The Priority of Inbound Calls over Outbound Calls Modeled as a Discrete-Time Retrial/Delay System

Rein Nobel(✉) and Maik Dekker

Department of Econometrics and Operations Research, Vrije Universiteit Amsterdam, Amsterdam, Netherlands
r.d.nobel@vu.nl, m.h.r.dekker@student.vu.nl

Abstract. A one-server discrete-time queueing model is studied with two arrival streams. Both arrival streams are in batches and we distinguish between a stream of low-priority customers, who are put in a queue which is served on a first-come-first-served basis, and a stream of (primary) high-priority customers, who are served uninterruptedly when the batch of high-priority customers finds the server idle upon arrival. High-priority customers are treated as retrial customers, but once in the orbit they lose their high-priority status. The Late Arrival Setup is chosen with Delayed Access. The high-priority retrial customers can be interpreted as inbound calls, and the low-priority customers as outbound calls in a call-center. The joint steady-state distribution of the queue length of the low-priority customers and the orbit size of secondary retrial customers is studied using probability generating functions. Several performance measures will be calculated, such as the mean queue length of the low-priority customers and the orbit size of the secondary retrial customers.

Keywords: Inbound and outbound calls · Discrete-time retrial queue
Priority customers · Generating functions

1 Introduction

In call-centers inbound calls have priority over outbound calls. Outbound calls will be handled only when after the end of a call no inbound calls are coming in, i.e. when a server would stay idle if he would not start answering outbound calls. Inbound calls do not wait in a queue and when upon arrival they find a busy tone they will try to call again some random time later. Outbound calls, for instance in the form of e-mails sent to the call-center with a request to be called back, will be handled by the center in the order of their arrival, when time is available due to the absence of incoming calls.

To model this priority-scheme for inbound calls over requests for being called back by the center we study a mixed retrial/delay model in discrete time with

This paper is based on the second author's Bachelor thesis [4].

W. Yue et al. (Eds.): QTNA 2017, LNCS 10591, pp. 199–216, 2017.
https://doi.org/10.1007/978-3-319-68520-5_13

one server. More specifically, we consider a one-server queueing model in discrete time with two types of customers. Time is divided in slots, and all events [arrivals, start of a service and departures] are considered to occur at the slot boundaries only. The high-priority customers [primary inbound calls] arrive in batches following a general probability distribution. When upon arrival of a batch of high-priority customers the server is idle, the complete batch is accepted for an uninterrupted (batch-)service. When upon arrival of a batch of high-priority customers the server is busy, the complete batch will be sent into orbit, and the individual customers lose their high-priority-status. They will approach the server individually [so-called secondary arrivals] some random time later, independently from the other customers in the orbit.

The low-priority customers [outbound calls] also arrive in batches, possibly following a different probability distribution, and they are put in a queue which is served in the order of arrival [within a batch in random order]. The low-priority customers are served individually and a low-priority customer is selected for service only when the server is idle and no batch of primary high-priority customers arrives in the idle slot. In case neither primary high-priority customers arrive nor low-priority customers are present in the queue, then a possible secondary arrival is selected for [an individual] service. The non-selected secondary customers are resent into the orbit. When neither low-priority customers are present in the queue at the end of the idle slot, nor any primary or secondary retrial customers will have arrived in the idle slot, the server stays idle also the following slot.

Notice that the modeling assumption is made that in the time slot following a (batch-)service completion the server always stays idle, even when the queue of low-priority customers is not empty, to enable the start of the service of an incoming batch of high-priority primary customers.

The service times of the high-priority [inbound calls] and the low-priority [outbound calls] customers are all independent and follow [possibly] a different general distribution. To resolve the conflict of simultaneous arrivals and departures we have chosen for the *late arrival setup with delayed access*, i.e. arrivals have precedence over departures and a service of newly arrived customers can only start at the time slot following the slot of the arrival at the earliest. For an overview of discrete-time retrial queues with the late arrival setup we refer to Nobel [7] and for the most complete monograph on retrial queues we refer to Artalejo and Gómez-Corral [1].

So, in this paper we will extend the classical discrete-time one-server retrial model of Nobel and Moreno [9] by adding a second type of customers [the outbound calls] who upon arrival are put in a queue. These low-priority customers will be served one by one on a first-come-first-served basis. The retrial primary customers [inbound calls] are given *non-preemptive* priority over the queued customers [the outbound calls]. Rejected inbound calls lose their priority, but they continue to act as retrial customers, and their service time remains unaltered. In Sharkawy [10] the high-priority retrial customers maintained their high-priority status in the orbit, but it turned out to be impossible to derive a closed-form expression for the probability generating function of the joint steady-state distribution of the queue size and the orbit size. For this technical reason in this paper we made the modeling assumption that high-priority customers lose their priority status once sent into the orbit.

In previous papers (Nobel and Moreno [8] and Nobel [6]) the priority has been mainly modeled the other way around: non-preemptive priority of the queued customers over the retrial customers. This is a natural hierarchy in mobile telephony for modeling handover calls [high priority] versus new calls [low priority] competing for the same target channel, see Nobel [6]. As pointed out before, giving priority to the retrial customers over the queued customers leads to an intractable model (see again Sharkawy [10]), and only to guarantee tractability we made the admittedly somewhat awkward assumption that high-priority retrial customers lose their priority once they have been sent into the orbit. In Artalejo et al. [2] a [somewhat simplified] continuous-time counterpart of our model with single arrivals is discussed, in which the retrial customers do not lose their high-priority status once they are sent into the orbit, but the authors only consider exponential service times, introduce a finite buffer size for the low-priority customers and, most importantly, they give *preemptive* priority to the retrial customers. These three characteristics of their model enable an algorithmic analysis. We think that in a call center outbound calls should not be interrupted by incoming inbound calls, and for that reason we have chosen for *non-preemptive* priority for the inbound calls, but to get an analytic solution we have to pay a price! Of course, it is also possible to give a practical application in which our modeling assumption that the high-priority retrial customers lose their priority status is more natural than in the call-center environment. Take for instance a small military field hospital with one operation unit where regularly scheduled patients [outbound calls!] and incoming emergency patients [inbound calls!] have to be operated. When an ambulance with a group of emergency patients arriving from the battlefield finds the operation unit busy they will be sent away (maybe after some necessary minimal treatment), and subsequently they will compete individually with the regular patients, i.e. they lose their high-priority [emergency] status. Although we had in mind a call-center application when we started this paper, the above hospital example illustrates that our technical assumption is quite realistic in another environment!

A discrete-time model with the easier priority setup, i.e. the queued customers have priority over the retrial customers, has been studied in Choi and Kim [3], but also they discuss only single arrivals and all customers follow the same service-time distribution. Further, they have chosen the early arrival setup. A continuous-time retrial model with priority for the queued customers has been studied by Falin et al. [5], but also in that paper only single arrivals have been considered. The model discussed in this paper can be seen both as an extension and as the discrete-time counterpart of that model, but above all as a first attempt to reverse the priority of retrial customers versus queued customers. As already indicated above, this reversed priority-scheme is mainly motivated by the priority of inbound calls over outbound calls in a call-center.

In the sections below we will study the joint steady-state distribution of the length of the queue of low-priority customers [outbound calls] and the size of the orbit with high-priority customers [inbound calls who lost their priority]. Not surprisingly, the mathematical analysis of our mixed retrial/delay model differs greatly from the analysis of the models discussed in the papers Choi and Kim [3] and Nobel and Moreno [8]. The analysis is also more involved than the analysis presented in Nobel [6].

As usual, we will derive the generating function of the joint steady-state distribution of the number of low-priority customers in the queue, the number of high-priority customers in the orbit and the residual service time of the (batch of) customer(s) in service. Notice that we do not keep track of the type of the ongoing service in the analysis. This generating function will be used to calculate several performance measures, e.g. the mean queue length and the mean orbit size. In Sect. 2 we describe the model in detail. In Sects. 3 and 4 we discuss the steady-state distributions and the first moment of the orbit size and the queue length. In Sect. 5 we will present some numerical results.

2 Description of the Model

For a detailed description of the discrete-time setup with late arrivals and delayed access [LAS/DA] we refer to Nobel and Moreno [9]. Recall that due to this LAS/DA setup in this classical retrial model the time slot after a departure the server always stays idle for at least one slot, because arrivals have precedence over departures. For the mixed retrial/delay model to be discussed in this paper we make the technical assumption that the slot following the completion of a (batch-)service the server always stays idle, *also in case low-priority customers are waiting in the queue*. Imposing this idle slot guarantees the priority of the (primary) retrial customers over the queued customers, by triggering the start of the batch-service of any incoming batch of high-priority customers in this idle slot at the start of the next slot, and so automatically blocking the possible start of the service of a (queued) low-priority customer, or a secondary arrival from the orbit. Only in case no primary batch of high-priority customers arrives during the idle slot, the service of the longest waiting low-priority customer will start his individual service the next slot. If no low-priority customers are present in the queue or no batch of low-priority customers will have arrived during the idle slot, then possibly a secondary arrival will start his individual service, and in case there are no secondary arrivals, the server stays idle also the next slot.

We will now give the precise description of the discrete-time mixed retrial/delay queueing model with one server and priorities for the primary retrial customers. In each time slot primary high-priority customers [inbound calls] arrive in batches. The batch sizes are mutually independent and follow a general probability distribution $\left\{a_k^{(H)}\right\}_{k=0}^{\infty}$ with probability generating function (p.g.f.)

$$\mathcal{A}_H(z) = \sum_{k=0}^{\infty} a_k^{(H)} z^k.$$

In every time slot also low-priority customers [outbound calls] arrive in batches. These batch sizes follow a general probability distribution $\left\{a_i^{(L)}\right\}_{i=0}^{\infty}$ with p.g.f.

$$\mathcal{A}_L(y) = \sum_{i=0}^{\infty} a_i^{(L)} y^i.$$

These batch sizes are again mutually independent and they are also independent of the batch sizes of the high-priority customers. Each individual high-priority customer requires a service time, measured as a number of time slots, which follows the discrete probability distribution $\left\{b_j^{(H)}\right\}_{j=1}^{\infty}$ with p.g.f.

$$\mathcal{B}_H(w) = \sum_{j=1}^{\infty} b_j^{(H)} w^j.$$

Similarly, every low-priority customer requires a generally distributed service time with distribution $\left\{b_j^{(L)}\right\}_{j=1}^{\infty}$ and p.g.f.

$$\mathcal{B}_L(w) = \sum_{j=1}^{\infty} b_j^{(L)} w^j.$$

All service times are mutually independent and they are also independent of the batch sizes of the arriving customers. A service time requires at least one time slot, so $b_0^{(H)} = b_0^{(L)} = 0$. As said before, the low-priority customers are placed in a queue, and are served individually on a first-come-first-served basis. Also primary high-priority customers are served individually, but *uninterruptedly as a batch-service*, i.e. after every individual service completion, the next customer of the batch starts his service *immediately* in the next slot. Only at the service completion of the last customer of the batch the server stays idle the next slot, even if low-priority customers are present in the queue, to enable the start of a batch-service in case a new batch of high-priority customers arrives in this idle slot. Rejected high-priority customers behave as the customers in the classical retrial queue, with the only difference that all incoming customers [inbound calls] from the orbit have lost their high-priority status. They even have lower-priority than the queued customers [outbound calls]. In each time slot retrial customers in the orbit [inbound calls who have lost their high-priority] try to reenter the system individually and independently with the so-called retrial probability r $[0 < r \leq 1]$.

We are interested in the joint steady-state distribution of the number of low-priority customers in the queue, the number of high-priority customers [strictly speaking a misnomer, because customers once in the orbit have lost their high-priority status] in orbit, and the residual service time of the (batch of) customer(s) currently in service. To analyze the mixed retrial/delay queueing model, we define a discrete-time Markov chain (DTMC) by observing the system at the epochs $k-$, that is at the start of the time slots k just after, possibly, a service of a (low- or high-priority) customer has started, but before the arrivals during time slot k have occurred. We define the following random variables,

$$R_k = \text{the residual service time of the ongoing (batch-)service at time } k-,$$
$$L_k = \text{the number of low-priority customers present in the queue at time } k-,$$
$$Q_k = \text{the number of high-priority customers in orbit at time } k- .$$

We define $R_k = 0$ when at epoch $k-$ the server is idle. Notice that the type of the residual service time is not part of the state description. Introduce the offered load

$$\varrho := \mathcal{A}'_L(1)\mathcal{B}'_L(1) + \mathcal{A}'_H(1)\mathcal{B}'_H(1).$$

Then, the stochastic process $\{(R_k, L_k, Q_k) : k = 0, 1, 2, \ldots\}$ is an irreducible aperiodic DTMC which is positive recurrent under the stability condition

$$a_0^{(H)}[1-\varrho] - \mathcal{A}'_L(1) - \varrho\mathcal{A}'_H(1) > 0.$$

This complicated stability condition is due to the modeling assumption that a batch of primary high-priority customers is served uninterruptedly, imposing only one forced idle slot after the completion of the last customer of the batch, whereas all the other customers [low-priority and secondary customers arriving from the orbit] force the server to stay idle after each [individual] service. So the *total used capacity*, i.e. the fraction of time that the server is busy or waiting for a possible arrival of a batch of high-priority customers, say σ, should be smaller than 1, i.e.

$$\sigma := \varrho + \mathcal{A}'_L(1) + \varrho\mathcal{A}'_H(1) + \left(1 - a_0^{(H)}\right)(1-\varrho) < 1.$$

A formal proof of this stability condition can be given using Foster's criterion [see Nobel and Moreno [9] for the details].

3 The Joint Distribution of Queue Length and Orbit Size

In this section we will derive the joint probability generating function [p.g.f.] of the steady-state distribution of the DTMC $\{(R_k, L_k, Q_k) : k = 0, 1, 2, \ldots\}$. Under the stability condition we can define the following limiting joint distribution of this DTMC

$$\pi(j, m, n) = \lim_{k\to\infty} \mathbf{P}(R_k = j; L_k = m; Q_k = n), \quad j, m, n = 0, 1, 2, \ldots,$$

with its associated three-dimensional generating function

$$\Pi(w, y, z) = \sum_{j=0}^{\infty}\sum_{m=0}^{\infty}\sum_{n=0}^{\infty} \pi(j, m, n) w^j y^m z^n.$$

In the following it is convenient to introduce also the partial generating functions,

$$\Pi_{jm}(z) = \sum_{n=0}^{\infty} \pi(j,m,n)z^n \quad \text{and}$$
$$\Pi_j(y,z) = \sum_{m=0}^{\infty}\sum_{n=0}^{\infty} \pi(j,m,n)y^m z^n = \sum_{m=0}^{\infty} \Pi_{jm}(z)y^m.$$

To find the p.g.f. $\Pi(w,y,z)$ we write down the system of balance equations,

$$\begin{aligned}\pi(0,m,n) = {} & \mathbf{I}_{\{m=0\}} a_0^{(L)} a_0^{(H)} (1-r)^n \pi(0,0,n) \\ & + \sum_{i=0}^{m} a_i^{(L)} \sum_{k=0}^{n} a_k^{(H)} \pi(1, m-i, n-k), \\ & m,n = 0,1,\ldots,\end{aligned} \tag{1}$$

$$\begin{aligned}\pi(j,m,n) = {} & \sum_{i=0}^{m} a_i^{(L)} \sum_{k=0}^{n} a_k^{(H)} \pi(j+1, m-i, n-k) \\ & + a_0^{(H)} \sum_{i=0}^{m+1} a_i^{(L)} \pi(0, m+1-i, n) b_j^{(L)} \\ & + \sum_{i=0}^{m} a_i^{(L)} \sum_{k=1}^{j} a_k^{(H)} \pi(0, m-i, n) b_j^{(H)(*k)} \\ & + \mathbf{I}_{\{m=0\}} a_0^{(L)} a_0^{(H)} \left(1-(1-r)^{n+1}\right) \pi(0,0,n+1) b_j^{(H)}, \\ & j = 1,2,\ldots; \quad m,n = 0,1,2,\ldots.\end{aligned} \tag{2}$$

Notice how our technical assumption that after the completion of a [batch-] service the server stays idle for at least one time slot plays its role in these balance equations.

From Eq. (1) we get by multiplying both sides with z^n and summing over $n = 0,1,\ldots$, and subsequently multiplying both sides of the result by y^m and summing over $m = 0,1,\ldots$,

$$\Pi_0(y,z) = a_0^{(L)} a_0^{(H)} \Pi_{00}((1-r)z) + \mathcal{A}_L(y)\mathcal{A}_H(z)\Pi_1(y,z). \tag{3}$$

From Eq. (2) we get, acting similarly,

$$\begin{aligned}\Pi_j(y,z) = {} & \mathcal{A}_L(y)\mathcal{A}_H(z)\Pi_{j+1}(y,z) \\ & + \frac{a_0^{(H)} b_j^{(L)}}{y} \left[\mathcal{A}_L(y)\Pi_0(y,z) - a_0^{(L)} \Pi_{00}(z)\right] \\ & + \mathcal{A}_L(y) \sum_{k=1}^{j} a_k^{(H)} b_j^{(H)(*k)} \Pi_0(y,z) \\ & + \frac{a_0^{(L)} a_0^{(H)} b_j^{(H)}}{z} \left[\Pi_{00}(z) - \Pi_{00}((1-r)z)\right].\end{aligned} \tag{4}$$

Next, multiplying Eq. (4) by w^j and summing over $j = 1, 2, \ldots$ gives after some simple algebra, using Eq. (3) to get rid of $\Pi_1(y, z)$,

$$\begin{aligned} & yz(w - \mathcal{A}_L(y)\mathcal{A}_H(z))\Pi(w, y, z) \\ & \quad = \mathcal{A}_L(y)z\left[a_0^{(H)}w(\mathcal{B}_L(w) - y) + y(w\mathcal{A}_H(\mathcal{B}_H(w)) - \mathcal{A}_H(z))\right]\Pi_0(y, z) \\ & \qquad + a_0^{(L)}a_0^{(H)}w\left[y\mathcal{B}_H(w) - z\mathcal{B}_L(w)\right]\Pi_{00}(z) \\ & \qquad + a_0^{(L)}a_0^{(H)}wy\left[z - \mathcal{B}_H(w)\right]\Pi_{00}((1 - r)z). \end{aligned} \tag{5}$$

So, the problem is to find the unknown partial generating functions $\Pi_0(y, z)$ and $\Pi_{00}(z)$. Firstly, take $w = \mathcal{A}_L(y)\mathcal{A}_H(z)$ in (5) to make the left-hand side zero. This gives

$$\begin{aligned} & \mathcal{A}_L(y)z\begin{bmatrix} a_0^{(H)}\omega(y, z)[\mathcal{B}_L(\omega(y, z)) - y] \\ +y[\omega(y, z)\mathcal{A}_H(\mathcal{B}_H(\omega(y, z))) - \mathcal{A}_H(z)] \end{bmatrix}\Pi_0(y, z) \\ & \quad = a_0^{(L)}a_0^{(H)}\omega(y, z)\left[z\mathcal{B}_L(\omega(y, z)) - y\mathcal{B}_H(\omega(y, z))\right]\Pi_{00}(z) + \\ & \qquad -a_0^{(L)}a_0^{(H)}\omega(y, z)y\left[z - \mathcal{B}_H(\omega(y, z))\right]\Pi_{00}((1 - r)z). \end{aligned} \tag{6}$$

where $\omega(y, z) := \mathcal{A}_L(y)\mathcal{A}_H(z)$. Now consider the coefficient of $\Pi_0(y, z)$. Let

$$\psi(y, z) := \begin{matrix} a_0^{(H)}\omega(y, z)[\mathcal{B}_L(\omega(y, z)) - y] \\ +y[\omega(y, z)\mathcal{A}_H(\mathcal{B}_H(\omega(y, z))) - \mathcal{A}_H(z)] \end{matrix}$$

Then we have

$$\forall z \exists! y : \ \psi(y, z) = 0.$$

For real $z \in (0, 1)$ this follows immediately

$$\psi(0, z) = a_0^{(H)}a_0^{(L)}\mathcal{A}_H(z)\mathcal{B}_L\left(a_0^{(L)}\mathcal{A}_H(z)\right) > 0.$$

$$\psi(1, z) = \mathcal{A}_H(z)\left[a_0^{(H)}\mathcal{B}_L(\mathcal{A}_H(z)) + \mathcal{A}_H(\mathcal{B}_H(\mathcal{A}_H(z))) - \left(1 + a_0^{(H)}\right)\right] \leq 0$$

with equality only for $z = 1$. Notice that $\psi(1, 1) = 0$.

Let $y^*(z)$ be the unique solution, i.e. $\psi(y^*(z), z) = 0$ and introduce

$$\phi(z) := \omega(y^*(z), z) = \mathcal{A}_L(y^*(z))\mathcal{A}_H(z).$$

Notice that from $\psi(y^*(z), z) = 0$ we get

$$y^*(z) = \frac{a_0^{(H)}\phi(z)\mathcal{B}_L(\phi(z))}{a_0^{(H)}\phi(z) + \mathcal{A}_H(z) - \phi(z)\mathcal{A}_H(\mathcal{B}_H(\phi(z)))}. \tag{7}$$

It is easy to see that $y^*(1) = 1$ and so also $\phi(1) = 1$. Now from (6) and using (7) we find the recursion

$$\Pi_{00}(z) = \frac{y^*(z)[\mathcal{B}_H(\phi(z)) - z]}{y^*(z)\mathcal{B}_H(\phi(z)) - z\mathcal{B}_L(\phi(z))}\Pi_{00}((1-r)z)$$

$$= \frac{\frac{a_0^{(H)}\phi(z)\mathcal{B}_L(\phi(z))}{a_0^{(H)}\phi(z)+\mathcal{A}_H(z)-\phi(z)\mathcal{A}_H(\mathcal{B}_H(\phi(z)))}[\mathcal{B}_H(\phi(z)) - z]}{\frac{a_0^{(H)}\phi(z)\mathcal{B}_L(\phi(z))}{a_0^{(H)}\phi(z)+\mathcal{A}_H(z)-\phi(z)\mathcal{A}_H(\mathcal{B}_H(\phi(z)))}\mathcal{B}_H(\phi(z)) - z\mathcal{B}_L(\phi(z))}\Pi_{00}((1-r)z).$$

Some algebra leads to a simple recursion,

$$\Pi_{00}(z) = \frac{a_0^{(H)}\phi(z)[z-\mathcal{B}_H(\phi(z))]}{a_0^{(H)}\phi(z)[z - \mathcal{B}_H(\phi(z))] + z[\mathcal{A}_H(z) - \phi(z)\mathcal{A}_H(\mathcal{B}_H(\phi(z)))]}\Pi_{00}((1-r)z).$$

Now introduce the so-called *retrial function* for the primary batch-service model

$$\mathcal{R}_b(z) = \frac{a_0^{(H)}\phi(z)[z - \mathcal{B}_H(\phi(z))]}{a_0^{(H)}\phi(z)[z - \mathcal{B}_H(\phi(z))] + z[\mathcal{A}_H(z) - \phi(z)\mathcal{A}_H(\mathcal{B}_H(\phi(z)))]}.$$

Notice that $\mathcal{R}_b(0) = 1$ and after using L'Hôpital we find that

$$\mathcal{R}_b(1) = \frac{a_0^{(H)}[1 - \mathcal{B}'_H(1)\phi'(1)]}{a_0^{(H)}[1 - \mathcal{B}'_H(1)\phi'(1)] + \mathcal{A}'_H(1) - \phi'(1)[\mathcal{A}'_H(1)\mathcal{B}'_H(1) + 1]}$$

$$= \frac{a_0^{(H)}[1 - \mathcal{A}'_L(1)\mathcal{B}'_L(1) - \mathcal{A}'_H(1)\mathcal{B}'_H(1)] - \mathcal{A}'_L(1)}{a_0^{(H)}[1 - \mathcal{A}'_L(1)\mathcal{B}'_L(1) - \mathcal{A}'_H(1)\mathcal{B}'_H(1)] - \mathcal{A}'_L(1) - [\mathcal{A}'_L(1)\mathcal{B}'_L(1) + \mathcal{A}'_H(1)\mathcal{B}'_H(1)]\mathcal{A}'_H(1)}.$$

In the denominator we recognize the stability condition!

Now we get by iteration

$$\Pi_{00}(z) = \prod_{k=0}^{n-1} \mathcal{R}_b\left((1-r)^k z\right)\Pi_{00}\left((1-r)^n z\right).$$

Next, sending n to infinity we find

$$\Pi_{00}(z) = \prod_{k=0}^{\infty} \mathcal{R}_b\left((1-r)^k z\right)\Pi_{00}(0). \tag{8}$$

The problem is to calculate $\Pi_{00}(0)$, the steady-state probability that the system is empty!

From (6) we find

$$\Pi_0(y,z) = \frac{a_0^{(L)}a_0^{(H)}\omega(y,z)\left\{\begin{array}{c}[z\mathcal{B}_L(\omega(y,z)) - y\mathcal{B}_H(\omega(y,z))]\,\Pi_{00}(z)\\ -y\,[z - \mathcal{B}_H(\omega(y,z))]\,\Pi_{00}((1-r)z)\end{array}\right\}}{\mathcal{A}_L(y)z\left[\begin{array}{c}a_0^{(H)}\omega(y,z)[\mathcal{B}_L(\omega(y,z)) - y]\\ +y[\omega(y,z)\mathcal{A}_H(\mathcal{B}_H(\omega(y,z))) - \mathcal{A}_H(z)]\end{array}\right]}.$$

and using the recursion $\Pi_{00}(z) = \mathcal{R}_b(z)\Pi_{00}((1-r)z)$ we get

$$\Pi_0(y,z) = \frac{a_0^{(L)} a_0^{(H)} \mathcal{A}_H(z) \left\{ \begin{array}{c} [z\mathcal{B}_L(\omega(y,z)) - y\mathcal{B}_H(\omega(y,z))]\,\mathcal{R}_b(z) \\ -y\,[z - \mathcal{B}_H(\omega(y,z))] \end{array} \right\}}{z \left[\begin{array}{c} a_0^{(H)} \omega(y,z)[\mathcal{B}_L(\omega(y,z)) - y] \\ +y[\omega(y,z)\mathcal{A}_H(\mathcal{B}_H(\omega(y,z))) - \mathcal{A}_H(z)] \end{array} \right]} \times \Pi_{00}((1-r)z). \tag{9}$$

We know that $\Pi_0(1,1) = 1 - \mathcal{A}'_L(1)\mathcal{B}'_L(1) - \mathcal{A}'_H(1)\mathcal{B}'_H(1)$. So we can find $\Pi_{00}(1-r)$, again using L'Hôpital, from (9),

$$1 - \mathcal{A}'_L(1)\mathcal{B}'_L(1) - \mathcal{A}'_H(1)\mathcal{B}'_H(1) = \Pi_0(1,1) = \lim_{y\to 1} \Pi_0(y,1)$$

$$= \lim_{y\to 1} \frac{a_0^{(L)} a_0^{(H)} \left\{ \begin{array}{c} [\mathcal{B}_L(\mathcal{A}_L(y)) - y\mathcal{B}_H(\mathcal{A}_L(y))]\,\mathcal{R}_b(1) + \\ -y\,[1 - \mathcal{B}_H(\mathcal{A}_L(y))] \end{array} \right\}}{\begin{array}{c} a_0^{(H)} \mathcal{A}_L(y)[\mathcal{B}_L(\mathcal{A}_L(y)) - y] \\ +y[\mathcal{A}_L(y)\mathcal{A}_H(\mathcal{B}_H(\mathcal{A}_L(y))) - 1] \end{array}} \times \Pi_{00}(1-r)$$

$$= \frac{a_0^{(L)} a_0^{(H)} \{\mathcal{R}_b(1)[\mathcal{B}'_L(1)\mathcal{A}'_L(1) - 1 - \mathcal{B}'_H(1)\mathcal{A}'_L(1)] + \mathcal{B}'_H(1)\mathcal{A}'_L(1)\}}{a_0^{(H)}[\mathcal{B}'_L(1)\mathcal{A}'_L(1) - 1] + \mathcal{A}'_L(1) + \mathcal{A}'_H(1)\mathcal{B}'_H(1)\mathcal{A}'_L(1)} \Pi_{00}(1-r)$$

$$= \frac{a_0^{(L)} a_0^{(H)} [1 - \mathcal{A}'_L(1)\mathcal{B}'_L(1) - \mathcal{A}'_H(1)\mathcal{B}'_H(1)]}{a_0^{(H)}[1 - \mathcal{A}'_L(1)\mathcal{B}'_L(1) - \mathcal{A}'_H(1)\mathcal{B}'_H(1)] - \mathcal{A}'_L(1) - [\mathcal{A}'_L(1)\mathcal{B}'_L(1) + \mathcal{A}'_H(1)\mathcal{B}'_H(1)]\mathcal{A}'_H(1)} \Pi_{00}(1-r).$$

So, using the *offered load* $\varrho = \mathcal{A}'_L(1)\mathcal{B}'_L(1) + \mathcal{A}'_H(1)\mathcal{B}'_H(1)$ we find

$$\Pi_{00}(1-r) = \frac{a_0^{(H)}(1-\varrho) - \mathcal{A}'_L(1) - \varrho\mathcal{A}'_H(1)}{a_0^{(L)} a_0^{(H)}}$$

and this leads to

$$\Pi_{00}(z) = \prod_{k=0}^{\infty} \mathcal{R}_b\left((1-r)^k z\right) \Pi_{00}(0)$$

$$= \frac{a_0^{(H)}(1-\varrho) - \mathcal{A}'_L(1) - \varrho\mathcal{A}'_H(1)}{a_0^{(L)} a_0^{(H)}} \mathcal{R}_b(1) \prod_{k=0}^{\infty} \frac{\mathcal{R}_b\left((1-r)^k z\right)}{\mathcal{R}_b\left((1-r)^k\right)}$$

$$= \frac{a_0^{(H)}[1-\varrho] - \mathcal{A}'_L(1)}{a_0^{(L)} a_0^{(H)}} \prod_{k=0}^{\infty} \frac{\mathcal{R}_b\left((1-r)^k z\right)}{\mathcal{R}_b\left((1-r)^k\right)}.$$

Now we can move to the next step in our search for a 'closed form formula' for $\Pi(w,y,z)$. Recall from (9) and the definition of $\mathcal{R}_b(z)$

$$\Pi_0(y,z) = \frac{a_0^{(L)} a_0^{(H)} \mathcal{A}_H(z) \left\{ \begin{array}{c} [z\mathcal{B}_L(\omega(y,z)) - y\mathcal{B}_H(\omega(y,z))]\,\mathcal{R}_b(z) \\ -y\,[z - \mathcal{B}_H(\omega(y,z))] \end{array} \right\}}{z \left[\begin{array}{c} a_0^{(H)} \omega(y,z)[\mathcal{B}_L(\omega(y,z)) - y] \\ +y[\omega(y,z)\mathcal{A}_H(\mathcal{B}_H(\omega(y,z))) - \mathcal{A}_H(z)] \end{array} \right]} \times \Pi_{00}((1-r)z) \tag{10}$$

$$= \left(\frac{a_0^{(L)} a_0^{(H)} \mathcal{A}_H(z) \left\{ \begin{array}{c} a_0^{(H)} \phi(z)\,[z\mathcal{B}_L(\omega(y,z)) - y\mathcal{B}_H(\omega(y,z))]\,[z - \mathcal{B}_H(\phi(z))] + \\ -y\,[z - \mathcal{B}_H(\omega(y,z))] \\ \times\{a_0^{(H)}\phi(z)[z - \mathcal{B}_H(\phi(z))] + z[\mathcal{A}_H(z) - \phi(z)\mathcal{A}_H(\mathcal{B}_H(\phi(z)))]\} \end{array} \right\}}{z \left[\begin{array}{c} a_0^{(H)} \omega(y,z)[\mathcal{B}_L(\omega(y,z)) - y] \\ +y[\omega(y,z)\mathcal{A}_H(\mathcal{B}_H(\omega(y,z))) - \mathcal{A}_H(z)] \end{array} \right] \times\{a_0^{(H)}\phi(z)[z - \mathcal{B}_H(\phi(z))] + z[\mathcal{A}_H(z) - \phi(z)\mathcal{A}_H(\mathcal{B}_H(\phi(z)))]\}} \right) \times \frac{a_0^{(H)}[1-\varrho] - \mathcal{A}_L'(1) - \varrho\mathcal{A}_H'(1)}{a_0^{(L)} a_0^{(H)}} \prod_{k=1}^{\infty} \frac{\mathcal{R}_b\left((1-r)^k z\right)}{\mathcal{R}_b\left((1-r)^k\right)}.$$

Finally, we can find the full p.g.f. $\Pi(w,y,z)$! Recall (5) and use the result (10) for $\Pi_0(y,z)$ and $\Pi_{00}((1-r)z)$,

$$\Pi(w,y,z)$$

$$= \frac{\begin{array}{c} \mathcal{A}_L(y)z\left[a_0^{(H)} w(\mathcal{B}_L(w) - y) + y(w\mathcal{A}_H(\mathcal{B}_H(w)) - \mathcal{A}_H(z))\right] \Pi_0(y,z) \\ +a_0^{(L)} a_0^{(H)} w\,[y\mathcal{B}_H(w) - z\mathcal{B}_L(w)]\,\Pi_{00}(z) \\ +a_0^{(L)} a_0^{(H)} wy\,[z - \mathcal{B}_H(w)]\,\Pi_{00}((1-r)z) \end{array}}{yz(w - \mathcal{A}_L(y)\mathcal{A}_H(z))}.$$

Substitution of our previous results gives

$$\Pi(w,y,z) = \left(a_0^{(H)}[1-\varrho] - \mathcal{A}_L'(1) - \varrho\mathcal{A}_H'(1)\right) \prod_{k=1}^{\infty} \frac{\mathcal{R}_b\left((1-r)^k z\right)}{\mathcal{R}_b\left((1-r)^k\right)}$$
$$\times \frac{\begin{array}{c} \mathcal{A}_L(y)z\left[a_0^{(H)} w(\mathcal{B}_L(w) - y) + y(w\mathcal{A}_H(\mathcal{B}_H(w)) - \mathcal{A}_H(z))\right] \\ \times \left[\frac{\mathcal{A}_H(z)\left\{ \begin{array}{c} [z\mathcal{B}_L(\omega(y,z)) - y\mathcal{B}_H(\omega(y,z))]\,\mathcal{R}_b(z) \\ -y\,[z - \mathcal{B}_H(\omega(y,z))] \end{array} \right\}}{z\left[\begin{array}{c} a_0^{(H)} \omega(y,z)[\mathcal{B}_L(\omega(y,z)) - y] \\ +y[\omega(y,z)\mathcal{A}_H(\mathcal{B}_H(\omega(y,z))) - \mathcal{A}_H(z)] \end{array} \right]} \right] \\ +w\,[y\mathcal{B}_H(w) - z\mathcal{B}_L(w)]\,\mathcal{R}_b(z) \\ +wy\,[z - \mathcal{B}_H(w)] \end{array}}{yz(w - \omega(y,z))}. \tag{11}$$

4 The Queue Size and the Orbit Size

From expression (11) we find the marginal p.g.f.'s $\mathcal{L}(y) := \Pi(1, y, 1)$ and $\mathcal{Q}(z) := \Pi(1, 1, z)$ of the limiting distribution of the queue length and the orbit size, respectively. After some simplifications we find

$$\mathcal{L}(y) = \left(a_0^{(H)}[1-\varrho] - \mathcal{A}_L'(1) - \varrho\mathcal{A}_H'(1)\right)\left(\frac{1-y}{1-\mathcal{A}_L(y)}\right) \times \frac{a_0^{(H)}\mathcal{A}_L(y)(\mathcal{R}_b(1)-1)[1-\mathcal{B}_H(\mathcal{A}_L(y))] + \mathcal{R}_b(1)[1-\mathcal{A}_L(y)\mathcal{A}_H(\mathcal{B}_H(\mathcal{A}_L(y)))]}{a_0^{(H)}\mathcal{A}_L(y)[\mathcal{B}_L(\mathcal{A}_L(y)) - y] + y[\mathcal{A}_L(y)\mathcal{A}_H(\mathcal{B}_H(\mathcal{A}_L(y))) - 1]}$$

$$\mathcal{Q}(z) = \left(a_0^{(H)}[1-\varrho] - \mathcal{A}_L'(1) - \varrho\mathcal{A}_H'(1)\right)\prod_{k=1}^{\infty}\frac{\mathcal{R}_b\left((1-r)^k z\right)}{\mathcal{R}_b\left((1-r)^k\right)} \times \left[\frac{\mathcal{R}_b(z)[z\mathcal{B}_L(\mathcal{A}_H(z)) - \mathcal{B}_H(\mathcal{A}_H(z))] + \mathcal{B}_H(\mathcal{A}_H(z)) - z}{z\left\{a_0^{(H)}[\mathcal{B}_L(\mathcal{A}_H(z)) - 1] + \mathcal{A}_H(\mathcal{B}_H(\mathcal{A}_H(z))) - 1\right\}} + \frac{(1-z)\left(\mathcal{R}_b(z)-1\right)}{z\left(1-\mathcal{A}_H(z)\right)}\right].$$

To find the *mean queue length* $\overline{\mathcal{L}} = \mathcal{L}'(1)$ we write

$$\mathcal{L}(y) \ = \ (1-\sigma) \times F(y) \times \frac{N(z)}{D(z)}$$

where σ is again the total used capacity

$$\sigma = \varrho + \mathcal{A}_L'(1) + \varrho\mathcal{A}_H'(1) + \left(1 - a_0^{(H)}\right)(1-\varrho)$$

and

$$F(y) = \frac{1-y}{1-\mathcal{A}_L(y)}$$
$$N(y) = a_0^{(H)}\mathcal{A}_L(y)(\mathcal{R}_b(1)-1)[1-\mathcal{B}_H(\mathcal{A}_L(y))] + \mathcal{R}_b(1)[1-\mathcal{A}_L(y)\mathcal{A}_H(\mathcal{B}_H(\mathcal{A}_L(y)))]$$
$$D(y) = a_0^{(H)}\mathcal{A}_L(y)[\mathcal{B}_L(\mathcal{A}_L(y)) - y] + y[\mathcal{A}_L(y)\mathcal{A}_H(\mathcal{B}_H(\mathcal{A}_L(y))) - 1].$$

Differentiating $\mathcal{L}(y)$ gives

$$\mathcal{L}'(y) \ = \ (1-\sigma)\left(F(y)\cdot\frac{N'(y)D(y) - N(y)D'(y)}{[D(y)]^2} + F'(y)\cdot\frac{N(y)}{D(y)}\right).$$

So we want to calculate

$$\mathcal{L}'(1) \ = \ (1-\sigma)\lim_{y\to 1}\left(F(y)\cdot\frac{N'(y)D(y) - N(y)D'(y)}{[D(y)]^2} + F'(y)\cdot\frac{N(y)}{D(y)}\right).$$

After tedious calculations using L'Hôpital we find for the mean queue length

$$\overline{\mathcal{L}} = \mathcal{L}'(1) = \frac{1-\sigma}{\mathcal{A}'_L(1)} \cdot \frac{N''(1)}{2D'(1)} - \frac{D''(1)}{2D'(1)} - \frac{\mathcal{A}''_L(1)}{2\mathcal{A}'_L(1)}$$

where

$$\begin{aligned}
N''(1) &= a_0^{(H)}\left(1-\mathcal{R}_b(1)\right)\left(2\mathcal{A}'_L(1)^2\mathcal{B}'_H(1) + \mathcal{A}'_L(1)^2\mathcal{B}''_H(1) + \mathcal{A}''_L(1)\mathcal{B}'_H(1)\right) + \\
&\quad -\mathcal{R}_b(1)\left(2\mathcal{A}'_L(1)^2\mathcal{B}'_H(1)\mathcal{A}'_H(1) + \mathcal{A}'_L(1)^2\mathcal{B}''_H(1)\mathcal{A}'_H(1) + \mathcal{A}''_L(1)\mathcal{B}'_H(1)\mathcal{A}'_H(1)\right. \\
&\quad \left. + \mathcal{A}''_L(1) + \mathcal{A}'_L(1)^2\mathcal{A}''_H(1)\mathcal{B}'_H(1)^2\right) \\
D'(1) &= a_0^{(H)}\left(\mathcal{B}'_L(1)\mathcal{A}'_L(1) - 1\right) + \mathcal{A}'_L(1) + \mathcal{A}'_H(1)\mathcal{B}'_H(1)\mathcal{A}'_L(1) \\
D''(1) &= 2a_0^{(H)}\mathcal{A}'_L(1)\left(\mathcal{A}'_L(1)\mathcal{B}'_L(1) - 1\right) + a_0^{(H)}\mathcal{A}''_L(1)\mathcal{B}'_L(1) + a_0^{(H)}\mathcal{A}'_L(1)^2\mathcal{B}''_L(1) \\
&\quad + 2\mathcal{A}'_L(1)^2\mathcal{A}'_H(1)\mathcal{B}'_H(1) + \mathcal{A}'_L(1)^2\mathcal{B}''_H(1)\mathcal{A}'_H(1) + \mathcal{A}''_L(1)\mathcal{A}'_H(1)\mathcal{B}'_H(1) + \mathcal{A}''_L(1) \\
&\quad + \mathcal{A}'_L(1)^2\mathcal{A}''_H(1)\mathcal{B}'_H(1)^2 + 2\mathcal{A}'_L(1) + 2\mathcal{A}'_L(1)\mathcal{A}'_H(1)\mathcal{B}'_H(1).
\end{aligned}$$

To calculate $\overline{\mathcal{Q}} = \mathcal{Q}'(1)$ first rewrite $\mathcal{Q}(z)$ as

$$\mathcal{Q}(z) \;=\; (1-\sigma)\prod_{k=1}^{\infty}\frac{\mathcal{R}_b((1-r)^k z)}{\mathcal{R}_b((1-r)^k)}\left(\frac{N_1(z)}{D_1(z)} + \frac{N_2(z)}{D_2(z)}\right)$$

with

$$\begin{aligned}
N_1(z) &= \left[z\mathcal{B}_L(\mathcal{A}_H(z)) - \mathcal{B}_H(\mathcal{A}_H(z))\right]\mathcal{R}_b(z) + \mathcal{B}_H(\mathcal{A}_H(z)) - z \\
D_1(z) &= z\left[a_0^{(H)}(\mathcal{B}_L(\mathcal{A}_H(z)) - 1) + \mathcal{A}_H(\mathcal{B}_H(\mathcal{A}_H(z))) - 1\right] \\
N_2(z) &= (1-z)\left[\mathcal{R}_b(z) - 1\right] \\
D_2(z) &= z\left[1 - \mathcal{A}_H(z)\right].
\end{aligned}$$

Differentiating $\mathcal{Q}(z)$ gives

$$\begin{aligned}
\frac{\mathcal{Q}'(z)}{1-\sigma} &= \prod_{k=1}^{\infty}\frac{\mathcal{R}_b((1-r)^k z)}{\mathcal{R}_b((1-r)^k)}\left(\frac{N'_1(z)D_1(z) - N_1(z)D'_1(z)}{D_1(z)^2} + \frac{N'_2(z)D_2(z) - N_1(z)D'_2(z)}{D_2(z)^2}\right) \\
&\quad + \sum_{k=1}^{\infty}\frac{(1-r)^k\mathcal{R}'_b((1-r)^k z)}{\mathcal{R}_b((1-r)^k)}\prod_{i\neq k}\frac{\mathcal{R}_b((1-r)^i z)}{\mathcal{R}_b((1-r)^i)}\left(\frac{N_1(z)}{D_1(z)} + \frac{N_2(z)}{D_2(z)}\right)
\end{aligned}$$

and we need to calculate

$$\begin{aligned}
\frac{\mathcal{Q}'(1)}{1-\sigma} &= \lim_{z\to 1}\left(\frac{N'_1(z)D_1(z) - N_1(z)D'_1(z)}{D_1(z)^2} + \frac{N'_2(z)D_2(z) - N_1(z)D'_2(z)}{D_2(z)^2}\right) \\
&\quad + \sum_{k=1}^{\infty}\frac{(1-r)^k\mathcal{R}'_b((1-r)^k)}{\mathcal{R}_b((1-r)^k)}\lim_{z\to 1}\left(\frac{N_1(z)}{D_1(z)} + \frac{N_2(z)}{D_2(z)}\right).
\end{aligned}$$

Again after tedious calculations we find

$$\mathcal{Q}'(1) = \frac{(1-\sigma)N_1''(1)}{2D_1'(1)} - \frac{(1-\varrho)D_1''(1)}{2D_1'(1)} + \frac{(1-\sigma)N_2''(1)}{2D_2'(1)} - \frac{\varrho D_2''(1)}{2D_2'(1)} + \sum_{k=1}^{\infty} \frac{(1-r)^k \mathcal{R}_b'((1-r)^k)}{\mathcal{R}_b((1-r)^k)} \tag{12}$$

with

$$\begin{aligned} N_1''(1) &= \mathcal{R}_b(1)\big(\mathcal{B}_L'(1)\mathcal{A}_H''(1) + \mathcal{A}_H'(1)^2\mathcal{B}_L''(1) + 2\mathcal{B}_L'(1)\mathcal{A}_H'(1) + \\ &\qquad -\mathcal{B}_H'(1)\mathcal{A}_H''(1) - \mathcal{A}_H'(1)^2\mathcal{B}_H''(1)\big) \\ &\quad +\mathcal{R}_b'(1)\big(\mathcal{B}_L'(1)\mathcal{A}_H'(1) - \mathcal{B}_H'(1)\mathcal{A}_H'(1) + 1\big) + \mathcal{B}_H'(1)\mathcal{A}_H''(1) + \mathcal{B}_H''(1)\mathcal{A}_H'(1)^2 \\ D_1'(1) &= a_0^{(H)}\mathcal{B}_L'(1)\mathcal{A}_H'(1) + \mathcal{A}_H'(1)^2\mathcal{B}_H'(1) \\ D_1''(1) &= 2a_0^{(H)}\mathcal{B}_L'(1)\mathcal{A}_H'(1) + 2\mathcal{A}_H'(1)^2\mathcal{B}_H'(1) + a_0^{(H)}\mathcal{B}_L'(1)\mathcal{A}_H''(1) + a_0^{(H)}\mathcal{B}_L''(1)\mathcal{A}_H'(1)^2 \\ &\quad +\mathcal{A}_H'(1)\mathcal{B}_H'(1)\mathcal{A}_H''(1) + \mathcal{A}_H'(1)^3\mathcal{B}_H''(1) + \mathcal{A}_H'(1)^2\mathcal{B}_H'(1)^2\mathcal{A}_H''(1) \end{aligned}$$

and

$$\begin{aligned} N_2''(1) &= -2\mathcal{R}_b'(1) \\ D_2'(1) &= -\mathcal{A}_H'(1), \quad D_2''(1) = -\mathcal{A}_H''(1) - 2\mathcal{A}_H'(1). \end{aligned}$$

Of course, we also have to calculate the derivative $\mathcal{R}_b'(1)$. Recall that

$$\begin{aligned} \mathcal{R}_b(z) &= \frac{a_0^{(H)}\phi(z)[z - \mathcal{B}_H(\phi(z))]}{a_0^{(H)}\phi(z)[z - \mathcal{B}_H(\phi(z))] + z[\mathcal{A}_H(z) - \phi(z)\mathcal{A}_H(\mathcal{B}_H(\phi(z)))]} \\ &= \frac{y^*(z)[\mathcal{B}_H(\phi(z)) - z]}{y^*(z)\mathcal{B}_H(\phi(z)) - z\mathcal{B}_L(\phi(z))} \end{aligned}$$

where

$$\phi(z) = \omega(y^*(z), z), \quad \omega(y, z) = \mathcal{A}_L(y)\mathcal{A}_H(z)$$

and $y = y^*(z)$ is the *unique solution* of the equation

$$a_0^{(H)}\omega(y,z)[\mathcal{B}_L(\omega(y,z)) - y] + y[\omega(y,z)\mathcal{A}_H(\mathcal{B}_H(\omega(y,z))) - \mathcal{A}_H(z)] = 0.$$

Now, introduce

$$\begin{aligned} N(z) &= y^*(z)[\mathcal{B}_H(\phi(z)) - z] \\ D(z) &= y^*(z)\mathcal{B}_H(\phi(z)) - z\mathcal{B}_L(\phi(z)). \end{aligned}$$

Then differentiation gives

$$\begin{aligned} \mathcal{R}_b'(z) &= \frac{D(z)N'(z) - N(z)D'(z)}{D(z)^2} \\ \mathcal{R}_b'(1) &= \lim_{z\to 1} \mathcal{R}_b(z) = \frac{N''(1) - \mathcal{R}_b(1)D''(1)}{2D'(1)} \end{aligned}$$

where we have

$$\begin{aligned} N''(1) &= 2y^{*\prime}(1)\,(1-\mathcal{B}_H'(1)\phi'(1)) - \mathcal{B}_H'(1)\phi''(1) - \mathcal{B}_H''(1)\phi'(1)^2 \\ D'(1) &= 1+\mathcal{B}_L'(1)\phi'(1) - \mathcal{B}_H'(1)\phi'(1) - y^{*\prime}(1) \\ D''(1) &= \phi''(1)\,(\mathcal{B}_L'(1)-\mathcal{B}_H'(1))\,\phi'(1)^2\,(\mathcal{B}_L''(1)-\mathcal{B}_H''(1)) \\ &\quad + 2\mathcal{B}_L'(1)\phi'(1) - 2\mathcal{B}_H'(1)y^{*\prime}(1)\phi'(1) - y^{*\prime\prime}(1), \end{aligned}$$

and we recall that

$$\mathcal{R}_b(1) = \frac{a_0^{(H)}[1-\mathcal{B}_H'(1)\phi'(1)]}{a_0^{(H)}[1-\mathcal{B}_H'(1)\phi'(1)] + \mathcal{A}_H'(1) - \phi'(1)[\mathcal{A}_H'(1)\mathcal{B}_H'(1)+1]}.$$

The expressions for $y^{*\prime}(1)$ and $\phi'(1)$ are given by

$$y^{*\prime}(1) = \frac{\mathcal{A}_H'(1)\left[a_0^{(H)}\mathcal{B}_L'(1) + \mathcal{A}_H'(1)\mathcal{B}_H'(1)\right]}{a_0^{(H)}\left[1-\mathcal{A}_L'(1)\mathcal{B}_L'(1)\right] - \mathcal{A}_L'(1)\left[1+\mathcal{A}_H'(1)\mathcal{B}_H'(1)\right]}$$

$$\phi'(1) = \frac{\mathcal{A}_H'(1)\left[a_0^{(H)} - \mathcal{A}_L'(1)\right]}{a_0^{(H)}\left[1-\mathcal{A}_L'(1)\mathcal{B}_L'(1)\right] - \mathcal{A}_L'(1)\left[1+\mathcal{A}_H'(1)\mathcal{B}_H'(1)\right]}$$

and after many further calculations we find

$$y^{*\prime\prime}(1) = \frac{\begin{array}{l} \left(a_0^{(H)}\mathcal{B}_L'(1) + \mathcal{A}_H'(1)\mathcal{B}_H'(1) + 1\right) \\ \times\left(2\mathcal{A}_L'(1)\mathcal{A}_H'(1)y^{*\prime}(1) + \mathcal{A}_H''(1) + \mathcal{A}_L''(1)y^{*\prime}(1)^2\right) \\ +\phi'(1)^2\left(2a_0^{(H)}\mathcal{B}_L'(1) + a_0^{(H)}\mathcal{B}_L''(1) + \mathcal{A}_H'(1)\mathcal{B}_H''(1)\right. \\ + \qquad\qquad \left.\mathcal{A}_H''(1)\mathcal{B}_H'(1)^2 + 2\mathcal{A}_H'(1)\mathcal{B}_H'(1)\right) + \\ -\mathcal{A}_H''(1) - 2y^{*\prime}(1)\left(\mathcal{A}_H'(1) + \phi'(1)\left(a_0^{(H)} - \mathcal{A}_H'(1)\mathcal{B}_H'(1) - 1\right)\right) \end{array}}{a_0^{(H)} - a_0^{(H)}\mathcal{A}_L'(1)\mathcal{B}_L'(1) - \mathcal{A}_L'(1)\mathcal{A}_H'(1)\mathcal{B}_H'(1) - \mathcal{A}_L'(1)}$$

$$\phi''(1) = \mathcal{A}_H''(1) + 2\mathcal{A}_L'(1)\mathcal{A}_H'(1)y^{*\prime}(1) + \mathcal{A}_L''(1)y^{*\prime}(1)^2 + \mathcal{A}_L'(1)y^{*\prime\prime}(1).$$

Plugging in all these results in (12) gives a closed form expression for $\overline{\mathcal{Q}} = \mathcal{Q}'(1)$.

5 Numerical Results

The starting position for our numerical results is

– All distributions geometric [batch size shifted to 0]

$$\mathcal{B}_H'(1) = \mathcal{B}_L'(1) = 2, \quad \mathcal{A}_H'(1) = 0.21, \;\; \mathcal{A}_L'(1) = 0.09,$$

– So $a_0^{(H)} = 0.826$,
– The offered load of L-customers is $\varrho_L := \mathcal{A}_L'(1)\mathcal{B}_L'(1) = 0.18$,

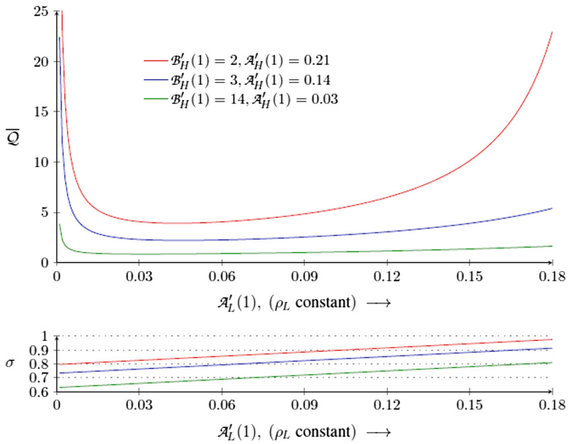

Fig. 1. The mean size of the orbit $\bar{Q}$ as a function of $\mathcal{A}'_L(1)$, with $r = 0.5$ and $\rho_L = 0.18$ constant.

- The offered loa4d of H-customers is $\varrho_H := \mathcal{A}'_H(1)\mathcal{B}'_H(1) = 0.42$,
- The total offered load is $\varrho = \varrho_L + \varrho_H = 0.60$,
- The total used capacity is

$$\sigma = \varrho + \mathcal{A}'_L(1) + \varrho\mathcal{A}'_H(1) + \left(1 - a_0^{(H)}\right)(1 - \varrho) = 0.885.$$

Firstly, we take the offered load of L-customers $\varrho_L := \mathcal{A}'_L(1)\mathcal{B}'_L(1)$ constant and $\mathcal{A}'_L(1)$ increasing. The numerical results are presented in Fig. 1.

Next, we keep the total offered load $\varrho = 0.6$ constant, $\varrho_H := \mathcal{A}'_H(1)\mathcal{B}'_H(1)$ increasing. The results are presented in Fig. 2.

Finally, we keep the total offered load $\varrho = 0.65$ constant, and again ϱ_H increasing. The results are presented in Fig. 3.

From the Figs. 1, 2 and 3 we can draw the following conclusions.

- Keeping both the offered load of inbound calls $\varrho_H = \mathcal{A}'_H(1)\mathcal{B}'_H(1)$ and the offered load of outbound calls $\varrho_L = \mathcal{A}'_L(1)\mathcal{B}'_L(1)$ constant we have seen that
  - Increasing the arrival intensity $\mathcal{A}'_H(1)$ of inbound calls [and simultaneously decreasing the mean service time $\mathcal{B}'_H(1)$] *decreases the mean queue length* $\overline{\mathcal{L}}$ of outbound calls and *increases the mean orbit size* $\overline{\mathcal{Q}}$ of inbound calls.
  - Increasing the arrival intensity $\mathcal{A}'_L(1)$ of outbound calls [and simultaneously decreasing the mean service time $\mathcal{B}'_L(1)$] *increases the mean queue length* $\overline{\mathcal{L}}$ of outbound calls and *first decreases and then increases the mean orbit size* $\overline{\mathcal{Q}}$ of inbound calls.

Keeping the total offered load $\varrho = \varrho_H + \varrho_L$ constant we have seen that *increasing the offered load* ϱ_H *of inbound calls* [and simultaneously decreasing the offered load ϱ_L of outbound calls]

- *decreases the mean queue length* $\overline{\mathcal{L}}$ of outbound calls and *increases the mean orbit size* $\overline{\mathcal{Q}}$ of inbound calls.

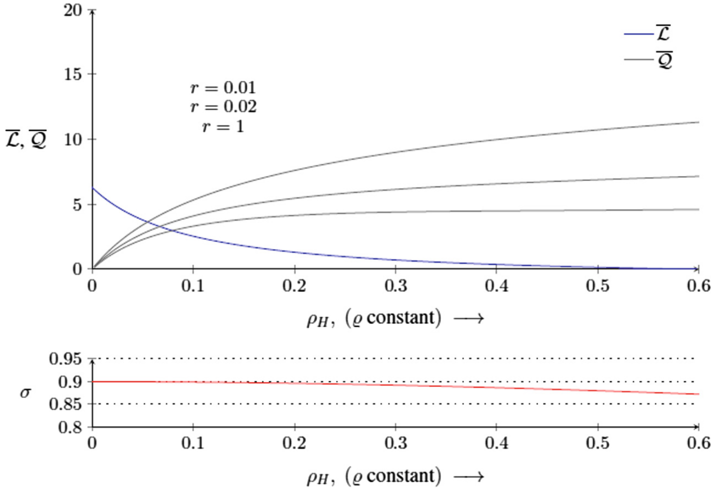

Fig. 2. The mean length of the queue $\bar{\mathcal{L}}$ and mean size of the orbit $\bar{\mathcal{Q}}$ as a function of ρ_H, with $\varrho = 0.6$ constant.

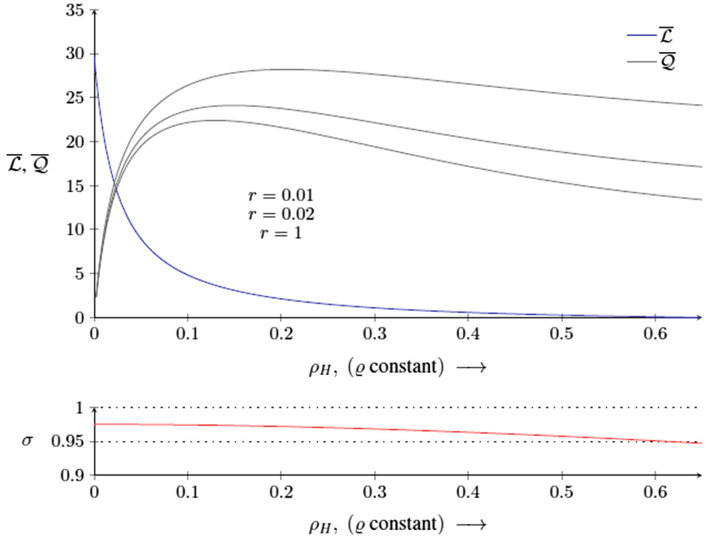

Fig. 3. The mean length of the queue $\bar{\mathcal{L}}$ and mean size of the orbit $\bar{\mathcal{Q}}$ as a function of ρ_H, with offered load $\varrho = 0.65$ constant.

– increases the mean orbit size $\overline{\mathcal{Q}}$ of inbound calls *for a moderate total offered load, say* $\varrho = 0.6$,
– first increases and then decreases the mean orbit size $\overline{\mathcal{Q}}$ of inbound calls *for a high total offered load, say* $\varrho = 0.65$.

For more numerical results we refer to Dekker [4].

References

1. Artalejo, J.R., Gómez-Corral, A.: Retrial Queueing Systems. Springer-Verlag, Heidelberg (2008). https://doi.org/10.1007/978-3-540-78725-9
2. Artalejo, J.R., Dudin, A.N., Klimenok, V.I.: Stationary analysis of a retrial queue with preemptive repeated attempts. Oper. Res. Lett. **28**, 173–180 (2001)
3. Choi, B.D., Kim, J.W.: Discrete-time $Geo_1, Geo_2/G/1$ retrial queueing systems with two types of calls. Comput. Math. Appl. **33**, 79–88 (1997)
4. Dekker, M.: Invloed van Uitgaande Gesprekken op Inkomende Gesprekken Bij Een Call-center [in Dutch], Bachelor thesis, Vrije Universiteit Amsterdam (2016)
5. Falin, G.I., Artalejo, J.R., Martin, M.: On the single server retrial queue with priority customers. Queueing Syst. **14**, 439–455 (1993)
6. Nobel, R.D.: A mixed discrete-time delay/retrial queueing model for handover calls and new calls competing for a targat channel, queueing theory and network applications. Adv. Intell. Syst. Comput. **383**, 173–185 (2015)
7. Nobel, R.D.: Retrial queueing models in discrete time: a short survey of some late arrival models. Ann. Oper. Res. **247**(1), 37–63 (2016)
8. Nobel, R.D., Moreno, P.: A discrete-time priority loss/retrial queueing model with two types of traffic. In: Choi, B.D. (ed.) Proceedings of the Korea-Netherlands Joint Conference on Queueing Theory and its Applications to Telecommunication Systems, Seoul, pp. 189–207 (2005)
9. Nobel, R.D., Moreno, P.: A discrete-time retrial queueing model with one server. Eur. J. Oper. Res. **189**(3), 1088–1103 (2008)
10. el Sharkawy, S.: Een Discrte-tijd Retrial/delay Wachtrijmodel Met Klanten Van Lage en Hoge Prioriteit [in Dutch], Bachelor thesis, Vrije Universiteit Amsterdam (2014)