EURANDOM PREPRINT SERIES

2015-004

January 29, 2015

Analysis and optimization of vacation and polling models with retrials

M. Abidini, O. Boxma, J. Resing
ISSN 1389-2355

Analysis and optimization of vacation and polling models with retrials ⋆

Murtuza Ali Abidini, Onno Boxma and Jacques Resing

EURANDOM and Department of Mathematics and Computer Science
Eindhoven University of Technology
P.O. Box 513, 5600 MB Eindhoven, The Netherlands
m.a.abidini@tue.nl,o.j.boxma@tue.nl,j.a.c.resing@tue.nl

Abstract. We study a vacation-type queueing model, and a single-server multi-queue polling model, with the special feature of retrials. Just before the server arrives at a station there is some deterministic glue period. Customers (both new arrivals and retrials) arriving at the station during this glue period will be served during the visit of the server. Customers arriving in any other period leave immediately and will retry after an exponentially distributed time. Our main focus is on queue length analysis, both at embedded time points (beginnings of glue periods, visit periods and switch- or vacation periods) and at arbitrary time points.

Keywords: vacation queue, polling model, retrials

1 Introduction

Queueing systems with retrials are characterized by the fact that arriving customers, who find the server busy, do not wait in an ordinary queue. Instead of that they go into an orbit, retrying to obtain service after a random amount of time. These systems have received considerable attention in the literature, see e.g. the book by Falin and Templeton [10], and the more recent book by Artalejo and Gomez-Corral [3].

Polling systems are queueing models in which a single server, alternatingly, visits a number of queues in some prescribed order. Polling systems, too, have been extensively studied in the literature. For example, various different service disciplines (rules which describe the server's behaviour while visiting a queue) and both models with and without switchover times have been considered. We refer to Takagi [24, 25] and Vishnevskii and Semenova [27] for some literature reviews and to Boon, van der Mei and Winands [5], Levy and Sidi [16] and Takagi [22] for overviews of the applicability of polling systems.

In this paper, motivated by questions regarding the performance modelling of optical networks, we consider vacation and polling systems with retrials. Despite

⋆ This is an invited, considerably extended version of [9]. The main additions are Subsections 2.4 and 2.5 and Section 3. These present respectively an alternative derivation of the mean number of customers and the optimal behaviour of a single queue system, and the performance analysis for a general number of queues.

the enormous amount of literature on both types of models, there are hardly any papers having both the features of retrials of customers and of a single server polling a number of queues. In fact, the authors are only aware of a sequence of papers by Langaris [12–14] on this topic. In all these papers the author determines the mean number of retrial customers in the different stations. In [12] the author studies a model in which the server, upon polling a station, stays there for an exponential period of time and if a customer asks for service before this time expires, the customer is served and a new exponential stay period at the station begins. In [13] the author studies a model with two types of customers: primary customers and secondary customers. Primary customers are all customers present in the station at the instant the server polls the station. Secondary customers are customers who arrive during the sojourn time of the server in the station. The server, upon polling a station, first serves all the primary customers present and after that stays an exponential period of time to wait for and serve secondary customers. Finally, in [14] the author considers a model with Markovian routing and stations that could be either of the type of [12] or of the type of [13].

In this paper we consider a polling station with retrials and so-called glue periods. Just before the server arrives at a station there is some deterministic glue period. Customers (both new arrivals and retrials) arriving at the station during this glue period "stick" and will be served during the visit of the server. Customers arriving in any other period leave immediately and will retry after an exponentially distributed time.

The study of queueing systems with retrials and glue periods is motivated by questions regarding the performance modelling and analysis of optical networks. Performance analysis of optical networks is a challenging topic (see e.g. Maier [17] and Rogiest [21]). In a telecommunication network, packets must be routed from source to destination, passing through a series of links and nodes. In copper-based transmission links, packets from different sources are time-multiplexed. This is often modeled by a single server polling system. Optical fibre offers some big advantages for communication w.r.t. copper cables: huge bandwidth, ultra-low losses, and an extra dimension – the wavelength of light. However, in an optical routing node, opposite to electronics, it is difficult to store photons, and hence buffering in optical routers can only be very limited. Buffering in these networks is typically realized by sending optical packets into fibre delay loops, i.e., letting them circulate in a fibre loop and extracting them after a certain number of circulations. This feature can be modelled by retrial queues. Recent experiments with 'slow light', where light is slowed down by significantly increasing the refractive index in waveguides, have up to now shown very modest buffering times [11]. It should be noted that with the very high speeds achievable in fibre, packet durations are very short, so that small buffering times may already allow sufficient storage of small packets. We represent the effect of slowing down light by introducing a glue period at a queue just before the server arrives.

The paper is organized as follows. In Section 2 we consider the case of a single queue with vacations and retrials; arrivals and retrials only "stick" during a glue period. We study this case separately because (i) it is of interest in its own right, (ii) it allows us to explain the analytic approach as well as the probabilistic meaning of the main components in considerable detail, (iii) it makes the analysis of the multi-queue case more accessible, and (iv) results for the one-queue case may serve as a first-order approximation for the behaviour of a particular queue in the N-queue case, switchover periods now also representing glue and visit periods at other queues. In Section 3 the N-queue case is analyzed. Section 4 presents some conclusions and suggestions for future research.

2 Queue length analysis for the single-queue case

2.1 Model description

In this section we consider a single queue Q in isolation. Customers arrive at Q according to a Poisson process with rate λ. The service times of successive customers are independent, identically distributed (i.i.d.) random variables (r.v.), with distribution $B(\cdot)$ and Laplace-Stieltjes transform (LST) $\tilde{B}(\cdot)$. A generic service time is denoted by B. After a visit period of the server at Q it takes a vacation. Successive vacation lengths are i.i.d. r.v., with S a generic vacation length, with distribution $S(\cdot)$ and LST $\tilde{S}(\cdot)$. We make all the usual independence assumptions about interarrival times, service times and vacation lengths at the queues. After the server's vacation, a *glue* period of deterministic (i.e., constant) length begins. Its significance stems from the following assumption. Customers who arrive at Q do not receive service immediately. When customers arrive at Q during a glue period G, they stick, joining the queue of Q. When they arrive in any other period, they immediately leave and retry after a retrial interval which is independent of everything else, and which is exponentially distributed with rate ν. The glue period is immediately followed by a visit period of the server at Q.

The service discipline at Q is gated: During the visit period at Q, the server serves all "glued" customers in that queue, i.e., all customers waiting at the end of the glue period (but none of those in orbit, and neither any new arrivals).

We are interested in the steady-state behaviour of this vacation model with retrials. We hence make the assumption that $\rho := \lambda \mathbb{E} B < 1$; it may be verified that this is indeed the condition for this steady-state behaviour to exist.

Some more notation:
G_n denotes the nth glue period of Q.
V_n denotes the nth visit period of Q (immediately following the nth glue period).
S_n denotes the nth vacation of the server (immediately following the nth visit period).
X_n denotes the number of customers in the system (hence in orbit) at the start of G_n.
Y_n denotes the number of customers in the system at the start of V_n. Notice that here we should distinguish between those who are queueing and those who are in orbit: We write $Y_n = Y_n^{(q)} + Y_n^{(o)}$, where q denotes queueing and o denotes in

orbit.
Finally,
Z_n denotes the number of customers in the system (hence in orbit) at the start of S_n.

2.2 Queue length analysis at embedded time points

In this subsection we study the steady-state distributions of the numbers of customers at the beginning of (i) glue periods, (ii) visit periods, and (iii) vacation periods. Denote by X a r.v. with as distribution the limiting distribution of X_n. Y and Z are similarly defined, and $Y = Y^{(q)} + Y^{(o)}$, the steady-state numbers of customers in queue and in orbit at the beginning of a visit period (which coincides with the end of a glue period). In the sequel we shall introduce several generating functions, throughout assuming that their parameter $|z| \leq 1$. For conciseness of notation, let $\beta(z) := \tilde{B}(\lambda(1-z))$ and $\sigma(z) := \tilde{S}(\lambda(1-z))$. Then it is easily seen that

$$\mathbb{E}[z^X] = \sigma(z)\mathbb{E}[z^Z], \tag{2.1}$$

since X equals Z plus the new arrivals during the vacation;

$$\mathbb{E}[z^Z] = \mathbb{E}[\beta(z)^{Y^{(q)}} z^{Y^{(o)}}], \tag{2.2}$$

since Z equals $Y^{(o)}$ plus the new arrivals during the $Y^{(q)}$ services; and

$$\mathbb{E}[z_q^{Y^{(q)}} z_o^{Y^{(o)}}] = \mathrm{e}^{-\lambda(1-z_q)G}\mathbb{E}[\{(1-\mathrm{e}^{-\nu G})z_q + \mathrm{e}^{-\nu G}z_o\}^X]. \tag{2.3}$$

The last equation follows since $Y^{(q)}$ is the sum of new arrivals during G and retrials who return during G; each of the X customers which were in orbit at the beginning of the glue period have a probability $1 - \mathrm{e}^{-\nu G}$ of returning before the end of that glue period.

Combining Equations (2.1)-(2.3), and introducing

$$f(z) := (1-\mathrm{e}^{-\nu G})\beta(z) + \mathrm{e}^{-\nu G}z, \tag{2.4}$$

we obtain the following functional equation for $\mathbb{E}[z^X]$:

$$\mathbb{E}[z^X] = \sigma(z)\mathrm{e}^{-\lambda(1-\beta(z))G}\mathbb{E}[f(z)^X].$$

Introducing $K(z) := \sigma(z)\mathrm{e}^{-\lambda(1-\beta(z))G}$ and $X(z) := \mathbb{E}[z^X]$, we have:

$$X(z) = K(z)X(f(z)). \tag{2.5}$$

This is a functional equation that naturally occurs in the study of queueing models which have a branching-type structure; see, e.g., [7] and [20]. Typically, one may view customers who newly arrive into the system during a service as children of the served customer ("branching"), and customers who newly arrive into the system during a vacation or glue period as immigrants. Such a functional equation may be solved by iteration, giving rise to an infinite product – where the jth term in the product typically corresponds to customers who descend from an ancestor of j generations before. In this particular case we have after n iterations:

$$X(z) = \prod_{j=0}^{n} K(f^{(j)}(z)) X(f^{(n+1)}(z)), \tag{2.6}$$

where $f^{(0)}(z) := z$ and $f^{(j)}(z) := f(f^{(j-1)}(z))$, $j = 1, 2, \ldots$. Below we show that this product converges for $n \to \infty$ iff $\rho < 1$, thus proving the following theorem:

Theorem 1. *If $\rho < 1$ then the generating function $X(z) = \mathbb{E}[z^X]$ is given by*

$$X(z) = \prod_{j=0}^{\infty} K(f^{(j)}(z)). \tag{2.7}$$

Proof. Equation (2.5) is an equation for a branching process with immigration, where the number of immigrants has generating function $K(z)$ and the number of children in the branching process has generating function $f(z)$. Clearly, $K'(1) = \lambda \mathbb{E}S + \lambda \rho G < \infty$ and $f'(1) = \mathrm{e}^{-\nu G} + \left(1 - \mathrm{e}^{-\nu G}\right) \rho < 1$, if $\rho < 1$. The result of the theorem now follows directly from the theory of branching processes with immigration (see e.g., Theorem 1 on page 263 in Athreya and Ney [4]). $\square$

Having obtained an expression for $\mathbb{E}[z^X]$ in Theorem 1, expressions for $\mathbb{E}[z^Z]$ and $\mathbb{E}[z_q^{Y^{(q)}} z_o^{Y^{(o)}}]$ immediately follow from (2.2) and (2.3). Moments of X may be obtained from Theorem 1, but it is also straightforward to obtain $\mathbb{E}X$ from Equations (2.1)-(2.3):

$$\mathbb{E}X = \lambda \mathbb{E}S + \mathbb{E}Z, \tag{2.8}$$
$$\mathbb{E}Z = \rho \mathbb{E}Y^{(q)} + \mathbb{E}Y^{(o)}, \tag{2.9}$$
$$\mathbb{E}Y^{(q)} = \lambda G + (1 - \mathrm{e}^{-\nu G}) \mathbb{E}X, \tag{2.10}$$
$$\mathbb{E}Y^{(o)} = \mathrm{e}^{-\nu G} \mathbb{E}X, \tag{2.11}$$

yielding

$$\mathbb{E}X = \frac{\lambda \mathbb{E}S + \lambda \rho G}{(1-\rho)(1 - \mathrm{e}^{-\nu G})}. \tag{2.12}$$

Hence

$$\mathbb{E}Y^{(q)} = \lambda G + (1 - \mathrm{e}^{-\nu G}) \frac{\lambda \mathbb{E}S + \lambda \rho G}{(1-\rho)(1 - \mathrm{e}^{-\nu G})} = \frac{\lambda \mathbb{E}S + \lambda G}{1 - \rho}, \tag{2.13}$$

$$\mathbb{E}Y^{(o)} = \mathrm{e}^{-\nu G} \frac{\lambda \mathbb{E}S + \lambda \rho G}{(1-\rho)(1 - \mathrm{e}^{-\nu G})}, \tag{2.14}$$

$$\mathbb{E}Z = \frac{\lambda \rho G + \lambda \mathbb{E}S[\rho(1 - \mathrm{e}^{-\nu G}) + \mathrm{e}^{-\nu G}]}{(1-\rho)(1 - \mathrm{e}^{-\nu G})}. \tag{2.15}$$

Notice that the denominators of the above expressions equal $1 - f'(1)$. Also notice that it makes sense that the denominators contain both the factor $1 - \rho$ and the probability $1 - \mathrm{e}^{-\nu G}$ that a retrial returns during a glue period.

In a similar way as the first moments of X, $Y^{(q)}$, $Y^{(o)}$ and Z have been obtained, we can also obtain their second moment. For further use we here calculate $\mathbb{E}[X(X-1)]$:

$$\begin{aligned} \mathbb{E}[X(X-1)] = & \frac{K''(1)}{(1-\rho)(1-\mathrm{e}^{-\nu G})(1+\rho(1-\mathrm{e}^{-\nu G})+\mathrm{e}^{-\nu G})} \\ & + \frac{K'(1)[2K'(1)(\rho(1-\mathrm{e}^{-\nu G})+\mathrm{e}^{-\nu G})+(1-\mathrm{e}^{-\nu G})\lambda^2\mathbb{E}B^2]}{(1-\rho)^2(1-\mathrm{e}^{-\nu G})^2(1+\rho(1-\mathrm{e}^{-\nu G})+\mathrm{e}^{-\nu G})}, \end{aligned} \tag{2.16}$$

where $K'(1) = \lambda \mathbb{E}S + \lambda\rho G$ and $K''(1) = \lambda^2\mathbb{E}S^2 + 2\rho\lambda^2 G\mathbb{E}S + \lambda^3 G\mathbb{E}B^2 + (\lambda G\rho)^2$.

Remark 1. Special cases of the above analysis are, e.g.:
(i) Vacations of length zero. Simply take $\sigma(z) \equiv 1$ and $\mathbb{E}S = 0$ in the above formulas.
(ii) $\nu = \infty$. Retrials now always return during a glue period. We then have $f(z) = \beta(z)$, which leads to minor simplifications.

Remark 2. It seems difficult to handle the case of non-constant glue periods, as it seems to lead to a process with complicated dependencies. If G takes a few distinct values $G_1, \dots, G_N$ with different probabilities, then one might still be able to obtain a kind of multinomial generalization of the infinite product featuring in Theorem 1. One would then have several functions $f_i(z) := (1-\mathrm{e}^{-\nu G_i})\beta(z) + \mathrm{e}^{-\nu G_i}z$, and all possible combinations of iterations $f_i(f_h(f_k(\dots(z))))$ arising in functions $K_i(z) := \sigma(z)\mathrm{e}^{-\lambda(1-\beta(z))G_i}$, $i = 1, 2, \dots, N$. By way of approximation, one might stop the iterations after a certain number of terms, the number depending on the speed of convergence (hence on $1-\rho$ and on $1-\mathrm{e}^{-\nu G_i}$).

2.3 Queue length analysis at arbitrary time points

Having found the generating functions of the number of customers at the beginning of (i) glue periods ($\mathbb{E}[z^X]$), (ii) visit periods ($\mathbb{E}[z_q^{Y^{(q)}} z_o^{Y^{(o)}}]$), and (iii) vacation periods ($\mathbb{E}[z^Z]$), we can also obtain the generating function of the number of customers at arbitrary time points.

Theorem 2. *If $\rho < 1$, we have the following results:*

a) The joint generating function, $R^{va}(z_q, z_o)$, of the number of customers in the queue and in the orbit at an arbitrary time point in a vacation period is given by

$$R_{va}(z_q, z_o) = \mathbb{E}[z_o^Z]\frac{1-\tilde{S}(\lambda(1-z_o))}{\lambda(1-z_o)\mathbb{E}S}. \tag{2.17}$$

b) The joint generating function, $R^{gl}(z_q, z_o)$, of the number of customers in the queue and in the orbit at an arbitrary time point in a glue period is given by

$$R_{gl}(z_q, z_o) = \int_{t=0}^{G} \mathrm{e}^{-\lambda(1-z_q)t} \mathbb{E}[\{(1-\mathrm{e}^{-\nu t})z_q + \mathrm{e}^{-\nu t} z_o\}^X] \frac{dt}{G}. \tag{2.18}$$

c) The joint generating function, $R^{vi}(z_q, z_o)$, of the number of customers in the queue and in the orbit at an arbitrary time point in a visit period is given by

$$R_{vi}(z_q, z_o) = \frac{z_q \left[\mathbb{E}[z_q^{Y^{(q)}} z_o^{Y^{(o)}}] - \mathbb{E}[\tilde{B}(\lambda(1-z_o))^{Y^{(q)}} z_o^{Y^{(o)}}]\right]}{\mathbb{E}[Y^{(q)}] \left(z_q - \tilde{B}(\lambda(1-z_o))\right)} \frac{1-\tilde{B}(\lambda(1-z_o))}{\lambda(1-z_o)\mathbb{E}B}. \tag{2.19}$$

d) The joint generating function, $R(z_q, z_o)$, of the number of customers in the queue and in the orbit at an arbitrary time point is given by

$$R(z_q, z_o) = \rho R_{vi}(z_q, z_o) + (1-\rho)\tfrac{G}{G+\mathbb{E}S} R_{gl}(z_q, z_o) + (1-\rho)\tfrac{\mathbb{E}S}{G+\mathbb{E}S} R_{va}(z_q, z_o). \tag{2.20}$$

Proof.

a) Follows from the fact that during vacation periods the number of customers in the queue is 0 and the fact that the number of customers at an arbitrary time point in the orbit is the sum of two independent terms: The number of customers at the beginning of the vacation period and the number that arrived during the past part of the vacation period. The generating function of the latter is given by

$$\frac{1-\tilde{S}(\lambda(1-z_o))}{\lambda(1-z_o)\mathbb{E}S}.$$

b) Follows from the fact that if the past part of the glue period is equal to t, the generating function of the number of new arrivals in the queue during this period is equal to $\mathrm{e}^{-\lambda(1-z_q)t}$ and each customer present in the orbit at the beginning of the glue period is, independent of the others, still in orbit with probability $\mathrm{e}^{-\nu t}$ and has moved to the queue with probability $1-\mathrm{e}^{-\nu t}$.

c) During an arbitrary point in time in a visit period the number of customers in the system consists of two parts:
   - the number of customers in the system at the beginning of the service time of the customer currently in service, leading to the term (see Remark 3 below):

$$\frac{z_q \left(\mathbb{E}[z_q^{Y^{(q)}} z_o^{Y^{(o)}}] - \mathbb{E}[\tilde{B}(\lambda(1-z_o))^{Y^{(q)}} z_o^{Y^{(o)}}]\right)}{\mathbb{E}[Y^{(q)}] \left(z_q - \tilde{B}(\lambda(1-z_o))\right)}; \tag{2.21}$$

   - the number of customers that arrived during the past part of the service of the customer currently in service, leading to the term

$$\frac{1-\tilde{B}(\lambda(1-z_o))}{\lambda(1-z_o)\mathbb{E}B}.$$

d) Follows from the fact that the fraction of time the server is visiting Q is equal to ρ, and if the server is not visiting Q, with probability $\mathbb{E}S/(G+\mathbb{E}S)$ the server is on vacation and with probability $G/(G+\mathbb{E}S)$ the system is in a glue phase.

□

Remark 3. A straightforward way to prove (2.21) is to condition on the number of customers, say, j, in queue at the end of a glue period, and to average the number of customers in queue and in orbit over the beginnings of the j services. A more elegant proof of Formula (2.21) uses the observation that typically in vacation and polling systems each time a visit beginning or a service completion occurs, this coincides with a service beginning or a visit completion. This observation yields (see, e.g., Boxma, Kella and Kosinski [8])

$$\gamma\mathcal{V}_b(z_q,z_o)+\mathcal{S}_c(z_q,z_o)=\mathcal{S}_b(z_q,z_o)+\gamma\mathcal{V}_c(z_q,z_o).$$

Here, $\mathcal{V}_b(z_q,z_o)$ and $\mathcal{V}_c(z_q,z_o)$ are the joint generating functions of the number of customers in the queue and in the orbit at visit beginnings and visit completions, respectively. Similarly, $\mathcal{S}_b(z_q,z_o)$ and $\mathcal{S}_c(z_q,z_o)$ are the joint generating functions of the number of customers in the queue and in the orbit at service beginnings and service completions, respectively. Finally, γ is the reciprocal of the mean number of customers served per visit. Clearly,

$$\gamma=\frac{1}{\mathbb{E}[Y^{(q)}]},\quad \mathcal{V}_b(z_q,z_o)=\mathbb{E}[z_q^{Y^{(q)}}z_o^{Y^{(o)}}],\quad \mathcal{V}_c(z_q,z_o)=\mathbb{E}[\tilde{B}(\lambda(1-z_o))^{Y^{(q)}}z_o^{Y^{(o)}}],$$

and

$$\mathcal{S}_c(z_q,z_o)=\frac{\mathcal{S}_b(z_q,z_o)}{z_q}\tilde{B}(\lambda(1-z_o)),$$

which yields that $\mathcal{S}_b(z_q,z_o)$ is given by (2.21).

From Theorem 2, we now can obtain the steady-state mean number of customers in the system at arbitrary time points in vacation periods ($\mathbb{E}[R_{va}]$), in glue periods ($\mathbb{E}[R_{gl}]$), in visit periods ($\mathbb{E}[R_{vi}]$) and in arbitrary periods ($\mathbb{E}[R]$). These are given by

$$\begin{aligned}
\mathbb{E}[R_{va}]&=\mathbb{E}[Z]+\lambda\tfrac{\mathbb{E}[S^2]}{2\mathbb{E}[S]},\\
\mathbb{E}[R_{gl}]&=\mathbb{E}[X]+\lambda\tfrac{G}{2},\\
\mathbb{E}[R_{vi}]&=1+\lambda\tfrac{\mathbb{E}[B^2]}{2\mathbb{E}[B]}+\tfrac{\mathbb{E}[Y^{(q)}Y^{(o)}]}{\mathbb{E}[Y^{(q)}]}+\tfrac{(1+\rho)\mathbb{E}[Y^{(q)}(Y^{(q)}-1)]}{2\mathbb{E}[Y^{(q)}]},\\
\mathbb{E}[R]&=\rho\mathbb{E}[R_{vi}]+(1-\rho)\tfrac{G}{G+\mathbb{E}S}\mathbb{E}[R_{gl}]+(1-\rho)\tfrac{\mathbb{E}S}{G+\mathbb{E}S}\mathbb{E}[R_{va}]. \qquad (2.22)
\end{aligned}$$

Remark that the quantities $\mathbb{E}[Y^{(q)}Y^{(o)}]$ and $\mathbb{E}[Y^{(q)}(Y^{(q)}-1)]$ can be obtained using (2.3):

$$\begin{aligned}
\mathbb{E}[Y^{(q)}Y^{(o)}]&=\lambda G\mathrm{e}^{-\nu G}\mathbb{E}[X]+\left(1-\mathrm{e}^{-\nu G}\right)\mathrm{e}^{-\nu G}\mathbb{E}[X(X-1)],\\
\mathbb{E}[Y^{(q)}(Y^{(q)}-1)]&=(\lambda G)^2+\left(1-\mathrm{e}^{-\nu G}\right)^2\mathbb{E}[X(X-1)]+2\lambda G\left(1-\mathrm{e}^{-\nu G}\right)\mathbb{E}[X].
\end{aligned}$$

By combining these relations with (2.22), (2.8), (2.12) and (2.16), we obtain – after tedious calculations – the following relatively simple expression for the mean number of customers $\mathbb{E}[R]$:

$$\mathbb{E}[R] = \rho + \frac{\lambda^2 \mathbb{E}[B^2]}{2(1-\rho)} + \frac{\lambda \mathbb{E}[(G+S)^2]}{2\mathbb{E}[G+S]} + \frac{\lambda\rho\mathbb{E}[G+S]}{1-\rho} + \lambda(\rho G + \mathbb{E}[S]) \frac{\mathrm{e}^{-\nu G}}{(1-\rho)(1-\mathrm{e}^{-\nu G})}, \tag{2.23}$$

which we rewrite for later purposes as

$$\mathbb{E}[R] = \rho + \frac{\lambda^2 \mathbb{E}[B^2]}{2(1-\rho)} + \lambda \frac{\mathbb{E}[S]}{G+\mathbb{E}[S]} \frac{\mathbb{E}[S^2]}{2\mathbb{E}[S]} + \lambda \frac{\mathbb{E}[S]}{G+\mathbb{E}[S]} G + \lambda \frac{G}{G+\mathbb{E}[S]} \frac{G}{2}$$
$$+ \frac{\lambda\rho(G+\mathbb{E}[S])}{1-\rho} + \lambda(\rho G + \mathbb{E}[S]) \frac{\mathrm{e}^{-\nu G}}{(1-\rho)(1-\mathrm{e}^{-\nu G})}. \tag{2.24}$$

Remark 4. (i). It should be noticed that the first two terms in the righthand side of (2.23) together represent the mean number of customers in the ordinary $M/G/1$ queue, without vacations and glue periods. The third term represents the mean number of arrivals during the residual part of a vacation plus glue period. The fourth term can be interpreted as the mean number of arrivals during a visit period of the server (since the mean length of one cycle, i.e., visit plus vacation plus glue period, is via a balance argument seen to equal $\mathbb{E}[C] = \frac{G+\mathbb{E}[S]}{1-\rho}$, while a mean visit period equals $\rho\mathbb{E}[C]$). The fifth term is the only term involving the retrial rate ν. In particular, that term disappears when $\nu \to \infty$. In the latter case, our model reduces to an $M/G/1$ queue with gated vacations, with vacation lengths $G + S$. The resulting expression for the mean number of customers coincides with formula (5.23) of [23] (see also formula (3.2.6) of [26]).
(ii). A more interesting limiting operation is to simultaneously let $\nu \to \infty$ and $G \downarrow 0$, such that νG remains constant. The resulting model is an $M/G/1$ queue with vacations and binomially gated service; see, e.g., Levy [15]. Here, each customer who is present at the end of a vacation, will be served in the next visit period with probability $p = 1 - \mathrm{e}^{-\nu G}$. In this case, the mean number of customers in the system is given by

$$\mathbb{E}[R] = \rho + \frac{\lambda^2 \mathbb{E}[B^2]}{2(1-\rho)} + \frac{\lambda \mathbb{E}[S^2]}{2\mathbb{E}[S]} + \frac{\lambda\rho\mathbb{E}[S]}{1-\rho} + \frac{\lambda\mathbb{E}[S](1-p)}{p(1-\rho)}. \tag{2.25}$$

This formula coincides with the results obtained in [15] (see e.g., formula (7.1) with $N = 1$ in [15] for the mean sojourn time of customers in this model). Observe that our formula, after application of Little's formula, does *not* fully agree with the mean delay expression (5.50b) in [23] and with a similar formula on p. 90 of [26]. Those mean delay expressions for the binomial gated model seem to refer to the case where customers who are not served during a visit (w.p. $1 - p$) are lost; hence factors like $\frac{1}{1-p\rho}$ in those mean delay expressions.
(iii). Formula (2.24) immediately shows how the mean number of customers behaves for very small and for very large values of the glue period length G:

$$\mathbb{E}[R] \sim \frac{\lambda\mathbb{E}[S]}{G\nu(1-\rho)}, \quad G \downarrow 0, \tag{2.26}$$

and

$$\mathbb{E}[R] \sim \frac{\lambda(1+\rho)}{2(1-\rho)} G, \quad G \to \infty. \tag{2.27}$$

In subsection 2.5 we explore the effect of G on $\mathbb{E}[R]$ more deeply.

2.4 An alternative derivation of the mean number of customers

Alternatively, equation (2.24) can be obtained using mean value analysis. Let T denote the steady-state sojourn time of customers in the system. The mean value relations for the system are given by Little's law

$$\mathbb{E}[R] = \lambda \mathbb{E}[T] \tag{2.28}$$

and the arrival relation

$$\begin{aligned} \mathbb{E}[T] = {} & (\mathbb{E}[R] - \rho)\,\mathbb{E}[B] + \rho \frac{\mathbb{E}[B^2]}{2\mathbb{E}[B]} + \mathbb{E}[B] \\ & + \rho \frac{\mathbb{E}[S] + G}{1 - e^{-\nu G}} + (1-\rho) \frac{G}{G + \mathbb{E}[S]} \frac{G}{2} \\ & + (1-\rho) \frac{\mathbb{E}[S]}{G + \mathbb{E}[S]} \left(\frac{\mathbb{E}[S^2]}{2\mathbb{E}[S]} + G + e^{-\nu G} \frac{G + \mathbb{E}[S]}{1 - e^{-\nu G}} \right). \end{aligned} \tag{2.29}$$

Combining (2.28) and (2.29), one readily arrives at (2.24). The main ideas to obtain (2.29) are the following:

1. The steady-state sojourn time can be written as the sum of the busy time of the server during the sojourn time and the idle time of the server during the sojourn time.
2. The mean busy time of the server during the sojourn time is given by

   $$(\mathbb{E}[R] - \rho)\,\mathbb{E}[B] + \rho \frac{\mathbb{E}[B^2]}{2\mathbb{E}[B]} + \mathbb{E}[B].$$

   This follows by changing the order of service in the system into the First Come First Serve order (remark that changing the order of service does not change the mean busy time during a sojourn time).
3. If a customer arrives during a service period, the mean idle of the server during its sojourn time is given by a geometric number of vacation plus glue periods. This gives the $\rho \frac{\mathbb{E}[S]+G}{1-e^{-\nu G}}$ in (2.29).
4. If a customer arrives during a glue period, the mean idle of the server during its sojourn time is given by $G/2$. This gives the $(1-\rho)\frac{G}{G+\mathbb{E}[S]}\frac{G}{2}$ in (2.29).
5. If a customer arrives during a vacation period, the mean idle of the server during its sojourn time is given by a residual vacation period plus a glue period plus, with probability $e^{-\nu G}$, a geometric number of vacation plus glue periods. This gives the term

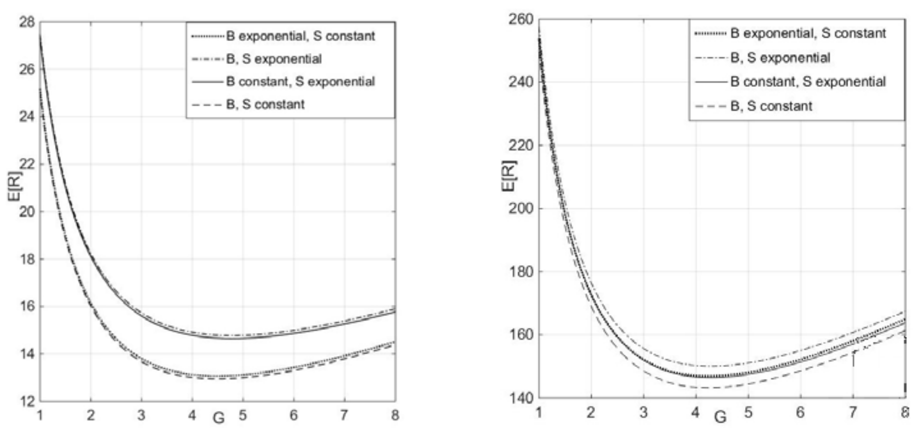
   $$(1-\rho) \frac{\mathbb{E}[S]}{G + \mathbb{E}[S]} \left(\frac{\mathbb{E}[S^2]}{2\mathbb{E}[S]} + G + e^{-\nu G} \frac{G + \mathbb{E}[S]}{1 - e^{-\nu G}} \right)$$

   in (2.29).

2.5 Optimizing the length of the glue period

The results of the previous subsections can, e.g., be used to determine the value of G which minimizes the mean number of customers in the system at any arbitrary time point. The mean sojourn time of an arbitrary customer follows from Little's formula. Therefore we can find the value of G which minimizes the mean sojourn time of an arbitrary customer.

Let us first present a sample of numerical results that we obtained for $\mathbb{E}[R]$ as a function of G. We consider four cases: (i) the service time distribution and vacation time distribution are exponential, (ii) the service time and vacation time are constant, (iii) the service time distribution is exponential and the vacation time is constant and (iv) the service time is constant and the vacation time distribution is exponential. For all these cases we assume that the expected service time $\mathbb{E}B = 1$, expected vacation time $\mathbb{E}S = 10$ and retrial rate $\nu = 0.5$. We plot G vs $\mathbb{E}[R]$ for $\lambda = 0.5, 0.9$.



(a) $\lambda = 0.5$ (b) $\lambda = 0.9$

Fig. 1: Exponential service and vacation time distributions

From the examples in Fig 1 we observe the following results:

- The mean number of customers at any arbitrary point seems to be convex w.r.t. glue period length, i.e., there exists a glue period G_{min} where the system has minimum mean number of customers $E[R_{min}]$ and hence minimum mean sojourn time.
- For a very small G, $\mathbb{E}[R]$ decreases exponentially with G (as confirmed by (2.26)).
- For a large G, $\mathbb{E}[R]$ increases linearly with G (as confirmed by (2.27)).

Indeed, if G is very small, the number of customers joining the queue in each glue period is very small and thereby the efficiency of the system is decreased.

On the other hand, a large G means the system stays in the glue period for a long time and this decreases the efficiency of the system. Hence it is logical to have a G_{min} which optimizes the system.

We will now prove that $E[R]$ is indeed convex in G. Differentiating the expression for $\mathbb{E}[R]$ in Equation (2.23) w.r.t. G gives

$$\frac{d}{dG}\mathbb{E}[R] = \lambda - \frac{\lambda\mathbb{E}[(G+S)^2]}{2\mathbb{E}[G+S]^2} + \frac{\lambda\rho}{(1-\rho)(1-\mathrm{e}^{-\nu G})} - \lambda(\rho G+\mathbb{E}[S])\frac{\nu\mathrm{e}^{-\nu G}}{(1-\rho)(1-\mathrm{e}^{-\nu G})^2}. \tag{2.30}$$

Differentiating (2.30) again w.r.t. G gives

$$\frac{d^2}{dG^2}\mathbb{E}[R] = \frac{\lambda}{\mathbb{E}[G+S]}\left[\frac{\mathbb{E}[(G+S)^2]}{\mathbb{E}[G+S]^2} - 1\right] + \frac{\lambda\nu\mathrm{e}^{-\nu G}}{(1-\rho)(1-\mathrm{e}^{-\nu G})^2}\left[\nu(\rho G+\mathbb{E}[S])\frac{(1+\mathrm{e}^{-\nu G})}{(1-\mathrm{e}^{-\nu G})} - 2\rho\right]. \tag{2.31}$$

The first term in the righthand side of (2.31) is nonnegative, because $\mathbb{E}[(G+S)^2] \geq \mathbb{E}[G+S]^2$.

Let

$$Q(G) := \nu(\rho G+\mathbb{E}[S])\frac{(1+\mathrm{e}^{-\nu G})}{(1-\mathrm{e}^{-\nu G})} - 2\rho.$$

We can see $Q(G) \longrightarrow \infty$ as $G \longrightarrow 0$ or $G \longrightarrow \infty$. Let $Q(G)$ attain its minimum at $G = g$. Hence at $G = g$,

$$\frac{d}{dG}Q(G) = 0,$$

which implies

$$\rho g + \mathbb{E}[S] = \frac{\rho}{2\nu\mathrm{e}^{-\nu g}}(1-\mathrm{e}^{-2\nu g}).$$

Therefore

$$Q(g) = \frac{\rho}{2}[\mathrm{e}^{\nu g} - 2 + \mathrm{e}^{-\nu g}] \geq 0.$$

We observe that the minimum value of $Q(G)$ is always nonnegative. Since both terms of (2.31) are nonnegative, $\frac{d^2}{dG^2}\mathbb{E}[R] \geq 0$.

Hence $\mathbb{E}[R]$, the mean number of customers at an arbitrary point of time in the system, is convex in G. So the system can improve the quality of service by setting an optimal value G for the fixed glue period.

In Table 1 we analyze the behaviour of G_{min} and $\mathbb{E}[R_{min}]$ as we increase $\mathbb{E}[S]$ for an exponential distribution $B(\cdot)$ with $\mathbb{E}B = 1$, arrival rate $\lambda = 0.5$ and retrial rate $\nu = 0.5$.

Table 1 suggests that, in the case under consideration,

- G_{min} and $\mathbb{E}[R_{min}]$ increase when $\mathbb{E}[S]$ increases.
- G_{min} and $\mathbb{E}[R_{min}]$ increase when the variance of S becomes larger.
- When $\mathbb{E}S$ approaches 0, G_{min} also approaches 0. When there is no customer in the queue, the system will then have a series of very short glue periods, and when a customer arrives or returns from orbit, it can almost instantaneously be taken into service. In this case, the system reduces to an ordinary $M/G/1$ retrial queue; indeed, Formula (2.23) reduces to $\mathbb{E}[R] = \rho + \frac{\lambda^2\mathbb{E}[b^2]}{2(1-\rho)} + \frac{\lambda\rho}{\nu(1-\rho)}$ which is in agreement with Formula (1.37) of [10].

Table 1: Exponential service time distribution, $\mathbb{E}B = 1$, $\lambda = 0.5$ and $\nu = 0.5$

$\mathbb{E}S$	G_{min} for exponential $S(\cdot)$	G_{min} for constant $S(\cdot)$	$\mathbb{E}[R_{min}]$ for exponential $S(\cdot)$	$\mathbb{E}[R_{min}]$ for constant $S(\cdot)$
0	$\epsilon, \quad \epsilon \longrightarrow 0$	$\epsilon, \quad \epsilon \longrightarrow 0$	~ 2	~ 2
0.1	0.608	0.606	2.476	2.473
0.5	1.334	1.320	3.419	3.385
1	1.846	1.801	4.259	4.170
5	3.694	3.508	9.276	8.549
10	4.785	4.495	14.778	13.070
50	7.783	7.225	56.027	45.156
100	9.167	8.521	106.606	83.634

3 Queue length analysis for the N-queue case

3.1 Model description

In this section we consider a one-server polling model with multiple queues, Q_i, $i = 1, \cdots, N$. Customers arrive at Q_i according to a Poisson process with rate λ_i; they are called type-i customers, $i = 1, \cdots, N$. The service times at Q_i are i.i.d. r.v., with B_i denoting a generic service time, with distribution $B_i(\cdot)$ and LST $\tilde{B}_i(\cdot)$, $i = 1, \cdots, N$. The server follows cyclic polling, thus after a visit of Q_i, it switches to Q_{i+1}. Successive switchover times from Q_i to Q_{i+1} are i.i.d. r.v., with S_i a generic switchover time, with distribution $S_i(\cdot)$ and LST $\tilde{S}_i(\cdot)$, $i = 1, \cdots, N$. We make all the usual independence assumptions about interarrival times, service times and switchover times at the queues. After a switch of the server to Q_i, there first is a deterministic (i.e., constant) glue period G_i, before the visit of the server at Q_i begins, $i = 1, \cdots, N$. As in the one-queue case, the significance of the glue period stems from the following assumption. Customers who arrive at Q_i do not receive service immediately. When customers arrive at Q_i during a glue period G_i, they stick, joining the queue of Q_i. When they arrive in any other period, they immediately leave and retry after a retrial interval which is independent of everything else, and which is exponentially distributed with rate ν_i, $i = 1, \cdots, N$.

The service discipline at all queues is gated: During the visit period at Q_i, the server serves all "glued" customers in that queue, i.e., all type-i customers waiting at the end of the glue period – but none of those in orbit, and neither any new arrivals.

We are interested in the steady-state behaviour of this polling model with retrials. We hence assume that the stability condition $\sum_{i=1}^{N} \rho_i < 1$ holds, where $\rho_i := \lambda_i \mathbb{E}B_i$.

Some more notation:
G_{ni} denotes the nth glue period of Q_i.
V_{ni} denotes the nth visit period of Q_i.
S_{ni} denotes the nth switch period out of Q_i, $i = 1, \cdots, N$.

$(X^{(i)}_{n1}, X^{(i)}_{n2}, \cdots, X^{(i)}_{nN})$ denotes the vector of numbers of customers of type 1 to type N in the system (hence in orbit) at the start of G_{ni}, $i = 1, \cdots, N$.
$(Y^{(i)}_{n1}, Y^{(i)}_{n2}, \cdots, Y^{(i)}_{nN})$ denotes the vector of numbers of customers of type 1 to type N in the system at the start of V_{ni}, $i = 1, \cdots, N$. We distinguish between those who are queueing in Q_i and those who are in orbit for Q_i: We write $Y^{(i)}_{ni} = Y^{(iq)}_{ni} + Y^{(io)}_{ni}$, $i = 1, \cdots, N$, where q denotes queueing and o denotes in orbit.
Finally,
$(Z^{(i)}_{n1}, Z^{(i)}_{n2}, \cdots, Z^{(i)}_{nN})$ denotes the vector of numbers of customers of type 1 to type N in the system (hence in orbit) at the start of S_{ni}, $i = 1, \cdots, N$.

3.2 Queue length analysis

In this section we study the steady-state joint distribution of the numbers of customers in the system at beginnings of glue periods. This will also immediately yield the steady-state joint distributions of the numbers of customers in the system at the beginnings of visit periods and of switch periods. We follow a similar generating function approach as in the one-queue case, throughout making the following assumption regarding the parameters of the generating functions: $|z_i| \leq 1$, $|z_{iq}| \leq 1$, $|z_{io}| \leq 1$. Observe that the generating function of the vector of numbers of arrivals at Q_1 to Q_N during a type-i service time B_i is $\beta_i(z_1, z_2, \cdots, z_N) := \tilde{B}_i(\sum_{j=1}^N \lambda_j(1 - z_j))$. Similarly, the generating function of the vector of numbers of arrivals at Q_1 to Q_N during a type-i switchover time S_i is $\sigma_i(z_1, z_2, \cdots, z_N) := \tilde{S}_i(\sum_{j=1}^N \lambda_j(1 - z_j))$. Since the server serves Q_{i+1} after Q_i we can successively express (in terms of generating functions) $(X^{(i+1)}_{n1}, X^{(i+1)}_{n2}, \cdots, X^{(i+1)}_{nN})$ into $(Z^{(i)}_{n1}, Z^{(i)}_{n2}, \cdots, Z^{(i)}_{nN})$, $(Z^{(i)}_{n1}, Z^{(i)}_{n2}, \cdots, Z^{(i)}_{nN})$ into $(Y^{(i)}_{n1}, Y^{(i)}_{n2}, \cdots, Y^{(iq)}_{ni}, Y^{(io)}_{ni}, \cdots, Y^{(i)}_{nN})$, and $(Y^{(i)}_{n1}, Y^{(i)}_{n2}, \cdots, Y^{(iq)}_{ni}, Y^{(io)}_{ni}, \cdots, Y^{(i)}_{nN})$ into $(X^{(i)}_{n1}, X^{(i)}_{n2}, \cdots, X^{(i)}_{nN})$; etc. Denote by $(X^{(i)}_1, X^{(i)}_2, \cdots, X^{(i)}_N)$ the vector with as distribution the limiting distribution of $(X^{(i)}_{n1}, X^{(i)}_{n2}, \cdots, X^{(i)}_{nN})$, $i = 1, \cdots, N$, and similarly introduce $(Z^{(i)}_1, Z^{(i)}_2, \cdots, Z^{(i)}_N)$ and $(Y^{(i)}_1, Y^{(i)}_2, \cdots, Y^{(i)}_N)$, with $Y^{(i)}_i = Y^{(iq)}_i + Y^{(io)}_i$, for $i = 1, \cdots, N$. We have:

$$\mathbb{E}[z_1^{X_1^{(i+1)}} z_2^{X_2^{(i+1)}} \cdots z_N^{X_N^{(i+1)}}] = \sigma_i(z_1, z_2, \cdots, z_N)\mathbb{E}[z_1^{Z_1^{(i)}} z_2^{Z_2^{(i)}} \cdots z_N^{Z_N^{(i)}}]. \tag{3.1}$$

$$\mathbb{E}[z_1^{Z_1^{(i)}} z_2^{Z_2^{(i)}} \cdots z_N^{Z_N^{(i)}} | Y_1^{(i)} = h_1, Y_2^{(i)} = h_2, \cdots, Y_i^{(iq)} = h_{iq}, Y_i^{(io)} = h_{io}, \cdots, Y_N^{(i)} = h_N]$$
$$= \Big(\prod_{j \neq i} z_j^{h_j}\Big) z_i^{h_{io}} [\beta_i(z_1, z_2, \cdots, z_N)]^{h_{iq}},$$

yielding

$$\mathbb{E}[z_1^{Z_1^{(i)}} z_2^{Z_2^{(i)}} \cdots z_N^{Z_N^{(i)}}] = \mathbb{E}[[\beta_i(z_1, z_2, \cdots, z_N)]^{Y_i^{(iq)}} z_1^{Y_1^{(i)}} z_2^{Y_2^{(i)}} \cdots z_i^{Y_i^{(io)}} \cdots z_N^{Y_N^{(i)}}]. \tag{3.2}$$

Furthermore,

$$\begin{aligned}&\mathbb{E}[z_1^{Y_1^{(i)}} z_2^{Y_2^{(i)}} \cdots z_{iq}^{Y_i^{(iq)}} z_{io}^{Y_i^{(io)}} \cdots z_N^{Y_N^{(i)}} | X_1^{(i)} = a_1, X_2^{(i)} = a_2, \cdots, X_N^{(i)} = a_N] \\ &= \Big(\prod_{j\neq i} z_j^{a_j} \mathrm{e}^{-\lambda_j(1-z_j)G_i}\Big) \mathrm{e}^{-\lambda_i(1-z_{iq})G_i} [(1-\mathrm{e}^{-\nu_i G_i}) z_{iq} + \mathrm{e}^{-\nu_i G_i} z_{io}]^{a_i},\end{aligned}$$

yielding

$$\begin{aligned}\mathbb{E}[z_1^{Y_1^{(i)}} z_2^{Y_2^{(i)}} \cdots z_{iq}^{Y_i^{(iq)}} z_{io}^{Y_i^{(io)}} \cdots z_N^{Y_N^{(i)}}] &= \Big(\prod_{j\neq i} \mathrm{e}^{-\lambda_j(1-z_j)G_i}\Big) \mathrm{e}^{-\lambda_i(1-z_{iq})G_i} \\ &\times \mathbb{E}[\Big(\prod_{j\neq i} z_j^{X_j^{(i)}}\Big) [(1-\mathrm{e}^{-\nu_i G_i}) z_{iq} + \mathrm{e}^{-\nu_i G_i} z_{io}]^{X_i^{(i)}}].\end{aligned} \tag{3.3}$$

It follows from (3.1), (3.2) and (3.3), with

$$f_i(z_1, z_2, \cdots, z_N) := (1-\mathrm{e}^{-\nu_i G_i})\beta_i(z_1, z_2, \cdots, z_N) + \mathrm{e}^{-\nu_i G_i} z_i,$$

that

$$\begin{aligned}\mathbb{E}[z_1^{X_1^{(i+1)}} z_2^{X_2^{(i+1)}} \cdots z_N^{X_N^{(i+1)}}] &= \sigma_i(z_1, z_2, \cdots, z_N) \Big(\prod_{j\neq i} \mathrm{e}^{-\lambda_j(1-z_j)G_i}\Big) \mathrm{e}^{-\lambda_i(1-\beta_i(z_1,z_2,\cdots,z_N))G_i} \\ &\times \mathbb{E}[\Big(\prod_{j\neq i} z_j^{X_j^{(i)}}\Big) [f_i(z_1, z_2, \cdots, z_N)]^{X_i^{(i)}}].\end{aligned} \tag{3.4}$$

Let z denote vector $(z_1, z_2, \cdots, z_N)$; further

$$h_i(z) := f_i(z_1, \cdots, z_i, h_{i+1}(z), \cdots, h_N(z)) \tag{3.5}$$

and

$$h_N(z) := f_N(z_1, \cdots, z_N). \tag{3.6}$$

Since the server moves to Q_1 after Q_N, substituting $i = N$ in (3.4), we have

$$\begin{aligned}\mathbb{E}[z_1^{X_1^{(1)}} z_2^{X_2^{(1)}} \cdots z_N^{X_N^{(1)}}] &= \sigma_N(z_1, z_2, \cdots, z_N) \Big(\prod_{j\neq N} \mathrm{e}^{-\lambda_j(1-z_j)G_N}\Big)\ \mathrm{e}^{-\lambda_N(1-\beta_N(z_1,z_2,\cdots,z_N))G_N} \\ &\times \mathbb{E}[\Big(\prod_{j\neq N} z_j^{X_j^{(N)}}\Big) [h_N(z)]^{X_N^{(N)}}].\end{aligned} \tag{3.7}$$

From (3.4) we have

$$\begin{aligned}\mathbb{E}[\Big(\prod_{j\neq N} z_j^{X_j^{(N)}}\Big)[h_N(z)]^{X_N^{(N)}}] &= \sigma_{N-1}(z_1, z_2, \cdots, z_{N-1}, h_N(z)) \Big(\prod_{j=1}^{N-2} \mathrm{e}^{-\lambda_j(1-z_j)G_{N-1}}\Big) \\ &\times \mathrm{e}^{-\lambda_{N-1}(1-\beta_{N-1}(z_1,z_2,\cdots,z_{N-1},h_N(z)))G_{N-1}}\quad \mathrm{e}^{-\lambda_N(1-h_N(z))G_{N-1}} \\ &\times \mathbb{E}[\Big(\prod_{j=1}^{N-2} z_j^{X_j^{(N-1)}}\Big) [h_{N-1}(z)]^{X_{N-1}^{(N-1)}} [h_N(z)]^{X_N^{(N-1)}}].\end{aligned} \tag{3.8}$$

From (3.7) and (3.8) we have

$$\begin{aligned}
\mathbb{E}[z_1^{X_1^{(1)}} z_2^{X_2^{(1)}} \cdots z_N^{X_N^{(1)}}] &= \sigma_N(z_1, z_2, \cdots, z_N)\Big(\prod_{j\neq N} \mathrm{e}^{-\lambda_j(1-z_j)G_N}\Big)\mathrm{e}^{-\lambda_N(1-\beta_N(z_1,z_2,\cdots,z_N))G_N} \\
&\quad \times\sigma_{N-1}(z_1, \cdots, z_{N-1}, h_N(z))\Big(\prod_{j=1}^{N-2} \mathrm{e}^{-\lambda_j(1-z_j)G_{N-1}}\Big)\mathrm{e}^{-\lambda_{N-1}(1-\beta_{N_1}(z_1,\cdots,z_{N-1},h_N(z)))G_N} \\
&\quad \times\mathrm{e}^{-\lambda_N(1-h_N(z))G_{N-1}}\mathbb{E}[\Big(\prod_{j=1}^{N-2} z_j^{X_j^{(N-1)}}\Big)[h_{N-1}(z)]^{X_{N-1}^{(N-1)}}[h_N(z)]^{X_N^{(N-1)}}]. \\
&= \Big[\prod_{i=N-1}^{N} \sigma_i(z_1, \cdots, z_i, h_{i+1}(z), \cdots, h_N(z)) \\
&\quad \times\Big(\prod_{j=1}^{i-1} \mathrm{e}^{-\lambda_j(1-z_j)G_i}\Big)\Big(\mathrm{e}^{-\lambda_i(1-\beta_i(z_1,\cdots,z_i,h_{i+1}(z),\cdots,h_N(z)))G_i}\Big)\Big(\prod_{j=i+1}^{N} \mathrm{e}^{-\lambda_j(1-h_j(z))G_i}\Big)\Big] \\
&\quad \times\mathbb{E}[\Big(\prod_{j=1}^{N-2} z_j^{X_j^{(N-1)}}\Big)[h_{N-1}(z)]^{X_{N-1}^{(N-1)}}[h_N(z)]^{X_N^{(N-1)}}].
\end{aligned}$$

By recursively substituting as above we get

$$\begin{aligned}
\mathbb{E}[z_1^{X_1^{(1)}} z_2^{X_2^{(1)}} \cdots z_N^{X_N^{(1)}}] &= \prod_{i=1}^{N} \sigma_i(z_1, \cdots, z_i, h_{i+1}(z), \cdots, h_N(z)) \prod_{i=1}^{N} \mathrm{e}^{-G_i D_i(z)} \\
&\quad \times\mathbb{E}\big[[h_1(z)]^{X_1^{(1)}}, [h_2(z)]^{X_2^{(1)}}, \cdots, [h_N(z)]^{X_N^{(1)}}\big],
\end{aligned} \tag{3.9}$$

where

$$D_i(z) = \sum_{j=1}^{i-1} \lambda_j(1-z_j)+\lambda_i\Big(1-\beta_i(z_1, \cdots, z_i, h_{i+1}(z), \cdots, h_N(z)\Big)+\sum_{j=i+1}^{N} \lambda_j(1-h_j(z)).$$

Equation (3.9) can be divided into three factors, representing the switchover period, glue period and visit period respectively. The first factor, for a particular value of i, represents the arrivals during the switchover time after the visit of Q_i. The second factor represents the arrivals during the glue period before a visit of Q_i. It is further divided into three generating functions. First are the arrivals of type $j < i$; these don't have any further effect on the system. Then the arrivals of type i, these are served during the following visit and produce new children (i.e., arrivals during their service) of each type. Finally those of type $j > i$ which may or may not be served in future visits and if served produce new children of each type. These two factors are taken for all $i = 1, \cdots, N$. The third factor represents the descendants (arrivals during services, arrivals during services of customers who arrived during services, etc.) of $(X_1^{(1)}, \cdots, X_N^{(1)})$.

Consider

$$X(z) = \mathbb{E}[\Big(\prod_{j=1}^{N} z_j^{X_j^{(1)}}\Big)],$$

with an obvious definition of $K(z)$, we can rewrite (3.9) into

$$X(z) = K(z)X(h(z)) \tag{3.10}$$

where

$$h(z) := (h_1(z), \cdots, h_N(z))$$

and $h_i(z)$ are as defined in (3.5) and (3.6).

Define $h_i^{(0)}(z) = z_i$, $h_i^{(n)}(z) = h_i(h_1^{(n-1)}(z), h_2^{(n-1)}(z), \cdots, h_N^{(n-1)}(z))$, $i = 1, \cdots, N$.

Theorem 3. *If $\sum_i \rho_i < 1$, then the generating function $X(z)$ is given by*

$$X(z) = \prod_{m=0}^{\infty} K(h_1^{(m)}(z), h_2^{(m)}(z), \cdots, h_N^{(m)}(z)). \tag{3.11}$$

Proof. Equation (3.11) follows from (3.10) by iteration. We still need to prove that the infinite product converges if $\sum_i \rho_i < 1$. Equation (3.10) is an equation for a multi-type branching process with immigration, where the number of immigrants of different types has generating function $K(z)$ and the number of children of different types of a type i individual in the branching process has generating function $h_i(z)$, $i = 1, \cdots, N$. An important role in the analysis of such a process is played by the mean matrix M of the branching process,

$$M = \begin{pmatrix} m_{11} \cdots m_{1N} \\ \vdots \quad \ddots \quad \vdots \\ m_{N1} \cdots m_{NN} \end{pmatrix}, \tag{3.12}$$

where m_{ij} represents the mean number of children of type j of a type i individual. The elements of the matrix M are the same as given in Section 5 of Resing [20], which is

$$m_{ij} = f_{ij} \cdot 1[j \le i] + \sum_{k=i+1}^{N} f_{ik} m_{kj}, \tag{3.13}$$

where

$m_{ij} = \frac{\partial h_i}{\partial z_j}(1, 1, \cdots, 1)$ and $f_{ij} = \frac{\partial f_i}{\partial z_j}(1, 1, \cdots, 1)$.

We observe that the equation for m_{ij} is the sum of two terms. First the children of type $j \le i$, who do not affect the system in the future. Next the children of type j produced by the children of type $k > i$ in the subsequent visits.

The theory of multi-type branching processes with immigration (see Quine [19] and Resing [20]) now states that if (i) the expected total number of immigrants in a generation is finite and (ii) the maximal eigenvalue λ_{max} of the mean matrix M satisfies $\lambda_{max} < 1$, then the generating function of the steady state distribution of the process is given by (3.11). To complete the proof of Theorem 3, we shall now verify (i) and (ii).

Ad (i): The expected total number of immigrants in a generation is

$$\begin{aligned}
&\lambda_1(G_1(\sum_j m_{1j}) + \sum_{j=1}^{N} \mathbb{E}S_j + \sum_{j=2}^{N} G_j) \\
&+\lambda_2(((G_1+S_1)(1-\mathrm{e}^{-\nu_2 G_2}) + G_2)(\sum_j m_{2j}) + (G_1+S_1)\mathrm{e}^{-\nu_2 G_2} + \sum_{j=2}^{N} \mathbb{E}S_j + \sum_{j=3}^{N} G_j) \\
&+\cdots \\
&+\lambda_N(((\sum_{j=1}^{N-1}(G_j+S_j))(1-\mathrm{e}^{-\nu_N G_N}) + G_N)(\sum_j m_{Nj}) + \sum_{j=1}^{N-1}(G_j+S_j)\mathrm{e}^{-\nu_N G_N} + \mathbb{E}S_N) \\
=&\sum_i \left(\lambda_i\Big(((\sum_{j=1}^{i-1}(G_j+S_j))(1-\mathrm{e}^{-\nu_i G_i}) + G_i)(\sum_j m_{ij}) + \sum_{j=1}^{i-1}(G_j+S_j)\mathrm{e}^{-\nu_i G_i} + \sum_{j=i}^{N} \mathbb{E}S_j + \sum_{j=i+1}^{N} G_j\Big)\right).
\end{aligned} \tag{3.14}$$

Since the above equation is a finite sum/product of finite terms it is indeed finite.

Here, the term $\lambda_1(G_1(\sum_j m_{1j}))$ corresponds to the type 1 customers arriving during the glue period of Q_1 and their subsequent children of all types. The term $\lambda_1(\sum_{j=1}^{N} \mathbb{E}S_j + \sum_{j=2}^{N} G_j)$ corresponds to the type 1 customers arriving during the glue periods of Q_j, $j = 2, \cdots, N$, and switchover periods after Q_j, $j = 1, \cdots, N$. These customers arrive after the visit of Q_1 and hence do not get served or produce children. The term $\lambda_2(((G_1+S_1)(1-\mathrm{e}^{-\nu_2 G_2})+G_2)(\sum_j m_{2j})+(G_1+S_1)\mathrm{e}^{-\nu_2 G_2})$ corresponds to the type 2 customers arriving during the glue period of Q_1, Q_2, the switchover period after Q_1 and their subsequent children. The term $\lambda_2(\sum_{j=2}^{N} \mathbb{E}S_j + \sum_{j=3}^{N} G_j)$ corresponds to the type 2 customers arriving during the glue periods of Q_j, $j = 3, \cdots, N$, and switchover periods after Q_j, $j = 2, \cdots, N$. These customers do not produce any children. Similarly the term $\lambda_N(((\sum_{j=1}^{N-1}(G_j + S_j))(1 - \mathrm{e}^{-\nu_N G_N}) + G_N)(\sum_j m_{Nj}) + \sum_{j=1}^{N-1}(G_j + S_j)\mathrm{e}^{-\nu_N G_N})$ corresponds to the type N customers arriving during the glue period of $Q_1, \cdots, Q_N$, the switchover periods after $Q_1, \cdots, Q_{N-1}$ and their subsequent children. The term $\lambda_N \mathbb{E}S_N$ corresponds to the type N customers arriving during the switchover period after Q_N, which do not produce any children.

Ad (ii): Define the matrix

$$H = \begin{pmatrix}
\mathrm{e}^{-\nu_1 G_1} + \left(1-\mathrm{e}^{-\nu_1 G_1}\right)\rho_1 & \left(1-\mathrm{e}^{-\nu_1 G_1}\right)\lambda_2 \mathbb{E}B_1 & \cdots & \left(1-\mathrm{e}^{-\nu_1 G_1}\right)\lambda_N \mathbb{E}B_1 \\
\left(1-\mathrm{e}^{-\nu_2 G_2}\right)\lambda_1 \mathbb{E}B_2 & \mathrm{e}^{-\nu_2 G_2} + \left(1-\mathrm{e}^{-\nu_2 G_2}\right)\rho_2 & \cdots & \left(1-\mathrm{e}^{-\nu_2 G_2}\right)\lambda_N \mathbb{E}B_2 \\
\vdots & \vdots & \cdots & \vdots \\
\left(1-\mathrm{e}^{-\nu_N G_N}\right)\lambda_1 \mathbb{E}B_N & \left(1-\mathrm{e}^{-\nu_N G_N}\right)\lambda_2 \mathbb{E}B_N & \cdots & \mathrm{e}^{-\nu_N G_N} + \left(1-\mathrm{e}^{-\nu_N G_N}\right)\rho_N
\end{pmatrix}, \tag{3.15}$$

where the elements h_{ij} of the matrix H represent the mean number of type j customers that replace a type i customer during a visit period of Q_i (either new arrivals if the customer is served, or the customer itself if it is not served). We have that

$$H\begin{pmatrix}\mathbb{E}B_1\\ \mathbb{E}B_2\\ \vdots\\ \mathbb{E}B_n\end{pmatrix}=\begin{pmatrix}\left[\mathrm{e}^{-\nu_1G_1}+\left(1-\mathrm{e}^{-\nu_1G_1}\right)\left(\sum_j\rho_j\right)\right]\mathbb{E}B_1\\ \left[\mathrm{e}^{-\nu_2G_2}+\left(1-\mathrm{e}^{-\nu_2G_2}\right)\left(\sum_j\rho_j\right)\right]\mathbb{E}B_2\\ \vdots\\ \left[\mathrm{e}^{-\nu_NG_N}+\left(1-\mathrm{e}^{-\nu_NG_N}\right)\left(\sum_j\rho_j\right)\right]\mathbb{E}B_N\end{pmatrix}<\begin{pmatrix}\mathbb{E}B_1\\ \mathbb{E}B_2\\ \vdots\\ \mathbb{E}B_n\end{pmatrix}\qquad(3.16)$$

if and only if $\sum_j\rho_j<1$. Using this result and following the same line of proof as in Section 5 of Resing [20], we can show that the stability condition $\sum_j\rho_j<1$ implies that also the maximal eigenvalue λ_{max} of the mean matrix M satisfies $\lambda_{max}<1$. This concludes the proof. □

We can now obtain the moments, $\mathbb{E}X_j^{(i+1)}$, either from (3.11) or in a similar way as in Section 2.2, in terms of $\mathbb{E}X_j^{(i)}$ and $\mathbb{E}X_i^{(i)}$:

$$\mathbb{E}X_j^{(i+1)}=\lambda_j\mathbb{E}S_i+\mathbb{E}Z_j^{(i)}.$$

When $j\neq i$,

$$\mathbb{E}Z_j^{(i)}=\lambda_j\mathbb{E}B_i\mathbb{E}Y_i^{(iq)}+\mathbb{E}Y_j^{(i)},$$

else

$$\mathbb{E}Z_i^{(i)}=\lambda_i\mathbb{E}B_i\mathbb{E}Y_i^{(iq)}+\mathbb{E}Y_i^{(io)}.$$

Further

$$\begin{aligned}\mathbb{E}Y_j^{(i)}&=\lambda_jG_i+\mathbb{E}X_j^{(i)},\\ \mathbb{E}Y_i^{(iq)}&=\lambda_iG_i+(1-\mathrm{e}^{-\nu_iG_i})\mathbb{E}X_i^{(i)},\\ \mathbb{E}Y_i^{(io)}&=\mathrm{e}^{-\nu_iG_i}\mathbb{E}X_i^{(i)}.\end{aligned}$$

From the above equations we get, when $j\neq i$:

$$\mathbb{E}X_j^{(i+1)}=\lambda_j\mathbb{E}S_i+\lambda_j(1+\rho_i)G_i+\lambda_j\mathbb{E}B_i(1-\mathrm{e}^{-\nu_iG_i})\mathbb{E}X_i^{(i)}+\mathbb{E}X_j^{(i)},$$

and

$$\mathbb{E}X_i^{(i+1)}=\lambda_i\mathbb{E}S_i+\lambda_i\rho_iG_i+(\rho_i(1-\mathrm{e}^{-\nu_iG_i})+\mathrm{e}^{-\nu_iG_i})\mathbb{E}X_i^{(i)}.$$

Using flow balance arguments (mean number of customers of type i served per cycle equals mean number of type i customers arriving per cycle) and the obvious fact that the mean cycle time equals $\mathbb{E}C:=\sum_i(\mathbb{E}S_i+G_i)/(1-\rho)$, we obtain

$$\mathbb{E}Y_i^{(iq)}=\frac{\lambda_i}{1-\rho}\sum_j(\mathbb{E}S_j+G_j).\qquad(3.17)$$

We can also use a similar argument for mean number of type i customers leaving the orbit, $(1-\mathrm{e}^{-\nu_i G_i})\mathbb{E}X_i^{(i)}$, to equal the mean number of type i customers entering it, $\lambda_i(\mathbb{E}C - G_i)$, per cycle, yielding

$$\mathbb{E}X_i^{(i)} = \frac{\lambda_i}{1-\mathrm{e}^{-\nu_i G_i}}[\frac{\sum_j(\mathbb{E}S_j + G_j)}{1-\rho} - G_i]. \tag{3.18}$$

We can observe that (3.17) and (3.18) satisfy the above relation of $\mathbb{E}Y_i^{(iq)}$ and $\mathbb{E}X_i^{(i)}$. Further for each i, cyclically substituting we get all $\mathbb{E}X_j^{(i)}$ and therefore $\mathbb{E}Y_j^{(i)}$ and $\mathbb{E}Z_j^{(i)}$.

The second moments of $X_j^{(1)}$ and the various terms $\mathbb{E}[X_j^{(i)}X_k^{(i)}]$ can be obtained by solving a set of equations which is derived by twice differentiating (3.10) w.r.t. z_j and z_k, $j,k = 1,\cdots,N$, and calculating the value at $z = (1,1,\cdots,1)$. Since the system is cyclic, once we obtain $\mathbb{E}{X_j^{(1)}}^2$, $j = 1,\cdots,N$, we can similarly obtain $\mathbb{E}{X_j^{(i)}}^2$, $i = 1,\cdots,N$, by changing indices. It is not difficult to develop an efficient procedure for determining higher moments in polling systems with a branching discipline, cf. [20].

3.3 Queue length analysis at arbitrary time points

In the previous section we have given the procedure for finding the distribution of the number of customers at the beginning of glue periods ($\mathbb{E}[\prod_j z_j^{X_j^{(i)}}]$), (ii) visit periods ($\mathbb{E}[\Big(\prod_{j\neq i} z_j^{Y_j^{(i)}}\Big) z_{iq}^{Y^{(iq)}} z_{io}^{Y^{(io)}}]$), and (iii) switchover periods ($\mathbb{E}[\prod_j z_j^{Z_j^{(i)}}]$), for $i = 1,\cdots,N$. Similar to the single queue case, we now obtain the generating function of the number of customers at arbitrary time points.

Theorem 4. *If $\sum_i \rho_i < 1$ and $(z_q, z_o) := (z_{1q}, z_{1o}, \cdots, z_{Nq}, z_{No})$, we have the following results:*

a) The joint generating function, $R_{sw}^{(i)}(z_q, z_o)$, of the number of customers in the queue and in the orbit at an arbitrary time point in a switchover period after Q_i is given by

$$R_{sw}^{(i)}(z_q, z_o) = \mathbb{E}[\prod_j z_{jo}^{Z_j^{(i)}}]\frac{1-\tilde{S}_i(\sum_j \lambda_j(1-z_{jo}))}{\Big(\sum_j \lambda_j(1-z_{jo})\Big)\mathbb{E}S_i}. \tag{3.19}$$

b) The joint generating function, $R_{gl}^{(i)}(z_q, z_o)$, of the number of customers in the queue and in the orbit at an arbitrary time point in a glue period of Q_i is given by

$$R_{gl}^{(i)}(z_q, z_o) = \Big(\prod_{j\neq i}\mathbb{E}[z_{jo}^{X_j^{(i)}}]\Big)\int_{t=0}^{G_i}\Big(\prod_{j\neq i}\mathrm{e}^{-\lambda_j(1-z_{jo})t}\Big)\mathrm{e}^{-\lambda_i(1-z_{iq})t}\mathbb{E}[\{(1-\mathrm{e}^{-\nu_i t})z_{iq} + \mathrm{e}^{-\nu_i t}z_{io}\}^{X_i^{(i)}}]\,\frac{dt}{G_i}. \tag{3.20}$$

c) The joint generating function, $R^{(i)}_{vi}(z_q, z_o)$, of the number of customers in the queue and in the orbit at an arbitrary time point in a visit period of Q_i is given by

$$R^{(i)}_{vi}(z_q, z_o) = \frac{z_{iq}\left(\mathbb{E}[z_{iq}^{Y_i^{(iq)}}\left(\prod_{j\neq i} z_{jo}^{Y_j^{(i)}}\right)z_{io}^{Y_i^{(io)}}] - \mathbb{E}[\tilde{B}_i(\sum_j \lambda_j(1-z_{jo}))^{Y_i^{(iq)}}\left(\prod_{j\neq i} z_{jo}^{Y_j^{(i)}}\right)z_{io}^{Y_i^{(io)}}]\right)}{\mathbb{E}[Y_i^{(iq)}]\left(z_{iq} - \tilde{B}_i(\sum_j \lambda_j(1-z_{jo}))\right)} \times \frac{1-\tilde{B}_i(\sum_j \lambda_j(1-z_{jo}))}{\left(\sum_j \lambda_j(1-z_{jo})\right)\mathbb{E}B_i}. \tag{3.21}$$

d) The joint generating function, $R(z_q, z_o)$, of the number of customers in the queue and in the orbit at an arbitrary time point is given by

$$R(z_q, z_o) = \sum_i \left(\rho_i R^{(i)}_{vi}(z_q, z_o) + (1-\rho_i)\frac{G_i}{\sum_j (G_j + \mathbb{E}S_j)} R^{(i)}_{gl}(z_q, z_o) + (1-\rho_i)\frac{\mathbb{E}S_i}{\sum_j (G_j + \mathbb{E}S_j)} R^{(i)}_{sw}(z_q, z_o)\right). \tag{3.22}$$

Proof. The proof follows the same lines as the proof of Theorem 2, in particular for parts a and d. We restrict ourselves here to an outline of the proof of parts b and c.

b) Follows from the fact that if the past part of the glue period is equal to t, the generating function of the number of new arrivals of type i in the queue during this period is equal to $\mathrm{e}^{-\lambda_i(1-z_{iq})t}$ and each type i customer present in the orbit at the beginning of the glue period is, independent of the others, still in orbit with probability $\mathrm{e}^{-\nu_i t}$ and has moved to the queue with probability $1-\mathrm{e}^{-\nu_i t}$. Further the generating function of the number of new arrivals of any type $j \neq i$ in the queue during this period is equal to $\mathrm{e}^{-\lambda_j(1-z_{jo})t}$.

c) During an arbitrary point in time in a visit period the number of customers in the system consists of two parts:
   - the number of customers in the system at the beginning of the service time of the customer currently in service, leading to the term

$$\frac{z_{iq}\left(\mathbb{E}[z_{iq}^{Y_i^{(iq)}}\left(\prod_{j\neq i} z_{jo}^{Y_j^{(i)}}\right)z_{io}^{Y_i^{(io)}}] - \mathbb{E}[\tilde{B}_i(\sum_j \lambda_j(1-z_{jo}))^{Y_i^{(iq)}}\left(\prod_{j\neq i} z_{jo}^{Y_j^{(i)}}\right)z_{io}^{Y_i^{(io)}}]\right)}{\mathbb{E}[Y_i^{(iq)}]\left(z_{iq} - \tilde{B}_i(\sum_j \lambda_j(1-z_{jo}))\right)};$$

     (see Remark 3).
   - the number of customers that arrived during the past part of the service of the customer currently in service, leading to the term

$$\frac{1-\tilde{B}_i(\sum_j \lambda_j(1-z_{jo}))}{\left(\sum_j \lambda_j(1-z_{jo})\right)\mathbb{E}B_i}.$$

$\square$

From Theorem 4, we now can obtain the steady-state mean number of customers in the system at arbitrary time points in switchover periods ($\mathbb{E}[R_{sw}^{(i)}]$) after Q_i, in glue periods ($\mathbb{E}[R_{gl}^{(i)}]$) and in visit periods ($\mathbb{E}[R_{vi}^{(i)}]$) of Q_i , for $i = 1, \cdots, N$, and at any arbitrary time point ($\mathbb{E}[R]$). These are given by

$$\mathbb{E}[R_{sw}^{(i)}] = \sum_j (\mathbb{E}[Z_j^{(i)}] + \lambda_j \tfrac{\mathbb{E}[S_i^2]}{2\mathbb{E}[S_i]}),$$

$$\mathbb{E}[R_{gl}^{(i)}] = \sum_j (\mathbb{E}[X_j^{(i)}] + \lambda_j \tfrac{G_i}{2}),$$

$$\mathbb{E}[R_{vi}^{(i)}] = 1 + (\sum_j \lambda_j) \tfrac{\mathbb{E}[B_i^2]}{2\mathbb{E}[B_i]} + \tfrac{\mathbb{E}[Y_i^{(iq)} Y_i^{(io)}]}{\mathbb{E}[Y_i^{(iq)}]} + (\sum_{j \neq i} \tfrac{\mathbb{E}[Y_i^{(iq)} Y_j^{(i)}]}{\mathbb{E}[Y_i^{(iq)}]}) + \tfrac{(1+\mathbb{E}B_i \sum_j \lambda_j)\mathbb{E}[Y_i^{(iq)}(Y_i^{(iq)}-1)]}{2\mathbb{E}[Y_i^{(iq)}]},$$

$$\mathbb{E}[R] = \sum_i \Big(\rho_i \mathbb{E}[R_{vi}^{(i)}] + (1-\rho_i) \tfrac{G_i}{\sum_j (G_j + \mathbb{E}S_j)} \mathbb{E}[R_{gl}^{(i)}] + (1-\rho_i) \tfrac{\mathbb{E}S_i}{\sum_j (G_j + \mathbb{E}S_j)} \mathbb{E}[R_{sw}^{(i)}] \Big).$$

The mean number of type k customers in the system at arbitrary time points in a switchover period after Q_i and a glue period before Q_i are given by the values of $\mathbb{E}[R_{sw}^{(i)}]$ and $\mathbb{E}[R_{gl}^{(i)}]$ at $j = k$. The mean number of type i customers in the system at arbitrary time points in a visit period of Q_i is given by

$$1 + \lambda_i \tfrac{\mathbb{E}[B_i^2]}{2\mathbb{E}[B_i]} + \tfrac{\mathbb{E}[Y_i^{(iq)} Y_i^{(io)}]}{\mathbb{E}[Y_i^{(iq)}]} + \tfrac{(1+\rho_i)\mathbb{E}[Y_i^{(iq)}(Y_i^{(iq)}-1)]}{2\mathbb{E}[Y_i^{(iq)}]}.$$

The quantities $\mathbb{E}[Y_i^{(iq)} Y_i^{(io)}]$, $\mathbb{E}[Y_i^{(iq)} Y_j^{(i)}]$ and $\mathbb{E}[Y_i^{(iq)}(Y_i^{(iq)} - 1)]$ can be obtained using (3.3).

4 Conclusions and suggestions for future research

In this paper we have studied vacation queues and N-queue polling models with the gated service discipline and with retrials. Motivated by optical communications, we have introduced a glue period just before a server visit; during such a glue period, new customers and retrials "stick" instead of immediately going into orbit. For both the vacation queue and the N-queue polling model, we have derived steady-state queue length distributions at an arbitrary epoch and at various specific epochs. This was accomplished by establishing a relation to branching processes.

In future studies, we would like to consider other service disciplines. Furthermore, the following model variants seem to fall within our framework: (i) customers may *not* retry with a certain probability; (ii) the arrival rates may be different for visit, switchover and glue periods; (iii) one might allow that new arrivals during a glue period are already served during that glue period. We also wish to take non-constant glue periods into account. We believe that a workload decomposition and pseudoconservation law, as discussed in [6], can be derived for these variants and generalizations, and they may be exploited for analysis and optimization purposes. We shall then also try to explore the following observation: One may view our N-queue model as a polling model with a new variant of binomial gated, with adaptive probability p_i of serving a customer at a visit of Q_i; $p_i = 1$ when the customer arrived in the preceding glue period, and $p_i = 1 - \mathrm{e}^{-\nu_i G_i}$ otherwise.

We would also like to explore the possibility to study the heavy traffic behavior of these models via the relation to branching processes, cf. [18].

Finally, we would like to point out an important advantage of optical fibre: the wavelength of light. A fibre-based network node may thus route incoming packets not only by switching in the time-domain, but also by wavelength division multiplexing. In queueing terms, this gives rise to *multiserver* polling models, each server representing a wavelength. We refer to [1] for the stability analysis of multiserver polling models, and to [2] for a mean field approximation of large passive optical networks. Therefore we would like to study multiserver polling models with the additional features of retrials and glue periods.

Acknowledgment

The authors gratefully acknowledge fruitful discussions with Kevin ten Braak and Tuan Phung-Duc about retrial queues and with Ton Koonen about optical networks. The research is supported by the IAP program BESTCOM, funded by the Belgian government, and by the Gravity program NETWORKS, funded by the Dutch government.

References

1. N. Antunes, Chr. Fricker and J. Roberts (2011). Stability of multi-server polling system with server limits. *Queueing Systems*, **68**, 229-235.
2. N. Antunes, Chr. Fricker, Ph. Robert and J. Roberts (2010). Traffic capacity of large WDM Passive Optical Networks. Proceedings 22nd International Teletraffic Congress (ITC 22), Amsterdam, September 2010.
3. J.R. Artalejo and A. Gomez-Corral (2008). Retrial Queueing Systems: A Computational Approach (Springer-Verlag, Berlin).
4. K.B. Athreya and P.E. Ney (1972). Branching Processes (Springer-Verlag, Berlin).
5. M.A.A. Boon, R.D. van der Mei, and E.M.M. Winands (2011). Applications of polling systems. *SORMS*, **16**, 67-82.
6. O.J. Boxma (1989). Workloads and waiting times in single-server systems with multiple customer classes. *Queueing Systems*, **5**, 185-214.
7. O.J. Boxma and J.W. Cohen (1991). The $M/G/1$ queue with permanent customers. *IEEE J. Sel. Areas in Commun.*, **9**, 179-184.
8. O.J. Boxma, O. Kella and K.M. Kosinski (2011). Queue lengths and workloads in polling systems. *Operations Research Letters*, **39**, 401-405.
9. O.J. Boxma and J.A.C. Resing (2014). Vacation and polling models with retrials. 11th European Workshop on Performance Engineering (EPEW 11), Florence, September 2014.

10. G.I. Falin and J.G.C. Templeton (1997). Retrial Queues (Chapman and Hall, London).
11. A.M.J. Koonen (2014). Personal communication.
12. C. Langaris (1997). A polling model with retrial of customers. *Journal of the Operations Research Society of Japan*, **40**, 489-507.
13. C. Langaris (1999). Gated polling models with customers in orbit. *Mathematical and Computer Modelling*, **30**, 171-187.
14. C. Langaris (1999). Markovian polling system with mixed service disciplines and retrial customers. *Top*, **7**, 305-322.
15. H. Levy (1991). Binomial-gated service: A method for effective operation and optimization of polling systems. *IEEE Trans. Commun.*, **39**, 1341-1350.
16. H. Levy and M. Sidi (1990). Polling models: applications, modeling and optimization. *IEEE Trans. Commun.*, **38**, 1750-1760.
17. M. Maier (2008). Optical Switching Networks (Cambridge University Press, Cambridge).
18. T.L. Olsen and R.D. van der Mei (2003). Periodic polling systems in heavy traffic: distribution of the delay. *Journal of Applied Probability* **40**, 305-326.
19. M.P. Quine (1970). The multitype Galton-Watson process with immigration. *Journal of Applied Probability* **7**, 411-422.
20. J.A.C. Resing (1993). Polling systems and multitype branching processes. *Queueing Systems* **13**, 409-426.
21. W. Rogiest (2008). Stochastic Modeling of Optical Buffers. Ph.D. Thesis, Ghent University, Ghent, Belgium.
22. H. Takagi (1991). Application of polling models to computer networks. *Comput. Netw. ISDN Syst.*, **22**, 193-211.
23. H. Takagi (1991). Queueing Analysis: A Foundation of Performance Evaluation. Volume 1: Vacation and Priority Systems (Elsevier Science Publishers, Amsterdam).
24. H. Takagi (1997). Queueing analysis of polling models: progress in 1990-1994. In J.H. Dshalalow, editor, Frontiers in Queueing: Models, Methods and Problems, pages 119-146. CRC Press, Boca Raton, 1997.
25. H. Takagi (2000). Analysis and application of polling models. In G. Haring, C. Lindemann, and M. Reiser, editors, Performance Evaluation: Origins and Directions, volume 1769 of Lecture Notes in Computer Science, pages 424-442. Springer, Berlin, 2000.
26. N. Tian and Z.G. Zhang (2006). Vacation Queueing Models: Theory and Applications (Springer, New York).
27. V.M. Vishnevskii and O.V. Semenova (2006). Mathematical methods to study the polling systems. *Autom. Remote Control*, **67**, 173-220.