Gated polling with stationary ergodic walking times, Markovian routing and random feedback

Dieter Fiems · Eitan Altman

Abstract We investigate a gated polling system with semi-linear feedback and Markovian routing. We thereby relax the classical independence assumption on the walking times; the walking times constitute a sequence of stationary ergodic random variables. It is shown that the dynamics of this polling system can be described by semi-linear stochastic recursive equations in a Markovian environment. We obtain expressions for the first and second order moments of the workload and queue content at polling instants and for the mean queue content and workload at random instants.

Keywords Polling system · Branching process · Semi-linear process · Feedback · Random routing

1 Introduction

In a polling system, a single server visits multiple stations in cyclic order. If customers are waiting at a station, the server serves one or more of them in accordance with a polling discipline and then moves to the next station. The most common polling disciplines are the gated, exhaustive and 1-limited disciplines. All customers that are present upon arrival of the server at a station are served if the discipline is gated, the server remains with the station until there are no more customers present for the exhaustive discipline and a single customer is served prior to moving to the next station for the 1-limited discipline. Moreover, moving from station to station usually is not instantaneously, a walking time is required.

Many real-world queueing systems in telecommunications and manufacturing can be described by polling systems. Hence, there has been a continuing interest in polling systems since the 1950s. Polling systems are also interesting from a mathematical point of view.

D. Fiems (✉)
Ghent University, Gent, Belgium
e-mail: Dieter.Fiems@UGent.be

E. Altman
INRIA, Sophia-Antipolis, France
e-mail: Eitan.Altman@inria.fr

These systems are one of the few multidimensional queueing systems for which explicit solutions for the expected waiting times have been available. The considerable literature on polling systems prior to 1985 is surveyed in Takagi and Kleinrock (1985) and Takagi (1987). A more recent survey on polling systems is Vishnevskii and Semenova (2006). The main characteristics of the polling system at hand and the related literature are summarised below.

Markovian routing For classic polling systems, it is assumed that the server visits the different stations in a deterministic order. Most often the stations are visited in cyclic order. There have been, however, a growing number of applications that can be modelled by polling systems in which the visit order of the server is random. This type of server routing is also considered in this paper.

Polling systems with random server routing were introduced in Kleinrock and Levy (1988), and pseudo-conservation laws for the waiting times were established in Boxma and Weststrate (1989). The random routing discipline was motivated by distributed network control, and in particular by slotted ALOHA. The authors considered exhaustive, gated and 1-limited discrete-time polling systems. Also Levy and Sidi (1990) mention random-access schemes in shared channel networks as a possible application for polling systems with random server routing. Polling systems with random routing and finite buffers were studied in Chung et al. (1994) and Lee and Sunjaya (1996), Lee (2003), the latter papers extending the former paper by allowing arrival correlation. In all cases, each buffer can only hold a single customer. In Lee (2003) the buffer can also hold an additional customer while another customer receives service.

Polling systems with random routing can be used to assess demand-based multiple access protocols. Already in geostationary satellite communication, we find demand-based multiple access protocols in which there is a first phase of random access that is used by terminals to send a reservation request packet. The satellite then responds by assigning a dedicated channel for the amount of data that is requested in the reservation. An example is the SPADE (single-channel-per carrier PCM multiple-access demand assignment equipment) system operated by INTELSAT described in Evans (1999, p. 137). A similar mechanism is available in the WiMAX metropolitan area network, standardised by IEEE (known as IEEE 802.16 standard), for access based on a contention phase in which a station is selected at random followed by a transmission period for the terminal that received access for a time limited to the amount of data it has to send (Staehle and Pries 2007). This type of access is used for elastic traffic. Yet WiMAX has also a polling based mechanism in which the visit order is determined by the base station (Chang et al. 2008). The standard does not specify how to choose the service order.

In addition, polling systems with random routing can also be applied in manufacturing. For example, Srinivasan (1991) investigates a polling system with state-dependent Markovian random routing—the transition matrix of the server routing is different if the server finds an empty queue—and 1-limited polling and applies it to the performance analysis of an Automated-Guided-Vehicle (AGV) materials handling system. A single vehicle moves loads from one machine to another. When the AGV delivers a load, it inspects (polls) the output buffer of that machine to determine if there are any loads waiting to be transported. As such, the polling order is determined by the destination of the loads. The same polling system can also be used to model and assess delivery of packets by a "messenger boy" who picks up new packets upon delivery of some other packet. Fayolle and Lasgouttes (1995) studies a similar polling system, thereby investigating the waiting times when the routing probabilities are state-dependent.

Feedback The concept of queues with feedback was introduced by Takács ([1963](#)). In such a queueing system, a customer has some fixed probability to rejoin the queue upon service completion. In the context of polling systems, a customer may either rejoin the same queue, join another queue or leave the polling system. For polling systems with Bernoulli feedback, customers cannot join other queues upon service completion whereas this restriction does not apply for polling systems with Markovian feedback.

Takagi ([1987](#)) and de Moraes ([1990](#)) consider 1-limited polling systems with Bernoulli feedback. This type of polling model is applied in the context of token-ring networks where a single message is segmented into a (geometrically distributed) number of packets. The feedback mechanism corresponds to the message segmentation: a message virtually reenters the queue until all its packets are transmitted. The same queueing system can also be applied to model packet retransmissions in case a packet contains errors upon reception. Apart from the 1-limited polling discipline, Bernoulli feedback is also investigated for gated and exhaustive polling systems (Takine et al. [1991](#)). Limiting the polling system to two stations with exhaustive service (such a system is sometimes referred to as a queueing system with alternating priorities), Gupta and Buzacott ([1990](#)) investigate Bernoulli feedback in the context of manufacturing. In contrast to the preceding papers, the customers rejoin the queue at the head of the queue and receive service immediately.

Polling models with Markovian feedback also naturally arise in the context of token ring networks. Sidi et al. ([1992](#)) note that work at some station can generate work at other queues, e.g. a file request from a node in the token ring network results in a file transmission initiated from another node. These authors investigate exhaustive and gated polling disciplines. This is also the case in Hirayama ([2005](#)) where a multi-class polling system (gated or exhaustive) with cyclic service, FCFS or priority scheduling at each polling station and Markovian feedback is studied.

All the former contributions assume cyclic server routing in contrast to the Markovian routing which is studied here. Furthermore, we complement the classic feedback mechanism—a customer may generate new customer arrivals upon service completion—by a conceptually different one: feedback is expressed in terms of workload and the dynamics of the feedback to the different stations while serving at a particular station are modelled by semi-linear processes.

Semi-linear processes The class of semi-linear processes includes both linear processes and branching processes and were introduced in Altman ([2009](#)). For both linear processes and branching processes, the first and second order moments can be expressed as linear forms of the process parameter. Recalling the definition from Altman ([2009](#)), an $\mathbb{R}^K$-valued stochastic process $\mathbf{A}(\mathbf{x})$ is semi-linear if (i) for $\mathbf{x} = \mathbf{x}_1 + \mathbf{x}_2 + \cdots + \mathbf{x}_k \in \mathbb{R}_+^K$, $\mathbf{A}(\mathbf{x})$ has the following representation,

$$\mathbf{A}(\mathbf{x}) = \sum_{l=1}^{k} \hat{\mathbf{A}}^{(l)}(\mathbf{x}_l),$$

whereby $\hat{\mathbf{A}}^{(l)}(\cdot)$, $l = 1, \ldots, k$, are identically distributed, but not necessarily independent, with the same distribution as $\mathbf{A}(\cdot)$. Moreover, (ii) $\mathbf{A}(\cdot)$ is linear in the mean,

$$\mathsf{E}\left[\mathbf{A}(\mathbf{x})\right] = \mathcal{A}\mathbf{x}, \quad \mathbf{x} \in \mathbb{R}_+^K, \tag{1}$$

and the correlation matrix of $\mathbf{A}(\mathbf{x})$ is linear in $\mathbf{x}\mathbf{x}'$ and $\mathbf{x}$. For all $\mathbf{x} = [x^{(1)}, \ldots, x^{(K)}] \in \mathbb{R}_+^K$, we have the following representation,

$$\mathsf{E}\left[\mathbf{A}(\mathbf{x})\mathbf{A}'(\mathbf{x})\right] = \Upsilon\left(\mathbf{x}\mathbf{x}'\right) + \sum_{j=1}^{K} x^{(j)}\mathcal{D}_j. \tag{2}$$

In the former expressions, $\mathcal{A}$ is fixed $K \times K$ matrix, Υ is a linear operator that maps $K \times K$ non-negative definite matrices on $K \times K$ non-negative definite matrices and satisfies $\Upsilon(0) = 0$ and $\{\mathcal{D}_j, j = 1, \ldots, K\}$ is a set of fixed $K \times K$ matrices.

Semi-linear processes being a generalisation of branching processes, the close connection between branching processes and polling systems is retained. Resing (1993) showed that multitype branching processes with migration capture the dynamics of the number of customers at polling instants. Moreover, a similar branching structure with a continuous state space was shown to describe the so called "station times" of polling systems (Altman and Fiems 2007). A station time is the time spent at the various queues including the walking time to the next queue. This structure was used to compute the expected waiting times of polling systems with up to two queues by reducing the state evolution to a one-dimensional branching process (Groenevelt and Altman 2005).

By the introduction of semi-linear feedback, the workload at polling instants cannot be described by a pure branching process: the dynamics are described by a semi-linear process with migration. Moreover, this process operates in a Markovian environment to account for the random routing. Such a process was already studied in the context of packet forwarding in delay-tolerant networks (Fiems and Altman 2009a).

Organisation of the paper   The remainder of this paper is organised as follows. The details of the polling models at hand are outlined in the next section. Some specific applications of the models are highlighted there as well. Sections 3 and 4 are then concerned with the analysis. In Sect. 3 the state of the polling system at polling instants is described in terms of workload, this is a station times approach. In contrast, the state is expressed in terms of customers in Sect. 4, this is a buffer occupancy approach. In either case, it is shown that the dynamics of the polling system can be described by semi-linear processes with migration in a Markovian environment. For self-containment, the main results of these processes are summarised in the appendix. Finally, conclusions are drawn in Sect. 5.

2 Polling model

We consider a polling system with K infinite capacity buffers, each buffer implementing a first-come-first-served service discipline. The polling system adheres to a gated polling discipline—the server stays at a buffer till all workload present upon arrival of the server at that buffer is processed—and the server moves randomly between the different buffers. The modelling assumptions are described in detail below and are summarised in Fig. 1. We consider two complementary polling systems: a system with continuous state space (Sect. 2.1) and a system with a discrete state space (Sect. 2.2). In the former system, arriving load is expressed in terms of workload whereas the load is expressed in terms of customers and their service times in the latter system. Before proceeding to the analysis, some example applications of the polling systems at hand are discussed in the Sects. 2.3 and 2.4.

2.1 Modelling assumptions: continuous state space

Service routing   Let $\xi_n \in \{1, 2, \ldots, K\}$ denote the nth buffer being polled. The consecutive ξ_n are not independent. Let θ be a (deterministic) surjective mapping from $\{1, 2, \ldots, N\}$ onto $\{1, 2, \ldots, K\}$. The nth queue being polled is $\xi_n = \theta(Y_n)$ whereby $\{Y_n\}$ denotes an ergodic Markov chain on the finite state space $\Omega = \{1, 2, \ldots, N\}$. The transition probabilities of the chain are denoted by p_{ij} and the corresponding $N \times N$ transition matrix is denoted
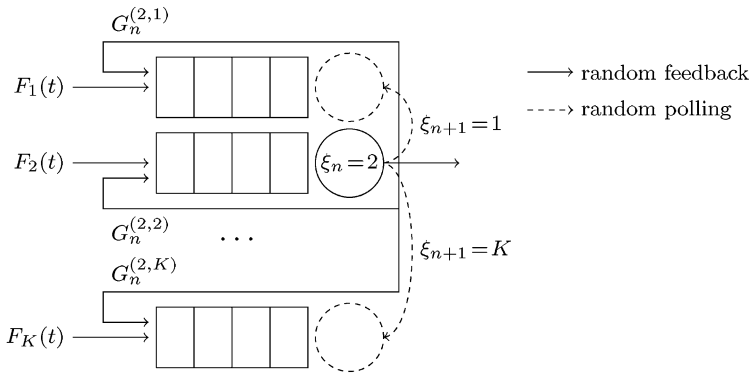
Fig. 1 Schematic representation of the polling system under investigation

by $\mathcal{P} = [p_{ij}]$. The assumptions above imply that the state of the Markov chain determines which buffer is polled but the buffer being polled does not necessarily determine the state of the Markov chain. These assumptions obviously include the case where the consecutive ξ_n constitute a Markov chain or are modulated by a Markov chain. Finally, notice that all buffers are being polled since the Markov chain is ergodic and the mapping θ is surjective.

Walking times   Moving between buffers incurs a walking time which does not only depend on the buffer being polled; the consecutive walking times are dependent random variables. In particular, let $\{\mathbf{W_n}\}$ denote a sequence of stationary ergodic random N-dimensional column vectors. The walking time between the nth and $n + 1$st polling instant is then given by $W_n^{(Y_n)}$, the Y_nth element of $\mathbf{W}_n$. Example applications of stationary ergodic walking times include signalling and switching time in mobile networks, where the signalling or switching speed depends on the wireless channel conditions which can be correlated in time, see (Altman and Fiems 2007).

For further use, let $\mathbf{w} = \mathsf{E}\,\mathbf{W}_0$ denote the mean walking time vector, $w^{(i)}$ being the ith element of this vector. Moreover, let $\mathcal{W}_n = \mathsf{E}\,\mathbf{W}_0\mathbf{W}_n'$ denote the autocorrelation matrix at lag n and let $\mathcal{W}_n^{(i,j)}$ denote the element at position (i, j) in this matrix. In line with the notation introduced above, the ith element of any vector $\mathbf{x}$ is denoted by $x^{(i)}$ in the remainder. Similarly, for any matrix $\mathcal{X}$, $\mathcal{X}^{(i,j)}$ denotes the element at position (i, j).

In the above setting, the walking time distribution from queue $\theta(Y_n)$ to queue $\theta(Y_{n+1})$ does not depend in an explicit way on Y_{n+1}. However, a direct transformation allows us to include such dependence in our model. Indeed, we can augment Y_n and consider the Markov chain $Z_n := (Y_n, Y_{n+1})$ instead. If the initial Markov chain is ergodic, then this is also the case for the new one. Furthermore, notice that the sequence of consecutive walking times $W_n^{(Y_n)}$ is a stationary ergodic sequence of random variables as both sequences $\mathbf{W}_n$ and Y_n are stationary ergodic.

Arrivals   The arrival process in the different buffers is modelled as an $\mathbb{R}_+^K$-valued Lévy process $\mathbf{F}(t)$. The kth element $F^{(k)}(t)$ of $\mathbf{F}(t)$ denotes the amount of work arriving at the kth buffer in the interval $(0, t]$. For ease of notation, let $\mathbf{F}_{n,m}(t)$ denote a doubly-indexed sequence of independent Lévy processes, distributed as $\mathbf{F}(t)$. Moreover, for Lévy processes, the mean value and autocovariance of $\mathbf{F}(t)$ can be expressed as follows,

$$\mathsf{E}\,\mathbf{F}(t) = \mathbf{f}t, \quad \text{cov}\,\mathbf{F}(t) = \mathcal{F}t,$$

with $\mathbf{f}$ a fixed column vector and $\mathcal{F}$ a fixed $K \times K$ matrix.

Feedback traffic.   Processing workload in a buffer may lead to additional workload in other buffers. In other words, workload can be fed back to the same buffer as well as to the other buffers. Let $\mathbf{G}_n(t)$ denote the feedback process corresponding to the nth polling instant. The process $\mathbf{G}_n(t)$ is a semi-linear process, its distribution depending on the buffer being polled. Semi-linear processes were introduced in Altman ([2009](#)), and their main properties are summarised in [Appendix](#) for completeness. The kth element $G_n^{(k)}(t)$ of $\mathbf{G}_n(t)$ is the amount of work that is fed back to queue k. For ease of notation and for each n, let $\mathbf{G}_{n,m}(t)$ denote a sequence of independent semi-linear processes, distributed as $\mathbf{G}_n(t)$. Moreover, by the semi-linearity, mean and autocovariance of $\mathbf{G}_n(t)$ (conditioned on the buffer being polled) can be expressed as follows,

$$\mathsf{E}\big[\mathbf{G}_0(t)|\xi_0 = k\big] = \mathbf{g}_k\, t, \quad \mathrm{cov}\big[\mathbf{G}_0(t)|\xi_0 = k\big] = \mathcal{G}_k t + \mathcal{H}_k t^2.$$

2.2 Modelling assumptions: discrete state space

For the discrete state space model, we retain the assumptions on server routing and walking times. The arrival and feedback processes adhere to the assumptions below. Since there is no risk for confusion, much of the notation of the continuous state space model is reused here.

Arrivals   For the discrete state space model, the arrival process at the different buffers is modelled as a K-dimensional batch Poisson process $\mathbf{F}(t)$. The kth element $F^{(k)}(t)$ of $\mathbf{F}(t)$ denotes the number of customers arriving in the kth buffer in the interval $(0, t]$. For ease of notation, let $\mathbf{F}_{n,m}(t)$ denote a doubly-indexed sequence of independent Poisson processes, distributed as $\mathbf{F}(t)$. Also, let $\mathbf{f} = \mathsf{E}\,\mathbf{F}(1)$ and $\mathcal{F} = \mathrm{cov}\,\mathbf{F}(1)$ denote the mean vector and covariance matrix of the number of customer arrivals in a unit-length interval. The mean and covariance of $\mathbf{F}(t)$ then equal $t\mathbf{f}$ and $t\mathcal{F}$, respectively.

Service and feedback   The customer service times constitute a sequence of independent random variables, their distribution depending on the buffer. Let $S_n(i)$ denote the service time of the ith customer at the nth polling instant. Moreover, let $s_k = \mathsf{E}\,[S_0(1)|\xi_0 = k]$ and $s_{2,k} = \mathsf{E}\,[(S_0(1))^2|\xi_0 = k]$ denote the first and second moment of the service times in buffer k. Note that there is no distinction between customers that enter the system for the first time and customers that are fed back to the system.

Processing customers in a buffer may trigger additional customer arrivals in other buffers. In other words, customers can be fed back to the same buffer as well as to the other buffers. Let $\mathbf{G}_n(t)$ denote the feedback process corresponding to the nth polling instant. Like the arrival process, $\mathbf{G}_n(t)$ is a Poisson process, its distribution depending on the buffer being polled. The kth element $G_n^{(k)}(t)$ of $\mathbf{G}_n(t)$ equals the number of customers that are fed back to queue k. For ease of notation and for each n, let $\mathbf{G}_{n,m}(t)$ denote a sequence of independent Poisson processes, distributed as $\mathbf{G}_n(t)$. Moreover, mean and autocovariance of $\mathbf{G}_n(t)$ (conditioned on the buffer being polled) can be expressed as follows,

$$\mathsf{E}\big[\mathbf{G}_0(t)|\xi_0 = k\big] = \mathbf{g}_k\, t, \quad \mathrm{cov}\big[\mathbf{G}_0(t)|\xi_0 = k\big] = \mathcal{G}_k t.$$

The former type of feedback is similar to the type of feedback that was assumed for the continuous state space model. In addition, we also allow for classical feedback: upon service

completion, a number of customers can enter the different buffers. Let $R_n^{(k)}(i)$ denote the number of customers that enter the kth buffer upon service completion of the ith customer at the nth polling instant. In addition, let $\mathbf{R}_n(i)$ denote the random vector with elements $R_n^{(k)}(i)$. The random vectors $\mathbf{R}_n(i)$ are independent, their distribution depending on the buffer being polled. The following notation is introduced for further use: $\mathbf{r}_k = \mathsf{E}\,[\mathbf{R}_0(1)|\xi_0 = k]$ and $\mathcal{R}_k = \text{cov}[\mathbf{R}_0(1)|\xi_0 = k]$.

2.3 Example of random service order

2.3.1 Bernoulli polling

An example application of our model is the Bernoulli polling model introduced in Altman and Yechiali (1993). The server moves cyclically between the stations in the order $1, 2, \ldots, K, 1, 2, \ldots, K$, etc. When arriving at queue i, a coin is flipped. With probability q the server decides to poll this queue and with probability $1 - q$ it moves to the next queue, etc. Moving from queue i to queue $i + 1$ requires a walking time distributed like a random variable D_i. If the server decides to poll queue $i + 1$ then an additional switching time of S_{i+1} is incurred. It was shown in (Altman and Yechiali 1993) that the expected waiting times under this visit order may be strictly smaller than with any cyclic order polling system.

We next show that this can be formulated within our framework. Let q_{ij} denote the probability that queue j is polled next, given that the server is at queue i. In view of the model description above, we find,

$$q_{ij} = \begin{cases} \dfrac{(1-q)^{i-j-1}q}{1 - (1-q)^K} & \text{for } j + 1 \leq i \leq K, \\[3mm] \dfrac{(1-q)^{i-j-1+K}q}{1 - (1-q)^K} & \text{for } 1 \leq i \leq j. \end{cases}$$

We can now determine the walking time from queue i to queue j as follows. For each $i \in \{1, \ldots, K\}$, let $D_{i,n}$ denote a sequence of independent random variables, distributed like D_i and let C_n be,

$$C_n = \sum_{i=1}^{K} D_{i,n}.$$

Obviously, C_n distributes as the time it takes to move from station i to station i without actually polling any queue. For $j > i$, the total walking time from station i to station j includes (i) the walking times from station i to station $i + 1$, from station $i + 1$ to station $i + 2, \ldots$, and from station $j - 1$ to station j, (ii) a geometrically distributed number G of walking times from station j to station j, and (iii) the switch-over time to station j. Hence the (total) walking time from station i to station j is distributed as,

$$W_{i,j} = D_{i,0} + \cdots + D_{j-1,0} + \sum_{n=1}^{G} C_n + S_j.$$

For $j \leq i$, similar observations show that the walking time is distributed as,

$$W_{i,j} = D_{i,0} + \cdots + D_{K,0} + D_{1,0} + \cdots + D_{j-1,0} + \sum_{n=1}^{G} C_n + S_j.$$

From the former expressions, it is clear that the walking times depend on the current and the next buffer being polled. Such a dependence fits our framework by extending the state space of the modulating Markov chain (the state includes the current buffer and the next buffer) as was already noted in Sect. 2.1.

2.3.2 Random walk polling

Another example is the random walk: when completing service at a station i, the server goes either to station $i + 1$ or to station $i - 1$, with probability q_i and $1 - q_i$, respectively.

2.4 Examples of feedback

We present various types of feedback that fit the present modelling assumptions.

2.4.1 Retransmission of packets

Consider a single queue with a gated polling (or vacation) discipline. Customers arrive in accordance with a Poisson process. The ith customer to be served has a service time of S_i time units, the service times being independent and identically distributed. Upon service completion, the customer reenters the queue with probability q or leaves with probability $1 - q$. Such a reentrance may represent a retransmission of a packet in some network.

2.4.2 Acknowledgements

Consider a system with two queues whereby customers arrive at queue 1 in accordance with a Poisson process. A customer served in queue 1 is routed to queue 2 with probability q and served there, or it leaves the system with probability $1 - q$. q can be viewed as the successful packet transmission probability, and the routing to queue 2 can be interpreted as acknowledgements that are sent back to the source and that are triggered by each successful transmission.

2.4.3 Compression

Consider two queues. There is a Lévy arrival process at queue 1. During the nth busy period, a constant percent α_n of the workload that leaves queue 1 is rerouted to queue 2. This traffic can be interpreted as a compressed version of the initial traffic.

For example, by means of H.264 scalable video coding (Schwartz et al. 2007), multiple decodable substreams can be extracted from a single scalable video stream. If a receiver only accepts a certain substream of the complete stream (e.g. at a lower spatial or temporal resolution), this substream has to be extracted in the network. Hence, at this network node, both the complete stream and the substream (modelled by feedback to a second queue) have to be transmitted.

2.4.4 Random response

Consider a single user (a single queue) that receives new jobs in accordance with a Poisson arrival process. During the busy period of the queue, there are additional jobs, modelled as another Poisson process. These jobs join the same queue. For example, an arrival corresponds to an email and its service time represents the time for reading and responding. The feedback models extra emails sent by other users who may react to the responses of the user.

3 Analysis: continuous state space model

Let $V_n^{(k)}$ denote the workload in queue k at the nth polling instant and let $\mathbf{V}_n$ denote the vector with elements $V_n^{(k)}$, $k = 1, \ldots, K$. The workload in the different buffers at polling instant $n + 1$ includes (i) all workload that was present in the different buffers, excluding buffer ξ_n which is being polled, (ii) all workload that is fed back into the system while processing the workload in buffer ξ_n, and (iii) all workload that arrived during the station time at buffer ξ_n and during the walking time from station ξ_n to station ξ_{n+1}. This observation yields the following set of recursions,

$$V_{n+1}^{(k)} = V_n^{(k)} + G_n^{(k)}\big(V_n^{(\xi_n)}\big) + F_n^{(k)}\big(V_n^{(\xi_n)} + W_n^{(Y_n)}\big), \quad \text{for } k \neq \xi_n,$$
$$V_{n+1}^{(\xi_n)} = G_n^{(\xi_n)}\big(V_n^{(\xi_n)}\big) + F_n^{(\xi_n)}\big(V_n^{(\xi_n)} + W_n^{(Y_n)}\big) \tag{3}$$

By application of the divisibility property of semi-linear processes, and in view of the mapping between Y_n and ξ_n, this set of recursive equations further simplifies to,

$$V_{n+1}^{(k)} = \underbrace{V_n^{(k)} + G_n^{(k)}\big(V_n^{(\theta(Y_n))}\big) + F_{n,1}^{(k)}\big(V_n^{(\theta(Y_n))}\big)}_{A_n^{(k)}(\mathbf{V}_n, Y_n)} + \underbrace{F_{n,2}^{(k)}\big(W_n^{(Y_n)}\big)}_{B_n^{(k)}(Y_n)}, \tag{4}$$

for $k \neq \theta(Y_n)$ and,

$$V_{n+1}^{(\theta(Y_n))} = \underbrace{G_n^{(\theta(Y_n))}\big(V_n^{(\theta(Y_n))}\big) + F_{n,1}^{(\theta(Y_n))}\big(V_n^{(\theta(Y_n))}\big)}_{A_n^{(\theta(Y_n))}(\mathbf{V}_n, Y_n)} + \underbrace{F_{n,2}^{(\theta(Y_n))}\big(W_n^{(Y_n)}\big)}_{B_n^{(\theta(Y_n))}(Y_n)}. \tag{5}$$

To simplify notation, we collect the terms depending on $\mathbf{V}_n$ and those that do not into two random functions as indicated above. In other words, we introduce the following compact notation for the set of recursive equations (4) and (5),

$$\mathbf{V}_{n+1} = \mathbf{A}_n(\mathbf{V}_n, Y_n) + \mathbf{B}_n(Y_n), \tag{6}$$

with,

$$\mathbf{A}_n(\mathbf{v}, i) = \mathbf{G}_n\big(v^{(\theta(i))}\big) + \mathbf{F}_{n,1}\big(v^{(\theta(i))}\big) + \mathcal{K}_{\theta(i)}\mathbf{v}, \tag{7}$$

and,

$$\mathbf{B}_n(i) = \mathbf{F}_{n,2}\big(W_n^{(i)}\big). \tag{8}$$

Here, $\mathcal{K}_i$ is a diagonal matrix whose ith diagonal element equals 0 while its other diagonal elements equal 1. Note that for fixed i, $\mathbf{A}_n(\cdot, i)$ is indeed semi-linear as it is the sum of semi-linear processes.

By definition, for each n and for fixed i, all terms on the right-hand side of (7) are semi-linear which implies that $\mathbf{A}_n$ is semi-linear as well. Moreover, in view of the independence assumptions on $\mathbf{F}$ and $\mathbf{G}$, the consecutive $\mathbf{A}_n$ are independent random processes. By the measurability of $\mathbf{F}_{n,2}$ and the ergodicity of the walking times, it is further observed that the sequence $\mathbf{B}_n$ is stationary ergodic. We conclude that the framework of Appendix applies. Expressions for the moments of $\mathbf{A}_0$ and $\mathbf{B}_0$ are established in the following subsection.

3.1 Moments of $\mathbf{A}_0$ and $\mathbf{B}_0$

Taking expectations in (7) immediately yields,

$$\mathsf{E}\,\mathbf{A}_0(\mathbf{v}, i) = v^{(\theta(i))}(\mathbf{g}_{\theta(i)} + \mathbf{f} - \mathbf{e}_{\theta(i)}) + \mathbf{v}. \tag{9}$$

for a non-random vector $\mathbf{v} \in \mathbb{R}_+^K$ and a non-random $i \in \Omega$. Here $\mathbf{e}_x$ is a column vector whose xth entry equals 1 while its other entries equal zero. The former expression can be expressed as a matrix product, that is,

$$\mathsf{E}\,\mathbf{A}_0(\mathbf{v}, i) = \mathcal{A}_i \mathbf{v}, \tag{10}$$

with,

$$\mathcal{A}_i = \begin{bmatrix} 1 & 0 & \dots & g_{\theta(i)}^{(1)} + f^{(1)} & \dots & 0 \\ 0 & 1 & \dots & g_{\theta(i)}^{(2)} + f^{(2)} & \dots & 0 \\ \vdots & & \ddots & \vdots & & \\ 0 & 0 & \dots & g_{\theta(i)}^{(\theta(i))} + f^{(\theta(i))} & \dots & 0 \\ \vdots & & & \vdots & \ddots & \\ 0 & 0 & \dots & g_{\theta(i)}^{(K)} + f^{(K)} & \dots & 1 \end{bmatrix}. \tag{11}$$

By the independence of G_0 and $\mathbf{F}_{0,1}$, the following expression for the autocovariance of $\mathbf{A}_0$ is found,

$$\operatorname{cov}\mathbf{A}_0(\mathbf{v}, i) = v^{(\theta(i))}\mathcal{G}_{\theta(i)} + \left(v^{(\theta(i))}\right)^2 \mathcal{H}_{\theta(i)} + v^{(\theta(i))}\mathcal{F}. \tag{12}$$

We then have the following expression for the second order vector,

$$\mathsf{E}\,\mathbf{A}_0(\mathbf{v}, i)\mathbf{A}_0(\mathbf{v}, i)' = v^{(\theta(i))}\underbrace{(\mathcal{G}_{\theta(i)} + \mathcal{F})}_{\mathcal{D}_{i,\theta(i)}} + \underbrace{\left(v^{(\theta(i))}\right)^2 \mathcal{H}_{\theta(i)} + \mathcal{A}_i \mathbf{v}\mathbf{v}'\mathcal{A}_i'}_{\Upsilon_i(\mathbf{v}\mathbf{v}')}$$

$$= \Upsilon_i\left(\mathbf{v}\mathbf{v}'\right) + \sum_{j=1}^{K} v^{(j)}\mathcal{D}_{i,j},$$

with,

$$\Upsilon_i\left(\mathbf{v}\mathbf{v}'\right) = \left(v^{(\theta(i))}\right)^2 \mathcal{H}_{\theta(i)} + \mathcal{A}_i \mathbf{v}\mathbf{v}'\mathcal{A}_i', \quad \mathcal{D}_{i,j} = \mathbb{1}\{j = \theta(i)\}(\mathcal{G}_j + \mathcal{F}). \tag{13}$$

Here, $\mathbb{1}\{\}$ is the standard indicator function.

Having established expressions for the moments of $\mathbf{A}_0$, we now focus on the moments of $\mathbf{B}_0$. Taking expectations in (8), we immediately find,

$$\mathsf{E}\,\mathbf{B}_0(i) = w^{(i)}\mathbf{f}. \tag{14}$$

Moreover, the second order moment can be expressed as,

$$\mathsf{E}\left[\mathbf{B}_0(i)\mathbf{B}_n(j)'\right] = \mathsf{E}\left[W_0^{(i)}W_n^{(j)}\right]\mathbf{f}\mathbf{f}' = \mathcal{W}_n^{(i,j)}\mathbf{f}\mathbf{f}', \tag{15}$$

for all $i, j = 1, \dots, N$ and $n \neq 0$, whereas for $n = 0$, we have,

$$\mathsf{E}\left[\mathbf{B}_0(i)\mathbf{B}_0(i)'\right] = \mathsf{E}\left[W_0^{(i)}\right]\mathcal{F} + \mathsf{E}\left[\left(W_0^{(i)}\right)^2\right]\mathbf{f}\mathbf{f}' = w^{(i)}\mathcal{F} + \mathcal{W}_0^{(i,i)}\mathbf{f}\mathbf{f}'. \tag{16}$$

3.2 Workload

With the expressions of the moments of $\mathbf{A}$ and $\mathbf{B}$ at hand, Theorem 1 of the Appendix establishes the existence of a stationary solution of the workload at polling instants. Moreover, Theorem 2 provides expressions for the first two moments of the workload in the different queues at polling instants. Note that the stability condition of Theorem 2 is expressed in terms of the maximal eigenvalue of a matrix of first order moments of the arrival and feedback processes. We were not able to further simplify the stability conditions.

Recall that $\boldsymbol{\mu}_i = \mathsf{E}\,[\mathbf{V}_0^* 1\{Y_0 = i\}]$ and $\boldsymbol{\Omega}_i = \mathsf{E}\,[\mathbf{V}_0^* (\mathbf{V}_0^*)' 1\{Y_0 = i\}]$. We now express various performance measures in terms of these vectors and matrices and in terms of the steady state vector $\boldsymbol{\pi}$ of the Markov chain Y_k. That is, $\boldsymbol{\pi}$ is the unique solution of $\boldsymbol{\pi}\mathcal{P} = \boldsymbol{\pi}$ and $\boldsymbol{\pi}\mathbf{e} = 1$, $\mathbf{e}$ being a column vector of ones. In the remainder, the superscript asterisks that indicate the stationary solution are dropped for ease of notation.

We first consider the average workload in the system. Let $\hat{\mathbf{v}}$ denote the mean workload (vector) at a random time when the polling system has reached steady state. This vector can be obtained by averaging over a station time. We have the following expression,

$$\hat{\mathbf{v}} = \frac{\mathsf{E}\,[\int_0^{V_0^{(\theta(Y_0))}+W_0^{(Y_0)}} \mathbf{V}_0(t)dt]}{\mathsf{E}\,[V_0^{(\theta(Y_0))} + W_0^{(Y_0)}]}. \tag{17}$$

Here, with a slight abuse of notation, $\mathbf{V}_0(t)$ denotes the workload in the different queues at time t during cycle 0. $\mathbf{V}_0(t)$ includes (i) the workload that was present at the stations at the 0th polling instant, excluding the work processed at the station being polled (ii) all work that arrived since the polling instant, and (iii) all work that has been fed back to the stations since the polling instant.

By conditioning on Y_0, the denominator in (17) can be expressed in terms of $\boldsymbol{\mu}_i$,

$$\mathsf{E}\,[V_0^{(\theta(Y_0))} + W_0^{(Y_0)}] = \sum_{i=1}^{N} \mathsf{E}\,[(V_0^{(\theta(i))} + W_0^{(i)})1\{Y_0 = i\}]$$

$$= \sum_{i=1}^{N} (\mu_i^{(\theta(i))} + \pi^{(i)} w^{(i)}). \tag{18}$$

Calculation of the numerator is more involved. Again, we first condition on Y_0. The cycle consists of the service time at station Y_0 and the consecutive walking time. During the service time, the mean queue contents at all stations but station $\theta(Y_0)$ grow linearly, while the mean queue content at station $\theta(Y_0)$ decreases linearly. During the walking time, the mean queue contents at all stations grow linearly. In view of these observations, we find the following expression,

$$\mathsf{E}\left[\int_0^{V_0^{(\theta(Y_0))}+W_0^{(Y_0)}} \mathbf{V}_0(t)dt\right]$$

$$= \sum_{i=1}^{N} \mathsf{E}\left[\int_0^{V_0^{(\theta(i))}} (\mathbf{V}_0 + (\mathbf{f} + \mathbf{g}_{\theta(i)} - \mathbf{e}_{\theta(i)})t)dt\, 1\{Y_0 = i\}\right]$$

$$+ \sum_{i=1}^{N} \mathsf{E}\left[\int_0^{W_0^{(i)}} (\mathbf{V}_0 + (\mathbf{f} + \mathbf{g}_{\theta(i)} - \mathbf{e}_{\theta(i)})V_0^{(\theta(i))} + \mathbf{f}t)dt\, 1\{Y_0 = i\}\right], \tag{19}$$

which further simplifies to,

$$\mathsf{E}\left[\int_0^{V_0^{(\theta(Y_0))}+W_0^{(Y_0)}} \mathbf{V}_0(t)dt\right]$$

$$= \frac{1}{2}\sum_{i=1}^N \mathsf{E}\left[\left(\mathbf{V}_0 + \left(\mathbf{V}_0 + (\mathbf{f}+\mathbf{g}_{\theta(i)} - \mathbf{e}_{\theta(i)})V_0^{(\theta(i))}\right)\right)V_0^{(\theta(i))}1\{Y_0=i\}\right]$$

$$+ \sum_{i=1}^N \mathsf{E}\left[\left(\left(\mathbf{V}_0 + (\mathbf{f}+\mathbf{g}_{\theta(i)} - \mathbf{e}_{\theta(i)})V_0^{(\theta(i))}\right)W_0^{(i)} + \frac{1}{2}\mathbf{f}\left(W_0^{(i)}\right)^2\right)1\{Y_0=i\}\right], \quad (20)$$

and,

$$\mathsf{E}\left[\int_0^{V_0^{(\theta(Y_0))}+W_0^{(Y_0)}} \mathbf{V}_0(t)dt\right] = \frac{1}{2}\sum_{i=1}^N \left(\mathbf{\Omega}_i\mathbf{e}_{\theta(i)} + \mathcal{A}_i\mathbf{\Omega}_i\mathbf{e}_{\theta(i)} + 2\mathcal{A}_i\mathbf{m}_i + \pi^{(i)}\mathbf{f}\mathcal{W}_0^{(i,i)}\right). \quad (21)$$

The last expression is obtained by noting that $V_0^{(\theta(i))} = \mathbf{e}'_{\theta(i)}\mathbf{V}_0 = \mathbf{V}'_0\mathbf{e}_{\theta(i)}$ and by equations (9) and (10). In this last expression, the unknown vectors $\mathbf{m}_i = \mathsf{E}[\mathbf{V}_0 W_0^{(i)}1\{Y_0=i\}]$ can be determined as follows. Following the lines of the proof of Theorem 2 in (Fiems and Altman 2009b), it is easy to show that

$$\mathsf{E}\left[\mathbf{V}_0\mathbf{B}_0(i)'1\{Y_0=i\}\right] = \mathcal{C}_i \quad (22)$$

with $\mathcal{C}_i$ the ith block diagonal element of $\hat{\mathcal{C}}$ as defined in (47). In view of equation (8), we further find,

$$\mathcal{C}_i = \mathbf{m}_i\mathbf{f}', \quad (23)$$

which uniquely defines the unknown expectation. That is, $\mathbf{m}_i$ equals the kth column of $\mathcal{C}_i$ divided by $f^{(k)}$ ($k = 1,\dots,K$).

Summarising, we have the following expression for the mean workload in the different buffers on a random time instant,

$$\hat{\mathbf{v}} = \frac{\sum_{i=1}^N(\mathbf{\Omega}_i\mathbf{e}_{\theta(i)} + \mathcal{A}_i\mathbf{\Omega}_i\mathbf{e}_{\theta(i)} + 2\mathcal{A}_i\mathbf{m}_i + \pi^{(i)}\mathbf{f}\mathcal{W}_0^{(i,i)})}{2\sum_{i=1}^N(\mu_i^{(\theta(i))} + \pi^{(i)}w^{(i)})}. \quad (24)$$

4 Analysis: discrete state space model

Let $V_n^{(k)}$ denote the number of customers in queue k at the nth polling instant and let $\mathbf{V}_n$ denote the vector with elements $V_n^{(k)}$, $k = 1,\dots,K$. At polling instant $n+1$, the buffers contain (i) all customers that were present in the different buffers, excluding buffer ξ_n which is polled, (ii) all customers that are fed back into the system while processing at buffer ξ_n, and (iii) all customers that arrived during the station time at buffer ξ_n and during the walking time from station ξ_n to station ξ_{n+1}. This observation yields the following set of recursions,

$$V_{n+1}^{(k)} = V_n^{(k)} + \sum_{i=1}^{V_n^{(\xi_n)}}\left(R_n^{(k)}(i) + G_n^{(k)}\left(\sum_{i=1}^{V_n^{(\xi_n)}} S_n(i)\right) + F_n^{(k)}\left(\sum_{i=1}^{V_n^{(\xi_n)}} S_n(i) + W_n^{(Y_n)}\right)\right),$$

for $k \neq \xi_n$,

$$V_{n+1}^{(\xi_n)} = \sum_{i=1}^{V_n^{(\xi_n)}} \left(R_n^{(\xi_n)}(i) + G_n^{(\xi_n)}\left(\sum_{i=1}^{V_n^{(\xi_n)}} S_n(i)\right) + F_n^{(\xi_n)}\left(\sum_{i=1}^{V_n^{(\xi_n)}} S_n(i) + W_n^{(Y_n)}\right)\right). \quad (25)$$

By application of the divisibility property of Poisson processes, and in view of the mapping between Y_n and ξ_n, this set of recursive equations further simplifies to,

$$V_{n+1}^{(k)} = \underbrace{V_n^{(k)} + \sum_{i=1}^{V_n^{(\theta(Y_n))}} \left(R_n^{(k)}(i) + G_{n,i}^{(k)}\big(S_n(i)\big) + F_{n,i}^{(k)}\big(S_n(i)\big)\right)}_{A_n^{(k)}(\mathbf{V}_n, Y_n)} + \underbrace{F_{n,0}^{(k)}\big(W_n^{(Y_n)}\big)}_{B_n^{(k)}(Y_n)},$$

for $k \neq \theta(Y_n)$, and

$$V_{n+1}^{(\theta(Y_n))} = \underbrace{\sum_{i=1}^{V_n^{(\theta(Y_n))}} \left(R_n^{(\theta(Y_n))}(i) + G_{n,i}^{(\theta(Y_n))}\big(S_n(i)\big) + F_{n,i}^{(\theta(Y_n))}\big(S_n(i)\big)\right)}_{A_n^{(\theta(Y_n))}(\mathbf{V}_n, Y_n)} + \underbrace{F_{n,0}^{(\theta(Y_n))}\big(W_n^{(Y_n)}\big)}_{B_n^{(\theta(Y_n))}(Y_n)}.$$

To simplify notation, we collect the terms depending on $\mathbf{V}_n$ and those that do not into two random functions as indicated above. In other words, we introduce the following compact notation for the former set of recursive equations,

$$\mathbf{V}_{n+1} = \mathbf{A}_n(\mathbf{V}_n, Y_n) + \mathbf{B}_n(Y_n), \quad (26)$$

with,

$$\mathbf{A}_n(\mathbf{v}, i) = \mathcal{K}_{\theta(i)}\mathbf{v} + \sum_{j=1}^{v^{(\theta(i))}} \big(\mathbf{R}_n(j) + \mathbf{G}_{n,j}\big(S_n(j)\big) + \mathbf{F}_{n,j}\big(S_n(j)\big)\big), \quad (27)$$

and,

$$\mathbf{B}_n(i) = \mathbf{F}_{n,0}\big(W_n^{(i)}\big). \quad (28)$$

Recall that $\mathcal{K}_i$ is a diagonal matrix whose ith diagonal element equals 0 while its other diagonal elements equal 1. As for the continuous state space model, it is easy to show that the framework of the appendix applies. The moments of $\mathbf{A}_0$ and $\mathbf{B}_0$ are determined in the next Subsection.

4.1 Moments

As for the continuous state space model, we determine the moments of $\mathbf{A}_0$ and $\mathbf{B}_0$. By taking expectations in (27), we immediately find,

$$\mathsf{E}\,\mathbf{A}_0(\mathbf{v}, i) = v^{(\theta(i))}s_{\theta(i)}(\mathbf{g}_{\theta(i)} + \mathbf{f}) + v^{(\theta(i))}\mathbf{r}_{\theta(i)} + \mathcal{K}_{\theta(i)}\mathbf{v} = \mathcal{A}_i\mathbf{v}, \quad (29)$$

for a non-random vector $\mathbf{v} \in \mathbb{N}^K$ and a non-random $i \in \Omega$ with,

$$
\mathcal{A}_i = \begin{bmatrix}
1 & 0 & \ldots & s_{\theta(i)}(g_{\theta(i)}^{(1)} + f^{(1)}) + r_{\theta(i)}^{(1)} & \ldots & 0 \\
0 & 1 & \ldots & s_{\theta(i)}(g_{\theta(i)}^{(2)} + f^{(2)}) + r_{\theta(i)}^{(2)} & \ldots & 0 \\
\vdots & & \ddots & \vdots & & \\
0 & 0 & \ldots & s_{\theta(i)}(g_{\theta(i)}^{(\theta(i))} + f^{(\theta(i))}) + r_{\theta(i)}^{(\theta(i))} & \ldots & 0 \\
\vdots & & & \vdots & \ddots & \\
0 & 0 & \ldots & s_{\theta(i)}(g_{\theta(i)}^{(K)} + f^{(K)}) + r_{\theta(i)}^{(K)} & \ldots & 1
\end{bmatrix}. \tag{30}
$$

Moreover, by the independence assumptions, we find the following expression for the co-variance,

$$
\text{cov}\, \mathbf{A}_0(\mathbf{v}, i) = v^{(\theta(i))}\big(\mathcal{R}_{\theta(i)} + s_{\theta(i)}(\mathcal{G}_{\theta(i)} + \mathcal{F}) + \big(s_{2,\theta(i)} - s_{\theta(i)}^2\big)(\mathbf{f} + \mathbf{g}_{\theta(i)})\big(\mathbf{f}' + \mathbf{g}_{\theta(i)}'\big)\big), \tag{31}
$$

which implies,

$$
\begin{aligned}
&\mathsf{E}\, \mathbf{A}_0(\mathbf{v}, i)\mathbf{A}_0(\mathbf{v}, i)' \\
&\quad = v^{(\theta(i))}\big(\mathcal{R}_{\theta(i)} + s_{\theta(i)}(\mathcal{G}_{\theta(i)} + \mathcal{F}) + \big(s_{2,\theta(i)} - s_{\theta(i)}^2\big)(\mathbf{f} + \mathbf{g}_{\theta(i)})\big(\mathbf{f}' + \mathbf{g}_{\theta(i)}'\big)\big) + \mathcal{A}_i \mathbf{v}\mathbf{v}' \mathcal{A}_i' \\
&\quad = \Upsilon_i\big(\mathbf{v}\mathbf{v}'\big) + \sum_{j=1}^{K} v^{(j)} \mathcal{D}_{i,j}. \tag{32}
\end{aligned}
$$

The last equality holds by introducing the following notation:

$$
\begin{aligned}
\Upsilon_i\big(\mathbf{v}\mathbf{v}'\big) &= \mathcal{A}_i \mathbf{v}\mathbf{v}' \mathcal{A}_i', \\
\mathcal{D}_{i,j} &= 1\big\{j = \theta(i)\big\}\big(\mathcal{R}_{\theta(i)} + s_{\theta(i)}(\mathcal{G}_{\theta(i)} + \mathcal{F}) + \big(s_{2,\theta(i)} - s_{\theta(i)}^2\big)(\mathbf{f} + \mathbf{g}_{\theta(i)})\big(\mathbf{f}' + \mathbf{g}_{\theta(i)}'\big)\big).
\end{aligned}
$$

Having established expressions for the moments of $\mathbf{A}_0$, we now focus on the moments of $\mathbf{B}_0$. Taking expectations in (28), we immediately find,

$$
\mathsf{E}\, \mathbf{B}_0(i) = w^{(i)}\mathbf{f}. \tag{33}
$$

Moreover, the second order moment can be expressed as,

$$
\mathsf{E}\big[\mathbf{B}_0(i)\mathbf{B}_n(j)'\big] = \mathsf{E}\big[W_0^{(i)} W_n^{(j)}\big]\mathbf{f}\mathbf{f}' = \mathcal{W}_n^{(i,j)}\,\mathbf{f}\mathbf{f}', \tag{34}
$$

for all $i, j = 1, \ldots, N$ and $n \neq 0$, whereas for $n = 0$, we have,

$$
\mathsf{E}\big[\mathbf{B}_0(i)\mathbf{B}_0(i)'\big] = \mathsf{E}\big[W_0^{(i)}\big]\mathcal{F} + \mathsf{E}\big[W_0^{(i)}\big]^2\mathbf{f}\mathbf{f}' = w^{(i)}\mathcal{F} + \mathcal{W}^{(i,i)}\mathbf{f}\mathbf{f}'. \tag{35}
$$

4.2 Queue content

As for the continuous state space model, we again rely on the theorems of the appendix, whereby expressions of the moments of $\mathbf{A}_0$ and $\mathbf{B}_0$ were derived in the preceding subsection. Now, Theorem 1 establishes the existence of a stationary solution of the queue content at polling instants while Theorem 2 provides expressions for the first two moments of this solution. Recall that $\boldsymbol{\mu}_i = \mathsf{E}[\mathbf{V}_0^* 1\{Y_0 = i\}]$ and $\boldsymbol{\Omega}_i = \mathsf{E}[\mathbf{V}_0^*(\mathbf{V}_0^*)' 1\{Y_0 = i\}]$. We now express various performance measures in terms of these vectors and matrices and in term of the

steady state vector $\boldsymbol{\pi}$ of the Markov chain Y_k. Again, we suppress the superscript asterisks that indicate the stationary solution for ease of notation.

For ease of notation, let $L_0 = \sum_{i=1}^{V_0^{(\theta(Y_0))}} S_0(i)$. Moreover, with a slight abuse of notation, let $\mathbf{V}_0(t)$ denote the queue content at time t during the 0th station time. The mean queue content at random points in time can then be expressed as follows,

$$\hat{\mathbf{v}} = \frac{\mathsf{E}\left[\int_0^{L_0+W_0^{(Y_0)}} \mathbf{V}_0(t)dt\right]}{\mathsf{E}\left[L_0 + W_0^{(Y_0)}\right]}.$$

We can easily calculate the denominator as follows,

$$\mathsf{E}\left[L_0 + W_0^{(Y_0)}\right] = \sum_{i=1}^{N} \mathsf{E}\left[\left(\sum_{j=1}^{V_0^{(\theta(i))}} S_0(j) + W_0^{(i)}\right) 1\{Y_0 = i\}\right]$$

$$= \sum_{i=1}^{N} \left(\mu_i^{(\theta(i))} s_{\theta(i)} + \pi^{(i)} w^{(i)}\right). \tag{36}$$

Calculation of the numerator is more involved. We have,

$$\mathsf{E}\left[\int_0^{L_0+W_0^{(Y_0)}} \mathbf{V}_0(t)\, dt\right] = \sum_{i=1}^{N} \mathsf{E}\left[\int_0^{L_0+W_0^{(i)}} \mathbf{V}_0(t)\, dt\, 1\{Y_0 = i\}\right]$$

$$= \sum_{i=1}^{N} \mathsf{E}\left[\left(\mathbf{V}_0 L_0 + \frac{L_0^2}{2}(\mathbf{f} + \mathbf{g}_{\theta(i)})\right) 1\{Y_0 = i\}\right]$$

$$+ \frac{1}{2} \sum_{i=1}^{N} \mathsf{E}\left[V_0^{(\theta(i))}\left(V_0^{(\theta(i))} - 1\right) s_{\theta(i)}(\mathbf{r}_{\theta(i)} - \mathbf{e}_{\theta(i)}) 1\{Y_0 = i\}\right]$$

$$+ \sum_{i=1}^{N} \mathsf{E}\left[\left(\mathbf{V}_0(L_0) W_0^{(i)} + \frac{(W_0^{(i)})^2}{2}\mathbf{f}\right) 1\{Y_0 = i\}\right]$$

The first term takes into account the new arrivals during service, as well as the feedback during service. The second term corresponds to the possible feedback at service completion epochs and to the departures. The last term accounts for the new arrivals during the walking time. Note that for the first and third term there were similar expressions in the continuous state space model while there is no corresponding term for the second term. Noting that,

$$\mathsf{E}\left[L_0^2 | \mathbf{V}_0, Y_0 = i\right] = V_0^{\theta(i)} s_{2,\theta(i)} + s_{\theta(i)}^2 V_0^{\theta(i)}\left(V_0^{\theta(i)} - 1\right),$$

and,

$$\mathsf{E}\left[\mathbf{V}_0(L_0) | \mathbf{V}_0, Y_0 = i\right] = \mathbf{V}_0 + L_0(\mathbf{f} + \mathbf{g}_{\theta(i)}) + (\mathbf{r}_{\theta(i)} - \mathbf{e}_{\theta(i)}) V_0^{(\theta(i))},$$

we find,

$$
\mathsf{E}\left[\int_0^{L_0+W_0^{(Y_0)}} \mathbf{V}_0(t)\,dt\right]
$$

$$
= \sum_{i=1}^N \mathsf{E}\left[\left(\mathbf{V}_0 s_{\theta(i)} + \frac{1}{2}\big(s_{2,\theta(i)} + s_{\theta(i)}^2\big(V_0^{(\theta(i))} - 1\big)\big)(\mathbf{f} + \mathbf{g}_{\theta(i)})\right) V_0^{(\theta(i))} 1\{Y_0 = i\}\right]
$$

$$
+ \frac{1}{2}\sum_{i=1}^N \mathsf{E}\left[V_0^{(\theta(i))}\big(V_0^{(\theta(i))} - 1\big) s_{\theta(i)}(\mathbf{r}_{\theta(i)} - \mathbf{e}_{\theta(i)}) 1\{Y_0 = i\}\right]
$$

$$
+ \sum_{i=1}^N \mathsf{E}\left[\left(\big(\mathbf{V}_0 + (\mathbf{f} + \mathbf{g}_{\theta(i)}) s_{\theta(i)} V_0^{(\theta(i))}\big)\right.\right.
$$

$$
\left.\left. + (\mathbf{r}_{\theta(i)} - \mathbf{e}_{\theta(i)}) V_0^{(\theta(i))}\right) W_0^{(i)} + \frac{1}{2}\big(W_0^{(i)}\big)^2 \mathbf{f}\right) 1\{Y_0 = i\}\right].
$$

Again, by the equality $V_0^{(\theta(i))} = \mathbf{e}'_{\theta(i)}\mathbf{V}_0 = \mathbf{V}'_0 \mathbf{e}_{\theta(i)}$ and in view of equation (29), the former expression further simplifies to,

$$
\mathsf{E}\left[\int_0^{L_0+W_0^{(Y_0)}} \mathbf{V}_0(t)\,dt\right]
$$

$$
= \frac{1}{2}\sum_{i=1}^N s_{\theta(i)}(\mathcal{I} + \mathcal{A}_i)\mathbf{\Omega}_i \mathbf{e}_{\theta(i)}
$$

$$
+ \frac{1}{2}\sum_{i=1}^N \big(s_{2,\theta(i)}(\mathbf{f} + \mathbf{g}_{\theta(i)})\mathbf{e}'_{\theta(i)} - s_{\theta(i)}(\mathcal{A}_i - \mathcal{I})\big)\boldsymbol{\mu}_i + \sum_{i=1}^N \mathcal{A}_i \mathbf{m}_i + \frac{1}{2}\sum_{i=1}^N \pi^{(i)}\mathcal{W}_0^{(i,i)}\mathbf{f}.
$$

where $\mathcal{I}$ denotes the identity matrix. As for the continuous state space model, the unknown vector $\mathbf{m}_i = \mathsf{E}\left[\mathbf{V}_0 W_0^{(i)} 1\{Y_0 = i\}\right]$ can be determined from the following equation,

$$
\mathcal{C}_i = \mathbf{m}_i \mathbf{f}'.
$$

Summarising, we have the following expression for the mean number of customers in the different buffers at random time instants,

$$
\hat{\mathbf{v}} = \frac{\sum_{i=1}^N [s_{\theta(i)}(\mathcal{I} + \mathcal{A}_i)\mathbf{\Omega}_i \mathbf{e}_{\theta(i)} + (s_{2,\theta(i)}(\mathbf{f} + \mathbf{g}_{\theta(i)})\mathbf{e}'_{\theta(i)} - s_{\theta(i)}(\mathcal{A}_i - \mathcal{I}))\boldsymbol{\mu}_i + 2\mathcal{A}_i \mathbf{m}_i + \pi^{(i)}\mathcal{W}_0^{(i,i)}\mathbf{f}]}{2\sum_{i=1}^N [\mu_i^{(\theta(i))} s_{\theta(i)} + \pi^{(i)} w^{(i)}]}.
$$

$$
(37)
$$

5 Conclusions

We considered polling systems with random routing and various types of feedback. Two complementary systems were analysed. The continuous state space model introduced semi-linear feedback. The discrete state space model combined classical feedback with Poisson feedback. Moreover, in either model, we relaxed the independence assumption on the walking times. The walking times constitute a sequence of stationary ergodic random variables. Our methodology relies on the framework of semi-linear equations in a random environment. We obtained the first and second order moments of workload and queue content at polling instants and the first moment of these quantities at random time instants.

Appendix: Semi-linear recursions in a random environment

Consider the sequence of random column vectors $\mathbf{V}_n \in \mathbb{R}^K$, adhering to the stochastic recursive equation,

$$\mathbf{V}_{n+1} = \mathbf{A}_n(\mathbf{V}_n, Y_n) + \mathbf{B}_n(Y_n), \quad n \in \mathbb{Z}. \tag{38}$$

Here $Y = \{Y_n\}$ denotes a Markov chain, taking values on a finite state-space $\Theta = \{1, 2, \ldots, N\}$ whereas $\mathbf{A}_n$ and $\mathbf{B}_n$ denote random vector-valued functions with domain $\mathbb{R}_+^K \times \Theta$ and Θ, respectively. The Markov chain Y_n is time-homogeneous, stationary ergodic and independent of $\mathbf{A}_n$ and $\mathbf{B}_n$. Let $p_{ij} = \Pr[Y_1 = j | Y_0 = i]$ denote the transition probability of the chain and let $P = [p_{ij}]$ denote the corresponding transition matrix. The steady-state probability vector is denoted by $\boldsymbol{\pi} = [\pi^{(i)}]_{i=1,\ldots,N}$. The consecutive $\mathbf{A}_n$ constitute a sequence of independent and identically distributed processes with the following properties:

- For each $i \in \Theta$, $\mathbf{A}_n(\cdot, i)$ has a divisibility property. Let $\mathbf{x} = \mathbf{x}_1 + \mathbf{x}_2 + \cdots + \mathbf{x}_k \in \mathbb{R}_+^K$, then $\mathbf{A}_n(\mathbf{x}, i)$ has the following representation,

$$\mathbf{A}_n(\mathbf{x}, i) = \sum_{l=1}^{k} \hat{\mathbf{A}}_n^{(l)}(\mathbf{x}_l, i), \tag{39}$$

 whereby $\hat{\mathbf{A}}_n^{(l)}(\cdot, i)$, $l = 1, \ldots, k$, are identically distributed, but not necessarily independent, with the same distribution as $\mathbf{A}_n(\cdot, i)$.
- For each $i \in \Theta$, $\mathbf{A}_n(\cdot, i)$ is linear in the mean,

$$\mathsf{E}\left[\mathbf{A}_n(\mathbf{x}, i)\right] = \mathcal{A}_i\, \mathbf{x}, \quad \mathbf{x} \in \mathbb{R}_+^K. \tag{40}$$

Here $\{\mathcal{A}_i, i \in \Theta\}$ is a set of fixed $K \times K$ matrices. Further, for each $i \in \Theta$, the correlation matrix of $\mathbf{A}_n(\mathbf{x}, i)$ is linear in $\mathbf{x}\mathbf{x}'$ and $\mathbf{x}$. For all $\mathbf{x} = [x^{(1)}, \ldots, x^{(K)}] \in \mathbb{R}_+^K$, we have the following representation,

$$\mathsf{E}\left[\mathbf{A}_n(\mathbf{x}, i)\mathbf{A}_n'(\mathbf{x}, i)\right] = \Upsilon_i\left(\mathbf{x}\mathbf{x}'\right) + \sum_{j=1}^{K} x^{(j)} \mathcal{D}_{i,j}. \tag{41}$$

For each $i \in \Theta$, Υ_i is a linear operator that maps $K \times K$ non-negative definite matrices on $K \times K$ non-negative definite matrices and satisfies $\Upsilon_i(0) = 0$. Further, $\{\mathcal{D}_{i,j}, i \in \Theta, j = 1, \ldots, K\}$ is a set of fixed $K \times K$ matrices.

Hence, for fixed n and i, $\mathbf{A}_n(\cdot, i)$ is a semi-linear process. Finally, the process $\{\mathbf{B}_n, n \in \mathbb{Z}\}$ is stationary ergodic. We introduce the following notation for its first and second order moments,

$$\mathbf{b}_i = \mathsf{E}\left[\mathbf{B}_0(i)\right], \quad \mathcal{B}_{ij}^{(n)} = \mathsf{E}\left[\mathbf{B}_0(i)\mathbf{B}_n(j)'\right]$$

Finally, the following block matrices and block vector are defined to simplify further notation,

$$\hat{\mathcal{A}} = \begin{bmatrix} \mathcal{A}_1 p_{11} & \mathcal{A}_2 p_{21} & \dots & \mathcal{A}_N p_{N1} \\ \mathcal{A}_1 p_{12} & \mathcal{A}_2 p_{22} & \dots & \mathcal{A}_N p_{N2} \\ \vdots & & \ddots & \vdots \\ \mathcal{A}_1 p_{1N} & \mathcal{A}_2 p_{2N} & \dots & \mathcal{A}_N p_{NN} \end{bmatrix}, \quad \hat{\mathbf{b}} = \sum_{i \in \Theta} \pi^{(i)} \begin{bmatrix} p_{i1} \\ p_{i2} \\ \vdots \\ p_{iN} \end{bmatrix} \otimes \mathbf{b}_i, \quad (42)$$

and,

$$\hat{\mathcal{B}}^{(n)} = \sum_{i \in \Theta} \pi^{(i)} \begin{bmatrix} \mathcal{B}_{i1}^{(n)} p_{i1} & \mathcal{B}_{i2}^{(n)} p_{i1} & \dots & \mathcal{B}_{iN}^{(n)} p_{i1} \\ \mathcal{B}_{i1}^{(n)} p_{i2} & \mathcal{B}_{i2}^{(n)} p_{i2} & \dots & \mathcal{B}_{iN}^{(n)} p_{i2} \\ \vdots & & \ddots & \\ \mathcal{B}_{i1}^{(n)} p_{iN} & \mathcal{B}_{i2}^{(n)} p_{iN} & \dots & \mathcal{B}_{iN}^{(n)} p_{iN} \end{bmatrix}. \quad (43)$$

For ease of notation, the following operator will prove useful: for any $\mathbf{x} \in \mathbb{R}^K$, let $\bigotimes_{i=n}^{k} \mathbf{A}_i(\mathbf{x}, Y_i) = \mathbf{x}$ for $k < n$ whereas, for $k \geq n$, this operator is defined by the following recursion,

$$\bigotimes_{i=n}^{k} \mathbf{A}_i(\mathbf{x}, Y_i) = \mathbf{A}_k \left(\bigotimes_{i=n}^{k-1} \mathbf{A}_i(\mathbf{x}, Y_i), Y_k \right).$$

The operator above can be applied likewise on $\hat{\mathbf{A}}_i^{(l)}$ for each l. We now state the stability theorem.

Theorem 1 *Assume that* (i) $\mathbf{b}_i < \infty$ *component-wise for all* $i \in \Theta$; *and* (ii) *that all the eigenvalues of the matrix* $\hat{\mathcal{A}}$ *are within the open unit disk. Then, there exist a unique stationary solution* $\mathbf{V}_n^*$, *distributed like,*

$$\mathbf{V}_n^* =_d \sum_{j=0}^{\infty} \bigotimes_{i=n-j}^{n-1} \hat{\mathbf{A}}_i^{(n-j)} (\mathbf{B}_{n-j-1}(Y_{n-j-1}), Y_i), \quad (44)$$

for $n \in \mathbb{Z}$. *The sum on the right side of the former expression converges absolutely almost surely. Furthermore, one can construct a probability space such that* $\lim_{n \to \infty} \|\mathbf{V}_n - \mathbf{V}_n^*\| = 0$, *almost surely, for any initial value* $\mathbf{V}_0$.

Assume that one of the eigenvalues of $\hat{\mathcal{A}}$ *is outside the closed unit disk. Then the system is unstable*: $\|\mathsf{E}[\mathbf{V}_n]\|$ *converges to infinity for some initial states.*

Proof The sufficiency condition for stability is established in (Fiems and Altman 2009b). We establish the necessary condition for stability. Define $\boldsymbol{\mu}_i(n) \triangleq \mathsf{E}[\mathbf{V}_n 1\{Y_n = i\}]$, $i \in \Theta$. Then it is shown in (Fiems and Altman 2009b) that

$$\boldsymbol{\mu}(n+1) = \hat{\mathcal{A}} \boldsymbol{\mu}(n) + \hat{\mathbf{b}}.$$

$\hat{\mathcal{A}}$ is a positive matrix. By the Perron-Frobenius theorem, the eigenvalue of $\hat{\mathcal{A}}$ with the largest norm (we call this eigenvalue ζ) is real and positive. By the assumption in the second part of the Theorem, $\zeta > 1$. The Perron-Frobenius theorem also implies that $\hat{\mathcal{A}}$ has a corresponding left eigenvector whose entries are all non-negative with at least one of them strictly positive.

If we start with that initial state or with any larger one, then $\mu(n)$ diverges. This also implies the divergence for some initial state of the form $\mathbf{V}_0$ which has just one nonzero element. This establishes the necessary condition for stability. □

With the stability conditions established, we now focus on expressions for the first and second moments of $\mathbf{V}_0^*$, conditioned on the state of the Markov chain Y_0. Let $\hat{\boldsymbol{\mu}}$, the conditional first moment vector, be the block column vector with elements $\boldsymbol{\mu}_i \triangleq \mathsf{E}\,[\mathbf{V}_0^* 1\{Y_0 = i\}]$, $i \in \Theta$. Analogously, let $\hat{\boldsymbol{\Omega}}$, the conditional second moment matrix, be the block column vector with elements $\boldsymbol{\Omega}_i \triangleq \mathsf{E}\,[\mathbf{V}_0^*(\mathbf{V}_0^*)' 1\{Y_0 = i\}]$, $i \in \Theta$. The following theorem provides expressions for these vectors.

Theorem 2 *Assume that the stability conditions of Theorem 1 are satisfied. The conditional first moment vector is then given by,*

$$\hat{\boldsymbol{\mu}} = (\mathcal{I} - \hat{\mathcal{A}})^{-1}\hat{\mathbf{b}}. \tag{45}$$

Under the additional assumption that the second moments of $\mathbf{B}_0(i)$ are finite, $i \in \Theta$, the elements $\boldsymbol{\Omega}_i$ of the conditional second moment matrix of $\mathbf{V}_0^*$ are the unique solution of the system of equations,

$$\boldsymbol{\Omega}_l = \sum_{k \in \Theta} \left(\Upsilon_k(\boldsymbol{\Omega}_k) + \sum_{j=1}^{K} \mu_k^{(j)} \mathcal{D}_{k,j} + \mathcal{B}_{kk}^{(0)} \pi^{(k)} + \mathcal{A}_k \mathcal{C}_k + \mathcal{C}_k' \mathcal{A}_k' \right) p_{kl}, \tag{46}$$

$l \in \Theta$, with $\mu_k^{(j)}$ the jth element of $\boldsymbol{\mu}_k$ and where $\mathcal{C}_k$ denotes the kth diagonal (block) element of

$$\hat{\mathcal{C}} = \sum_{j=0}^{\infty} \hat{\mathcal{A}}^j \hat{\mathcal{B}}^{(j+1)}. \tag{47}$$

References

Altman, E. (2009). Semi-linear stochastic difference equations. *Discrete Event Dynamic Systems*, *19*, 115–136.

Altman, E., & Fiems, D. (2007). Expected waiting time in polling systems with correlated vacations. *Queueing Systems*, *56*(3–4), 241–253.

Altman, E., & Yechiali, U. (1993). Cyclic Bernoulli polling. *ZOR Zeitschrift Für Operations-Research*, *38*(1), 55–76.

Boxma, O. J., & Weststrate, J. A. (1989). Waiting times in polling systems with Markovian server routing. *Informatik-Fachberichte*, *218*, 89–104.

Chang, B. J., Chou, C. M., & Liang, Y. H. (2008). Markov chain analysis of uplink subframe in polling-based WiMAX networks. *Computer Communications*, *31*, 2381–2390.

Chung, H., Un, C., & Jung, W. (1994). Performance analysis of Markovian polling systems with single buffers. *Performance Evaluation*, *19*(4), 303–315.

de Moraes, L. (1990). Comments on 'Analysis and applications of a multiqueue cyclic service system with feedback'. *IEEE Transactions on Communications*, *38*, 148–149.

Evans, B. G. (1999). *Satellite communication systems* (3rd ed.). Stevenage: IET.

Fayolle, G., & Lasgouttes, J. (1995). A state-dependent polling model with Markovian routing. In F. P. Kelly & R. R. Williams (Eds.), *The IMA volumes in mathematics and its applications* (Vol. 71, pp. 283–312). Berlin: Springer.

Fiems, D., & Altman, E. (2009a). Applying branching processes to delay-tolerant networks. In *Proceedings of bionetics 2009*, Avignon, France.

Fiems, D., & Altman, E. (2009b). *Markov-modulated stochastic recursive equations with applications to delay-tolerant networks* (Rapport de Recherche, RR-6872). INRIA.

Groenevelt, R., & Altman, E. (2005). Analysis of alternating-priority queueing models with (cross) correlated switchover times. In *Proceedings of IEEE infocom 2005*, Miami.

Gupta, D., & Buzacott, J. (1990). A production system with two job classes, changeover times and revisitation. *Queueing Systems*, *6*, 353–368.

Hirayama, T. (2005). Multiclass polling systems with Markovian feedback: mean sojourn times in gated and exhaustive systems with local priority and FCFS service orders. *Journal of the Operations Research Society of Japan*, *48*(3), 226–255.

Kleinrock, L., & Levy, H. (1988). The analysis of random polling systems. *Operations Research*, *36*(5), 716–732.

Lee, T. Y. S. (2003). Analysis of random polling system with an infinite number of coupled servers and correlated input process. *Computers & Operations Research*, *30*(13), 2003–2020.

Lee, T. Y. S., & Sunjaya, J. (1996). Exact analysis of asymmetric random polling systems with single buffers and correlated input process. *Queueing Systems*, *23*(1–4), 131–156.

Levy, H., & Sidi, M. (1990). Polling systems: applications, modeling, and optimization. *IEEE Transactions on Communications*, *38*(10), 1750–1760.

Resing, J. (1993). Polling systems and multi-type branching processes. *Queueing Systems*, *13*, 409–426.

Schwartz, H., Marpe, D., & Wiegand, T. (2007). Overview of the scalable video coding extension of the H.264/AVC standard. *IEEE Transactions on Circuits and Systems for Video Technology*, *17*(9), 1103–1120.

Sidi, M., Levy, H., & Fuhrmann, S. (1992). A queueing network with a single cyclically roving server. *Queueing Systems*, *11*(1–2), 121–144.

Srinivasan, M. M. (1991). Nondeterministic polling systems. *Management Science*, *37*(6), 667–681.

Staehle, D., & Pries, R. (2007). Comparative study of the IEEE 802.16 random access mechanisms. In *Proceedings of the international conference on next generation mobile applications, services and technologies (NGMAST 2007)*.

Takács, L. (1963). A single server queue with feedback. *The Bell System Technical Journal*, *42*, 505–519.

Takagi, H., & Kleinrock, L. (1985). *A tutorial on the analysis of polling systems* (Technical Report CSD-850005). University of California.

Takagi, H. (1987). A survey of queueing analysis of polling models. In *Proceedings of the third IFIP international conference on data communication systems and their performance*, Rio de Janeiro, Brazil, 22–25 June.

Takagi, H. (1987). Analysis and applications of a multiqueue cyclic service system with feedback. *IEEE Transactions on Communications*, *35*(2), 248–250.

Takine, T., Takagi, H., & Hasegawa, T. (1991) Sojourn times in vacation and polling systems with Bernoulli feedback. *Journal of Applied Probability*, *28*(2), 422–432.

Vishnevskii, V. M., & Semenova, O. V. (2006). Mathematical methods to study the polling systems. *Automation and Remote Control*, *2*, 3–56.