The Israeli Queue with Finite and Infinite Number of Families

Nir Perel and Uri Yechiali

September 12, 2012

Department of Statistics and Operations Research,
School of Mathematical Sciences, Tel-Aviv University, Tel-Aviv, Israel.
[perelnir@post.tau.ac.il] [uriy@post.tau.ac.il]

Abstract

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 having the most senior waiting customer. The case with finite number of queues (families) was introduced in Boxma, Wal and Yechiali [2]. In this paper we extend the model to the case with a (possibly) infinite number of families. We analyze the $M/M/1$, the $M/M/c$, and the $M/M/1/N$ - 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 polling system (see e.g. Takagi [6], Boon et al. [1], Yechiali [8]) with N queues and a single server. The 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 a random time B_i, independent of the size of the batch. This model was first introduced and studied by Wal and Yechiali [7] where various server's dynamic visit-order rules were analyzed, leading, for 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, Wal and Yechiali [2] for the Exhaustive, Gated and Globally-Gated service disciplines. Furthermore, in [2] the following server's visit-order rule was studied: After completion of service in queue i, say, the next queue to be served is the one where the first customer in line there 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', illustrated vividly as follows: In the so called 'Israeli Queue', each arriving customer first searches for a 'friend' or a 'family member' already standing in line. If he ('he' stands for 'she' as well) finds such a family member, he joins him and his group (family) and waits with all his family members to be served in a batch mode. That is, the whole group is served at one service period, while, as indicated, the service duration is not affected by the size of the batch.

In this work we extend this model to the case where there is no bound on the number of different families that can be present simultaneously in the system. We assume that the probability for an arriving job to know another job (group leader) standing in line is p, with the same p for all families. For example, suppose there are n groups (different family types or different family leaders) in the system, including the one in service. Then the probability that an arriving customer will join the group standing in the k-th place is $(1-p)^{k-1}p$ for $1 \leq k \leq n$. Clearly, if an arriving customer finds no friends among any of the n group leaders, then he will create a new group, with probability $(1-p)^n$. So, as the number of groups in the system increases, the growth rate of new groups decreases. In what follows we present an extensive probabilistic analysis of the 'Israeli Queue' for the following cases: (i) an $M/M/1$-type queue with an un-restricted number of families; (ii) an $M/M/c$-type queue, also with an un-limited number of families; (iii) an $M/M/1/N$-type queue; and (iv) a two-class (VIP and ordinary) priority model where the VIP's can form at most M Israeli queue-type families. 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.

The paper is organized as follows: In section 2 we analyze the infinite-buffer single server system, where the arrival stream follows a homogeneous Poisson process, and the batch service time is Exponentially distributed. In section 3 we study a similar model but with c servers. In section 4 we analyze a system with at most N families, and we conclude in section 5 with a priority model.

2 The $M/M/1$-type model

2.1 Model description

We consider a single server queue with an infinite buffer. Every customer arriving to a non-empty queue searches for a 'friend' and if he finds one, he joins him and his group (family) and receives service in a batch mode together with his friend's group.

We assume that the arrival process is Poisson with rate λ, while a service time of a batch, independent of its size, is Exponentially distributed with parameter μ.

Groups are formed as follows: each group has a "leader" or a "head" - the first one of the group to arrive to the system. New arrivals see only the head of each group. The probability of being friend with a group leader is p, and the groups are statistically independent. We assume that an arriving customer can also join the group that is being served. That is, if there are n groups in the system (including the one in service), than the probability to join the $k-th$ group, for $1 \leq k \leq n$, is $(1-p)^{k-1}p$, while the probability of creating a new group is $(1-p)^n$. 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 family types) in the system at time t. A transition rate diagram of $L(t)$ is depicted in Figure 1.

In the next subsections we 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 $k-th$ position, the mean length of a busy period, and the number of groups being bypassed by a newly arriving customer.

2.2 Balance equations

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:

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

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

Equation (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 equation (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 equation (2.3) is a q-hypergeometric series ($q = 1-p$) defined in the appendix (see also [3]) .

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 equation (2.2) by z^n for all $n \geq 1$, summing

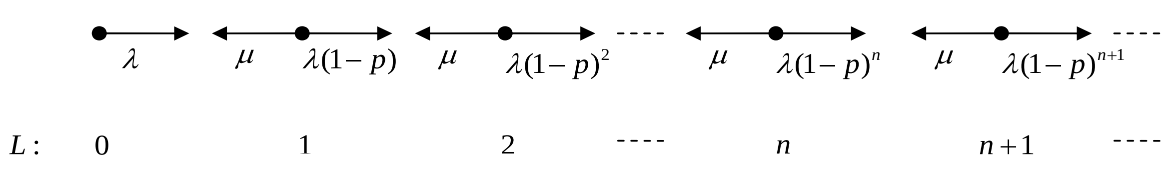

Figure 1: Transition rate diagram of the process $\{L(t), t \geq 0\}$.

with equation (2.1) and dividing by $(1-z)$ we get

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

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 of creating a new group in the system: $\sum_{n=0}^{\infty} \pi_n (1-p)^n = G(1-p)$. Furthermore, $\lambda_{eff} := \lambda G(1-p)$ 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. Equation (2.5) then gives

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

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

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

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 $\pi_0 = \frac{\mu}{\lambda+\mu}$ and $\pi_1 = \frac{\lambda}{\lambda+\mu}$, respectively. It follows that $\mathbb{E}[L|_{p=1}] = \frac{\lambda}{\mu}\pi_0 = \frac{\lambda}{\lambda+\mu}$.
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 $\mathrm{Exp}(\lambda)$, we get

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

resulting in

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

$\mathbb{E}[L]$ is plotted in Figure 2 as a function of p. The upper straight line depicts $\mathbb{E}[L]\left(= \frac{\lambda}{\mu-\lambda}\right)$ 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

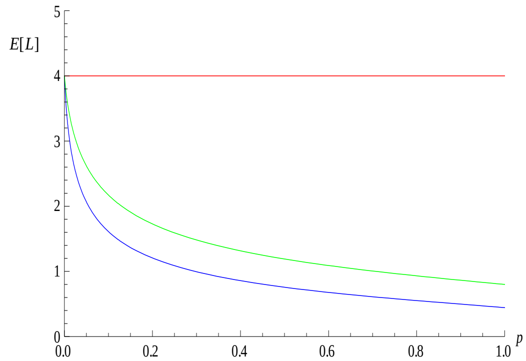

Figure 2: Plots of $\mathbb{E}[L]$ as a function of p

with arrival rate $= 4G(1-p)$, and $\mu = 5$. The bottom line depicts $\mathbb{E}[L]$ for the Israeli Queue with $\lambda = 4$, $\mu = 5$.

2.3 Sojourn times and waiting times

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,

$$\begin{aligned}
\widetilde{W}(s) = \mathbb{E}\left[e^{-sW}\right] &= \mathbb{E}_L\left[\mathbb{E}\left[e^{-sW}|L \text{ in the system and a new arrival does not find a friend}\right]\right] \\
&= \frac{1}{G(1-p)}\left[\pi_0\frac{\mu}{\mu+s} + \pi_1(1-p)\left(\frac{\mu}{\mu+s}\right)^2 + \ldots + \pi_n(1-p)^n\left(\frac{\mu}{\mu+s}\right)^{n+1} + \ldots\right] \\
&= \frac{1}{G(1-p)} \cdot \frac{\mu}{\mu+s} \cdot G\left(\frac{\mu}{\mu+s}(1-p)\right). \qquad (2.9)
\end{aligned}$$

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

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

and by using Little's Law we have

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

which coincides with equation (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,

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

Consequently,

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

Define W^a as the total sojourn time of an *arbitrary* customer in the system, and $\widetilde{W}^a(\cdot)$ as its LST. We get

$$\begin{aligned}
\widetilde{W}^a(s) &= \mathbb{E}\left[e^{-sW^a}\right] = \mathbb{E}_L\left[\mathbb{E}\left[e^{-sW^a}|L\right]\right] = \\
&= \pi_0 \cdot \frac{\mu}{\mu+s} + \pi_1\left(p \cdot \frac{\mu}{\mu+s} + (1-p)\left(\frac{\mu}{\mu+s}\right)^2\right) + \dots \\
&+ \pi_n\left(p \cdot \frac{\mu}{\mu+s} + \dots + p(1-p)^{n-1}\left(\frac{\mu}{\mu+s}\right)^n + (1-p)^n\left(\frac{\mu}{\mu+s}\right)^{n+1}\right) + \dots \\
&= \pi_0 \cdot \frac{\mu}{\mu+s} + \sum_{n=1}^{\infty}\pi_n p\frac{\mu}{\mu+s}\sum_{k=0}^{n-1}\left(\frac{\mu}{\mu+s}(1-p)\right)^k + \sum_{n=1}^{\infty}\pi_n(1-p)^n\left(\frac{\mu}{\mu+s}\right)^{n+1} \\
&= \sum_{n=1}^{\infty}\pi_n p\frac{\mu}{\mu+s}\sum_{k=0}^{n-1}\left(\frac{\mu}{\mu+s}(1-p)\right)^k + \sum_{n=0}^{\infty}\pi_n(1-p)^n\left(\frac{\mu}{\mu+s}\right)^{n+1} \\
&= \frac{\mu p}{\mu p+s}\left(1 - G\left(\frac{\mu}{\mu+s}(1-p)\right)\right) + \frac{\mu}{\mu+s}G\left(\frac{\mu}{\mu+s}(1-p)\right),
\end{aligned}$$

and

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

Clearly,

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

and by applying L'hopital's rule we get

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

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 get

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

Clearly, when $p=0$, $\mathbb{E}[L^{total}] = \mathbb{E}[L] = \frac{\lambda}{\mu-\lambda}$.
Another important service measure is the waiting time of an arbitrary customer, denoted by W_q^a.

Similarly to the derivation of $\widetilde{W}^a(s)$ we obtain

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

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 i-th 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 k-th place (for $k \geq 1$), at the moment of service completion ($k = 1$ 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 k-th place (for $k \geq 1$) among the Y arrivals, assuming that the k-th group exists.

Then, from the above, $D^{(k)} = D^{(k+1)} + \xi^{(k)}$, 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 $\mathbb{P}(Y = m) = \left(\frac{\lambda}{\lambda+\mu}\right)^m \left(\frac{\mu}{\lambda+\mu}\right)$, we have

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

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 k-th position, with rate $\lambda(1-p)^{k-1}p$, and the departure process, with rate μ. Having exactly j jobs joining the k-th group during a service duration means that the arriving process to the k-th group 'wins' j times before the end of a single service duration.
Consequently,

$$\mathbb{E}\left[\xi^{(k)}\right] = \frac{\lambda}{\mu}(1-p)^{k-1}p. \tag{2.16}$$

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

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

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

Define by $D_i^{(1)}$ the size of the group that has just completed its service, given that this group was formed in the i-th place. Then, $D_i^{(1)} = 1+\sum_{k=1}^{i}\xi^{(i+1-k)}$ with probability $\frac{\pi_{i-1}(1-p)^{i-1}}{\sum_{i=1}^{\infty}\pi_{i-1}(1-p)^{i-1}}$, or equivalently, with probability $\frac{\pi_{i-1}(1-p)^{i-1}}{G(1-p)}$. Using this fact, and equation (2.16), yields,

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

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

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

and together with (2.5) we obtain

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

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

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

where in the last equality we utilize equation (2.17), and the fact that the variables $\xi^{(i+1-k)}$, for $k = 1, \dots, i$, 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 k-th place at the moment of service completion, given that it was formed in the i-th place, for $i \geq k \geq 1$. Then, with probability $\frac{\pi_{i-1}(1-p)^{i-1}}{\sum_{j=k}^{\infty} \pi_{j-1}(1-p)^{j-1}}$, $D_i^{(k)} = 1 + \sum_{m=1}^{i-k+1} \xi^{(i+1-m)}$. Using the fact that $\pi_{i-1}(1-p)^{i-1} = \pi_i \frac{\mu}{\lambda}$ we get

$$
\begin{aligned}
\mathbb{E}\left[D^{(k)}\right] &= \sum_{i=k}^{\infty} \mathbb{E}\left[D_i^{(k)}\right] \frac{\pi_{i-1}(1-p)^{i-1}}{\sum_{j=k}^{\infty} \pi_{j-1}(1-p)^{j-1}} \\
&= \sum_{i=k}^{\infty} \left(1 + \sum_{m=1}^{i-k+1} \mathbb{E}\left[\xi^{(i+1-m)}\right]\right) \frac{\pi_{i-1}(1-p)^{i-1}}{\sum_{j=k}^{\infty} \pi_{j-1}(1-p)^{j-1}} \\
&= 1 + \sum_{i=k}^{\infty} \frac{\pi_i}{\sum_{j=k}^{\infty} \pi_j} \sum_{m=1}^{i-k+1} \frac{\lambda}{\mu}(1-p)^{i-m} p \\
&= 1 + \frac{\lambda}{\mu} \sum_{i=k}^{\infty} \frac{\pi_i}{\sum_{j=k}^{\infty} \pi_j} \left((1-p)^{k-1} - (1-p)^i\right) \\
&= 1 + \frac{\lambda}{\mu}(1-p)^{k-1} - \frac{\sum_{i=k}^{\infty} \pi_{i+1}}{\sum_{j=k}^{\infty} \pi_j} \\
&= 1 + \frac{\lambda}{\mu}(1-p)^{k-1} - 1 + \frac{\pi_k}{1 - \sum_{j=0}^{k-1} \pi_j} \\
&= \frac{\lambda}{\mu}(1-p)^{k-1} + \frac{\pi_k}{1 - \sum_{j=0}^{k-1} \pi_j}.
\end{aligned} \tag{2.22}
$$

We can now verify equation (2.13) by using equation (2.22):

$$
\begin{aligned}
\mathbb{E}\left[L^{total}\right] &= \sum_{j=1}^{\infty} \pi_j \sum_{k=1}^{j} \mathbb{E}\left[D^{(k)}\right] = \sum_{j=1}^{\infty} \pi_j \sum_{k=1}^{j} \left(\frac{\lambda}{\mu}(1-p)^{k-1} + \frac{\pi_k}{1 - \sum_{j=0}^{k-1} \pi_j}\right) \\
&= \frac{\lambda}{\mu} \sum_{j=1}^{\infty} \pi_j \frac{1-(1-p)^j}{p} + \sum_{j=1}^{\infty} \pi_j \sum_{k=1}^{j} \frac{\pi_k}{1 - \sum_{j=0}^{k-1} \pi_j} \\
&= \frac{\lambda}{\mu} \cdot \frac{1 - G(1-p)}{p} + \sum_{k=1}^{\infty} \pi_k \sum_{j=k}^{\infty} \frac{\pi_j}{\sum_{i=k}^{\infty} \pi_i} \\
&= \frac{\lambda}{\mu} \cdot \frac{1 - G(1-p)}{p} + \sum_{k=1}^{\infty} \pi_k = \frac{\lambda(1 - G(1-p))}{\mu p} + 1 - \pi_0.
\end{aligned} \tag{2.23}
$$

Substituting $1 - \pi_0 = \frac{\lambda}{\mu} G(1-p)$ in (2.23) yields equation (2.13).

2.5 The position of a new arrival and the number of groups being bypassed

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

$$\mathbb{P}(X = k) = \pi_{k-1}(1-p)^{k-1} + \sum_{n=k}^{\infty} \pi_n (1-p)^{k-1} p, \quad k \geq 1. \tag{2.24}$$

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

$$\begin{aligned}
\hat{X}(z) &= \sum_{k=1}^{\infty} z^k \left(\pi_{k-1}(1-p)^{k-1} + (1-p)^{k-1} p \sum_{n=k}^{\infty} \pi_n \right) \\
&= zG\left((1-p)z\right) + \sum_{n=1}^{\infty} \pi_n pz \sum_{k=0}^{n-1} \left(z(1-p)\right)^{k-1} \\
&= zG\left((1-p)z\right) + \sum_{n=1}^{\infty} \pi_n pz \left(\frac{1-(z(1-p))^n}{1-z(1-p)} \right) \\
&= zG\left((1-p)z\right) + \frac{pz}{1-z(1-p)} (1 - G\left((1-p)z\right)).
\end{aligned} \tag{2.25}$$

From the PGF of X we obtain

$$\begin{aligned}
\mathbb{E}[X] = \hat{X}'(z)|_{z=1} &= G(1-p) + \frac{1}{p}\left(1 - G(1-p)\right) \\
&= \frac{1}{p}\left(1 - (1-p)G(1-p)\right).
\end{aligned} \tag{2.26}$$

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

We also define the random variable N_B, as the number of groups that are being bypassed by an arriving customer. Then, by conditioning on the number of different families present in the system, we get that the probability distribution function of N_B is,

$$\begin{aligned}
\mathbb{P}(N_B = 0) &= \sum_{n=1}^{\infty} \pi_n p(1-p)^{n-1} + \sum_{n=0}^{\infty} \pi_n (1-p)^n = \frac{p}{1-p}\left(G(1-p) - \pi_0\right) + G(1-p), \\
\mathbb{P}(N_B = k) &= \sum_{n=k+1}^{\infty} \pi_n p(1-p)^{n-k-1}, \quad k \geq 1.
\end{aligned} \tag{2.27}$$

From equation (2.27) we get

$$\begin{aligned}
\mathbb{E}[N_B] &= \sum_{k=0}^{\infty} k \sum_{n=k+1}^{\infty} \pi_n p(1-p)^{n-k-1} \\
&= \sum_{n=1}^{\infty} \pi_n \sum_{k=0}^{n-1} kp(1-p)^{n-k-1} \\
&= \sum_{n=1}^{\infty} \pi_n \left(n - \frac{1}{p} + \frac{(1-p)^n}{p} \right) \\
&= \mathbb{E}[L] - \frac{1}{p}\left(1 - G(1-p)\right). \qquad (2.28)
\end{aligned}$$

The PGF of N_B is obtained as follows:

$$\begin{aligned}
\mathbb{E}\left[z^{N_B}\right] \equiv \hat{N}_B(z) &= G(1-p) + \sum_{k=0}^{\infty} z^k \sum_{n=k+1}^{\infty} \pi_n p(1-p)^{n-k-1} \\
&= G(1-p) + p\sum_{n=1}^{\infty} \pi_n (1-p)^{n-1} \sum_{k=0}^{n-1} \left(\frac{z}{1-p}\right)^k \\
&= G(1-p) + \frac{p}{1-p-z}\left(\sum_{n=1}^{\infty} \pi_n (1-p)^n - \sum_{n=1}^{\infty} \pi_n z^n\right) \\
&= G(1-p) + \frac{p}{1-p-z}\left(G(1-p) - \pi_0 - G(z) + \pi_0\right) \\
&= \frac{1-z}{1-p-z} G(1-p) - \frac{p}{1-p-z} G(z). \qquad (2.29)
\end{aligned}$$

Indeed, differentiating $\hat{N}_B(z)$ and setting $z = 1$ yields result (2.28).

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 get (2.28).

3 The $M/M/c$-type Model

3.1 Model description

We consider the case where the system is comprised of c identical servers. Service time for each server is an Exponentially distributed random variable with parameter μ. Customers arrive at the system according to a Poisson process with rate λ. As described in section 2, an arriving customer automatically searches for a friend. We assume that even if there is at least one vacant server among the c servers, the arriving customer nevertheless searches for a friend among the groups standing in front of the busy servers. The probability for an arriving customer to know another customer (group leader) standing in line is p, with the same p for all customers. The system is represented by a continuous-time Markov chain $\{L_c(t), t \geq 0\}$, where $L_c(t)$ is the total number of different group leaders (different family types) in the system at time t. A transition rate diagram is depicted in Figure 3.

3.2 Balance equations and steady-state probabilities

As for the $M/M/1$ model, we define the steady-state probabilities $\pi_n = P(L_c = n)$, for $n \geq 0$, where $L_c = \lim_{t\to\infty} L_c(t)$. That is, π_n, $n \geq 0$, is the probability that n different groups are present in the system. By considering horizontal 'cuts' on Figure 3, we have the following relations:

$$\lambda\pi_0 = \mu\pi_1 \quad \Rightarrow \quad \pi_1 = \frac{\lambda}{\mu}\pi_0, \tag{3.1}$$

$$\lambda(1-p)\pi_1 = 2\mu\pi_2 \quad \Rightarrow \quad \pi_2 = \left(\frac{\lambda}{\mu}\right)^2 \frac{1}{2!}(1-p)\pi_0,$$

$$\vdots$$

$$\lambda(1-p)^{n-1}\pi_{n-1} = n\mu\pi_n \quad \Rightarrow \quad \pi_n = \left(\frac{\lambda}{\mu}\right)^n \frac{1}{n!}(1-p)^{\frac{n(n-1)}{2}}\pi_0\ , \quad 1 \leq n \leq c-1, \tag{3.2}$$

$$\vdots$$

$$\lambda(1-p)^{c+k-1}\pi_{c+k-1} = c\mu\pi_{c+k} \quad \Rightarrow \quad \pi_{c+k} = \left(\frac{\lambda}{c\mu}\right)^k \left(\frac{\lambda}{\mu}\right)^c \frac{1}{c!}(1-p)^{\frac{(c+k)(c+k-1)}{2}}\pi_0\ , \ k \geq 0. \tag{3.3}$$

From equations (3.2) and (3.3) we get

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

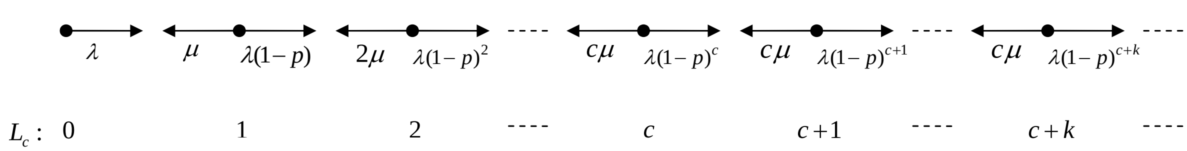

Figure 3: Transition rate diagram of the process $\{L_c(t), t \geq 0\}$.

Generating function of L_c The PGF of L_c is given by $G_c(z) = \sum_{n=0}^{\infty} \pi_n z^n$. Using equations (3.2) and (3.3) we have

$$G_c(z) = \left[\sum_{n=0}^{c-1} \left(\frac{\lambda z}{\mu}\right)^n \frac{1}{n!}(1-p)^{\frac{n(n-1)}{2}} + \sum_{k=0}^{\infty} \left(\frac{\lambda z}{c\mu}\right)^k \left(\frac{\lambda z}{\mu}\right)^c \frac{1}{c!}(1-p)^{\frac{(c+k)(c+k-1)}{2}} \right] \pi_0, \tag{3.5}$$

where π_0 is given by (3.4).
Apparently there is no closed form expression for $G_c(z)$, but a functional equation can be derived as follows: consider the steady-state balance equations obtained from Figure 3,

$$(\lambda(1-p)^n + n\mu)\pi_n = \lambda(1-p)^{n-1}\pi_{n-1} + (n+1)\mu\pi_{n+1} \quad , n = 1, 2, \ldots, c-1, \tag{3.6}$$
$$(\lambda(1-p)^n + c\mu)\pi_n = \lambda(1-p)^{n-1}\pi_{n-1} + c\mu\pi_{n+1} \quad , n \geq c. \tag{3.7}$$

Multiplying equations (3.6) and (3.7) by z^n (for the corresponding n), and summing over n together with equation (3.1) gives

$$\begin{aligned} &\sum_{n=0}^{\infty} \lambda(1-p)^n \pi_n z^n + \sum_{n=1}^{c-1} n\mu\pi_n z^n + \sum_{n=c}^{\infty} c\mu\pi_n z^n \\ &= \sum_{n=1}^{\infty} \lambda(1-p)^{n-1}\pi_{n-1} z^n + \sum_{n=0}^{c-1} (n+1)\mu\pi_{n+1} z^n + \sum_{n=c}^{\infty} c\mu\pi_{n+1} z^n. \end{aligned} \tag{3.8}$$

Define $H(z) = \sum_{n=0}^{c-1} \pi_n z^n$. Then, after some manipulations on equation (3.8) we get a functional equation for $G_c(z)$ as follows:

$$G_c(z) = \frac{\lambda}{c\mu} z G_c\left((1-p)z\right) - \frac{z}{c} H'(z) + H(z). \tag{3.9}$$

Substituting $z = 1$ in equation (3.9) gives

$$\lambda G_c(1-p) = \sum_{n=1}^{c-1} n\mu\pi_n + c\mu \sum_{n=c}^{\infty} \pi_n. \tag{3.10}$$

Clearly, when $c = 1$, equation (3.10) coincides with equation (2.5).
Equation (3.10) means that the rate at which new groups are created (left-hand side) is equal to the overall service rate of the groups present in the system (right-hand side).
In order to derive the mean total number of different groups in the system we differentiate $G_c(z)$ and set $z = 1$. This yields

$$\mathbb{E}[L_c] = G_c'(z)|_{z=1} = \frac{\lambda}{c\mu}\Big(G_c(1-p) + (1-p)G_c'(1-p)\Big) + H'(1) - \frac{H'(1) + H''(1)}{c} \tag{3.11}$$

3.3 Sojourn times and waiting times

Let W_c denote the total time a group leader spends in the system, i.e. the time from the moment he joins the queue (forming a new group) until his service completion. Let $\widetilde{W}_c(s) = \mathbb{E}\left[e^{-sW_c}\right]$ be the LST of W_c. To calculate $\widetilde{W}_c(s)$, we condition on L_c. If $L_c = 0$ (with probability π_0), then $\widetilde{W}_c(s)$ is $\frac{\mu}{\mu+s}$. If $L_c = 1 < c$, then with probability $(1-p)$ an arriving customer becomes a group leader, and $\widetilde{W}_c(s)$ is again $\frac{\mu}{\mu+s}$ since this customer immediately receives service. In general, if $L = n$, for $0 \leq n \leq c-1$, then with probability $(1-p)^n$ a new group is formed and $\widetilde{W}_c(s) = \frac{\mu}{\mu+s}$. However, if $L = c+k$, for $k \geq 0$, then a new group is formed with probability $(1-p)^{c+k}$, and $\widetilde{W}_c(s) = \left(\frac{c\mu}{c\mu+s}\right)^{k+1} \cdot \frac{\mu}{\mu+s}$. This follows since a customer that forms a new group in position $c+k+1$, needs to wait $k+1$ independent service completions (where the time between each completion is Exponential with rate $c\mu$) before reaching a server, and then his own Exponential (with rate μ) service time is added. This results in

$$\begin{aligned}
\widetilde{W}_c(s) &= \frac{1}{G_c(1-p)}\Bigg[\pi_0 \frac{\mu}{\mu+s} + \pi_1(1-p)\frac{\mu}{\mu+s} \ldots + \pi_{c-1}(1-p)^{c-1}\frac{\mu}{\mu+s} \\
&\quad + \pi_c(1-p)^c \frac{c\mu}{c\mu+s} \cdot \frac{\mu}{\mu+s} + \pi_{c+1}(1-p)^{c+1}\left(\frac{c\mu}{c\mu+s}\right)^2 \cdot \frac{\mu}{\mu+s} \\
&\quad + \ldots + \pi_{c+k}(1-p)^{c+k}\left(\frac{c\mu}{c\mu+s}\right)^{k+1} \cdot \frac{\mu}{\mu+s}\Bigg] \\
&= \frac{1}{G(1-p)}\left[\frac{\mu}{\mu+s}\sum_{n=0}^{c-1}\pi_n(1-p)^n + \frac{\mu}{\mu+s}\sum_{k=0}^{\infty}\pi_{c+k}(1-p)^{c+k}\left(\frac{c\mu}{c\mu+s}\right)^{k+1}\right] \\
&= \frac{1}{G_c(1-p)} \cdot \frac{\mu}{\mu+s}\left[H(1-p) + \left(\frac{c\mu}{c\mu+s}\right)^{1-c}\left(G_c\left(\frac{(1-p)c\mu}{c\mu+s}\right) - H\left(\frac{(1-p)c\mu}{c\mu+s}\right)\right)\right].
\end{aligned} \tag{3.12}$$

From equation (3.12) we get

$$\mathbb{E}[W_c] = -\widetilde{W}_c{}'(s)|_{s=0} = \frac{1}{c\mu}\left(1 + \frac{(c-1)H(1-p) + (1-p)\left(G'(1-p) - H'(1-p)\right)}{G(1-p)}\right) \tag{3.13}$$

Define W_c^a as the total sojourn time in the system of an *arbitrary* customer. Similar considerations to those described when calculating $\widetilde{W_c}(s)$ yield

$$\begin{aligned}
\widetilde{W_c}^a(s) &= \pi_0\frac{\mu}{\mu+s} + \pi_1\frac{\mu}{\mu+s} + \ldots + \pi_{c-1}\frac{\mu}{\mu+s} + \pi_c\left[(1-(1-p)^c)\frac{\mu}{\mu+s} + (1-p)^c\frac{c\mu}{c\mu+s}\cdot\frac{\mu}{\mu+s}\right] \\
&\quad + \pi_{c+1}\left[(1-(1-p)^c)\frac{\mu}{\mu+s} + (1-p)^c p\cdot\frac{c\mu}{c\mu+s}\cdot\frac{\mu}{\mu+s} + (1-p)^{c+1}\left(\frac{c\mu}{c\mu+s}\right)^2\cdot\frac{\mu}{\mu+s}\right] + \ldots \\
&\quad + \pi_{c+k}\left[(1-(1-p)^c)\frac{\mu}{\mu+s} + \sum_{j=0}^{k-1}(1-p)^{c+j}p\left(\frac{c\mu}{c\mu+s}\right)^{j+1}\frac{\mu}{\mu+s} + (1-p)^{c+k}\left(\frac{c\mu}{c\mu+s}\right)^{k+1}\frac{\mu}{\mu+s}\right] + \ldots \\
&= H(1)\frac{\mu}{\mu+s} + \frac{\mu}{\mu+s}(1-(1-p)^c)\sum_{j=0}^{\infty}\pi_{c+j} + \sum_{k=1}^{\infty}\pi_{c+k}\sum_{j=0}^{k-1}(1-p)^{c+j}p\left(\frac{c\mu}{c\mu+s}\right)^{j+1}\frac{\mu}{\mu+s} \\
&\quad + \sum_{k=0}^{\infty}\pi_{c+k}(1-p)^{c+k}\left(\frac{c\mu}{c\mu+s}\right)^{k+1}\frac{\mu}{\mu+s}. \qquad (3.14)
\end{aligned}$$

After some algebra equation (3.14) becomes

$$\begin{aligned}
\widetilde{W_c}^a(s) &= \frac{\mu}{\mu+s}\left[1 - \frac{(1-p)^c(1-H(1))s}{cp\mu+s}\right] \\
&\quad + \frac{\mu}{\mu+s}\left[G_c\left(\frac{(1-p)c\mu}{c\mu+s}\right) - H\left(\frac{(1-p)c\mu}{c\mu+s}\right)\right]\left[\left(\frac{c\mu}{c\mu+s}\right)^{1-c} - \frac{cp\mu}{cp\mu+s}\left(1+\frac{s}{c\mu}\right)^c\right] \qquad (3.15)
\end{aligned}$$

4 The $M/M/1/N$-type model

4.1 Model description and steady-state distribution

We consider the case where the queue of group leaders is restricted to a finite number N. That is, at most N different family types are allowed to be created in the system, but the number of members in each group is unbounded. In other words, the queue is composed of at most N groups, where each group can contain an unlimited number of customers. We also assume that no jobs are lost. That is, if there are $N-1$ groups and a new customer does not know any of the $N-1$ group leaders, he creates a new (N-th) group, while if there are N groups present and a new customer knows none of the first $N-1$ leaders, he joins the N-th group. Define $L^{(N)}(t)$ as the total number of different group leaders (different family types) in the system at time t, where $0 \leq L^{(N)}(t) \leq N$, $L^{(N)} = \lim_{t\to\infty} L^{(N)}(t)$, and $\pi_n = P(L^{(N)} = n)$, for $0 \leq n \leq N$. A transition rate diagram is depicted in Figure 4. Vertical cuts on Figure 4 immediately give

$$\pi_n = \left(\frac{\lambda}{\mu}\right)^n (1-p)^{\frac{n(n-1)}{2}}\pi_0, \quad 1 \leq n \leq N. \qquad (4.1)$$

Combining equation (4.1) and $\sum_{n=0}^{N} \pi_n = 1$, we obtain

$$\pi_0 = \left(\sum_{n=0}^{N} \left(\frac{\lambda}{\mu} \right)^n (1-p)^{\frac{n(n-1)}{2}} \right)^{-1}. \tag{4.2}$$

Rewriting the balance equations as

$$\lambda \pi_0 = \mu \pi_1, \tag{4.3}$$

and

$$\left(\lambda(1-p)^n + \mu \right) \pi_n = \lambda(1-p)^{n-1} \pi_{n-1} + \mu \pi_{n+1}, \quad 1 \le n \le N-1, \tag{4.4}$$

the PGF of $L^{(N)}$, $G^{(N)}(z) = \sum_{n=0}^{N} \pi_n z^n$, is given by

$$G^{(N)}(z) = \frac{\lambda}{\mu} z \left(G^{(N)}\left((1-p)z\right) - (1-p)^N \pi_N z^N \right) + \pi_0. \tag{4.5}$$

Substituting $z = 1$ in (4.5) gives

$$G^{(N)}(1-p) = \frac{\mu}{\lambda}(1-\pi_0) + (1-p)^N \pi_N. \tag{4.6}$$

4.2 Size of a batch

We now calculate the mean total size of the group standing at the k-th position ($1 \le k \le N$), right after the moment of service completion, and before the group moves forward to the $(k-1)$-st position (or leaves the system, when $k = 1$). As in section 2, we define the following variables, for $1 \le k \le N$:

- $\xi^{(k)}$ = number of customers who have joined the group in the k-th position during a service time, assuming that the k-th group exists.
- $D_i^{(k)}$ = size of the batch standing in the k-th position at the moment of service completion, given that it was formed in the i-th position, $i \ge k$.
- $D^{(k)}$ = size of the batch standing in the k-th position at the moment of service completion.

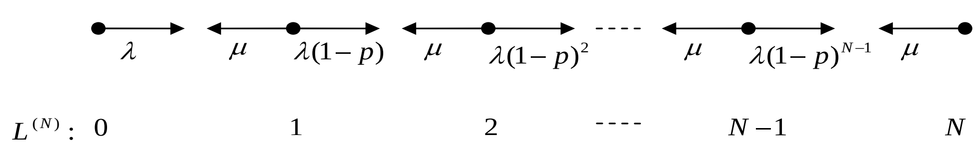

Figure 4: Transition rate diagram for $L^{(N)}$ in the $M/M/1/N$-type queue.

Following the calculations done in section 2, we get for $j \geq 0$,

$$\mathbb{P}(\xi^{(k)} = j) = \frac{\mu}{\lambda(1-p)^{k-1}p + \mu}\left(\frac{\lambda(1-p)^{k-1}p}{\lambda(1-p)^{k-1}p + \mu}\right)^j, \quad 1 \leq k \leq N-1,$$
$$\mathbb{P}(\xi^{(N)} = j) = \frac{\mu}{\lambda(1-p)^{N-1} + \mu}\left(\frac{\lambda(1-p)^{N-1}}{\lambda(1-p)^{N-1} + \mu}\right)^j. \tag{4.7}$$

From (4.7) we have

$$\mathbb{E}\left[\xi^{(k)}\right] = \frac{\lambda}{\mu}(1-p)^{k-1}p, \quad 1 \leq k \leq N-1,$$
$$\mathbb{E}\left[\xi^{(N)}\right] = \frac{\lambda}{\mu}(1-p)^{N-1}, \tag{4.8}$$

and

$$\mathbb{E}\left[z^{\xi^{(k)}}\right] = \frac{\mu}{\lambda(1-p)^{k-1}p(1-z) + \mu}, \quad 1 \leq k \leq N-1,$$
$$\mathbb{E}\left[z^{\xi^{(N)}}\right] = \frac{\mu}{\lambda(1-p)^{N-1}(1-z) + \mu}. \tag{4.9}$$

Now, for $i \geq k$, with probability $\frac{\pi_{i-1}(1-p)^{i-1}}{\sum_{j=k}^{N}\pi_{j-1}(1-p)^{j-1}}$ (the conditional probability for a new group to be formed in the i-th position), $D_i^{(k)} = 1 + \sum_{m=1}^{i-k+1}\xi^{(i+1-m)}$. Conditioning on i, and using the fact that $\pi_{i-1}(1-p)^{i-1} = \pi_i\frac{\mu}{\lambda}$ for $1 \leq i \leq N$ yield

$$\begin{aligned}
\mathbb{E}\left[D^{(k)}\right] &= \sum_{i=k}^{N}\mathbb{E}\left[D_i^{(k)}\right]\frac{\pi_{i-1}(1-p)^{i-1}}{\sum_{j=k}^{N}\pi_{j-1}(1-p)^{j-1}} = \sum_{i=k}^{N}\left(1 + \sum_{m=1}^{i-k+1}\mathbb{E}\left[\xi^{(i+1-m)}\right]\right)\frac{\pi_i}{\sum_{j=k}^{N}\pi_j} \\
&= 1 + \sum_{i=k}^{N-1}\frac{\pi_i}{\sum_{j=k}^{N}\pi_j}\sum_{m=1}^{i-k+1}\frac{\lambda}{\mu}(1-p)^{i-m}p + \frac{\pi_N}{\sum_{j=k}^{N}\pi_j}\left(\sum_{m=1}^{N-k}\frac{\lambda}{\mu}(1-p)^{N-1-m}p + \frac{\lambda}{\mu}(1-p)^{N-1}\right) \\
&= 1 + \frac{\lambda}{\mu}\left[\sum_{i=k}^{N-1}\frac{\pi_i}{\sum_{j=k}^{N}\pi_j}\left((1-p)^{k-1} - (1-p)^i\right) + \frac{\pi_N}{\sum_{j=k}^{N}\pi_j}(1-p)^{k-1}\right] \\
&= 1 + \frac{\lambda}{\mu}(1-p)^{k-1} - \frac{\lambda}{\mu}\frac{\sum_{i=k}^{N-1}(1-p)^i\pi_i}{\sum_{j=k}^{N}\pi_j} = \frac{\lambda}{\mu}(1-p)^{k-1} + \frac{\pi_k}{\sum_{j=k}^{N}\pi_j}.
\end{aligned} \tag{4.10}$$

The probabilistic explanation of this result is similar to the one given to equation (2.22). The mean total number of customers in the system, $\mathbb{E}\left[L^{total}\right]$, is then given by

$$\begin{aligned}\mathbb{E}\left[L^{total}\right] &= \sum_{j=1}^{N}\pi_j\sum_{k=1}^{j}\mathbb{E}\left[D^{(k)}\right] = \frac{\lambda(1-G^{(N)}(1-p))}{\mu p}+1-\pi_0 \\ &= \frac{\lambda}{\mu p}-\frac{(1-p)(1-\pi_0)}{p}+\frac{\lambda}{\mu p}(1-p)^N\pi_N \end{aligned} \tag{4.11}$$

where in the last transition we used equation (4.6).

5 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. So, if the number of high priority groups in the system is k, where $1 \leq k \leq M-1$, then he will either join the i-th group ($i \leq k$), 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 M-th 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 $0 \leq m \leq M$ and $n \geq 0$. A transition rate diagram of the two-dimensional process (L_1, L_2) is given in Figure 5. Define the marginal probabilities,

$$\begin{aligned}\pi_{m\bullet} &= \sum_{n=0}^{\infty}\pi_{m,n}, \quad 0 \leq m \leq M, \\ \pi_{\bullet n} &= \sum_{m=0}^{M}\pi_{m,n}, \quad n \geq 0.\end{aligned}$$

Considering the marginal probabilities $\pi_{m\bullet}$ and utilizing horizontal cuts in Figure 5, it is readily seen that $\pi_{m\bullet}$ is given by equations (4.1) and (4.2), 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

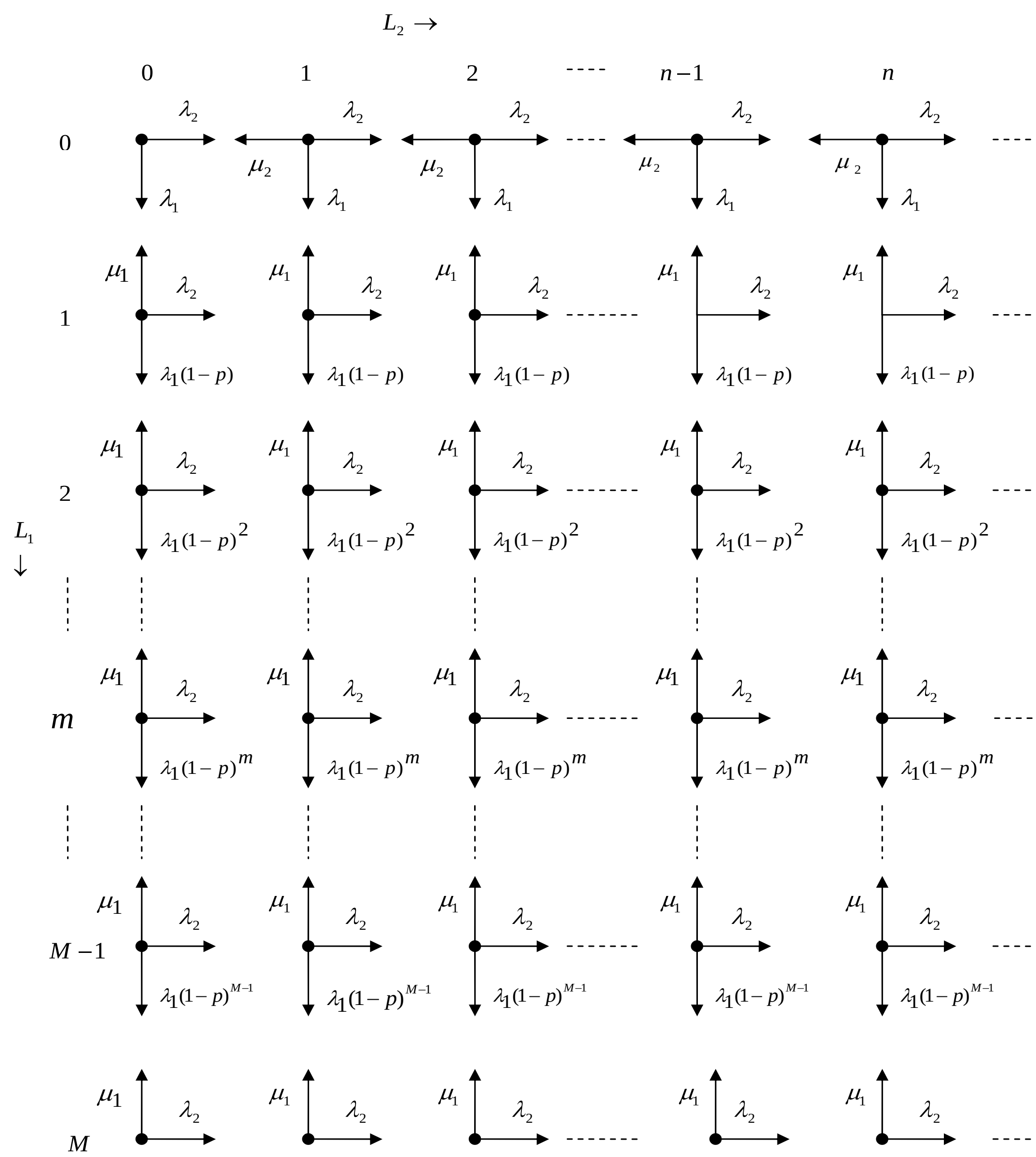

Figure 5: Transition rate diagram of (L_1, L_2) for model 5.

immediately follows that the stability condition is

$$\frac{\lambda_2}{\mu_2} < \pi_{0\bullet}, \tag{5.1}$$

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

5.1 Balance equations and generating functions

From Figure 5, we have for all $n \geq 0$,

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

Summing over n gives

$$\pi_{0,0} = \pi_{0\bullet} - \frac{\lambda_2}{\mu_2}. \tag{5.2}$$

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

$$(\lambda_1 + \lambda_2)\, \pi_{0,0} = \mu_1 \pi_{1,0} + \mu_2 \pi_{0,1}, \tag{5.3}$$

$$(\lambda_1 + \lambda_2 + \mu_2)\, \pi_{0,n} = \lambda_2 \pi_{0,n-1} + \mu_1 \pi_{1,n} + \mu_2 \pi_{0,n+1}, \quad n \geq 1. \tag{5.4}$$

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 equation (5.4) by z^n and summing over n together with (5.3) gives

$$(\lambda_1 z + \lambda_2 z(1-z) - \mu_2(1-z))\, G_0(z) - \mu_1 z G_1(z) = -\mu_2 \pi_{0,0}(1-z). \tag{5.5}$$

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

$$(\lambda_1 (1-p)^m + \lambda_2 + \mu_1)\, \pi_{m,0} = \lambda_1 (1-p)^{m-1} \pi_{m-1,0} + \mu_1 \pi_{m+1,0}, \tag{5.6}$$

$$(\lambda_1 (1-p)^m + \lambda_2 + \mu_1)\, \pi_{m,n} = \lambda_1 (1-p)^{m-1} \pi_{m-1,n} + \lambda_2 \pi_{m,n-1} + \mu_1 \pi_{m+1,n}, \quad n \geq 1. \tag{5.7}$$

Multiplying equation (5.7) by z^n and summing over n together with (5.6) leads to

$$(\lambda_1 (1-p)^m + \lambda_2(1-z) + \mu_1)\, G_m(z) - \lambda_1 (1-p)^{m-1} G_{m-1}(z) - \mu_1 G_{m+1}(z) = 0, \quad 1 \leq m \leq M-1. \tag{5.8}$$

Last, for $m = M$ we have

$$(\lambda_2 + \mu_1)\, \pi_{M,0} = \lambda_1 (1-p)^{M-1} \pi_{M-1,0}, \tag{5.9}$$

$$(\lambda_2 + \mu_1)\, \pi_{M,n} = \lambda_1 (1-p)^{m-1} \pi_{M-1,n} + \lambda_2 \pi_{M,n-1}, \quad n \geq 1. \tag{5.10}$$

Multiplying equation (5.10) by z^n and summing over n together with (5.9) yields

$$(\lambda_2(1-z)+\mu_1)\,G_M(z)=\lambda_1(1-p)^{m-1}G_{M-1}(z). \tag{5.11}$$

The set of equations (5.5), (5.8) and (5.11) can be written as

$$A(z)\cdot \underline{G}(z)=\underline{\Pi}(z), \tag{5.12}$$

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

$$a_{i,j}=\begin{cases}\lambda_1 z+\lambda_2(1-z)-\mu_2(1-z) & i=1,\ j=1\\ -\mu_1 z & i=1,\ j=2\\ -\lambda_1(1-p)^{i-2} & 2\le i\le M+1,\ j=i-1\\ \lambda_1(1-p)^{i-1}+\lambda_2(1-z)+\mu_1 & 2\le i\le M+1,\ j=i\\ -\mu_1 & 2\le i\le M,\ j=i+1\\ \lambda_2(1-z)+\mu_1 & i=M+1,\ j=M+1\\ 0 & \text{elsewhere}\end{cases} \tag{5.13}$$

$\underline{G}(z)=(G_0(z),G_1(z),\ldots,G_M(z))^T$ 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 PGF's 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:

$$\begin{aligned}\pi_{0\bullet}&=0.3286,\\ \pi_{1\bullet}&=0.3286,\\ \pi_{2\bullet}&=0.2301,\\ \pi_{3\bullet}&=0.1127,\\ \pi_{0,0}&=\pi_{0\bullet}-\frac{\lambda_2}{\mu_2}=0.0786.\end{aligned}$$

The probability generating functions are given by:

$$G_0(z) = \frac{0.8(z-6.1749)(z-3.61361)(z-1.59149)}{(z-5.8578)(z-3.17745)(z-1.14476)},$$

$$G_1(z) = \frac{-1.6(z-4.97327)(z-2.00673)}{(z-5.8578)(z-3.17745)(z-1.14476)},$$

$$G_2(z) = \frac{2.24(z-3)}{(z-5.8578)(z-3.17745)(z-1.14476)},$$

$$G_3(z) = \frac{-2.1952}{(z-5.8578)(z-3.17745)(z-1.14476)}.$$

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.

5.2 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 4. Let $\vec{\pi} = (\pi_0, \pi_1, \dots, \pi_M)$ be a stationary vector of the irreducible matrix A, i.e. $\vec{\pi}A = 0$ and $\vec{\pi} \cdot \vec{e} = 1$. Note that $\pi_m = \pi_{m,\bullet}$ for all $0 \leq m \leq M$, where $\pi_{m\bullet}$ is given by equations (4.1) and (4.2), where m replaces n.

The stability condition is ([5], page 83)

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

which immediately translates to

$$\mu_2 \pi_0 > \lambda_2.$$

This coincides with the condition given by equation (5.1).

Define for all $n \geq 0$ 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 [5]) that

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

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,

$$\begin{aligned} &\vec{P_0} B + \vec{P_1} A_2 = \vec{0}, \\ &\vec{P_0} A_0 + \vec{P_1}(A_1 + RA_2) = \vec{0}, \\ &\vec{P_0} \cdot \vec{e} + \vec{P_1} \left[I - R \right]^{-1} \cdot \vec{e} = 1. \end{aligned} \tag{5.14}$$

The derivation of the matrix R is based on *Theorem 8.5.1* in [4], 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

$$R = A_0 \left(I - A_1 - A_0 \vec{e} \cdot r \right)^{-1}. \tag{5.15}$$

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

$$\mathbb{E}[L_2] = \sum_{n=1}^{\infty} n\vec{P}_n \cdot \vec{e} = \sum_{n=1}^{\infty} n\vec{P}_1 R^{n-1} \cdot \vec{e} = \vec{P}_1 \left[I - R\right]^{-2} \cdot \vec{e} \tag{5.16}$$

6 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 by the server is the one having the most senior (waiting) 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 all 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 $k-th$ 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] 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.

[2] Boxma, O.J., Van der Wal, Y. & Yechiali, U. (2008). *Polling with batch service.* Stochastic Models 24(4): 604–625.

[3] Gasper, G. & Rahman, M. (2004). *Basic hypergeometric Series*, 2nd Edition, Cambridge University Press.

[4] Latouche, G. & Ramaswami, V. (1999). *Introduction to matrix analytic methods in stochastic modeling.* ASA-SIAM Series on Statistics and Applied Probability. SIAM, Philadelphia.

[5] Neuts, M.F. (1981). *Matrix-Geometric solutions in stochastic models. An algorithmic approach.* The Johns Hopkins University Press, Baltimore and London.

[6] Takagi, H. (1986). *Analysis of polling systems.* The MIT Press.

[7] 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.

[8] Yechiali, U. (1993). *Analysis and control of polling systems.* Performance Evaluation of Computer and Communication Systems", Springer-Verlag: 630–650.

A Appendix

The q-Hypergeometric series is defined as

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

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, equation (2.3) for π_0 can be expressed as

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

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

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