \tilde{M}$.
(iii) For arbitrary initial condition X_1, introduce a sequence $\{\psi_n\}$, $\psi_{n+1} > \psi_n$ a.s., where

$$\begin{aligned} \psi_1 &= \min\{i \geq 1 : V_i \leq \tilde{L}; \chi_i \leq \tilde{M}\} \\ \psi_{n+1} &= \min\{i \geq \psi_n + 1 : V_i \leq \tilde{L}; \chi_i \leq \tilde{M}\}, \qquad n \geq 1. \end{aligned}$$

Then there exists a constant $u' < \infty$, such that $E\{\psi_{n+1} - \psi_n|\mathcal{F}_{\psi_n}\} \leq u'$ a.s. for all n.

Remark 4.1 *The statement of Theorem 4.1 means, that the set $B \equiv [0, \tilde{L}] \times [0, \tilde{M}]$ is uniformly positive recurrent for the Markov chains $\{Y_i\}$ and $\{X_i\}$.*

Proof of Theorem 4.1: The proof follows from Corollary 3.4 by choosing ϵ such that $\rho + \epsilon(1 + \rho) < 1$, and by applying Tweedie's criterion [25] Section 6, using the Lyapunov function $f(X_i) = V_i + \chi_i$. ∎

Remark 4.2 *If $\rho > 1$, then $V_i \to \infty$ a.s. as $i \to \infty$. Indeed, this can easily be established by noting that the amount of work in the system can be lower-bounded by considering the corresponding polling system with zero walking times and an exhaustive policy. The amount of work in the latter is the same as in the standard GI/GI/1 queue, for which indeed the amount of work in the system is known to grow to infinity if $\rho > 1$ a.s.*

5 Concluding remarks

We presented conditions for positive recurrents of some compact sets. Combined with some extra standard well-known mixing conditions, these further yield ergodicity, existence and convergence of moments of the workload in the system, and coupling convergence (see [8, 19, 20, 26], and others). For instance, assume that

- (c1) $\mathbf{P}(\tau^1 > x) > 0$ for all x;
- (c2) there exist a constant $0 < p \leq 1$ and a probability measure $\mathbf{Q}$ on $\mathbf{R}_+$ such that $\mathbf{P}(\tau^1 - x \in \cdot | \tau^1 > x) \geq p \cdot \mathbf{Q}(\cdot)$ for all x;
- (c3) the function h (see "service policies") is "super-gated": for any $e \in \mathcal{E}$, $\vec{v} \in \mathcal{V}$, if $\vec{v}_e = (n; x_1, \ldots, x_n)$, then $h_e(\vec{v}) \geq n$.

Then a standard argument shows that there exists a constant $0 < p^{'} \leq 1$ such that

$$\mathbf{P}(N(\mathcal{T}_2) = 0, \chi(\mathcal{T}_2) \in \cdot | \mathcal{F}_0) \geq p^{'} \cdot \mathbf{Q}(\cdot)$$

a.s. on the event $\{(V_1, \chi_1) \in [0, \tilde{L}] \times [0, \tilde{M}]\}$. This implies the Harris ergodicity of X_n and Y_n. If, in addition, D has a non-lattice distribution, then the continuous-time process $\{X(t); t \geq 0\}$ is stable, i.e. these exists a probability distribution $\mathcal{P}$ such that $\mathbf{P}(X(t) \in \cdot) \to \mathcal{P}(\cdot)$ weakly as $t \to \infty$ [8]. Note that condition (c2) may be replaced by a similar condition on the distribution of D. Note that if we replace the inequality $h_e(\vec{v}) \geq n$ in (c3) by the more general condition $h_e(\vec{v}) \geq k(\vec{v}_e)$ where $k(\vec{v}_e) := \min\{l : \sum_1^l x_i \geq c \cdot \sum_1^n x_i\}$ and $0 < c \leq 1$ is a fixed constant, then one can choose an integer $N \geq 2$ such that the latter inequality is still true but for T_n instead of T; that still gives us the ergodic property of X_n and of Y_n. In [1] we present two more general examples of sufficient conditions, thus establishing the ergodicity and results on moments, based on the recurrence conditions which we obtained above.

References

[1] E. Altman and S. Foss, "Polling on a graph with general arrival and service time", INRIA report No. 1992.

[2] E. Altman, S. Foss, E. R. Riehl and S. Stidham, " Pathwise Stability and Performance Bounds for Generalized Vacation and Polling Systems", Research report UNC/OR TR93-8, Univ. of North Carolina, Chapel Hill, USA, 1993.

[3] E. Altman and D. Kofman, "Sample Path Bounds for Performance Measures of Token Rings", Research Report Telecom Paris No. 93C004, 1993.

[4] E. Altman, P. Konstantopoulos, Z. Liu, "Stability, Monotonicity and Invariant Quantities in General Polling Systems", *Queuing Systems, special issue on Polling Models*, Eds. Takagi and Boxma, 35-57, 1992.

[5] E. Altman and H. Levy, "Queueing in space", *Adv. Applied Probability* **26**, 1095-1116, 1994.

[6] E. Altman and Z. Liu, "Improving the stability characteristics of asynchronous traffic in FDDI token rings", *High Speed Networks and their performance*, Perros and Viniotis Eds., North Holland, 441-460, 1994.

[7] E. Altman and F. Spieksma, "Geometric Ergodicity and Moment Stability of Station Times in Polling Systems", to appear in *Stochastic Models*, 1995.

[8] S. Asmussen, *Applied Probability and Queues*, Wiley, 1987.

[9] A. A. Borovkov, *Ergodicity and stability of stochastic processes*, to appear.

[10] A. A. Borovkov, "Limit theorems for queueing networks", *Theory Prob. Appl.*, **31**, 413-427, 1986.

[11] A. A. Borovkov and R. Schassberger, "Ergodicity of a polling system", *Stochastic Processes and Applications* **50**, 253-262, 1994.

[12] L. Greorgiadis, W. Szpankowski "Stability of Token Passing Rings", *QUESTA, special issue on Polling Models*, Eds. Takagi and Boxma, 7-33, 1992.

[13] S. Foss and A. Rybko, "Stability of multiclass Jackson-type networks", Submitted to *Adv. Applied Probability*, 1995.

[14] C. Fricker and M.R. Jaibi, "Monotonicity and stability of periodic polling models", *QUESTA* **15**, 211-238, 1994.

[15] S. Karlin and H. M. Taylor, *A first course in stochastic processes*, Academic Press, 1975.

[16] D. P. Kroese and V. Schmidt, "A continuous Polling System with General Service Times", *Annals of Applied Probability* **2**, 906-927, 1992.

[17] D. P. Kroese and V. Schmidt, "Single-Server Queues with Spatially Distributed Arrivals", Submitted.

[18] H. Levy, M. Sidi, O. J. Boxma, "Dominance Relations in Polling Systems", *QUESTA*, **6**, 155-172, 1990.

[19] T. Lindvall. *Lecture on the Coupling Method*, Wiley, 1992.

[20] S. P. Meyn and R. L. Tweedie, "Stability of Markovian processes I: criteria for discrete time chains", *Adv Applied Probability*, **24**, 542-574, 1992.

[21] J. A. C. Resing, "Polling systems and multi-type branching processes", *QUESTA* **13**, 409-426, 1993.

[22] V. Sherma, "Stability and continuity of polling systems", *QUESTA* **16**, 115-137, 1994.

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

[24] H. Takagi, "Queuing Analysis of Polling Models: an Update", *Stoch. Anal. Computer and Communications Systems*, (Takagi, Ed.), Elsevier, 267-318, 1990.

[25] R. L. Tweedie, "Criteria for classifying general Markov chains", *Adv. Appl. Prob.* **8**, 737-771, 1976.

[26] R. L. Tweedie, "Criteria for rates of convergence of Markov chains, with application to queuing and storage theory", *Probability, Statistics and Analysis*, ed. Kingman and Renter, London Math Soc. Lecture Notes 79, Cambridge Univ. Press, 260-276, 1983.

[27] V. S. Zhdanov, E. A. Saksonov, "Conditions of Existence of Steady-State Modes in Cyclic Queuing Systems", *Automation and Remote Control*, **40**, 176-184, 1979.