Expected waiting time in symmetric polling systems with correlated walking times

Eitan Altman* Dieter Fiems†

June 9, 2007

Abstract Polling systems have been extensively studied, and have found many applications. They have often been used for studying wired local areas networks such as token passing rings and wireless local area networks such as bluetooth. In this contribution we relax one of the main restrictions on the statistical assumptions under which polling systems have been analyzed. Namely, we allow correlation between walking times. We consider (i) the gated regime where a gate closes whenever the server arrives at a queue. It then serves at that queue all customers who were present when the gate closes. (ii) the exhaustive regime in which the server remains at a queue till it empties.

Our analysis is based on stochastic recursive equations related to branching processes with migration with a random environment. In addition to our derivation of expected waiting times for polling systems with correlated walking times, we set the foundations for computing second order statistics of the general multi-dimensional stochastic recursions.

1 Introduction

Polling systems have been studied extensively over the last 20 years, and found many applications in various areas of performance evaluation. They have often been used for studying wired local areas networks such as token passing rings [1] and wireless local area networks such as bluetooth [2]. They have also been useful for analyzing access to a disk [3]. Polling systems are one of the few multidimensional queueing systems for which explicit solutions for the expected waiting times have been available. The reader is referred to Takagi's monograph and its supplement [4, 5] for analytical results and to Yechiali [6] and Lévy and Sidi [7] for surveys on applications.

In this paper we relax one of the main restrictions on the statistical assumptions under which polling systems have been analyzed. Namely, we allow correlation between walking times: the walking times constitute a stationary ergodic series of random variables and as such no restrictions are posed on the shape of the autocorrelation function of the walking times. As an example of systems that may have such correlation, consider a wireless LAN where an access point (the "server") polls mobiles according to some order. Assume that there is some signaling traffic between the access point and a mobile that is going to be polled, for example in order to receive the information of how many packets are awaiting for an uplink transmission from the mobile to the access point. (The signaling is thus used for reservations of the number of slots needed in order to transfer the packets present at the mobile.) Further signaling could be used at the end of a polling period of a mobile. Assume that the access point is aware of the radio channel state to each mobile and that the transmission rate of the signaling traffic is a function of the channel state. The duration of signaling could be modeled as part of the "walking times" that the server takes between periods of service of two consecutive mobiles. In this example there can be correlation between the radio conditions of a polled mobile and the radio condition of the next mobile to be polled (this is a spatial correlation). Further correlation can be due to the fact that switching time

*INRIA, BP93, 06902 Sophia Antipolis, France. Email: altman@sophia.inria.fr
†SMACS Research Group, Ghent University, Belgium. Email: Dieter.Fiems@UGent.be

to a mobile and switching from the same mobile (after its packets have been received) are likely to occur under similar radio conditions (temporal correlation).

We consider in this paper gated and exhaustive polling systems, i.e. systems in which the server remains with a queue until all customers are served that were present upon arrival of the server at the queue (gated) or until there are no more customers in the queue (exhaustive). In terms of our example of an access point that polls mobiles, the gated model is natural since the access point's information on the number of packets to expect from a mobile is based on the reservation signaling from that mobile that occurs just before transmission from the mobile starts.

Our analysis is based on stochastic recursive equations [8–10] of a form that is related to branching processes with migration in a random environment. Branching processes find their origins in the work of Bienaymé [11] and Galton and Watson [12], the first asymptotic result in the theory of branching processes being obtained by Kolmogorov [13]. The first reference on branching with migration is [14]. Multi-type branching processes in a varying environment without migration have been studied in [15]. Further, an overview on branching processes can be found in [16, 17].

There is a close connection between branching processes and polling systems. Already Resing [18] demonstrated the fact that the number of customers at polling instants can be described as a discrete multi-type branching processes with migration. A similar branching structure with a continuous state space was shown to describe the so called "station times" of polling systems. 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 [19]. However, this approach did not extend to more than two queues. The basic obstacle in extending the analysis to a polling system with more than two queues has been that expected waiting times require to derive second order properties of the stochastic recursive equations.

Branching processes have also been identified in other queueing models including infinite server queues [20] and processor sharing queues [21, 22]. Although some stochastic processes appearing in queueing models correspond to continuous branching processes (CBP) (an example is the station times in polling systems [23]), it seems that the queueing theory community has only been aware of the formalism of discrete branching processes. For background on CBP, see [24].

In this paper we compute the second moment and correlation in multi-type branching processes with a stationary ergodic migration process. Our framework has its origin in [25], which has studied a general form of stochastic recursive equations that applies in particular to the model in this paper. In our present contribution we make use of those results to derive the first two moments of multi-type branching processes with stationary ergodic migration. The second moment results we obtain here allow us to compute the expected waiting times in a polling system with any number of queues.

The contribution of this paper is thus not only in analyzing polling systems with correlated walking times but also in setting the foundations for computing second order statistics of the stochastic under non Markov setting.

The remainder of this contribution is organized as follows. In the next section we present the continuous state branching model with stationary ergodic migration, for which we obtain the first two moments in stationary regime. Some background for this section is delayed to the Appendix. We apply these results to a symmetric polling system with the gated regime in Section 3 and with an exhaustive regime in Section 4. Finally, conclusions are drawn in Section 5.

2 Branching model: two first moments

Our starting point is the following stochastic recursive equation,

$$Y_{n+1} = A_n(Y_n) + B_n \, . \tag{1}$$

For each given integer n, A_n is a nonnegative Additive Lévy field: it is an extension of a stochastic process where the "time" parameter is not a scalar but a vector in $\mathbb{R}_+^m$. For each A_n and for each

$y \in \mathbb{R}_+^m$, $A_n(y)$ takes values in $\mathbb{R}_+^m$. As a function of n, the random fields $A_n(.)$ are assumed to be independent and identically distributed (i.i.d.). The reader is referred to the appendix for the definition of Additive Lévy fields.

The series $\{B_n\}$ is a stationary ergodic series of m-dimensional column vectors whose entries take values in the nonnegative reals $\mathbb{R}^+$. Further, the process $\{B_n\}$ is assumed to be independent of the process $\{A_n\}$. We call the process $\{Y_n\}$ a continuous-state branching process (see [26] where related one-dimensional processes that have the above form are defined, as well as [24] and references therein). B_n is then called the migration process. Let $\mathcal{B}(k)$ be the matrix $\mathrm{E}[B_0(B_k)^T]$, where k is an integer and let $\hat{\mathcal{B}}(k)$ be defined as $\mathcal{B}(k) - \mathrm{E}[B_0]\,\mathrm{E}[B_0]^T$. Notice that in particular $\hat{\mathcal{B}}(0)$ equals the covariance matrix $\mathrm{cov}[B_0]$ of the random vector B_0.

Although the stochastic recursive equation (1) is not linear, it is linear in expectation. That is, for any $y \in \mathbb{R}_+^m$,

$$\mathrm{E}[A_n(y)] = \mathcal{A}y\,, \tag{2}$$

for some matrix $\mathcal{A}$. The latter matrix is defined in the Appendix, see expression (40). Moreover, we have for $j > 1$,

$$\mathrm{E}\left[\left(\bigotimes_{i=1}^{j} A_j\right)(y)\right] = \mathcal{A}^j\, y\,. \tag{3}$$

Here we understand $\bigotimes_{i=n}^{k} A_i(x) = x$ whenever $k < n$, and $\bigotimes_{i=n}^{k} A_i(x) = A_k(A_{k-1}(\ldots(A_n(x))))$ whenever $k > n$.

In the following Theorem, we obtain expressions for the mean vector and covariance matrix of the stationary solution. Let $\|\cdot\|$ denote the Euclidean norm.

Theorem 1. *Consider the stochastic recursive equation (1) where A_n are i.i.d. Additive Lévy processes, independent of the sequence B_n which is assumed to be stationary ergodic. Assume that all eigenvalues of $\mathcal{A}$ are within the unit disk and that the elements of B_0 have finite second order moments. Then, there exists a unique stationary solution $\{Y_n^*\}$ to (1); moreover, one can construct a probability space on which both $\{Y_n\}$ as well as $\{Y_n^*\}$ are defined such that $\lim_{n\to\infty} \|Y_n - Y_n^*\| = 0$, $P-a.s.$ The stationary limit Y_n^* satisfies the following:*

(i) the first moment of Y_0^* is given by

$$\mathrm{E}[Y_0^*] = (\mathcal{I} - \mathcal{A})^{-1}\,\mathrm{E}[B_0]\,, \tag{4}$$

where $\mathcal{I}$ denotes the $m \times m$ identity matrix.

(ii) $\mathrm{cov}(Y_0^*)$ is given as the unique solution of the following set of linear equations:

$$\mathcal{Z} = \sum_{j=1}^{m} \mathrm{E}[Y_0^j]\Gamma^{(j)} + \mathcal{A}\mathcal{Z}\mathcal{A}^T + \mathrm{cov}[B_0] + \sum_{j=1}^{\infty} \mathcal{A}^j\hat{\mathcal{B}}(j) + (\mathcal{A}^j\hat{\mathcal{B}}(j))^T\,, \tag{5}$$

where $\mathrm{E}[Y_0^j]$ denotes the jth element of the vector $\mathrm{E}[Y_0^*]$, $\mathcal{Z}$ is the $m \times m$ matrix to be solved for and the matrices $\Gamma^{(j)}$ are defined in eq. (38) in the Appendix.

Proof. The existence, uniqueness of the stationary regime and the convergence to it have been established in [25]. In the sequal we shall understand Y_n to denote Y_n^*.

Taking expectation in (1) we have

$$\mathrm{E}[Y_0] = \mathcal{A}\,\mathrm{E}[Y_0] + \mathrm{E}[B_0],$$

and we immediately obtain (4) since the fact that the eigenvalues of $\mathcal{A}$ are within the unit disk implies that $(\mathcal{I} - \mathcal{A})$ is a non-singular matrix.

Further, multiplying both sides of (1) by their transpose, taking expectation and using the stationarity yields,

$$\mathrm{E}[Y_0Y_0^T] = \mathrm{E}[A_0(Y_0)A_0^T(Y_0)] + \mathrm{E}[B_0B_0^T] + \mathrm{E}[A_0(Y_0)B_0^T] + \mathrm{E}[B_0A_0^T(Y_0)]\,.$$

The covariance matrix $\mathrm{cov}[Y_0]$ therefore equals,

$$\begin{aligned}\mathrm{cov}[Y_0] &= \mathrm{cov}[A_0(Y_0)] + \mathrm{cov}[B_0] + \mathrm{E}\left[A_0(Y_0)B_0^T\right] \\ &\quad -\mathcal{A}\,\mathrm{E}[Y_0]\,\mathrm{E}[B_0]^T + \mathrm{E}\left[B_0A_0(Y_0)^T\right] - \mathrm{E}[B_0](\mathcal{A}\,\mathrm{E}[Y_0])^T\,. \end{aligned} \tag{6}$$

In view of property (42) of Additive Lévy processes, we further find,

$$\mathrm{cov}[A_0(Y_0)] = \sum_{j=1}^{m} \mathrm{E}[Y_0^j]\Gamma^{(j)} + \mathcal{A}\,\mathrm{cov}[Y_0]\mathcal{A}^T\,. \tag{7}$$

The stationary solution of the recursive equation (1) is distributed as the right hand side of expression (46) of Theorem 5 in the appendix. Therefore we find,

$$\begin{aligned}\mathrm{E}[Y_0B_0^T] &= \sum_{j=0}^{\infty}\mathrm{E}\left\{\bigotimes_{i=-j}^{-1} A_{-j,i}(B_{-j-1})B_0^T\right\} \\ &= \sum_{j=0}^{\infty}\mathrm{E}\left(\mathrm{E}\left\{\bigotimes_{i=-j}^{-1} A_{-j,i}(B_{-j-1})B_0^T\right\}\middle|\,\mathbf{B_0^-}\right) \\ &= \sum_{j=0}^{\infty}\mathrm{E}\left(\mathcal{A}^jB_{-j-1}B_0^T\right) = \sum_{j=0}^{\infty}\mathcal{A}^j\mathcal{B}(j+1)\,, \end{aligned} \tag{8}$$

with $\mathbf{B_0^-} := (B_0, B_{-1}, B_{-2}, ...)$. Recall that $\mathcal{B}(k) = \mathrm{E}[B_0(B_k)^T]$. Notice that the sums in the last line are finite since the finiteness of the second moments of the elements of B_0 implies that $\mathcal{B}(j)$ is uniformly bounded and since all eigenvalues of $\mathcal{A}$ are within the unit disk. Finally, in view of the former expression, we compute,

$$\mathrm{E}[A_0(Y_0)B_0^T] = \mathrm{E}\left[\mathrm{E}\left[A_0(Y_0)B_0^T\,\middle|\,Y_0, B_0\right]\right] = \mathcal{A}\,\mathrm{E}\left[Y_0B_0^T\right] = \sum_{j=1}^{\infty}\mathcal{A}^j\mathcal{B}(j)\,,$$

or equivalently,

$$\begin{aligned}\mathrm{E}[A_0(Y_0)B_0^T] &= \sum_{j=1}^{\infty}\mathcal{A}^j\hat{\mathcal{B}}(j) + \sum_{j=1}^{\infty}\mathcal{A}^j\,\mathrm{E}[B_0]\,\mathrm{E}[B_0]^T \\ &= \sum_{j=1}^{\infty}\mathcal{A}^j\hat{\mathcal{B}}(j) + \mathcal{A}(\mathcal{I}-\mathcal{A})^{-1}\,\mathrm{E}[B_0]^T \\ &= \sum_{j=1}^{\infty}\mathcal{A}^j\hat{\mathcal{B}}(j) + \mathcal{A}\,\mathrm{E}[Y_0]\,\mathrm{E}[B_0]^T\,. \end{aligned} \tag{9}$$

Substitution of expressions (7) and (9) into expression (6) then yields (5).

Next, we show uniqueness. Let Z_1 and Z_2 be two solutions of (6) and define $Z = Z_1 - Z_2$. Then Z satisfies $Z = \mathcal{A}^TZ\mathcal{A}$ in view of equation 5. Iterating the former expression we obtain,

$$Z = \lim_{n\to\infty}\mathcal{A}^nZ(\mathcal{A}^T)^n = 0$$

where the last equality follows from the fact that all the eigenvalues of $\mathcal{A}$ are within the unit disk. This implies the uniqueness of the solution of (5). □

3 Symmetric gated polling systems

We now consider a polling system with a gated service discipline and with correlated walking times. The server polls m queues and the workload arrival processes into the different queues are modeled by means of independent subordinators – Lévy processes with increasing sample paths – distributed as some (generic) subordinator $\mathcal{R}(t),\ t \in \mathbb{R}_+$. For further use, let $\overline{\rho} = \mathrm{E}[\mathcal{R}(1)]$ and $\sigma^2 = \mathrm{var}[\mathcal{R}(1)]$ denote the mean and variance of $\mathcal{R}(1)$. Also, the Itô decomposition states that a subordinator decomposes into a Poisson process and a constant flow. Let λ and r denote the Poisson arrival intensity and the flow rate respectively and let p_1 and p_2 denote the first two moments of the Poisson jumps. Notice that $\overline{\rho} = r + \lambda p_1$. The walking times are assumed to constitute a stationary ergodic series $\{V_n\}$ of nonnegative random variables and the average walking time is denoted by $v = \mathrm{E}[V_0]$. For further use, let $\mathcal{V}(j) = \mathrm{E}[V_0 V_j]$ for some integer j and let $\hat{\mathcal{V}}(j) = \mathrm{E}[V_0 V_j] - v^2$. Notice that $\hat{\mathcal{V}}(0)$ equals the variance $\mathrm{var}[V_0]$ of the random variable V_0. Also, in the case that the consecutive walking times are independent, we have $\hat{\mathcal{V}}(j) = 0$ for all $j = 1, 2, \ldots$.

3.1 Sample-path modeling as a stochastic recursive equation

There are m queues visited by the server in a cyclic way: $1, 2, ..., m-1, m, 1, 2, ...$ When the server has completed all the work it found upon arrival at a queue, the server requires a walking time during which it idles. Then the server moves to the next queue.

We consider the polling system at polling instants. That is, at time instants where the server arrives at a queue. Let $S(n)$ denote the nth polling instant and let $I(n) = ((n-1) \mod m) + 1$ denote the queue that the server visits at the nth polling instant. Further, the walking time to move to queue $I(n+1)$ is denoted by V_n. Let

$$Y_n^i := S(n) - S(n-i), \qquad (i = 1, 2, ..., m)\,.$$

It is the time between the $(n-i)$th and the nth polling instant. The workload arrival process at queue i is described by a subordinator $\mathcal{R}^i(t)$ with time parameter $t \in \mathbb{R}_+$, which is distributed as $\mathcal{R}(t)$. Let $\mathcal{R}_n^i$ be i.i.d. copies of $\mathcal{R}^i$ for all integer n. We can then describe the dynamics of the gated polling system through the following set of equations:

$$\begin{aligned}
Y_{n+1}^1 &= S(n+1) - S(n) = \mathcal{R}_n^m(Y_n^m) + V_n\,, \\
Y_{n+1}^2 &= S(n+1) - S(n-1) = Y_n^1 + \mathcal{R}_n^m(Y_n^m) + V_n\,, \\
Y_{n+1}^3 &= S(n+1) - S(n-2) = Y_n^2 + \mathcal{R}_n^m(Y_n^m) + V_n\,, \\
&\vdots \\
Y_{n+1}^m &= S(n+1) - S(n-m+1) = Y_n^{m-1} + \mathcal{R}_n^m(Y_n^m) + V_n\,.
\end{aligned} \tag{10}$$

Equation (10) states that the time between $S(n)$ and $S(n+1)$ is the sum of the amount of work that the server finds in queue $I(n)$ at time $S(n)$ and the nth walking time. The server finds in queue $I(n)$ at time $S(n)$ all work that arrived in this queue since time $S(n-m)$, i.e., since the previous time that the server polled queue $I(n)$. Note that we here implicitly used the independent increments property of the workload arrival processes.

In vector notation we have

$$Y_{n+1} = A_n(Y_n) + B_n\,,$$

with

$$B_n = V_n \cdot (1, 1, ..., 1)^T \quad \text{and} \quad A_n(y) = A_n^{(1)}(y_1) + ... + A_n^{(m)}(y_m)\,, \tag{11}$$

for $y = (y_1, ..., y_m)^T \in \mathbb{R}_+^m$, and with

$$\begin{aligned} A_n^{(1)}(t) &= (0, t, 0, 0, ..., 0)^T, \\ A_n^{(2)}(t) &= (0, 0, t, 0, ..., 0)^T, \\ &\vdots \\ A_n^{(m-1)}(t) &= (0, 0, 0, ..., 0, t)^T, \\ A_n^{(m)}(t) &= \mathcal{R}_n^m(t)(1, 1, \ldots, 1)^T, \end{aligned} \tag{12}$$

for $t \in \mathbb{R}_+$.

Y_n can be viewed as the state variables of a Markov chain in the special case that the series $\{B_n\}$ is i.i.d. too. Different state variables have been used before in this Markovian case. Takagi [27] uses the "buffer occupancy" approach where the state is the number of customers at each queue at polling instants. Another well known alternative is the use of station times as states, where a station time is the time spent at a station plus the walking time from that station to the next one. The advantage in our choice of state vector is that one of its components equals the cycle time (see further), whose first two moments, as we shall see, are precisely what we need for computing the expected waiting time.

Finally, notice that the processes A_n are Additive Lévy processes and that the series B_n is a stationary ergodic series in $\mathbb{R}_+^m$ which implies that we can use the framework developed in the preceding section.

3.2 First and second moment

In accordance with the definition of the matrix $\mathcal{A}$ (see Appendix) and from equation (12) it follows that,

$$\mathcal{A} = \begin{pmatrix} 0 & 0 & 0 & 0 & \ldots & 0 & \overline{\rho} \\ 1 & 0 & 0 & 0 & \ldots & 0 & \overline{\rho} \\ 0 & 1 & 0 & 0 & \ldots & 0 & \overline{\rho} \\ 0 & 0 & 1 & 0 & \ldots & 0 & \overline{\rho} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \ldots & 0 & \overline{\rho} \\ 0 & 0 & 0 & 0 & \ldots & 1 & \overline{\rho} \end{pmatrix}. \tag{13}$$

We shall sometimes use the notation $\mathcal{A}(\overline{\rho})$ to stress the dependence on $\overline{\rho} = \mathrm{E}\,\mathcal{A}(1)$. The latter matrix then satisfies the following theorem.

Theorem 2. *A sufficient and necessary condition for all eigenvalues of $\mathcal{A}$ to be in the interior of the unit circle is*

$$\overline{\rho} < \frac{1}{m}.$$

Proof. $\mathcal{A}$ is known as the *Companion matrix*, and its eigenvalues are given as the m roots of the polynomial equation,

$$P_m(z) = z^m - \overline{\rho}(1 + z + ... + z^{m-1}) = 0, \tag{14}$$

see Horn and Johnson [28, pp. 146-147]. Choose some z with $|z| \geq 1$. If $\bar{\rho} < 1/m$, then

$$|P_m(z)| \geq |z|^m - \overline{\rho} \sum_{i=0}^{m-1} |z|^i > 0.$$

We conclude that $\overline{\rho} < 1/m$ is a sufficient condition for all eigenvalues of $\mathcal{A}$ to be in the interior of the unit circle.

If $\overline{\rho} \geq 1/m$ then at least one of the eigenvalues of $\mathcal{A}$ is not within the interior of the unit disk. To see that, we note that the matrix $\kappa = \mathcal{A}(1/m)$ is the transposed of a stochastic matrix and therefore has an eigenvalue of 1. For $\overline{\rho} \geq 1/m$, each entry of $\mathcal{A}(\overline{\rho})$ is greater than or equal to the corresponding entry of κ. We can then apply Theorem 8.4.5 of Horn and Johnson [28] to conclude that $\mathcal{A}(\overline{\rho})$ has an eigenvalue not contained in the interior of the unit disk. This establishes the necessity of the condition. $\square$

We conclude that the conditions of Theorem 1 hold if and only if $\overline{\rho} < 1/m$. The steady state expectation of Y_0 is then given by,

$$\mathrm{E}[Y_0] = (\mathcal{I} - \mathcal{A})^{-1}\,\mathrm{E}[B_0] = \frac{v}{1 - m\overline{\rho}} \cdot (1, 2, 3, ..., m)^T .$$

Recall that the covariance matrix of $A(y)$ is given by $\mathrm{cov}[A(y)] = \sum_{j=1}^{m} y_j \Gamma^{(j)}$, where $\Gamma^{(j)}$ is the corresponding covariance matrix of $A^{(j)}(1)$ and where y_j denotes the jth element of the vector y (see Appendix). From (12) one finds that for all $j \neq m$, $\Gamma^{(j)}$ is an $m \times m$ matrix whose elements are all zero. Indeed, this follows from,

$$\mathrm{E}[A^{(j)}(1) \cdot (A^{(j)}(1))^T] = \mathrm{diag}(0, 0, ..., 1, ..., 0) = \mathrm{E}[A^{(j)}(1)] \cdot \mathrm{E}[(A^{(j)}(1))^T] ,$$

where the 1 in the diagonal matrix is in position $j+1$. It remains to compute $\Gamma^{(m)}$. Clearly, from (12) we find, $A_n^{(m)}(1) = \mathcal{R}_n^m(1) \cdot (1, 1, \ldots, 1)^T$ and therefore,

$$\Gamma^{(m)} = \sigma^2 \mathcal{E} ,$$

where $\mathcal{E}$ denotes an $m \times m$ matrix with all elements equal to 1. Further, in view of equation (11), we find,

$$\mathrm{cov}[B_0] = \hat{\mathcal{V}}(0) \cdot \mathcal{E} , \quad \mathcal{B}(j) = \mathcal{V}(j) \cdot \mathcal{E} , \quad \hat{\mathcal{B}}(j) = \hat{\mathcal{V}}(j) \cdot \mathcal{E} .$$

Hence, (5) simplifies to

$$\mathrm{cov}[Y_0] = \frac{mv\sigma^2}{1 - m\overline{\rho}}\mathcal{E} + \mathcal{A}\,\mathrm{cov}[Y_0]\mathcal{A}^T + \hat{\mathcal{V}}(0)\mathcal{E} + \sum_{j=1}^{\infty} \hat{\mathcal{V}}(j) \left[\mathcal{A}^j\mathcal{E} + (\mathcal{A}^j\mathcal{E})^T\right] . \tag{15}$$

We conclude that the covariance of Y_0 in stationary regime is given by the unique solution of (15). Note that the last term in (15) disappears for the case of independent walking times.

3.3 Performance measures

We now find expressions for various performance measures of the polling system under consideration.

Cycle time and busy time. Let the cycle time be defined as the time between (the start of) two consecutive visits of the server to a queue. In particular, let the nth cycle time be defined as,

$$C_n = S(n + m) - S(n) .$$

C_n is thus the time between the arrival of the server at the nth queue that it visits, until the next time it arrives at that queue. Clearly, we have $C_n = Y_n^m$ and therefore,

$$\mathrm{E}[C_0] = \mathrm{E}[Y_0^m] = \frac{mv}{1 - m\overline{\rho}} , \quad \mathrm{var}[C_0] = \mathrm{var}[Y_0^m] , \quad \mathrm{E}[{C_0}^2] = \mathrm{var}[Y_0^m] + \mathrm{E}[Y_0^m]^2$$

with $\mathrm{var}[Y_0^m]$ the bottom right element of the matrix $\mathrm{cov}[Y_0]$.

Let the nth busy time G_n be defined as the part of the nth cycle time during which the server attends queue $I(n)$. The gated polling policy implies that the length of the nth busy time equals the time to serve the amount of work that arrived in the queue during cycle C_{n-m}. That is,

$$G_n = \mathcal{R}(C_{n-m}) .$$

Therefore, we find,

$$\mathrm{E}[G_0] = \overline{\rho}\,\mathrm{E}[C_0] , \quad \mathrm{var}[G_0] = \sigma^2\,\mathrm{E}[C_0] + \overline{\rho}^2\,\mathrm{var}[C_0] , \quad \mathrm{E}[{G_0}^2] = \sigma^2\,\mathrm{E}[C_0] + \overline{\rho}^2\,\mathrm{E}[{C_0}^2] .$$

Workload. Let the workload of a queue at a point in time be defined as the amount of time it takes the server to empty the queue under the assumption that there arrives no new work and under the assumption that the server remains with the queue. At any point in time, one may decompose the workload into two components: (i) the workload in front of the gate which will be served during the next cycle and (ii) the workload behind the gate which is served during the current cycle.

The expected workload in front of the gate at a queue grows from 0 linearly in time (see equation (38)) during the cycle. The average time since the start of the ongoing cycle from the vantage point of a random point in time is given by (see a.o. Baccelli and Brémaud [29]),

$$\mathrm{E}^s[C_p] = \frac{\mathrm{E}[{C_0}^2]}{2\,\mathrm{E}[C_0]}\,.$$

The expectation E^s in the left hand side of the above equation is with respect to the stationary regime. The expecation operator in the right hand side, involved in the expected cycle and expected busy duration per cycle which we have just computed, are to be undersstood as the those corresponding to the Palm probabilities (see Baccelli and Brémaud [29]). We next consider the expected workload, where this time the expectation is with respect to the stationary probability distribution.

In order for the paper to be self contained, we briefly recall what are the stationarny and the Palm distributions at the end of the section. We shall assume that the cycle durations are strictly positive w.p.1.

The expected stationary average workload in front of the gate which we denote by $\mathrm{E}^s[U_f]$ is given by,

$$\mathrm{E}^s[U_f] = \overline{\rho}\,\mathrm{E}^s[C_p] = \overline{\rho}\frac{\mathrm{E}[{C_0}^2]}{2\,\mathrm{E}[C_0]}\,.$$

The expected workload behind the gate diminishes linearly in time during the busy period and equals 0 during the walking time. The time until the end of the ongoing busy period (the residual busy period) as seen from the vantage point of a random point in time during a busy period is given by (see a.o. Baccelli and Brémaud [29]),

$$\mathrm{E}^s[G_r] = \frac{\mathrm{E}[{G_0}^2]}{2\,\mathrm{E}[G_0]}\,.$$

(A more formal discussion is given at the remark at the end of the Section. It explains what we mean by random point in time.) Since a random point in time is part of the busy period with probability $\overline{\rho}$, we find that the expected stationary average workload behind the gate $\mathrm{E}^s[U_a]$ is given by,

$$\mathrm{E}^s[U_a] = \overline{\rho}^2\frac{\mathrm{E}[{G_0}^2]}{2\,\mathrm{E}[G_0]} = \overline{\rho}\frac{\sigma^2\,\mathrm{E}[C_0] + \overline{\rho}^2\,\mathrm{E}[C_0^2]}{2\,\mathrm{E}[C_0]}\,.$$

The total expected workload in the queue therefore equals,

$$\mathrm{E}^s[U] = \mathrm{E}^s[U_f] + \mathrm{E}^s[U_a] = \overline{\rho}\frac{(\overline{\rho}^2+1)\,\mathrm{E}[{C_0}^2] + \sigma^2\,\mathrm{E}[C_0]}{2\,\mathrm{E}[C_0]}\,.$$

Virtual waiting time. Since the arrival process is described in terms of work streams and not in terms of customer arrival instants, one cannot consider customer waiting times. We may however consider the "virtual" waiting time of an infinitely small amount of work – a virtual customer – that arrives in the system. That is, let virtual waiting time be defined as the amount of time that it takes to empty the queue upon arrival of a (virtual) customer, given that there are no future arrivals.

The waiting time of a tagged virtual customer can be decomposed into the following three terms: (i) the expectation of the residual cycle time C_r upon arrival, (ii) the time to serve all the workload in front of the gate present at the queue upon arrival, i.e. the workload that arrived since the cycle began; the later duration is denoted by C_p, and (iii) the amount of work that arrives at the same epoch but before the tagged infinitesimal amount.

A random infinitesimal amount of work arrives at a random point in time. To see this, note that a subordinator combines a deterministic arrival stream and a batch Poisson process. Therefore the infinitesimal amount is either part of the stream or of a batch. Clearly, if the infinitesimal amount is part of the stream, it arrives at a random point in time. If the infinitesimal amount is part of a batch, we may apply the well-known PASTA (Poisson Arrivals See Time Averages) property of Poisson processes. Summerising we have (also see Baccelli and Brémaud [29]),

$$\mathrm{E}^s[C_r] = \mathrm{E}^s[C_p] = \frac{\mathrm{E}[{C_0}^2]}{2\,\mathrm{E}[C_0]}\,. \tag{16}$$

By (38) and since the workload arrival process has independent increments, the expectation of term (ii) equals $\overline{\rho}\,\mathrm{E}^s[C_p]$. Further, an infinitesimal amount of work is part of a Poisson jump with probability $\lambda p_1/(r + \lambda p_1)$. If this is the case, the average amount of work that arrives before the tagged amount of work equals $p_2/2p_1$ (see a.o. Baccelli and Brémaud [29]) and therefore the expected amount of work that arrives at the same epoch but before a tagged virtual customer equals,

$$\mathrm{E}^s[U_j] = \frac{\lambda p_2}{2(r + \lambda p_1)}\,. \tag{17}$$

Recall that λ, p_i and r denote the Poisson arrival rate, the ith moment of the batch size distribution and the deterministic rate of the arrival process respectively. We conclude that the average waiting time is given by,

$$\mathrm{E}^s[W] = \frac{\mathrm{E}[(C_0)^2]}{2\,\mathrm{E}[C_0]}(1 + \overline{\rho}) + \frac{\lambda p_2}{2(r + \lambda p_1)}. \tag{18}$$

Remark 1. *We provide here some formal definition of the term "random point in time". Let T_1 be the first time a cycle starts at some time larger than 0. Thus $T_0 \leq 0 < T_n$. Call the cycle starting at T_1 the first cycle. Define T_n to be the instant when the nth cycle starts; we allow n to take both positive and negative integer values. Let P^s and E^s denote the stationary probability measure and the corresponding expectation. We define P^o to be the Palm probability corresponding to P^s where the Palm probability of an event A is the stationary probability of that event conditioned by the event $T_0 = 0$. Equations (16) and (18) then express the first moments of C_r, C_p and W at the stationary regime in terms of the first and second moments of the cycle time with respect to the Palm probability. C_p is defined as the time as observed at $t = 0$ since the current cycle began; it is thus given by $-T_0$ and its expectation is given by $\mathrm{E}^s[C_p] = -\,\mathrm{E}^s[T_0]$. The residual cycle time G_r is simply T_1 so its expectation is $\mathrm{E}^s[G_r] = \mathrm{E}^s[T_1]$. The state at a "random point in time" can be understood as the state at time zero given that the system is stationiary at that time. Unless the notation E^s is used, all expectations in this subsection are to be understood as corresponding to the Palm probability.*

4 Symmetric exhaustive polling systems

We now consider the exhaustive polling system: the server remains with the same queue as long as there is work in this queue. More precisely, the server remains with the queue as long as there is a sufficient amount of work such that the server can operate at full capacity. That is, it is possible that the server stops serving a queue when there is a steady stream of arriving work with a rate smaller then the service rate.

Regarding the arrival processes and walking times, we make the same assumptions as in the preceding section. We also continue using the notation introduced there.

4.1 Completion periods

Let the notion of a completion period correspond to the time that it takes the server to completely empty a queue and let $\theta(y)$ denote the completion time given that the amount of work at the start of the completion time ($t = 0$) equals y. Further, let $\mathcal{R}(t)$ denote the amount of work that has arrived in the queue up to time t. One may then express $\theta(y)$ in terms of $\mathcal{R}(t)$ as follows,

$$\theta(y) = \inf\{t \geq 0 : y + \mathcal{R}(t) - t \leq 0\} . \tag{19}$$

In particular, let $\mathcal{R}(t)$ denote a subordinator. The following theorem then allows us to retrieve various characteristics of the process $\theta(t)$. By analytic continuation, the Laplace exponent $\phi(\cdot)$ and the Lévy exponent $\psi(\cdot)$ of a subordinator (see Appendix) relate as $\phi(\zeta) = -\psi(-i\zeta)$.

Theorem 3. *Let $\mathcal{R}(t)$ denote a subordinator with drift smaller than 1 and with Laplace exponent $\phi(\zeta)$. Further, let $\kappa(0)$ denote the largest solution of $\kappa(0) = \phi(\kappa(0))$. Then, the process $\theta(y)$ (as defined in expression (19)) is a subordinator killed at a rate $\kappa(0)$. Its Laplace exponent $\kappa(\zeta)$: $[0, \infty) \to [\kappa(0), \infty)$ is the unique solution of the functional equation $\kappa(\zeta) - \phi(\kappa(\zeta)) = \zeta$.*

Proof. Since $\mathcal{R}(t)$ is a subordinator with drift smaller than 1, $\mathcal{R}(t) - t$ can be decomposed into a subordinator with zero drift and a strictly negative drift. Further, the Itô decomposition of subordinators and the right continuity of the sample paths shows that this process crosses levels whenever it reaches levels from above for every sample path. The stated results then immediately follow from proposition 2.1 of [30]. □

By means of Hölders inequality, one finds for $\zeta_1, \zeta_2 \in [0, \infty)$,

$$e^{\phi\left(\frac{\zeta_1+\zeta_2}{2}\right)} = \mathrm{E}\left[\left(e^{\frac{\zeta_1}{2}\theta(1)}\right)\left(e^{\frac{\zeta_2}{2}\theta(1)}\right)\right] \leq \mathrm{E}\left[e^{\zeta_1\theta(1)}\right]^{1/2} \mathrm{E}\left[e^{\zeta_2\theta(1)}\right]^{1/2} = e^{\frac{\phi(\zeta_1)+\phi(\zeta_2)}{2}}$$

with equality if and only if $\zeta_1 = \zeta_2$. This then implies the strict convexity of the function $\zeta - \phi(\zeta)$ on $[0, \infty)$. Therefore the killing rate is strictly positive if and only if the derivative $1 - \phi'(0)$ is strictly negative. In other words, the subordinator $\theta(y)$ is never killed whenever $\mathrm{E}[\mathcal{R}(1)] = \overline{\rho} \leq 1$. Moreover, for $\overline{\rho} < 1$, the average completion time $\mathrm{E}[\theta(1)]$ and the corresponding variance $\mathrm{var}[\theta(1)]$ are given by,

$$\mathrm{E}[\theta(1)] = \overline{\theta} = \frac{1}{1 - \overline{\rho}} , \tag{20}$$

$$\mathrm{var}[\theta(1)] = \frac{\sigma^2}{(1 - \overline{\rho})^3} \tag{21}$$

which immediately follows from differentiation of the functional equation $\kappa(\zeta) - \phi(\kappa(\zeta)) = \zeta$.

4.2 Sample path modeling as a stochastic recursive equation

The former characterization of the completion times now allows us to follow an approach similar to the one that was used for the symmetric gated polling system with correlated walking times. Let m denote the number of queues visited by the server in a cyclic way. As for the gated polling system, let $I(n) = ((n-1) \mod m) + 1$ denote the queue that is visited at the nth polling instant and let V_n denote the walking time that the server takes after serving this queue.

As opposed to the gated polling system, we observe the polling system at time instants where the server leaves a queue. That is, let $S(n)$ denote the time epoch where the server leaves the queue that was polled at the nth polling instant. The polling system is observed at these instants since — in accordance with the exhaustive polling discipline — there is no work in queue $I(n)$ at time $S(n)$.

Further, let Y_n^i denote,

$$Y_n^i = S(n) - S(n - i) , \qquad (i = 1, 2, \ldots, m - 1) .$$

That is, Y_n^i is the time between the instants where the server leaves the $(n - i)$th and nth queue. The workload at queue i is described by a subordinator $\mathcal{R}^i(t)$ with parameter $t \in \mathbb{R}_+$ and which is distributed as $\mathcal{R}(t)$. Let $\mathcal{R}_n^i$ and $\hat{\mathcal{R}}_n^i$ denote series of independent copies of $\mathcal{R}^i$, $n = 1, 2, 3, \ldots$.

Similar, let $\theta^i(y)$ denote the completion process corresponding to $\mathcal{R}^i(t)$ and let θ_n^i and $\hat{\theta}_n^i$ denote independent copies of θ^i, $n = 1, 2, 3, \ldots$. The dynamics of the exhaustive polling system are then described by the following set of $m-1$ equations,

$$\begin{aligned}
Y_{n+1}^1 &= S(n+1) - S(n) = V_n + \theta_n^{m-1}(\mathcal{R}_n^{m-1}(Y_n^{m-1})) + \hat{\theta}_n^{m-1}(\hat{\mathcal{R}}_n^{m-1}(V_n)) \\
Y_{n+1}^2 &= S(n+1) - S(n-1) = Y_n^1 + V_n + \theta_n^{m-1}(\mathcal{R}_n^{m-1}(Y_n^{m-1})) + \hat{\theta}_n^{m-1}(\hat{\mathcal{R}}_n^{m-1}(V_n)) \\
Y_{n+1}^3 &= S(n+1) - S(n-2) = Y_n^2 + V_n + \theta_n^{m-1}(\mathcal{R}_n^{m-1}(Y_n^{m-1})) + \hat{\theta}_n^{m-1}(\hat{\mathcal{R}}_n^{m-1}(V_n)) \\
&\vdots \\
Y_{n+1}^{m-1} &= S(n+1) - S(n-m+2) = Y_n^{m-2} + V_n + \theta_n^{m-1}(\mathcal{R}_n^{m-1}(Y_n^{m-1})) + \hat{\theta}_n^{m-1}(\hat{\mathcal{R}}_n^{m-1}(V_n))
\end{aligned} \tag{22}$$

The former equations follow from the fact that at the beginning of the service period of the nth queue, the polling station finds all work in the queue that arrived since the last service period $(\mathcal{R}_n^{m-1}(Y_n^{m-1} + V_n))$. The corresponding completion period then corresponds to the time it takes to reduce the queue size to zero. The independent increments property finally leads to the former expressions.

The set of equations (22) can then be written in vector notation as follows,

$$Y_{n+1} = A_n(Y_n) + B_n \, .$$

Here B_n denotes the following vector of size $m-1$:

$$B_n = \gamma_n^m(V_n) \cdot (1, 1, \ldots, 1)^T \, , \tag{23}$$

with,

$$\gamma_n^m(x) = x + \hat{\theta}_n^{m-1}(\hat{\mathcal{R}}_n^{m-1}(x)) \, .$$

Notice that the processes γ_n^m are subordinators for all m, n since composition and summation of subordinators yields subordinators. Further, the processes A_n can be decomposed as,

$$A_n(y) = A_n^{(1)}(y^1) + \ldots + A_n^{(m-1)}(y^{m-1}) \, , \tag{24}$$

for $y = (y^1, \ldots, y^{m-1})^T \in \mathbb{R}_+^{m-1}$ with,

$$\begin{aligned}
A_n^{(1)} &= (0, t, 0, 0, \ldots, 0, 0)^T \\
A_n^{(2)} &= (0, 0, t, 0, \ldots, 0, 0)^T \\
&\vdots \\
A_n^{(m-2)} &= (0, 0, 0, 0, \ldots, 0, t)^T \\
A_n^{(m-1)} &= \theta_n^{m-1}(\mathcal{R}_n^{m-1}(t))\,(1, 1, \ldots, 1)^T
\end{aligned} \tag{25}$$

for $t \in \mathbb{R}_+$.

Clearly, the processes A_n constitute a series of independent and identically distributed Additive Lévy processes. We can further show that the series of vectors $\{B_k\}$ constitutes a stationary and ergodic series of random vectors. This immediately follows from the following theorem.

Theorem 4. *Let $\gamma_k(\cdot)$ denote a series of independent and identically distributed subordinators and let X_k denote a stationary ergodic series of random variables, then the series $\gamma_k(X_k)$ is also stationary ergodic.*

Proof. Let U_k denote an independent series of random variables, uniformly distributed on $[0, 1]$. The series (X_k, U_k) is then stationary ergodic and therefore this is also the case for the series $Y_k = f(X_k, U_k)$ for any Borel measurable function f (see e.g. Breiman [31]). In particular, let $f(x, y) = g_x^{-1}(y)$ with $g_x(y) = \Pr[\gamma(x) \le y]$. The Itô decomposition of subordinators implies that $f(x, y)$ is Borel measurable and the series Y_k is therefore stationary ergodic. Finally, from the definition of $f(x, y)$ it follows that the processes Y_k and $\gamma_k(X_k)$ share the same law and therefore $\gamma_k(X_k)$ is stationary ergodic. ☐

Summarizing, we find that the series A_n and B_n constitute a series of i.i.d. Additive Lévy processes in $\mathbb{R}_+^{m-1}$ and a series of stationary ergodic random vectors in $\mathbb{R}_+^{m-1}$ respectively. As such, we can use the framework of section 2.

4.3 First and second moments

From equation (25) and the definition of the matrix $\mathcal{A}$, we find,

$$\mathcal{A} = \begin{pmatrix} 0 & 0 & 0 & \dots & 0 & \overline{\theta}\overline{\rho} \\ 1 & 0 & 0 & \dots & 0 & \overline{\theta}\overline{\rho} \\ 0 & 1 & 0 & \dots & 0 & \overline{\theta}\overline{\rho} \\ 0 & 0 & 1 & \dots & 0 & \overline{\theta}\overline{\rho} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 1 & \overline{\theta}\overline{\rho} \end{pmatrix} \tag{26}$$

In view of Theorem 2, all eigenvalues of the former matrix are in the interior of the unit circle whenever,

$$\overline{\theta}\overline{\rho} < \frac{1}{m-1} \,. \tag{27}$$

For $\overline{\rho} < 1$, the latter condition is satisfied whenever (see equation (20)),

$$\overline{\rho} < \frac{1}{m} \,. \tag{28}$$

For $\overline{\rho} \geq 1$, this condition is never satisfied.

For $\overline{\rho} < 1/m$, the conditions of Theorem 1 hold and therefore the steady state expectation of Y_0 is given by,

$$\mathrm{E}[Y_0] = (\mathcal{I} - \mathcal{A})^{-1}\,\mathrm{E}[B] = \frac{(1+\overline{\rho}\overline{\theta})v}{1-(m-1)\overline{\rho}\overline{\theta}} \cdot (1,2,\dots,m-1)^T = \frac{v}{1-m\overline{\rho}} \cdot (1,2,\dots,m-1)^T \,. \tag{29}$$

The covariance matrix of $A(y)$ is given by $\mathrm{cov}[A(y)] = \sum_{j=1}^m y_j \Gamma^{(j)}$, where $\Gamma^{(j)}$ is the corresponding covariance matrix of $A^{(j)}(1)$ and where y_j denotes the jth element of the vector y (see Appendix). Clearly, for $j = 1 \dots m-2$, the covariance matrix $\Gamma^{(j)}$ is an $(m-1)\times(m-1)$ matrix whose elements are all zero. This follows from,

$$\mathrm{E}[A^{(j)}(1)A^{(j)}(1)^T] = \mathrm{diag}(0,0,\dots,1,\dots,0) = \mathrm{E}[A^{(j)}(1)]\,\mathrm{E}[A^{(j)}(1)]^T \,. \tag{30}$$

Here the 1 in the diagonal matrix is in position $j+1$. We now compute $\Gamma^{(m-1)}$. In view of the definition of $A^{(m-1)}(y)$ (see expression (25)), we find,

$$\Gamma^{(m-1)} = \mathrm{var}[\theta_0^{m-1}(\mathcal{R}_0^{m-1}(1))]\,\mathcal{E} = (\mathrm{var}[\theta_0^{m-1}(1)]\overline{\rho} + \overline{\theta}^2\sigma^2)\,\mathcal{E} = \frac{\sigma^2}{(1-\overline{\rho})^3}\,\mathcal{E} \,. \tag{31}$$

where $\mathcal{E}$ denotes an $(m-1)\times(m-1)$ matrix with all elements equal to 1.

In view of expression (23), we further find,

$$\begin{aligned} \mathrm{cov}[B_0] &= \mathrm{var}[\gamma_0^m(V_0)] \cdot \mathcal{E} = \frac{\sigma^2 v + (1-\overline{\rho})^2\hat{\mathcal{V}}(0)}{(1-\overline{\rho})^3} \cdot \mathcal{E} \,, \\ \hat{\mathcal{B}}(j) &= \left\{ \mathrm{E}\left[\gamma_0^m(V_0)\gamma_j^m(V_j)\right] - \frac{v^2}{(1-\overline{\rho})^2} \right\} \cdot \mathcal{E} = \frac{\hat{\mathcal{V}}(j)}{(1-\overline{\rho})^2} \cdot \mathcal{E} \,, \\ \mathcal{B}(j) &= \mathrm{E}\left[\gamma_0^m(V_0)\gamma_j^m(V_j)\right] \cdot \mathcal{E} = \frac{\mathcal{V}(j)}{(1-\overline{\rho})^2} \cdot \mathcal{E} \,. \end{aligned}$$

Finally, after plugging in the former expressions into equation (5) of Theorem 1, we find that the $(m-1)\times(m-1)$ matrix $\mathrm{cov}[Y_0]$ is the unique solution of,

$$\mathrm{cov}[Y_0] = \frac{mv\sigma^2}{(1-m\overline{\rho})(1-\overline{\rho})^2}\mathcal{E} + \mathcal{A}\,\mathrm{cov}[Y_0]\mathcal{A}^T + \frac{\hat{\mathcal{V}}(0)}{1-\overline{\rho}}\mathcal{E} + \sum_{j=1}^{\infty}\frac{\hat{\mathcal{V}}(j)}{(1-\overline{\rho})^2}\left[\mathcal{A}^j\mathcal{E} + (\mathcal{A}^j\mathcal{E})^T\right]. \quad (32)$$

Note that the last term in (32) disappears for the case of independent walking times.

4.4 Performance measures

We now find expressions for various performance measures of the polling system under consideration.

Busy and vacation times Let the nth busy time G_n be defined as the time between the arrival of the server at the nth queue and the start of the following walking time. Also, let the nth vacation time H_n be defined as the time between the end of the nth busy time and the time the server returns to the queue. In view of the definition of Y_n^i and since the server finds all work that arrived during the $(n-m)$th vacation time upon arrival at the nth queue, we find,

$$H_n = Y_{n+m-1}^{m-1} + V_{n+m-1}\,, \quad G_n = \theta(\mathcal{R}(H_{n-m}))\,.$$

The expected busy and vacation times therefore equal,

$$\begin{aligned}
\mathrm{E}[H_0] &= \mathrm{E}[Y_0^{m-1}] + \mathrm{E}[V_0] = \frac{vm(1-\overline{\rho})}{1-m\overline{\rho}}\,,\\
\mathrm{E}[G_0] &= \overline{\theta}\overline{\rho}\,\mathrm{E}[H_0] = \frac{vm\overline{\rho}}{1-m\overline{\rho}}\,.
\end{aligned}$$

Further, we find following expressions for the second moments of the busy and vacation times,

$$\begin{aligned}
\mathrm{E}[{H_0}^2] &= \mathrm{E}[(Y_0^{m-1})^2] + \mathrm{E}[(V_0)^2] + 2\,\mathrm{E}[Y_0^{m-1}V_0]\\
&= \mathrm{E}[(Y_0^{m-1})^2] + \mathcal{V}(0) + \frac{\mathrm{E}[Y_0^{m-1}B_0^{m-1}]}{1+\overline{\theta}\overline{\rho}}\\
\mathrm{E}[{G_0}^2] &= \mathrm{E}[\theta(\rho(H_0))^2] = \frac{\sigma^2\,\mathrm{E}[H_0] + (1-\overline{\rho})\overline{\rho}^2\,\mathrm{E}[{G_0}^2]}{(1-\overline{\rho})^3}\,.
\end{aligned}$$

Here $\mathrm{E}[Y_0^{m-1}B_0^{m-1}]$ is the $(m-1)$th diagonal element of the matrix (see equation (8)),

$$\mathrm{E}[Y_0B_0^T] = \sum_{j=0}^{\infty}\mathcal{A}^j\mathcal{B}(j+1) = \sum_{j=0}^{\infty}\mathcal{V}(j+1)\mathcal{A}^j\mathcal{E}\,,$$

and $\mathrm{E}[(Y_0^m)^2]$ equals the $(m-1)$th diagonal element of the matrix $\mathrm{cov}[Y_0] + \mathrm{E}[Y_0]\,\mathrm{E}[Y_0]^T$.

Workload. As before, let the workload in a queue be defined as the amount of time in takes to empty the queue under the assumption that there arrives no new work and that the server remains with the queue. Since there is no work in the (tagged) queue at the beginning of a vacation period and since work arrives at a rate $\overline{\rho}$ during the vacation period, we find following expression for the mean workload during vacations,

$$\mathrm{E}[U_v] = \overline{\rho}\frac{E\left[{H_0}^2\right]}{2\,\mathrm{E}[H_0]} \quad (33)$$

Further, let $\mathrm{E}[U_b]$ denote the mean workload during busy periods. The expectation of the remaining busy time then equals $\overline{\theta}\,\mathrm{E}[U_b]$ since the queue builds down at a rate $\overline{\theta}$ during service periods. Therefore we find,

$$\mathrm{E}[U_b] = \frac{\mathrm{E}[{G_0}^2]}{2\,\mathrm{E}[G_0]\overline{\theta}}\,. \quad (34)$$

Combining the former expressions and taking into account that the server is busy for a fraction $\mathrm{E}[G_0]/(\mathrm{E}[G_0]+\mathrm{E}[H_0])$ of the time then leads to the following expression for the expectation of the unfinished work,

$$\mathrm{E}[U] = \frac{1}{2}\frac{\overline{\rho}\,\mathrm{E}[{H_0}^2] + (1-\overline{\rho})\,\mathrm{E}[{G_0}^2]}{\mathrm{E}[G_0]+\mathrm{E}[H_0]}\,. \tag{35}$$

Expected waiting time Clearly, the expected waiting time of a (tagged) virtual customer that arrives during a vacation time equals the sum of (i) the expected remaining vacation time $E[H_0^2]/2\,\mathrm{E}[H_0]$ (see a.o. Baccelli and Brémaud [29]), (ii) the expected workload in the queue upon arrival (see equation (33)) and (iii) the amount of work that arrives at the same epoch but before the tagged virtual customer (see equation (17)). We find,

$$\mathrm{E}[W_v] = \frac{E\left[{H_0}^2\right]}{2\,\mathrm{E}[H_0]}(1+\overline{\rho}) + \frac{\lambda p_2}{2(r+\lambda p_1)}\,.$$

Further, the expected virtual waiting time during busy times equals the sum of (i) the expected workload upon arrival of the tagged virtual customer (see equation (34)) and (ii) the amount of work that arrives at the same epoch but before the tagged virtual customer (see equation (17)). We obtain the following expression for the expectation of the (virtual) waiting times during busy times,

$$\mathrm{E}[W_b] = \frac{\mathrm{E}[{G_0}^2]}{2\,\mathrm{E}[G_0]\overline{\theta}} + \frac{\lambda p_2}{2(r+\lambda p_1)}\,.$$

Taking into account that a fraction $\mathrm{E}[G_0]/(\mathrm{E}[G_0]+\mathrm{E}[H_0])$ of all work arrives during the busy time, we find the following expression of the expected virtual waiting time,

$$\mathrm{E}[W] = \frac{1}{2}\frac{(1+\overline{\rho})\,\mathrm{E}[{H_0}^2] + (1-\overline{\rho})\,\mathrm{E}[{G_0}^2]}{\mathrm{E}[H_0]+\mathrm{E}[G_0]} + \frac{\lambda p_2}{2(r+\lambda p_1)}\,. \tag{36}$$

5 Concluding comments

In this paper we have studied and used multi-type branching processes with a continuous state-space and derived their first two moments for the case of a (possibly non-Markovian) stationary ergodic migration process. The framework is then used to derive explicit formulas for the expected workload and waiting times in symmetric gated and exhaustive polling systems with correlated walking times.

Acknowledgment

This work was supported by the EURO FGI network of excellence.

References

[1] H. Takagi. Application of polling models to computer-networks. *Computer Networks and ISDN systems*, 22(3):193–211, 1991.

[2] G. Zussman, A. Segall, and U. Yechiali. Bluetooth time division duplex - analysis as a polling system. In *Proc. 1st IEEE Conference on Sensor and Ad Hoc Communications and Networks*, 2004.

[3] K. H. Yeung, K. P. Fung, and K. Y. Wong. Performance study on disk operations by broadcast polling model. *IEICE Transactions on Communications E86B*, 6:2012–2014, june 2003.

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

[5] H. Takagi. Queueing analysis of polling models: an update. In H. Takagi, editor, *Stochastic Analysis of Computer and Communication Systems*, pages 267–318. Elsevier Science Publishers, 1990.

[6] U. Yechiali. Analysis and control of polling systems. In *Performance Evaluation of Computer and Communication Systems*, pages 630–650. Springer-Verlag, 1993.

[7] H. Lévy and M. Sidi. Polling systems: Applications, modeling, and optimization. *IEEE Transactions on Communication*, 38(10):1750–1760, 1990.

[8] A. Brandt. The stochastic equation $y_{n+1} = a_n y_n + b_n$ with stationary coefficients. *Advances in Applied Probability*, 18:211–220, 1986.

[9] A. Brandt, P. Franken, and B. Lisek. *Stationary Stochastic Models.* Akademie-Verlag, Berlin, 1992.

[10] P. Glasserman and D. D. Yao. Stochastic vector difference equations with stationary coefficients. *Journal of Applied Probability*, 32:851–866, 1995.

[11] I. J. Bienaymé. De la loi de la multiplication et de la durée des familles. *Société Philomatique de Paris, Extraits*, 5:37–29, 1845.

[12] F. Galton and H. W. Watson. On the probability of the extinction of the families. *Journal of the Royal Anthropological Institute*, 4:138–144, 1874.

[13] A. N. Kolmogorov. On the solution of a biological problem. In *Proceedings of Tomsk University*, volume 2, pages 7–12, 1938.

[14] B. A. Sevastyanov. Limit theorem for branching processes of special form. *TPA*, 2:121–136, 1957.

[15] J. D. Biggins, H. Cohn, and O. Nerman. Multi-type branching in varying environment. *Stochastic Processes and their Applications*, 83:357–400, 1999.

[16] K. B. Athreya and A. N. Vidyashankar. Branching processes. In D. N. Shanbhag, editor, *Handbook of statistics 19: Stochastic Processes: Theory and Methods.* Elsevier, 2001.

[17] K. B. Athreya and P. Jaggers, editors. *Classical and Modern Branching Processes Series*, volume 84 of *The IMA Volumes in Mathematics and its Applications.* Springer Berlin Heidelberg New York, 1997.

[18] J. A. C. Resing. Polling systems and multi-type branching processes. *Queueing Systems, Theory and Applications*, 13:409–426, 1993.

[19] R. Groenevelt and E. Altman. Analysis of alternating-priority queueing models with (cross) correlated switchover times. In *Proceedings of IEEE Infocom 2005*, Miami, March 2005.

[20] E. Altman. On stochastic recursive equations and infinite server queues. In *Proceedings of IEEE Infocom 2005*, Miami, March 2005.

[21] S. A. Grishenchkin. On a relation between processor sharing queues and crump-mode-jager branching processes. *Advances in Applied Probability*, 24:653–698, 1992.

[22] R. Nunez-Queija. *Processor-Sharing Models for Integrated-Services Networks.* PhD thesis, Eindhoven University of Technology, 2000.

[23] Y. J. Aminetzah M. J. Ferguson. Exact results for nonsymmetric token ring systems. *IEEE Transactions on Communications*, 33(3):223–231, 1985.

[24] A. Lambert. The genealogy of continuous-state branching processes with immigration. *Journal of Probability Theory and Related Fields*, 122(1):42–70, 2002.

[25] E. Altman. Stochastic recursive equations with applications to queues with dependent vacations. *Annals of Operations Research*, 112(1):43–61, April 2002.

[26] S. R. Adke and V. G. Gadag. A new class of branching processes. In C.C. Heyde, editor, *Branching Processes: Proceedings of the First World Congress*, volume 99 of *Springer Lecture Notes*, pages 1–13, 1995.

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

[28] R. A. Horn and C. R. Johnson. *Matrix Analysis*. Cambridge University Press, 1996.

[29] F. Baccelli and P. Brémaud. *Elements of Queueing Theory*. Springer, second edition, 2003.

[30] J. Bertoin. Subordinators, Lévy processes with no negative jumps and branching processes. Lecture notes, Centre for Mathematical Physics and Stochastics, Department of Mathematical Sciences, University of Aarhus, 2000.

[31] L. Breiman. *Probability*. Number 7 in Classics in Applied Mathematics. Addison-Wesley, 1968.

[32] K. Sato. *Lévy Processes and Infinitely Divisible Distributions*. Cambridge University Press, 1999.

[33] J. Bertoin. *Lévy Processes*. Cambridge University Press, 2002.

[34] D. Khoshnevisan, Y. Xiao, and Y. Zhong. Local times of additive Lévy processes. *Stochastic Processes and their Applications*, 104:193–216, 2003.

[35] O. E. Barndorff-Nielsen, J. Federsen, and K. I. Sato. Multivariate subordination self-decomposability and stability. *Advances in Applied Probability*, 33:130–187, 2001.

[36] H. K. Gjessing, O. O. Aalen, and N.L. Hjort. Frailty models based on Lévy processes. *Advances in Applied Probability*, 35:532–550, 2003.

Appendix: Background material

We frequently encounter vector-valued stochastic processes where the "time" parameter is a scalar. We shall need however to use a vector valued parameter, which gives us a Lévy field. An example of a scalar field with a vector valued parameter is a black and white picture. The parameter space is two dimensional (the x and y directions) and for each value (x,y) of the parameter, we have a specific gray level which is given by a scalar. An example of a field where both the field's value at a point as well as the parameter space are vectors is a color photo, where for each point (x,y), a three dimensional vector is introduced, where the three components represent the intensity level of the blue, red and yellow colors.

We begin by recalling the definition of a K-parameter Lévy Field. Let K be a cone in $\mathbb{R}^d$ inducing an ordering $\leq_K$. A K-parameter Lévy process $\{A(s), s \in K\}$ on $\mathbb{R}^m$ is a collection of random variables on $\mathbb{R}^m$ satisfying the following properties.

(a) Independent increments;

(b) Stationarity in each direction in K;

(c) Continuity in probability: for each $s \in K$, $A(s') \to A(s)$ in probability as $|s' - s| \to 0$ with $s' \in K$;

(d) $A(0) = 0$ almost surely;

(e) Almost surely, $A(s)$ is K-right continuous with K-left limits in s.

For precise definitions of independent increments and stationarity, the reader is referred to Sato's monograph [32]. In the sequel, we shall consider $K = \mathbb{R}_+^d$.

The case $K = \mathbb{R}_+$

We first consider a multivariate (vector valued) Lévy process with a one-dimensional (scalar valued) time parameter t (i.e. $K = \mathbb{R}_+$) which we denote – with some abuse of notation – by $A(t)$. Let m denote the dimension of $A(t)$. The characteristic function of $A(t)$ is then given by (see a.o. [32–35]),

$$\mathrm{E}[e^{i<\xi, A(t)>}] = e^{-t\psi(\xi)},$$

for any $t \in \mathbb{R}_+$, where by the Lévy-Khintchine formula,

$$\psi(\xi) = i < a, \xi > + \int_{\mathbb{R}_+^m} \left[e^{i<x,\xi>} - 1\right] L(dx), \tag{37}$$

for all $\xi \in \mathbb{R}^m$ and for a given $a \in \mathbb{R}_+^m$. Here L is a finite measure on $\mathbb{R}^m$ concentrated on $\mathbb{R}_+^m - \{0\}$. ψ is called the Lévy exponent of A and L is the corresponding Lévy measure [34].

The expectation and covariance of a multivariate Lévy process have the following form:

$$E[A(t)] = t\mathcal{A}, \qquad \mathrm{cov}[A(t)] = t\Gamma \tag{38}$$

where $\mathcal{A}$ is an m-dimensional column vector and Γ is a symmetric $m \times m$ matrix. The values of $\mathcal{A}$ and of Γ can be obtained by differentiating (37) once and twice respectively. That is, the ith element of $\mathcal{A}$ and the ijth element of Γ are given by (see also [36]),

$$[\mathcal{A}]_i = \left.\frac{\partial \psi(\xi)}{i \partial \xi_i}\right|_{\xi=0} \qquad \text{and} \qquad [\Gamma]_{ij} = -\left.\frac{\partial \psi(\xi)^2}{\partial \xi_i \partial \xi_j}\right|_{\xi=0}.$$

We next present useful formulas for $m = 1$, for the mean and variance of A evaluated at a random time. Let τ be a nonnegative random variable, independent of A. The mean and variance of $A(\tau)$ are then given by,

$$E[A(\tau)] = E[\tau]\mathcal{A},$$

and,

$$\begin{aligned} \mathrm{var}[A(\tau)] &= \mathrm{E}[A(\tau)^2] - (\mathrm{E}[A(\tau)])^2 \\ &= \mathrm{E}\Big(\mathrm{E}\left[\mathrm{var}[A(\tau)] + (\mathcal{A}\tau)^2 \middle| \tau\right]\Big) - (\mathrm{E}[A(\tau)])^2 \\ &= \mathrm{E}[\tau]\Gamma + \mathrm{var}[\tau]\mathcal{A}^2, \end{aligned} \tag{39}$$

respectively.

Additive Lévy process

For the case of Lévy processes with an $\mathbb{R}_+^d$ valued "time" parameter (or Lévy fields), we shall focus on fields with a special structure: Additive Lévy fields. Let A denote a Lévy field and let $A^{(1)}, ..., A^{(d)}$ be d independent Lévy processes on $\mathbb{R}^m$ with scalar valued time parameters. We then assume that the random field A has the following decomposition:

$$A(y) = A^{(1)}(y_1) + ... + A^{(d)}(y_d),$$

for all $y = (y_1, ..., y_d) \in \mathbb{R}_+^d$. Let $\psi_1, ..., \psi_d$ be the Lévy exponents corresponding to $A^{(1)}, ..., A^{(d)}$. Then for any $y \in \mathbb{R}^d$, the characteristic function of $A(y) = \sum_{j=1}^d A^{(j)}(y_j)$ is given by

$$\mathrm{E}[e^{i<\xi, A(y)>}] = e^{-\sum_{j=1}^d y_j \psi_j(\xi)} = e^{-<y, \Psi(\xi)>}, \quad \xi \in \mathbb{R}^m.$$

where $\Psi = (\psi_1, ..., \psi_d)$.

The expectation of $A(y)$ is given by

$$\mathrm{E}[A(y)] = \sum_{j=1}^{d} y^j \mathcal{A}^{(j)} = \mathcal{A}y\,, \tag{40}$$

where $\mathcal{A}^{(j)} = \mathrm{E}[A^{(j)}(1)]$ denotes the expectation of $A^{(j)}(1)$ and where $\mathcal{A}$ is a matrix whose jth column equals $\mathcal{A}^{(j)}$. Similarly, the covariance matrix of $A(y)$ is given by,

$$\mathrm{cov}[A(y)] = \sum_{j=1}^{d} y_j \Gamma^{(j)}\,, \tag{41}$$

where $\Gamma^{(j)} = \mathrm{cov}[A^{(j)}(1)]$ is the corresponding covariance matrix of $A^{(j)}(1)$.

As for the scalar case, we derive the first and second moments of the process A at a random time $A(\tau)$. Here τ is a non-negative random variable in $\mathbb{R}_+^d$, which is independent of A and represented as a column vector. The mean vector and covariance matrix of $\mathcal{A}(\tau)$ are given by,

$$E[A(\tau)] = \sum_{j=1}^{m} \mathcal{A}^{(j)} E[\tau_j]\,,$$

and,

$$\mathrm{cov}[A(\tau)] = \sum_{j=1}^{d} \mathrm{E}[\tau_j]\Gamma^{(j)} + \mathcal{A}\,\mathrm{cov}[\tau]\mathcal{A}^T\,, \tag{42}$$

where τ_j is the jth entry of the vector τ. Similarly, we also have,

$$\begin{aligned} E[A(\tau)A(\tau)^T] &= \mathrm{E}\left\{\mathrm{E}\left[A(\tau)A(\tau)^T\right]\middle|\,\tau\right\} \\ &= \mathrm{E}\left\{\mathrm{E}\left(\mathrm{cov}[A(\tau)] + \mathcal{A}\tau\left(\mathcal{A}\tau\right)^T\middle|\,\tau\right)\right\} \\ &= \sum_{j=1}^{d} E[\tau_j]\Gamma^{(j)} + \mathcal{A}E[\tau\tau^T]\mathcal{A}^T\,. \end{aligned} \tag{43}$$

Stability and stationary distribution

Finally, we recall some properties of the stationary distribution of the stochastic recursive equation (1) where the A_n constitute a series of i.i.d. Additive Lévy processes in $\mathbb{R}_+^m$ with an $\mathbb{R}_+^m$ valued time parameter. The B_n constitute a series of stationary ergodic random variables in $\mathbb{R}_+^m$ and the process B_n is assumed to be independent of the processes A_n.

Additive Lévy processes have a divisibility property. For any integers n and k and for any $y(i) \in \mathbb{R}_+^m$, $i = 1, ..., k$, we have,

$$A_n\left(\sum_{i=0}^{k} y(i)\right) = \sum_{i=0}^{k} A_{n,i}(y(i))\,,$$

where for any fixed n, the processes $A_{n,i}$ are i.i.d. copies of the process A_n. Using the former property, we obtain for any integers k and n with $k < n$ by iterating (1),

$$Y_n = \sum_{j=k}^{n-1}\left(\bigotimes_{i=n-j}^{n-1} A_{n-j,i}\right)(B_{n-j-1}) + \left(\bigotimes_{i=k}^{n-1} A_{k,i}\right)(Y_k)\,. \tag{44}$$

Here we understand $\bigotimes_{i=n}^{k} A_i(x) = x$ whenever $k < n$, and $\bigotimes_{i=n}^{k} A_i(x) = A_k(A_{k-1}(\ldots(A_n(x))))$ whenever $k > n$.

Note that compositions of Lévy processes – as we have in equation (44) – are themselves Lévy processes. Moreover, if A_n and A_{n+1} are Additive Lévy processes in $\mathbb{R}_+^m$ then their composition is also an Additive Lévy process. Indeed, let A_n and A_{n+1} have the decomposition,

$$A_i(y) = A_i^{(1)}(y_1) + ... + A_i^{(m)}(y_m)\,,$$

for all $y = (y_1, ..., y_m) \in \mathbb{R}_+^m$ and for $i = n, n+1$ and where $A_n^{(1)}, ..., A_n^{(m)}$ and $A_{n+1}^{(1)}, ..., A_{n+1}^{(m)}$ are $2m$ independent Lévy processes on $\mathbb{R}_+^m$. We then have,

$$A_{n+1}(A_n(y)) = A_{n+1}\left(\sum_{i=1}^{m} A_n^{(i)}(y_i))\right) = \sum_{i=1}^{m} A_{n+1,i}(A_n^{(i)}(y_i)) = \sum_{i=1}^{m} \tilde{A}_n^{(i)}(y_i) \qquad (45)$$

where the processes $A_{n+1,i}$, $i = 1, \ldots, M$ are i.i.d. copies of the process A_{n+1} and where $\tilde{A}^{(i)} = A_{n+1,i}A_n^{(i)}$ is an independent Lévy process with a scalar valued time parameter.

As already mentioned, the equilibrium distribution of the dynamics equation (1) has been studied before in [25]. In particular, the following Theorem is a consequence of Lemma 1 and Theorem 2 in the latter contribution.

Theorem 5. *Assume that the sequence $\{(A_n(\cdot), B_n), -\infty < n < \infty\}$ is stationary ergodic, defined on some probability space $(\Omega, \mathcal{F}, P)$. For each n, let A_n be an Additive Lévy process and assume that the processes A_n constitute a series of i.i.d. random processes. Further, assume that all eigenvalues of the matrix $\mathcal{A}$ are all in the interior of the unit circle, and that* $\mathrm{E}[\max(\log\|B\|, 0)]$ *is finite for some norm $\|\cdot\|$.*

Then there is a unique stationary solution Y_n^* of (1), distributed like

$$Y_n^* =_d \sum_{j=0}^{\infty}\left(\bigotimes_{i=n-j}^{n-1} A_{n-j,i}\right)(B_{n-j-1}), \qquad n \in \mathbb{Z}, \qquad (46)$$

where for each integer i, $\{A_{j,i}(\cdot)\}_j$ are independent of each other and have the same distribution as $A_i(\cdot)$. The sum on the right side of (46) converges absolutely P-almost surely. Furthermore, for all initial conditions Y_0, $\|Y_n - Y_n^*\| \to 0$, P-almost surely on the same probability space. In particular, the distribution of Y_n converges to that of Y_0^* as $n \to \infty$.