Probability in the Engineering and Informational Sciences, **28**, 2014, 1–19.
doi:10.1017/S0269964813000296

THE ISRAELI QUEUE WITH INFINITE NUMBER OF GROUPS

Nir Perel and Uri Yechiali
Department of Statistics and Operations Research
School of Mathematical Sciences, Tel-Aviv University
Tel-Aviv, Israel
E-mail: perelnir@post.tau.ac.il; uriy@post.tau.ac.il

The so called "Israeli Queue" is a single server polling system with batch service of an unlimited size, where the next queue to be visited is the one in which the first customer in line has been waiting for the longest time. The case with finite number of queues (groups) was introduced by Boxma, Van der Wal and Yechiali [3]. In this paper we extend the model to the case with a (possibly) infinite number of queues. We analyze the $M/M/1$, $M/M/c$, and $M/M/1/N$—type queues, as well as a priority model with (at most) M high-priority classes and a single lower priority class. In all models we present an extensive probabilistic analysis and calculate key performance measures.

1. INTRODUCTION

Consider a single-server queue where arriving customers form groups. Each group is unrestricted in its size, and when served it is served in one batch. The service duration of a batch is independent of the group size. After a group is served, the next group to be attended is the one having the most "senior" customer (the one who has been waiting for the longest time). For example, this queue discipline represents a real situation (termed "The Israeli Queue") of a physical waiting line for buying tickets to a movie, theater, or a rock-concert performance. New arrivals see only one representative of each group (the most senior member or the group leader). A new arrival who is a friend with a group leader already standing in line joins the group. When the "leader" reaches the cashier he ("he" stands for "she" as well) buys tickets for the entire group. It is assumed that the buying process is (almost) not affected by the number of tickets purchased. If a new arrival does not find such a friend, he forms a new group, the last in line.

This model was first introduced and studied by Van der Wal and Yechiali [12]. They studied a polling system (see e.g., Takagi [10], Boon, Van der Mei, and Winands [2], Yechiali [13]) with N queues and a single server, where service at each queue is performed in unlimited-size batches. That is, when the server arrives at queue i, $i = 1, 2 \ldots, N$, it serves all customers (jobs) present there in one batch, where the duration of a batch-service is independent of the batch size. The motivation in [12] was to analyze a computer tape-reading problem in a system where large amounts of information are stored on tapes. Requests for data stored on one of these tapes arrive randomly and in order to read the data the tape

 0269-9648/14 $25.00

has to be mounted, read and then dismounted. If there are several requests to be read from a tape, they all can be read in (more or less) the same time, thus suggesting a modeling as a batch-service with unlimited batch size. Various server dynamic visit-order rules were analyzed, leading, under various objective functions, to surprisingly simple optimal index-type operating procedures. The probabilistic characteristics of the unlimited-size batch service polling system were further analyzed by Boxma et al. [3] for the Exhaustive, Gated and Globally-Gated service disciplines, where the server visits the queues in a cyclic order. Furthermore, in [3] the following server's visit-order rule was studied: after completion of service in a queue, the next queue to be served is the one where its first customer in line has been waiting for the longest time. That is, the criterion of selecting the next queue to visit and serve is an age-based one. This type of service discipline was termed "The Israeli Queue".

Unlimited batch-service models were also considered in the literature as application to videotex, telex and Time Division Multiple Access (TDMA) systems (see e.g. [1,4,8]). In addition, Van Oyen and Teneketzis [11] formulated a central data base system and an Automated Guided Vehicle as a polling system with an infinite capacity batch service.

In this work we extend the batch-service model with finite number of groups to the case where there is no bound on the number of different groups that can be present simultaneously in the system. We assume that the probability that an arriving job knows a group leader standing in line is p, independent of the group's size. Specifically, suppose there are n groups in the system, including the one in service. Then, the probability that an arriving customer will join the group standing in the kth position is $(1-p)^{k-1}p$ for $1 \leq k \leq n$ (the 1st position refers to the group in service, the 2nd refers to the group to be served next, and so on). Clearly, if an arriving customer finds no friends among any of the n group leaders, he will create a new group, with probability $(1-p)^n$. Hence, as the number of groups in the system increases, the growth rate of new groups decreases geometrically. Recently, He and Chavoushi [6] studied a queueing model with customer interjections, where customers are distinguished between normal and interjecting. All customers join a single queue. A normal customer joins the queue at its end, while an interjecting customer tries to cut into the queue following a geometric distribution. The waiting times of normal customers and of interjecting customers were studied.

In what follows we present an extensive probabilistic analysis of the "Israeli Queue' for the following cases: In Section 2, we analyze an $M/M/1$-type queue with an un-restricted number of groups, that is, a single server system, where the arrival stream follows a homogeneous Poisson process, and the batch service time is exponentially distributed. In Section 3, we briefly give the main results for the corresponding $M/M/c$-type and $M/M/1/N$-type queues. We conclude in Section 4 with a two-class (VIP and ordinary) priority model where the VIP's can form at most M Israeli Queue-type groups. For each model, a set of performance measures is derived for key parameters such as queue size, busy period, sojourn and waiting times, size of a batch, position of a new arrival, and number of groups being bypassed by an arriving customer.

2. THE $M/M/1$-TYPE MODEL

2.1. Model Description

We consider a single-server queue with an infinite buffer, where arriving customers form groups as follows: each group has a "leader" or a "head"—the first one of the group to arrive to the system. A new arrival sees only the head of each existing group, and the probability that he knows a group leader is p, independent of the group's size. That is, if there are n groups in the system (including the one in service), then the probability that

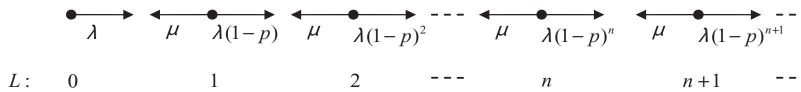

FIGURE 1. Transition rate diagram of the process $\{L(t), t \geq 0\}$.

a new arrival joins the kth group is $(1-p)^{k-1}p$, for $1 \leq k \leq n$, while the probability that he creates a new group is $(1-p)^n$. We assume that an arriving customer can also join the group that is being served. The arrival process is Poisson with rate λ, and the service is given in unlimited-size batches. That is, it takes one service duration to serve a group, independent of its size. We assume that a service duration of each group is exponentially distributed with parameter μ.

The underlying process of the system is represented by a continuous-time Markov chain $\{L(t), t \geq 0\}$, where $L(t)$ is the total number of different groups (different families) in the system at time t. A transition rate diagram of $L(t)$ is depicted in Figure 1.

We analyze this model and calculate various performance measures, such as the waiting time and the sojourn time of a group leader and of an arbitrary customer; the number of different family types in the system, the mean size of the served batch and of the batch in the kth position, the mean length of a busy period, and the number of groups being bypassed by a newly arriving customer.

2.2. Steady-State Probabilities, Queue Length, and Busy Period

Steady-state probabilities. Let π_n, $n \geq 0$, denote the equilibrium distribution of $L = \lim_{t\to\infty} L(t)$. That is, $\pi_n = P(L = n)$. Then, the balance equations of the model are given as follows:

$$\lambda\pi_0 = \mu\pi_1, \tag{2.1}$$

$$(\mu + \lambda(1-p)^n)\,\pi_n = \lambda(1-p)^{n-1}\pi_{n-1} + \mu\pi_{n+1}, \quad n \geq 1. \tag{2.2}$$

Eq. (2.2) together with $\sum_{n=0}^{\infty} \pi_n = 1$ gives

$$\pi_n = \pi_0 \left(\frac{\lambda}{\mu}\right)^n (1-p)^{\frac{n(n-1)}{2}}, \quad n \geq 1,$$

where π_0 is given by

$$\pi_0 = \left(\sum_{n=0}^{\infty} \left(\frac{\lambda}{\mu}\right)^n (1-p)^{\frac{n(n-1)}{2}}\right)^{-1}. \tag{2.3}$$

We note that the system is always stable for $0 < p \leq 1$, since the sum in Eq. (2.3) is finite. When $p = 0$ (regular $M/M/1$ queue) the stability condition is $\lambda < \mu$.

We also mention that the sum in Eq. (2.3) is a q-hypergeometric series ($q = 1 - p$) defined in the Appendix (see also [5]).

Probability generating function. Let $G(z) = \sum_{n=0}^{\infty} \pi_n z^n$ be the probability generating function (PGF) of L, the number of different group leaders in the system. Then, by multiplying Eq. (2.2) by z^n for all $n \geq 1$, summing with Eq. (2.1) and dividing by $(1-z)$

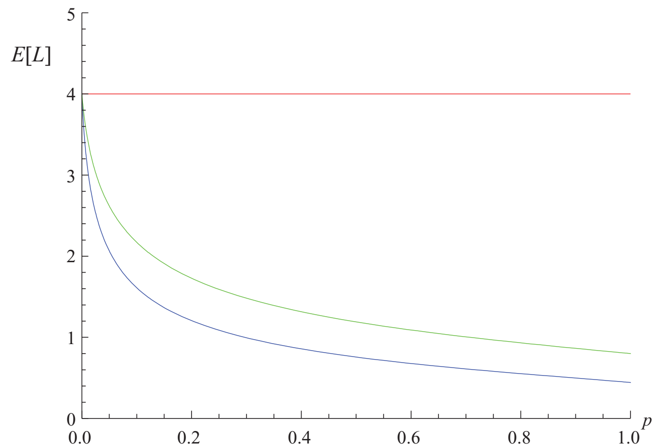

Figure 2. (Color online) Plots of $\mathbb{E}[L]$ as a function of p.

we obtain

$$G(z) = \frac{\lambda}{\mu} z G\left((1-p)z\right) + \pi_0. \tag{2.4}$$

Substituting $z = 1$ in (2.4) and using $G(1) = 1$ gives

$$G(1-p) = \frac{\mu}{\lambda}(1-\pi_0), \tag{2.5}$$

which is the probability that an arrival will lead to creating a new group in the system: $\sum_{n=0}^{\infty} \pi_n (1-p)^n = G(1-p)$. Furthermore, $\lambda_{\text{eff}} := \lambda G(1-p)$ is the average effective arrival rate (i.e. the rate at which new groups (families) are created), and $\mu(1-\pi_0)$ is the average rate of emptying the system. Eq. (2.5) then gives

$$\pi_0 = 1 - \frac{\lambda_{\text{eff}}}{\mu}. \tag{2.6}$$

Queue length. In order to calculate $\mathbb{E}[L]$, the expected total number of different groups in the system, we differentiate Eq. (2.4) and substitute $z = 1$. We obtain

$$\mathbb{E}[L] = \frac{\lambda}{\mu} G(1-p) + \frac{\lambda}{\mu}(1-p)G'(1-p). \tag{2.7}$$

Taking the limit $p \to 0$ in (2.7) leads to

$$\mathbb{E}[L]_{p=0} = \frac{\lambda}{\mu} + \frac{\lambda}{\mu}\mathbb{E}[L]_{p=0},$$

which gives

$$\mathbb{E}[L]_{p=0} = \frac{\lambda}{\mu - \lambda}, \tag{2.8}$$

which is the expected total number of customers in a regular $M/M/1$ queue.

For the case $p \to 1$, L can only be 0 or 1, with probabilities $\pi_0 = \mu/(\lambda+\mu)$ and $\pi_1 = \lambda/(\lambda+\mu)$, respectively. It follows that $\mathbb{E}[L]_{p=1} = (\lambda/\mu)\pi_0 = \lambda/(\lambda+\mu)$.

$\mathbb{E}[L]$ is plotted in Figure 2 as a function of p. The upper straight line depicts $\mathbb{E}[L]\,(=\lambda/(\mu-\lambda))$ of a regular $M/M/1$ queue with $\lambda=4$, $\mu=5$. The center line depicts $\mathbb{E}[L]$ for a regular $M/M/1$ queue with arrival rate $\lambda_{\text{eff}}=4G(1-p)$, and $\mu=5$. The bottom line depicts $\mathbb{E}[L]$ for the Israeli Queue with $\lambda=4$, $\mu=5$. It is readily seen, as expected, that the mean number of groups in the $M/M/1$-type Israeli Queue is considerably smaller than the mean number of individual customers in a regular $M/M/1$ queue. The same holds for waiting times.

Busy period. Let θ denote the busy period (the period of time during which the server is working continuously). Since the idle time of the server is $\text{Exp}(\lambda)$, we obtain

$$\frac{\mathbb{E}[\theta]}{(1/\lambda)+\mathbb{E}[\theta]}=1-\pi_0=\frac{\lambda G(1-p)}{\mu},$$

resulting in

$$\mathbb{E}[\theta]=\frac{G(1-p)}{\mu-\lambda G(1-p)}.$$

2.3. Sojourn Times and Waiting Times

Sojourn and waiting times of a group leader. Let W denote the total sojourn time of a group *leader* in the system and let $\widetilde{W}(\cdot)$ denote its Laplace–Stieltjes Transform (LST). Then,

$$\begin{aligned}\widetilde{W}(s)&=\mathbb{E}\left[e^{-sW}\right]=\mathbb{E}\left[\mathbb{E}\left[e^{-sW}|L\text{ in the system and a new arrival creates a new group}\right]\right]\\&=\frac{1}{G(1-p)}\sum_{n=0}^{\infty}\pi_n(1-p)^n\left(\frac{\mu}{\mu+s}\right)^{n+1}\\&=\frac{\mu}{\mu+s}\times\frac{1}{G(1-p)}\times G\left(\frac{\mu}{\mu+s}(1-p)\right).\end{aligned}\tag{2.9}$$

Note that we divide by $G(1-p)$, the probability that a new group is created. Furthermore,

$$\mathbb{E}[W]=-\widetilde{W}'(s)|_{s=0}=\frac{1}{\mu}\left(1+\frac{(1-p)G'(1-p)}{G(1-p)}\right),\tag{2.10}$$

and by using Little's Law we have

$$\mathbb{E}[L]=\lambda_{\text{eff}}\mathbb{E}[W]=\lambda G(1-p)\mathbb{E}[W],$$

which coincides with Eq. (2.7).

The waiting time of a group leader, W_q, is derived from $W=W_q+B$, where B denotes the service time. As a result of independence,

$$\widetilde{W}(s)=\widetilde{W}_q(s)\times\frac{\mu}{\mu+s}.$$

Consequently,

$$\widetilde{W}_q(s)=\frac{1}{G(1-p)}\times G\left(\frac{\mu}{\mu+s}(1-p)\right).\tag{2.11}$$

Sojourn and waiting times of an arbitrary customer. Define W^{a} as the total sojourn time of an *arbitrary* customer in the system, and $\widetilde{W}^{\text{a}}(\cdot)$ as its LST. Distinguishing between the events that a new arrival joins an existing group, and the event that he creates a new one, we write

$$\begin{aligned}
\widetilde{W}^{\mathrm{a}}(s) &= \mathbb{E}[e^{-sW^{\mathrm{a}}}] = \mathbb{E}[\mathbb{E}[e^{-sW^{\mathrm{a}}}|L]] = \sum_{n=0}^{\infty}\pi_n \mathbb{E}[e^{-sW^{\mathrm{a}}}|L=n] \\
&= \sum_{n=1}^{\infty}\pi_n p\frac{\mu}{\mu+s}\sum_{k=0}^{n-1}\left(\frac{\mu}{\mu+s}(1-p)\right)^k + \sum_{n=0}^{\infty}\pi_n(1-p)^n\left(\frac{\mu}{\mu+s}\right)^{n+1} \\
&= \sum_{n=1}^{\infty}\pi_n p\frac{\mu}{\mu+s}\sum_{k=0}^{n-1}\left(\frac{\mu}{\mu+s}(1-p)\right)^k + \sum_{n=0}^{\infty}\pi_n(1-p)^n\left(\frac{\mu}{\mu+s}\right)^{n+1} \\
&= \frac{\mu p}{\mu p+s}\left(1-G\left(\frac{\mu}{\mu+s}(1-p)\right)\right) + \frac{\mu}{\mu+s}G\left(\frac{\mu}{\mu+s}(1-p)\right),
\end{aligned}$$

and

$$\mathbb{E}[W^{\mathrm{a}}] = -\tilde{W}^{\mathrm{a}\prime}(s)|_{s=0} = \frac{1}{\mu p} - \frac{1-p}{\mu p}G(1-p) = \frac{1}{\mu p} - \frac{(1-p)(1-\pi_0)}{\lambda p}. \tag{2.12}$$

Clearly, as can be expected,

$$\lim_{p\to 1}\mathbb{E}[W^{\mathrm{a}}] = \frac{1}{\mu},$$

and by applying l'Hopital's rule we obtain

$$\lim_{p\to 0}\mathbb{E}[W^{\mathrm{a}}] = \frac{1}{\mu-\lambda},$$

which is the mean sojourn time of an arbitrary customer in a regular $M/M/1$ queue.

Define L^{total} to be the total number of customers in the system. Then, from Little's Law, we obtain

$$\begin{aligned}
\mathbb{E}[L^{\mathrm{total}}] = \lambda\mathbb{E}[W^{\mathrm{a}}] &= \frac{\lambda}{\mu p} - \frac{\lambda(1-p)G(1-p)}{\mu p} \\
&= \frac{\lambda}{\mu p} - \frac{(1-p)(1-\pi_0)}{p}.
\end{aligned} \tag{2.13}$$

Clearly, when $p\to 0$, $\mathbb{E}[L^{\mathrm{total}}] = (\lambda/(\mu-\lambda))$. Also, when $p\to 1$, only a single group is formed, and the process is reduced to a two-state process, where $\pi_0 = P(L=0) = (\mu/(\lambda+\mu))$ and $\pi_1 = P(L=1) = (\lambda/(\lambda+\mu))$. Hence, $\mathbb{E}[L^{\mathrm{total}}] = \pi_1(1+(\lambda/\mu)) = \lambda/\mu$. Alternatively, since the mean sojourn time of each customer is $1/\mu$, then by Little's Law, $\mathbb{E}[L^{\mathrm{total}}] = \lambda/\mu$.

Another important service measure is the waiting time of an arbitrary customer, denoted by W_q^a. Since $W^{\mathrm{a}} = W_q^{\mathrm{a}} + B$, from the expression for $\widetilde{W}^{\mathrm{a}}(s)$ we obtain

$$\widetilde{W_q^a}(s) = \frac{s(1-p)}{\mu p+s}\times G\left(\frac{\mu}{\mu+s}(1-p)\right) + \frac{p(\mu+s)}{\mu p+s}. \tag{2.14}$$

2.4. Size of a Batch

We first wish to calculate the mean size of the served batch. For that, we consider a group that was formed in the ith place ($i \geq 1$), and follow its progress up the queue. That is, we calculate the number of jobs joining this group during each service period, until this group completes its service and leaves the system. Let us define the following variables:

- $D^{(k)}$ = the size of the group standing in the kth place (for $k \geq 1$) at the moment of service completion, assuming that the kth group exists ($k = 1$ refers to the group that has just completed service).
- Y = the number of arrivals to the system during a single service duration.
- $\xi^{(k)}$ = the number of customers who joined the group in the kth place (for $k \geq 1$) among the Y arrivals, assuming that the kth group exists.

Then, from the above, $D^{(k)} \stackrel{d}{=} D^{(k+1)} + \xi^{(k)}$, for $k \geq 1$. Our goal is to calculate $\mathbb{E}\left[D^{(1)}\right]$.

First, note that for all $m \geq 0$ and $0 \leq j \leq m$,

$$\mathbb{P}(\xi^{(k)} = j | Y = m) = \binom{m}{j} \left((1-p)^{k-1}p\right)^j \left(1-(1-p)^{k-1}p\right)^{m-j}.$$

Second, for $k \geq 1$ and $j \geq 0$, and since $\mathbb{P}(Y = m) = (\lambda/(\lambda+\mu))^m (\mu/(\lambda+\mu))$ for $m \geq 0$, we have

$$\begin{aligned}\mathbb{P}(\xi^{(k)} = j) &= \sum_{m=j}^{\infty} \mathbb{P}(Y=m) \binom{m}{j} \left((1-p)^{k-1}p\right)^j \left(1-(1-p)^{k-1}p\right)^{m-j} \\ &= \frac{\mu}{\lambda(1-p)^{k-1}p+\mu} \left(\frac{\lambda(1-p)^{k-1}p}{\lambda(1-p)^{k-1}p+\mu}\right)^j. \qquad (2.15)\end{aligned}$$

In order to better understand the result given in (2.15), consider two Poisson processes competing with each other: the arriving process to the group standing in the kth position, with rate $\lambda(1-p)^{k-1}p$, and the departure process, with rate μ. Having exactly j jobs joining the kth group during a service duration means that the arriving process to the kth group "wins" j times before the end of a single service duration.

Consequently,

$$\mathbb{E}[\xi^{(k)}] = \frac{\lambda}{\mu}(1-p)^{k-1}p, \quad k \geq 1. \qquad (2.16)$$

Indeed, the mean number of arrivals during a service is $\mathbb{E}[Y] = \lambda/\mu$, while the probability of an arrival to join the kth group (when it exists) is $(1-p)^{k-1}p$.

Moreover, using Eq. (2.15), the PGF of $\xi^{(k)}$ is given by

$$\mathbb{E}[z^{\xi^{(k)}}] = \sum_{j=0}^{\infty} \mathbb{P}(\xi^{(k)} = j) z^j = \frac{\mu}{\lambda(1-p)^{k-1}p(1-z)+\mu}. \qquad (2.17)$$

Define by $D_i^{(1)}$ the size of the group that has just completed its service, given that this group was formed in the ith place. Then, $D^{(1)} = D_i^{(1)} = 1 + \sum_{k=1}^{i} \xi^{(i+1-k)}$ with probability $(\pi_{i-1}(1-p)^{i-1})/(\sum_{i=1}^{\infty} \pi_{i-1}(1-p)^{i-1})$, or equivalently, with probability

$(\pi_{i-1}(1-p)^{i-1})/(G(1-p))$. Using this fact, and Eq. (2.16), yields,

$$\begin{aligned}\mathbb{E}[D^{(1)}] &= \sum_{i=1}^{\infty}\mathbb{E}[D_i^{(1)}]\frac{\pi_{i-1}(1-p)^{i-1}}{G(1-p)} = \sum_{i=1}^{\infty}\left(1+\sum_{k=1}^{i}\mathbb{E}[\xi^{(i+1-k)}]\right)\frac{\pi_{i-1}(1-p)^{i-1}}{G(1-p)}\\ &= 1+\sum_{i=1}^{\infty}\frac{\pi_{i-1}(1-p)^{i-1}}{G(1-p)}\sum_{k=1}^{i}\frac{\lambda}{\mu}(1-p)^{k-1}p\\ &= 1+\frac{\lambda}{\mu}\sum_{i=1}^{\infty}\frac{\pi_{i-1}(1-p)^{i-1}}{G(1-p)}\left(1-(1-p)^i\right)\\ &= 1+\frac{\lambda}{\mu}-\frac{\lambda}{\mu}\left(\frac{(1-p)G\left((1-p)^2\right)}{G(1-p)}\right). \end{aligned} \tag{2.18}$$

In order to find $G\left((1-p)^2\right)$, we substitute $z=1-p$ in Eq. (2.4), which gives

$$G\left((1-p)^2\right) = \frac{\mu(\mu(1-\pi_0)-\lambda\pi_0)}{\lambda^2(1-p)}, \tag{2.19}$$

and together with (2.5) we obtain

$$\mathbb{E}[D^{(1)}] = \frac{\lambda}{\mu(1-\pi_0)} = \frac{1}{G(1-p)}. \tag{2.20}$$

The PGF of $D^{(1)}$, is given by

$$\begin{aligned}\mathbb{E}[z^{D^{(1)}}] &= \sum_{i=1}^{\infty}\mathbb{E}[z^{D_i^{(1)}}]\frac{\pi_{i-1}(1-p)^{i-1}}{G(1-p)}\\ &= \sum_{i=1}^{\infty}\mathbb{E}[z^{1+\sum_{k=1}^{i}\xi^{(i+1-k)}}]\frac{\pi_{i-1}(1-p)^{i-1}}{G(1-p)}\\ &= \sum_{i=1}^{\infty} z\frac{\pi_{i-1}(1-p)^{i-1}}{G(1-p)}\prod_{k=1}^{i}\frac{\mu}{\lambda(1-p)^{k-1}p(1-z)+\mu}, \end{aligned} \tag{2.21}$$

where in the last equality we utilize Eq. (2.17), and the fact that the variables $\xi^{(i+1-k)}$, for $k=1,\ldots,i$, are independent since they are generated in distinct service periods.

In the same manner, let us define $D_i^{(k)}$ to be the size of the batch standing in the kth place at the moment of service completion, given that it was formed in the ith place, for $i \ge k \ge 1$. Then, with probability $(\pi_{i-1}(1-p)^{i-1})/(\sum_{j=k}^{\infty}\pi_{j-1}(1-p)^{j-1})$, $D^{(k)} = D_i^{(k)} = 1+\sum_{m=1}^{i-k+1}\xi^{(i+1-m)}$. Using the fact that $\pi_{i-1}(1-p)^{i-1} = \pi_i(\mu/\lambda)$ we obtain

$$\begin{aligned}\mathbb{E}[D^{(k)}] &= \sum_{i=k}^{\infty}\mathbb{E}[D_i^{(k)}]\frac{\pi_{i-1}(1-p)^{i-1}}{\sum_{j=k}^{\infty}\pi_{j-1}(1-p)^{j-1}}\\ &= \sum_{i=k}^{\infty}\left(1+\sum_{m=1}^{i-k+1}\mathbb{E}[\xi^{(i+1-m)}]\right)\frac{\pi_{i-1}(1-p)^{i-1}}{\sum_{j=k}^{\infty}\pi_{j-1}(1-p)^{j-1}}\\ &= 1+\sum_{i=k}^{\infty}\frac{\pi_i}{\sum_{j=k}^{\infty}\pi_j}\sum_{m=1}^{i-k+1}\frac{\lambda}{\mu}(1-p)^{i-m}p \end{aligned}$$

$$= 1 + \frac{\lambda}{\mu} \sum_{i=k}^{\infty} \frac{\pi_i}{\sum_{j=k}^{\infty} \pi_j} \left((1-p)^{k-1} - (1-p)^i \right)$$

$$= 1 + \frac{\lambda}{\mu}(1-p)^{k-1} - \frac{\sum_{i=k}^{\infty} \pi_{i+1}}{\sum_{j=k}^{\infty} \pi_j} = \frac{\lambda}{\mu}(1-p)^{k-1} + \frac{\pi_k}{1 - \sum_{j=0}^{k-1} \pi_j}. \tag{2.22}$$

Note that, since $\lambda(1-p)^{k-1}\pi_{k-1} = \mu\pi_k$, Eq. (2.22) can be written as

$$\mathbb{E}[D^{(k)}] = \frac{\pi_k / \sum_{j=k}^{\infty} \pi_j}{\pi_{k-1} / \sum_{j=k-1}^{\infty} \pi_j} = \frac{\mathbb{P}(L = k | L \geq k)}{\mathbb{P}(L = k-1 | L \geq k-1)}. \tag{2.23}$$

Let $\hat{L}^{\text{total}}$ denote the total number of customers in the system at instant of service completion. Then,

$$\mathbb{E}[\hat{L}^{\text{total}}] = \sum_{j=1}^{\infty} \pi_j \sum_{k=1}^{j} \mathbb{E}[D^{(k)}] = \sum_{j=1}^{\infty} \pi_j \sum_{k=1}^{j} \left(\frac{\lambda}{\mu}(1-p)^{k-1} + \frac{\pi_k}{1 - \sum_{j=0}^{k-1} \pi_j} \right)$$

$$= \frac{\lambda}{\mu} \sum_{j=1}^{\infty} \pi_j \frac{1-(1-p)^j}{p} + \sum_{j=1}^{\infty} \pi_j \sum_{k=1}^{j} \frac{\pi_k}{1 - \sum_{j=0}^{k-1} \pi_j}$$

$$= \frac{\lambda}{\mu} \cdot \frac{1 - G(1-p)}{p} + \sum_{k=1}^{\infty} \pi_k \sum_{j=k}^{\infty} \frac{\pi_j}{\sum_{i=k}^{\infty} \pi_i}$$

$$= \frac{\lambda}{\mu} \cdot \frac{1 - G(1-p)}{p} + \sum_{k=1}^{\infty} \pi_k = \frac{\lambda(1 - G(1-p))}{\mu p} + 1 - \pi_0. \tag{2.24}$$

Substituting $1 - \pi_0 = (\lambda/\mu)G(1-p)$ in (2.24) we conclude that $\mathbb{E}[\hat{L}^{\text{total}}] = \mathbb{E}[L^{\text{total}}]$ (see Eq. (2.13)).

2.5. Number of Bypassed Groups

Let X denote the index of the group to which a new arrival joins, and $\hat{X}(z)$ its PGF. The distribution of X is given by

$$\mathbb{P}(X = k) = \pi_{k-1}(1-p)^{k-1} + \sum_{n=k}^{\infty} \pi_n (1-p)^{k-1} p, \ k \geq 1. \tag{2.25}$$

This follows since if there are $k-1$ groups, the new job creates a new group in the kth position with probability $(1-p)^{k-1}$, and if there are at least k different families (groups), the new job joins the kth group with probability $(1-p)^{k-1}p$. Hence,

$$\hat{X}(z) = \sum_{k=1}^{\infty} z^k \left(\pi_{k-1}(1-p)^{k-1} + (1-p)^{k-1} p \sum_{n=k}^{\infty} \pi_n \right)$$

$$= zG\left((1-p)z\right) + \sum_{n=1}^{\infty} \pi_n p z \sum_{k=0}^{n-1} (z(1-p))^k$$

$$= zG\left((1-p)z\right) + \sum_{n=1}^{\infty} \pi_n pz \left(\frac{1-(z(1-p))^n}{1-z(1-p)}\right)$$

$$= zG\left((1-p)z\right) + \frac{pz}{1-z(1-p)}(1-G\left((1-p)z\right)). \tag{2.26}$$

From the PGF of X we obtain

$$\mathbb{E}[X] = \hat{X}'(z)|_{z=1} = G(1-p) + \frac{1}{p}\left(1-G(1-p)\right)$$

$$= \frac{1}{p}\left(1-(1-p)G(1-p)\right). \tag{2.27}$$

Note that, multiplying $\mathbb{E}[X]$ by the mean service time, $1/\mu$, gives the mean total sojourn time of an arbitrary customer in the system, which coincides with Eq. (2.12).

We also define the random variable N_B as the number of groups that are being bypassed by an arriving customer. That is, if $L = n$ and the new arrival joins group $k \le n$, then he bypasses $N_B = n - k$ groups. By conditioning on the number of different families present in the system, we obtain that the probability distribution of N_B is,

$$\mathbb{P}(N_B = 0) = \sum_{n=1}^{\infty} \pi_n p(1-p)^{n-1} + \sum_{n=0}^{\infty} \pi_n (1-p)^n = \frac{p}{1-p}\left(G(1-p) - \pi_0\right) + G(1-p),$$

$$\mathbb{P}(N_B = k) = \sum_{n=k+1}^{\infty} \pi_n p(1-p)^{n-k-1}, \ k \ge 1. \tag{2.28}$$

From Eq. (2.28) we obtain

$$\mathbb{E}[N_B] = \sum_{k=0}^{\infty} k \sum_{n=k+1}^{\infty} \pi_n p(1-p)^{n-k-1} = \sum_{n=1}^{\infty} \pi_n \sum_{k=0}^{n-1} kp(1-p)^{n-k-1}$$

$$= \sum_{n=1}^{\infty} \pi_n \left(n - \frac{1}{p} + \frac{(1-p)^n}{p}\right) = \mathbb{E}[L] - \frac{1}{p}\left(1 - G(1-p)\right). \tag{2.29}$$

The PGF of N_B is obtained as follows:

$$\mathbb{E}[z^{N_B}] \equiv \hat{N}_B(z) = G(1-p) + \sum_{k=0}^{\infty} z^k \sum_{n=k+1}^{\infty} \pi_n p(1-p)^{n-k-1}$$

$$= G(1-p) + p\sum_{n=1}^{\infty} \pi_n (1-p)^{n-1} \sum_{k=0}^{n-1} \left(\frac{z}{1-p}\right)^k$$

$$= G(1-p) + \frac{p}{1-p-z}\left(\sum_{n=1}^{\infty} \pi_n (1-p)^n - \sum_{n=1}^{\infty} \pi_n z^n\right)$$

$$= G(1-p) + \frac{p}{1-p-z}\left(G(1-p) - \pi_0 - G(z) + \pi_0\right)$$

$$= \frac{1-z}{1-p-z}G(1-p) - \frac{p}{1-p-z}G(z). \tag{2.30}$$

Indeed, differentiating $\hat{N}_B(z)$ and setting $z = 1$ yields result (2.29).

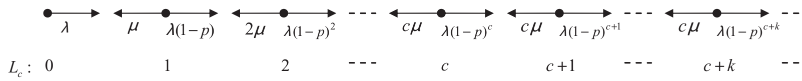

FIGURE 3. Transition rate diagram of the process $\{L_c(t), t \geq 0\}$.

Another way to calculate $\mathbb{E}[N_B]$ is the following. Suppose there are n different groups in the system, including the one in service. Then, the expected number of families being bypassed by an arriving customer is

$$\sum_{k=1}^{n}(n-k)p(1-p)^{k-1} = n - \frac{1}{p} + \frac{(1-p)^n}{p}.$$

By conditioning over all values of n we obtain (2.29).

3. THE $M/M/C$ AND $M/M/1/N$ MODELS

3.1. The $M/M/c$—Type Model

This case is similar to the $M/M/1$ model, where the state space is defined by $\{L_c(t), t \geq 0\}$, and its transition rate diagram is depicted in Figure 3.

Steady-state probabilities and generating function. With $L_c = \lim_{t\to\infty} L_c(t)$ we let $\pi_n = P(L_c = n)$, for $n \geq 0$. It is readily obtained that:

$$\pi_n = \left(\frac{\lambda}{\mu}\right)^n \frac{1}{n!}(1-p)^{\frac{n(n-1)}{2}}\pi_0, \quad 1 \leq n \leq c-1, \tag{3.1}$$

$$\pi_{c+k} = \left(\frac{\lambda}{c\mu}\right)^k \left(\frac{\lambda}{\mu}\right)^c \frac{1}{c!}(1-p)^{\frac{(c+k)(c+k-1)}{2}}\pi_0,\ k \geq 0. \tag{3.2}$$

where

$$\pi_0 = \left(\sum_{n=0}^{c-1}\left(\frac{\lambda}{\mu}\right)^n \frac{1}{n!}(1-p)^{\frac{n(n-1)}{2}} + \left(\frac{\lambda}{\mu}\right)^c \frac{1}{c!}\sum_{k=0}^{\infty}\left(\frac{\lambda}{c\mu}\right)^k (1-p)^{\frac{(c+k)(c+k-1)}{2}}\right)^{-1}. \tag{3.3}$$

The corresponding PGF is given by

$$G_c(z) = \left[\sum_{n=0}^{c-1}\left(\frac{\lambda z}{\mu}\right)^n \frac{1}{n!}(1-p)^{\frac{n(n-1)}{2}} + \sum_{k=0}^{\infty}\left(\frac{\lambda z}{c\mu}\right)^k \left(\frac{\lambda z}{\mu}\right)^c \frac{1}{c!}(1-p)^{\frac{(c+k)(c+k-1)}{2}}\right]\pi_0. \tag{3.4}$$

Apparently there is no closed form expression for $G_c(z)$. However, it satisfies the following functional equation,

$$G_c(z) = \frac{\lambda}{c\mu}zG_c\left((1-p)z\right) - \frac{z}{c}H'(z) + H(z), \tag{3.5}$$

where $H(z) = \sum_{n=0}^{c-1}\pi_n z^n$. Substituting $z = 1$ in Eq. (3.5) gives

$$\lambda G_c(1-p) = \sum_{n=1}^{c-1} n\mu\pi_n + c\mu\sum_{n=c}^{\infty}\pi_n. \tag{3.6}$$

Clearly, when $c = 1$, Eq. (3.6) coincides with Eq. (2.5), and its explanation is similar.

In addition,

$$\mathbb{E}[L_c] = G_c'(z)|_{z=1} = \frac{\lambda}{c\mu}(G_c(1-p) + (1-p)G_c'(1-p)) + H'(1) - \frac{H'(1) + H''(1)}{c} \quad \textbf{(3.7)}$$

Sojourn and waiting times. Let W_c and W_c^{a} denote the total time a group leader and an arbitrary customer spends in the system, with LST $\widetilde{W}_c(s)$ and $\widetilde{W}_c^{\mathrm{a}}(s)$, respectively. Using a similar approach as presented in Section 2, we obtain:

$$\begin{aligned}\widetilde{W}_c(s) &= \frac{1}{G_c(1-p)} \cdot \frac{\mu}{\mu+s} \\ &\quad \times \left[H(1-p) + \left(\frac{c\mu}{c\mu+s}\right)^{1-c} \left(G_c\left(\frac{(1-p)c\mu}{c\mu+s}\right) - H\left(\frac{(1-p)c\mu}{c\mu+s}\right)\right)\right], \end{aligned} \quad \textbf{(3.8)}$$

and

$$\begin{aligned}\widetilde{W}_c^{\mathrm{a}}(s) &= \frac{\mu}{\mu+s}\left[1 - \frac{(1-p)^c(1-H(1))s}{cp\mu+s}\right] + \frac{\mu}{\mu+s} \\ &\quad \times \left[G_c\left(\frac{(1-p)c\mu}{c\mu+s}\right) - H\left(\frac{(1-p)c\mu}{c\mu+s}\right)\right]\left[\left(\frac{c\mu}{c\mu+s}\right)^{1-c} - \frac{cp\mu}{cp\mu+s}\left(1+\frac{s}{c\mu}\right)^c\right]. \end{aligned} \quad \textbf{(3.9)}$$

3.2. The $M/M/1/N$-Type Model

Again, this case is similar to the previous models, but the state space $\{L^N(t), t \geq 0\}$ is finite. The only modification is that whenever a new customer finds the system in state $L = N$ he joins the Nth (last) group even if he does not know any of the existing group leaders. With $L^{(N)} = \lim_{t\to\infty} L^{(N)}(t)$, the stationary distribution $\pi_n = P(L^{(N)} = n)$ is given by

$$\pi_n = \left(\frac{\lambda}{\mu}\right)^n (1-p)^{\frac{n(n-1)}{2}} \pi_0, \quad 1 \leq n \leq N, \quad \textbf{(3.10)}$$

$$\pi_0 = \left(\sum_{n=0}^{N} \left(\frac{\lambda}{\mu}\right)^n (1-p)^{\frac{n(n-1)}{2}}\right)^{-1}. \quad \textbf{(3.11)}$$

The corresponding PGF satisfies the relation

$$G^{(N)}(z) = \frac{\lambda}{\mu} z \left(G^{(N)}((1-p)z) - (1-p)^N \pi_N z^N\right) + \pi_0, \quad \textbf{(3.12)}$$

and the mean batch size right after the moment of service completion is

$$\mathbb{E}[D^{(k)}] = \frac{\lambda}{\mu}(1-p)^{k-1} + \frac{\pi_k}{\sum_{j=k}^{N} \pi_j}, \quad k = 1, 2, \ldots, N. \quad \textbf{(3.13)}$$

4. A PRIORITY MODEL (SINGLE SERVER)

Assume that there is a single server and that the population of customers is comprised of two priority classes: a VIP (class 1), and an ordinary (class 2). There are M different families among the VIP class, and a class-1 customer, upon arrival, searches for a friend among the VIP groups present in a manner similar to previous sections. The size of each of the (possibly) M VIP groups is unlimited. The customers of class 2 form a single regular $M/M/1$-type queue (they do not search for a friend). The number of low priority customers in the system is unbounded. A preemptive priority rule is considered. That is, a higher priority customer, upon arrival, takes over the service if a lower priority customer is currently being served, or looks for a friend among the higher priority group leaders. Hence, if the number of high priority groups in the system is k, where $1 \le k \le M-1$, then he will either join the ith group ($i \le k$), or create a new group, in front of all the low priority customers (if present). If M high-priority groups are present, then an arriving class 1 customer will join the Mth group if he does not know any of the first $M-1$ group leaders (with probability $(1-p)^{M-1}$).

We define $L_1(t)$ as the number of class-1 groups in the system at time t, and by $L_2(t)$ as the total number of class-2 customers in the system at time t. Let $L_i = \lim_{t\to\infty} L_i(t)$, and $\pi_{m,n} = P(L_1 = m, L_2 = n)$, for $0 \le m \le M$ and $n \ge 0$. A transition rate diagram of the two-dimensional process (L_1, L_2) is given in Figure 4. Define the marginal probabilities,

$$\pi_{m\bullet} = \sum_{n=0}^{\infty} \pi_{m,n}, \quad 0 \le m \le M,$$

$$\pi_{\bullet n} = \sum_{m=0}^{M} \pi_{m,n}, \quad n \ge 0.$$

Considering the marginal probabilities $\pi_{m\bullet}$ and utilizing horizontal cuts in Figure 4, it is readily seen that $\pi_{m\bullet}$ is given by Eqs. (3.10) and (3.11), where m replaces n (and M replaces N), λ_1 and μ_1 replaces λ and μ, respectively.

Since the number of high-priority classes is bounded, and the service is in (unlimited) batches, it immediately follows that the stability condition is

$$\frac{\lambda_2}{\mu_2} < \pi_{0\bullet}, \tag{4.1}$$

as the queue of the low-priority customers empties out only when there are no high priority customers in the system.

4.1. Balance Equations and Generating Functions

From Figure 4, we have for all $n \ge 0$,

$$\lambda_2 \pi_{\bullet n} = \mu_2 \pi_{0,n+1}.$$

Summing over n gives

$$\pi_{0,0} = \pi_{0\bullet} - \frac{\lambda_2}{\mu_2}. \tag{4.2}$$

Now, for $m = 0$, the following relations hold,

$$(\lambda_1 + \lambda_2)\,\pi_{0,0} = \mu_1 \pi_{1,0} + \mu_2 \pi_{0,1}, \tag{4.3}$$

$$(\lambda_1 + \lambda_2 + \mu_2)\,\pi_{0,n} = \lambda_2 \pi_{0,n-1} + \mu_1 \pi_{1,n} + \mu_2 \pi_{0,n+1}, \quad n \ge 1. \tag{4.4}$$

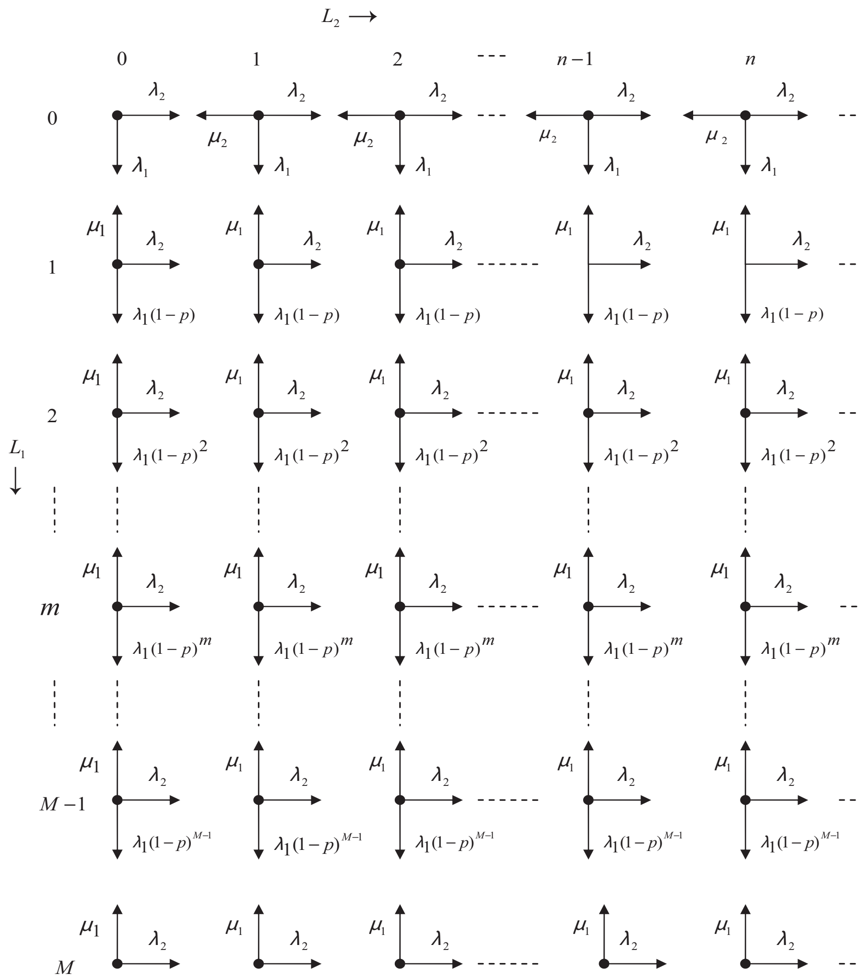

Figure 4. Transition rate diagram of (L_1, L_2) for priority model.

Define the marginal PGF, $G_m(z) = \sum_{n=0}^{\infty} \pi_{m,n} z^n$, for all $0 \leq m \leq M$. Then, multiplying Eq. (4.4) by z^n and summing over n together with (4.3) gives

$$(\lambda_1 z + \lambda_2 z(1-z) - \mu_2(1-z))\, G_0(z) - \mu_1 z G_1(z) = -\mu_2 \pi_{0,0}(1-z). \tag{4.5}$$

Moreover, for $1 \leq m \leq M-1$ we obtain

$$(\lambda_1(1-p)^m + \lambda_2 + \mu_1)\, \pi_{m,0} = \lambda_1(1-p)^{m-1}\pi_{m-1,0} + \mu_1 \pi_{m+1,0}, \tag{4.6}$$

$$(\lambda_1(1-p)^m + \lambda_2 + \mu_1)\, \pi_{m,n} = \lambda_1(1-p)^{m-1}\pi_{m-1,n} + \lambda_2 \pi_{m,n-1} + \mu_1 \pi_{m+1,n}, \quad n \geq 1. \tag{4.7}$$

Multiplying Eq. (4.7) by z^n and summing over n together with (4.6) leads to

$$(\lambda_1(1-p)^m + \lambda_2(1-z) + \mu_1)\, G_m(z) - \lambda_1(1-p)^{m-1} G_{m-1}(z) - \mu_1 G_{m+1}(z) = 0,$$
$$1 \leq m \leq M-1. \tag{4.8}$$

Last, for $m = M$ we have

$$(\lambda_2 + \mu_1)\,\pi_{M,0} = \lambda_1(1-p)^{M-1}\pi_{M-1,0}, \tag{4.9}$$

$$(\lambda_2 + \mu_1)\,\pi_{M,n} = \lambda_1(1-p)^{m-1}\pi_{M-1,n} + \lambda_2\pi_{M,n-1}, \quad n \geq 1. \tag{4.10}$$

Multiplying Eq. (4.10) by z^n and summing over n together with (4.9) yields

$$(\lambda_2(1-z) + \mu_1)\,G_M(z) = \lambda_1(1-p)^{m-1}G_{M-1}(z). \tag{4.11}$$

The set of Eqs. (4.5), (4.8), and (4.11) can be written as

$$A(z) \cdot \underline{G}(z) = \underline{\Pi}(z), \tag{4.12}$$

where $A(z) = [a_{i,j}]_{1 \leq i,j \leq M+1}$ is an $(M+1) \times (M+1)$ tridiagonal matrix with the following entries

$$a_{i,j} = \begin{cases} \lambda_1 z + \lambda_2(1-z) - \mu_2(1-z) & i = 1,\ j = 1 \\ -\mu_1 z & i = 1,\ j = 2 \\ -\lambda_1(1-p)^{i-2} & 2 \leq i \leq M+1,\ j = i-1 \\ \lambda_1(1-p)^{i-1} + \lambda_2(1-z) + \mu_1 & 2 \leq i \leq M,\ j = i \\ -\mu_1 & 2 \leq i \leq M,\ j = i+1 \\ \lambda_2(1-z) + \mu_1 & i = M+1,\ j = M+1 \\ 0 & \text{elsewhere} \end{cases} \tag{4.13}$$

$\underline{G}(z) = (G_0(z), G_1(z), \ldots, G_M(z))^T$ is a column vector (of size $M+1$) of the desired PGF's, and $\underline{\Pi}(z)$ is the column vector $\underline{\Pi}(z) = \left(-\mu_2\pi_{0,0}(1-z), \underbrace{0, \ldots, 0}_{M\text{times}}\right)^T$. From the structure of $\underline{\Pi}(z)$ we have that $\underline{G}(z)$ is equal to the product of the first column of the matrix $(A(z))^{-1}$ and $-\mu_2\pi_{0,0}(1-z)$. Once the PGFs are obtained, the mean total number of class-2 customers in the system, $\mathbb{E}[L_2]$, can be derived from

$$\mathbb{E}[L_2] = \sum_{m=0}^{M} G'_m(1).$$

A numerical example:
Consider the following set of parameters: $\lambda_1 = 2$, $\mu_1 = 2$, $\lambda_2 = 1$, $\mu_2 = 4$, $p = 0.3$ and $M = 3$. Calculations of the boundary probabilities give:

$$\pi_{0\bullet} = 0.3286,$$

$$\pi_{1\bullet} = 0.3286,$$

$$\pi_{2\bullet} = 0.2301,$$

$$\pi_{3\bullet} = 0.1127,$$

$$\pi_{0,0} = \pi_{0\bullet} - \frac{\lambda_2}{\mu_2} = 0.0786.$$

The PGFs are given by:

$$
\begin{aligned}
G_0(z) &= \frac{0.8(z-6.1749)(z-3.61361)(z-1.59149)}{(z-5.8578)(z-3.17745)(z-1.14476)}, \\
G_1(z) &= \frac{-1.6(z-4.97327)(z-2.00673)}{(z-5.8578)(z-3.17745)(z-1.14476)}, \\
G_2(z) &= \frac{2.24(z-3)}{(z-5.8578)(z-3.17745)(z-1.14476)}, \\
G_3(z) &= \frac{-2.1952}{(z-5.8578)(z-3.17745)(z-1.14476)}.
\end{aligned}
$$

The mean number of class-1 groups and class-2 customers are $\mathbb{E}[L_1] = 1.1269$ and $\mathbb{E}[L_2] = 5.2933$, respectively.

4.2. Matrix Geometric Method

An alternative approach to describe the model presented in this section is by constructing a finite Quasi Birth and Death (QBD) process, with $M+1$ phases and infinite number of levels. State (n, m) indicates that there are m different class-1 groups and n class-2 customers in the system, $n \geq 0$, $0 \leq m \leq M$. The infinitesimal generator of the QBD is denoted by Q, and is given by

$$
Q = \begin{pmatrix}
B & A_0 & \mathbf{0} & \mathbf{0} & \cdots \\
A_2 & A_1 & A_0 & \mathbf{0} & \cdots \\
\mathbf{0} & A_2 & A_1 & A_0 & \cdots \\
\vdots & \vdots & \vdots & \vdots &
\end{pmatrix},
$$

where B, A_0, A_1, and A_2 are all square matrices of order $M+1$, as follows:
$A_0 = \lambda_2 \mathbf{I}$,

$$
A_1 = \begin{pmatrix}
-(\lambda_1+\lambda_2+\mu_2) & \lambda_1 & 0 & \cdots & \cdots & 0 \\
\mu_1 & -(\lambda_1(1-p)+\lambda_2+\mu_1) & \lambda_1(1-p) & 0 & \cdots & \vdots \\
0 & \mu_1 & -(\lambda_1(1-p)^2+\lambda_2+\mu_1) & \lambda_1(1-p)^2 & \ddots & \vdots \\
\vdots & 0 & \ddots & \ddots & \ddots & 0 \\
\vdots & \ddots & \ddots & \ddots & -(\lambda_1(1-p)^{M-1}+\lambda_2+\mu_1) & \lambda_1(1-p)^{M-1} \\
0 & \cdots & 0 & 0 & \mu_1 & -(\lambda_2+\mu_1)
\end{pmatrix},
$$

$$A_2 = \begin{pmatrix} \mu_2 & 0 & 0 & \dots & 0 \\ 0 & 0 & 0 & \dots & 0 \\ \vdots & & & & \\ 0 & 0 & 0 & \dots & 0 \end{pmatrix}.$$

and $B = A_1 + A_2$.

The elements of A_1 correspond to transitions within a given level, while the elements of A_0 and A_2 correspond to transitions from level n to level $n+1$ and to level $n-1$, respectively. Furthermore, the matrix $A = A_0 + A_1 + A_2$ is the infinitesimal generator of the $M/M/1/N$-type model described in Section 3.2. Let $\vec{\pi} = (\pi_0, \pi_1, \dots, \pi_M)$ be a stationary vector of the irreducible matrix A, that is, $\vec{\pi}A = 0$ and $\vec{\pi} \cdot \vec{e} = 1$. Note that $\pi_m = \pi_{m,\bullet}$ for all $0 \le m \le M$, where $\pi_{m\bullet}$ is given by Eqs. (3.10) and (3.11), where m replaces n.

The stability condition is ([9], page 83)

$$\vec{\pi} A_0 \vec{e} < \vec{\pi} A_2 \vec{e},$$

which immediately translates to

$$\lambda_2 < \mu_2 \pi_0.$$

This coincides with the condition given by Eq. (4.1).

Define for all $n \ge 0$ the steady-state probability vector $\vec{P_n} = (\pi_{0,n}, \pi_{1,n}, \dots, \pi_{M,n})$. Then, a well known result is (see Theorem 3.1.1 in [9]) that

$$\vec{P_n} = \vec{P_1} R^{n-1}, \quad n \ge 1,$$

where R is the minimal non-negative solution of the matrix quadratic equation

$$A_0 + RA_1 + R^2 A_2 = 0.$$

The vectors $\vec{P_0}$ and $\vec{P_1}$ are derived by solving the following linear system,

$$\begin{aligned} &\vec{P_0} B + \vec{P_1} A_2 = \vec{0}, \\ &\vec{P_0} A_0 + \vec{P_1}(A_1 + RA_2) = \vec{0}, \\ &\vec{P_0} \cdot \vec{e} + \vec{P_1}[I - R]^{-1} \cdot \vec{e} = 1. \end{aligned} \tag{4.14}$$

The derivation of the matrix R is based on Theorem 8.5.1 in [7], which states that if the QBD is recurrent and $A_2 = c \cdot r$, where c is a column vector and r is a row vector normalized by $r \cdot \vec{e} = 1$, than the matrix R is explicitly given by

$$R = A_0 \left(-A_1 - A_0 \vec{e} \cdot r\right)^{-1}. \tag{4.15}$$

Indeed, in our case, the matrix A_2 may be represented as

$$A_2 = \begin{pmatrix} \mu_2 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & \cdots & 0 \end{pmatrix} = c \cdot r.$$

The expected total number of class-2 customers in the system is given by

$$\mathbb{E}[L_2] = \sum_{n=1}^{\infty} n \vec{P_n} \cdot \vec{e} = \sum_{n=1}^{\infty} n \vec{P_1} R^{n-1} \cdot \vec{e} = \vec{P_1} \left[I - R\right]^{-2} \cdot \vec{e} \tag{4.16}$$

5. CONCLUSION

In this paper we have studied the so called Israeli Queue, which originated from a polling model where the next queue to be served is the one having the most senior customer. We analyzed the $M/M/1$, $M/M/c$, and $M/M/1/N$ - type models, in which an arriving customer searches for a friend standing in line. If he finds a friend in the queue, he joins his group and together they receive service in a batch mode. We derived various performance measures, such as the waiting time and sojourn time of a group leader, and of an arbitrary customer; the number of different family types in the system, the mean size of the served batch and of the batch in the kth position; and the number of groups being bypassed by a newly arriving customer. In addition, a priority model has been studied, where the VIP (high priority) customers form the Israeli Queue, while the low priority customers form a single regular queue. We analyzed this model using both PGF and matrix geometric methods.

References

1. Ammar, M.H. & Wong, J.W. (1987). On the optimality o cyclic transmission in teletext systems. *IEEE Transactions on Communications COM* 35(1): 68–73.
2. Boon, M.M.A, Van der Mei, R.D & Winands, E.M.M. (2011). Applications of polling systems. *Surveys in Operations Research and Management Science* 16(2): 67–82.
3. Boxma, O.J., Van der Wal, Y. & Yechiali, U. (2008). Polling with batch service. *Stochastic Models* 24(4): 604–625.
4. Dykeman, H.D., Ammar, M.H. & Wong, J.W. (1986). Scheduling algorithms for videotex systems under broadcast delivery. In Proceedings of the International Conference on Communications (ICC'86), pp. 1847–1851.
5. Gasper, G. & Rahman, M. (2004). *Basic hypergeometric series*, 2nd ed., Cambridge: Cambridge University Press.
6. He, Q.-M. & Chavoushi, A.A. (2013). Analysis of queueing systems with customer interjections. *Queueing Systems* 73: 79–104.
7. Latouche, G. & Ramaswami, V. (1999). *Introduction to matrix analytic methods in stochastic modeling.* ASA-SIAM Series on Statistics and Applied Probability. Philadelphia: SIAM.
8. Liu, Z. & Nain, P. (1992). Optimal scheduling in some multiqueue single-server systems. *IEEE Transactions on Automatic Control* 37(2): 247–252.
9. Neuts, M.F. (1981). *Matrix-geometric solutions in stochastic models. An algorithmic approach.* Baltimore: The Johns Hopkins University Press.
10. Takagi, H. (1986). *Analysis of polling systems.* Cambridge: The MIT Press.
11. Van Oyen, M.P. & Teneketzis, D. (1996). Optimal batch service of a polling system under partial informtion. *Methods and Models in OR* 44(3): 401–419.
12. Van der Wal, J. & Yechiali, U. (2003). Dynamic visit-order rules for batch-service polling. *Probability in the Engineering and Informational Sciences* 17(3): 351–367.
13. Yechiali, U. (1993). *Analysis and control of polling systems.* Performance Evaluation of Computer and Communication Systems. Heidelberg: Springer-Verlag, 630–650.

APPENDIX

The q-hypergeometric series is defined as

$$ {}_r\phi_s(a_1, a_2, \ldots, a_r; b_1, b_2, \ldots, b_s; q, x) = \sum_{n=0}^{\infty} \frac{(a_1, a_2, \ldots, a_r; q)_n}{(q;q)_n (b_1, b_2, \ldots, b_s; q)_n} [(-1)^n q^{\binom{n}{2}}]^{1+s-r} x^n, $$

where r and s are non-negative integers, $(a_1, a_2, \ldots, a_r; q)_n = (a_1;q)_n (a_2;q)_n \cdots (a_r;q)_n$, $(a;q)_0 = 1$, and $(a;q)_n = \prod_{k=0}^{n-1} \left(1 - aq^k\right)$ for $n \geq 1$.

With the above notation, Eq. (2.3) for π_0 can be expressed as

$$\pi_0 = \left({}_1\phi_1 \left(1-p; 0; 1-p, -\frac{\lambda}{\mu} \right) \right)^{-1}.$$

Since $\pi_j = \pi_0 \left(\frac{\lambda}{\mu}\right)^j (1-p)^{\binom{n}{2}}$, $G(z)$ can be written as

$$G(z) = \sum_{n=0}^{\infty} \pi_n z^n = \frac{{}_1\phi_1(1-p; 0; 1-p, -\frac{\lambda z}{\mu})}{{}_1\phi_1(1-p; 0; 1-p, -\frac{\lambda}{\mu})}.$$