An Efficient Convolution Method to Compute the Stationary Transition Probabilities of the G/M/c Model and its Variants

Muhammad El-Taha
Department of Mathematics and Statistics
University of Southern Maine
96 Falmouth Street
Portland, ME 04104-9300
el-taha@maine.edu

***Abstract*—The multi-server G/M/c model is a fundamental building block in the analysis of many systems including computer, communication, and telecommunication networks. Consider this model with c servers, general arrivals and exponential service times. Inter-arrival times and service times are *i.i.d.* and independent of each other. The key contribution is developing a convolution method to compute transition probabilities exactly, thus avoiding numerical integration. We show how this can lead to a remarkably simple algorithm to derive an exact solution for the steady state system size probabilities.**

***Index Terms*—G/M/c model, batch arrivals, general arrival process, multi-server**

1. Introduction

The multi-server $G/M/c$ queue is a fundamental model in queueing theory that is widely used in computer, communication, and telecommunication networks. In this article we give a remarkably simple algorithm to derive an exact solution for the steady state system size probabilities. Our approach is based on developing and solving an imbedded Markov chain at arrival epochs coupled with a birth process over inter-arrival times.

Consider the standard multi-server $G/M/c$ model with c servers, general independent inter-arrival times, A_i that are $i.i.d.$ with distribution function $A(t)$ and mean $E[A] = 1/\lambda$, and exponential service times with distribution function $B(t)$. Service times, B_i, are $i.i.d.$ such that $E[B] = 1/\mu$, so that the state dependent service rate is given by $\mu_i = \min(i, c)\mu$, where i is the number of customers in the system. We study the system by focusing on the state as seen by arrivals. Let X_n be the number of customers in the system as seen by the n^{th} arrival, then $\{X_n, n = 0, 1, \cdots\}$ is a Markov chain with transition probabilities given by the following lemma (see for example Gross and Harris [6]).

Lemma 1.1: (i) for $j \leq i+1 \leq c$, the transition probabilities of $p(i,j)$ are given by

$$p(i,j) = \int_0^\infty \binom{i+1}{i-j+1} e^{-\mu t j}(1-e^{-\mu t})^{i-j+1} dA(t) \; ;$$

(ii) For $c \leq j \leq i+1$, the transition probabilities of $p(i,j)$ are given by

$$p(i,j) = \int_0^\infty \frac{e^{-c\mu t}(c\mu t)^{i-j+1}}{(i-j+1)!} dA(t);$$

(iii) for $j+1 \leq c \leq i$

$$\begin{aligned} p(i,j) \quad &= \quad \binom{c}{c-j} \frac{(c\mu)^{i-c+1}}{(i-c)!} \int_0^\infty \int_0^t e^{-\mu(t-v)j - c\mu v} \\ &\quad \times v^{i-c}(1-e^{-\mu(t-v)})^{c-j} dv dA(t) \; ; \end{aligned}$$

and $p(i,j) = 0$, otherwise.

Using these transition probabilities, specifically (iii), to compute the stationary distribution for the imbedded Markov chain is given by Takács [17] and [16]. But his method uses generating functions to develop a computational recursion. The derivation of this recursion is described by Gross and Harris [6], page 268, as "extremely long". Also Kleinrock [10] page 255, describes Takács method as "complex". Moreover, the recursive methods described by Kleinrock [10] pages 254 255, Gross and Harris [6], pages 266-267, and Medhi [12], pages 320-321 need evaluation of the expression in (iii) of Lemma 1.1, which requires numerical integration. See also Ross [14] page 556, and Tijms [18] page 400. For approximations of this model the reader may refer to Cosmetatos and Godsave [2]. Considering the wide applicability of the $G/M/c$ queueing model it is desirable to develop simpler methods to evaluate the the transition probabilities in Lemma 1.1(iii).

There are several approaches to dealing with this model and its variants, like the batch arrival with and without finite buffer. Using the supplemental variable approach Hokstad [7] studies the $G/M/c$ model with finite waiting room using pre-arrival, post-departure and time-average probabilities. Ferreira and Pacheco [4] use a uniformization technique to approximate values for the transition probabilities $p(i,j)$ of the imbedded Markov chain. Moreover, Grassmann and Tavakoli [5] give a review of several methods to compute the transition probabilities focusing on numerical stability issues. See also Yao et al [19] who uses level crossing methods to relate pre-arrival, post-departure and time-average probabilities. Matrix analytic methods pioneered by Neuts [13] are used to study this model and its variants. Another approach is to write the balance equations of the imbedded Markov chain resulting from observing the system at arrival epochs, then use transform methods to solve for the arrival probabilities. Kim and Chaudhry [9] study the $G/M/c/N$ and Chaudhry and Kim [1] study the $G^X/M/c$ models. Laxmi and Gupta [11] use the imbedded Markov chain approach to obtain the pre-arrival probabilities. By improving the computation of the transition probabilities, our results can enhance the computational approaches detailed in these articles as well.

All the different approaches dealing with this model and its variants share in common the need to compute some version of the one step transition probabilities at pre-arrival epochs. The contribution of this article is in the development of a convolution method to exactly compute the transition probabilities in Lemma 1.1(iii), and in proposing a stable efficient algorithm to compute the stationary distribution function of the system size probabilities. We greatly simply the expression for one step transition probabilities for case (iii) of Lemma 1.1 by converting an integration into a finite summation.

The rest of the article is organized as follows. In Section 2 we introduce a birth process and evaluate its transition probabilities. In Section 3 we determine the transition probabilities, the system size probabilities at arrival instants, the time average stationary distribution, and measures of performance of the $G/M/c$ model. Specifically, we propose an algorithm that solves for the stationary distribution of the imbedded Markov chain.

2. Pure Birth Process

Let $\{Y(t), t \geq 0\}$ be a pure birth process representing the number of births (service completions) during an inter-arrival time of length t. Birth rates, i.e. departure rates, depend on the number of customers in the system and are given by $\mu_i = \min(i, c)\mu$. That $\{Y(t), t \geq 0\}$ is a pure birth process is due to the fact that the time between service completions is exponential, possibly with different rates. Let $Q_{ij}(t); j \leq i$ be the transition probabilities of the birth process $\{Y(t), t \geq 0\}$ defined as

$$Q_{ij}(t) = p(Y(t) = i - j | Y(0) = i) \; ;$$

where $Y(0) = i$ represents the number of customers at time zero.

Let $\{X_i, i \geq 1\}$ be a sequence of independent random variables with distribution functions F^i. Here X_i represent inter-event times. Let $F_n(t)$ be the d.f. of $S_n = X_1 + \cdots + X_n$ which is the n-fold convolution of $F^1, \ldots, F^n$.

Lemma 2.1: Let $\{N(t), t \geq 0\}$ be a simple counting process such that $N(t) = \max\{n : S_n \leq t\}$ and $N(t) \to \infty$ as $t \to \infty$. Moreover, let times between events, X_i, be independent but not necessarily identical. Then $P(N(t) = n) = F_n(t) - F_{n+1}(t) = \bar{F}_{n+1}(t) - \bar{F}_n(t)$.

Proof. The difference between Lemma 2.1 and the well-known renewal case is that we relax the assumption that inter-event times are identically distributed. Noting that $\{N(t) \geq n\} \iff \{S_n \leq t\}$, we obtain

$$\begin{aligned} P(N(t) = n) &= P(N(t) \geq n) - P(N(t) \geq n+1)) \\ &= P(S_n \leq t) - P(S_{n+1} \leq t) \\ &= F_n(t) - F_{n+1}(t); \end{aligned}$$

this completes the proof of the lemma. ∎

We will also need the following well-known convolution result. Let $Y = X + Z$ where X and Z are non-negative random variables. Then

$$\bar{F}_Y(t) = \int_0^t f_X(s) \bar{F}_Z(t-s) ds \; . \quad (1)$$

We are interested in the birth process $\{Y(t), t \geq 0\}$ and its transition probabilities $Q_{ij}(t)$ when birth rates are not necessarily constant. Let $n = n_1 + n_2$ births, the first $n_1 \geq 1$ births are such that the time between events (births) is an exponential random variable with a constant rate θ, and the remaining n_2 inter-events are exponential random variables with non-constant rates $\lambda_k, k = 1, \cdots, n_2$. Specifically, let $\tilde{X}_{n_1} = X_1 + X_2 + \cdots + X_{n_1}$, where each $X_k \sim \exp(\theta), k = 1, 2, \cdots, n_1$; that is X_k is an exponential random variable with parameter θ; so that $\tilde{X}_{n_1} \sim$ Erlang(n_1, θ). Now, we state a fundamental result.

Lemma 2.2: Let $\theta > \lambda_k$ for all $k = 1, \cdots n_2 + 1$. Then

$$\begin{aligned} Q_{ij}(t) &= \sum_{k=1}^{n_2+1} \frac{\lambda_k C_{k,n_2+1}}{\lambda_{n_2+1}} \left(\frac{\theta}{\theta - \lambda_k} \right)^{n_1} \times \\ & \left[e^{-\lambda_k t} - \sum_{r=0}^{n_1-1} \frac{((\theta - \lambda_k)t)^r}{r!} e^{-\theta t} \right] \; ; \end{aligned}$$

where $C_{k,n_2+1} = \prod_{\substack{m=1 \\ m \neq k}}^{n_2+1} \frac{\lambda_m}{\lambda_m - \lambda_k}$, and $\prod$ and $\sum$ over empty sets are one and zero respectively.

Proof. One can see that $\tilde{X}_{n_1} \sim$ Erlang(n_1, θ), so that the pdf of $\tilde{X}_{n_1}$ is given by

$$f_{\tilde{X}_{n_1}}(t) = \frac{(\theta t)^{n_1 - 1} \theta e^{-\theta t}}{(n_1 - 1)!} \; .$$

Also $\tilde{Z}_{n_2} \sim$ hypo-exponential (see Ross [14], p 299), so that $\bar{F}_{\tilde{Z}_{n_2}}(t) = \sum_{k=1}^{n_2} C_{k,n_2} e^{-\lambda_k t}$, and $C_{k,n_2} = \prod_{\substack{m=1 \\ m \neq k}}^{n_2} \frac{\lambda_m}{\lambda_m - \lambda_k}$, where a product over an empty set is one. Now by Lemma 2.1 and equation (1)

$$\begin{aligned} Q_{ij}(t) &= P(N(t) = i - j = n_1 + n_2 | D(0) = i) \\ &= \bar{F}_{\tilde{X}_{n_1} + \tilde{Z}_{n_2+1}}(t) - \bar{F}_{\tilde{X}_{n_1} + \tilde{Z}_{n_2}}(t) \; . \text{ Now} \end{aligned}$$

$$
\begin{aligned}
&\bar{F}_{\tilde{X_{n_1}}+\tilde{Z}_{n_2+1}}(t)\\
&= \int_0^t \frac{(\theta s)^{n_1-1}\theta e^{-\theta s}}{(n_1-1)!} \times \sum_{k=1}^{n_2+1} C_{k,n_2+1} e^{-\lambda_k(t-s)} ds\\
&= \sum_{k=1}^{n_2+1} C_{k,n_2+1}\theta^{n_1} e^{-\lambda_k t}\\
&\quad \int_0^t \frac{s^{n_1-1}}{(n_1-1)!}\, e^{-(\theta-\lambda_k)s} ds\\
&= \sum_{k=1}^{n_2+1} \frac{C_{k,n_2+1}\theta^{n_1} e^{-\lambda_k t}}{(\theta-\lambda_k)^{n_1}} \times\\
&\quad \int_0^t \frac{((\theta-\lambda_k)s)^{n_1-1}(\theta-\lambda_k)}{(n_1-1)!} e^{-(\theta-\lambda_k)s} ds\\
&= \sum_{k=1}^{n_2+1} C_{k,n_2+1}\left(\frac{\theta}{\theta-\lambda_k}\right)^{n_1} e^{-\lambda_k t}\\
&\quad \left[1-\sum_{r=0}^{n_1-1}\frac{((\theta-\lambda_k)t)^r}{r!} e^{-(\theta-\lambda_k)t}\right]\\
&= \sum_{k=1}^{n_2+1} C_{k,n_2+1}\left(\frac{\theta}{\theta-\lambda_k}\right)^{n_1}\\
&\quad \left[e^{-\lambda_k t}-\sum_{r=0}^{n_1-1}\frac{((\theta-\lambda_k)t)^r}{r!} e^{-\theta t}\right] .
\end{aligned}
$$

Therefore

$$
\begin{aligned}
Q_{ij}(t) &= \bar{F}_{\tilde{X_{n_1}}+\tilde{Z}_{n_2+1}}(t)-\bar{F}_{\tilde{X_{n_1}}+\tilde{Z}_{n_2}}(t)\\
&= \sum_{k=1}^{n_2+1} C_{k,n_2+1}\left(\frac{\theta}{\theta-\lambda_k}\right)^{n_1} [\ e^{-\lambda_k t}\\
&\quad -\sum_{r=0}^{n_1-1}\frac{((\theta-\lambda_k)t)^r}{r!} e^{-\theta t}\ ] - \sum_{k=1}^{n_2} C_{k,n_2}\\
&\times\left(\frac{\theta}{\theta-\lambda_k}\right)^{n_1}\left[e^{-\lambda_k t}-\sum_{r=0}^{n_1-1}\frac{((\theta-\lambda_k)t)^r}{r!} e^{-\theta t}\right]\\
&= \sum_{k=1}^{n_2} C_{k,n_2+1}\left(\frac{\theta}{\theta-\lambda_k}\right)^{n_1} [\ e^{-\lambda_k t}\\
&\quad -\sum_{r=0}^{n_1-1}\frac{((\theta-\lambda_k)t)^r}{r!} e^{-\theta t}\ ] - \sum_{k=1}^{n_2} C_{k,n_2}\\
&\times\left(\frac{\theta}{\theta-\lambda_k}\right)^{n_1}\left[e^{-\lambda_k t}-\sum_{r=0}^{n_1-1}\frac{((\theta-\lambda_k)t)^r}{r!} e^{-\theta t}\right]\\
&\quad + C_{n_2+1,n_2+1}\left(\frac{\theta}{\theta-\lambda_{n_2+1}}\right)^{n_1}\\
&\times\left[e^{-\lambda_{n_2+1} t}-\sum_{r=0}^{n_1-1}\frac{((\theta-\lambda_{n_2+1})t)^r}{r!} e^{-\theta t}\right] ;
\end{aligned}
$$

which simplifies to

$$
\begin{aligned}
&Q_{ij}(t)\\
&= \sum_{k=1}^{n_2}(C_{k,n_2+1}-C_{k,n_2})\left(\frac{\theta}{\theta-\lambda_k}\right)^{n_1} [\ e^{-\lambda_k t}\\
&-\sum_{r=0}^{n_1-1}\frac{((\theta-\lambda_k)t)^r}{r!} e^{-\theta t}\ ] + C_{n_2+1,n_2+1}\left(\frac{\theta}{\theta-\lambda_{n_2+1}}\right)^{n_1}\\
&\times\left[e^{-\lambda_{n_2+1} t}-\sum_{r=0}^{n_1-1}\frac{((\theta-\lambda_{n_2+1})t)^r}{r!} e^{-\theta t}\right] .
\end{aligned}
$$

Note that for $k=1,\cdots,n_2$

$$
\begin{aligned}
C_{k,n_2+1}-C_{k,n_2} &= C_{k,n_2+1}-C_{k,n_2+1}\frac{\lambda_{n_2+1}-\lambda_k}{\lambda_{n_2+1}}\\
&= C_{k,n_2+1}[1-\frac{\lambda_{n_2+1}-\lambda_k}{\lambda_{n_2+1}}]\\
&= C_{k,n_2+1}\frac{\lambda_k}{\lambda_{n_2+1}} .
\end{aligned}
$$

Therefore

$$
\begin{aligned}
&Q_{ij}(t)\\
&= \sum_{k=1}^{n_2} C_{k,n_2+1}\frac{\lambda_k}{\lambda_{n_2+1}}\left(\frac{\theta}{\theta-\lambda_k}\right)^{n_1} [\ e^{-\lambda_k t}\\
&-\sum_{r=0}^{n_1-1}\frac{((\theta-\lambda_k)t)^r}{r!} e^{-\theta t}\ ] + \ C_{n_2+1,n_2+1}\left(\frac{\theta}{\theta-\lambda_{n_2+1}}\right)^{n_1}\\
&\left[e^{-\lambda_{n_2+1} t}-\sum_{r=0}^{n_1-1}\frac{((\theta-\lambda_{n_2+1})t)^r}{r!} e^{-\theta t}\right] ,
\end{aligned}
$$

which completes the proof. ∎

Now, we state a main result.

Lemma 2.3: Consider the $G/M/c$ queueing model. For $i \geq c-1$; $1 \leq j \leq c-1$,

$$
\begin{aligned}
Q_{ij}(t) &= \sum_{k=1}^{c-j}\frac{C_{k,c-j}(c-k)}{j}\left(\frac{c}{k}\right)^{i-c+1} e^{-c\mu t}\\
&\times\left[e^{k\mu t}-\sum_{r=0}^{i-c}\frac{(k\mu t)^r}{r!}\right] ;
\end{aligned}
$$

where $C_{k,c-j}=\prod_{m=1}^{k-1}\frac{c-m}{k-m}\times\prod_{m=k+1}^{c-j}\frac{c-m}{k-m}$, and $\prod$ and $\sum$ over empty sets are 1 and 0 respectively.

Proof. We use Lemma 2.2. Let $n_1=i-c+1, n_2=c-1-j, \theta=c\mu$, and $\lambda_k=(c-k)\mu$, so that $\lambda_{n_2+1}=j\mu, \theta-\lambda_k=k\mu$, $\frac{\theta}{\theta-\lambda_k}=\frac{c}{k}$, and $\frac{\theta}{\theta-\lambda_{n_2+1}}=\frac{c}{c-j}$. Thus, for $k=1,\cdots,c-j$

$$
\begin{aligned}
C_{k,c-j} &= \prod_{\substack{m=1\\ m\neq k}}^{c-j}\frac{\lambda_m}{\lambda_m-\lambda_k}=\prod_{\substack{m=1\\ m\neq k}}^{c-j}\frac{c-m}{c-m-(c-k)}\\
&= \prod_{\substack{m=1\\ m\neq k}}^{c-j}\frac{c-m}{k-m}=\prod_{m=1}^{k-1}\frac{c-m}{k-m}\times\prod_{m=k+1}^{c-j}\frac{c-m}{k-m} ,
\end{aligned}
$$

where a product over an empty set is 1. Substitute in Lemma 2.2 to obtain

$$\begin{aligned} & Q_{ij}(t) \\ = & \sum_{k=1}^{c-j} \frac{C_{k,c-j}(c-k)\mu}{j\mu} \left(\frac{c\mu}{c\mu-(c-k)\mu} \right)^{i-c+1} \\ & \left[e^{-(c-k)\mu t} - \sum_{r=0}^{i-c} \frac{(k\mu t)^r}{r!} e^{-c\mu t} \right] \\ = & \sum_{k=1}^{c-j} \frac{C_{k,c-j}(c-k)}{j} \left(\frac{c}{k} \right)^{i-c+1} \\ & \left[e^{-(c-k)\mu t} - \sum_{r=0}^{i-c} \frac{(k\mu t)^r}{r!} e^{-c\mu t} \right] . \end{aligned}$$

This complete the proof. ∎

Special Cases.

(i.) Let $i = c-1, j = c-1$, then $Q_{c-1,c-1}(t) = e^{-(c-1)\mu t}$.

(ii.) Let $i = c, j = c-1$, then $Q_{c,c-1}(t) = ce^{-c\mu t}[e^{\mu t} - 1] = ce^{-(c-1)\mu t}(1 - e^{-\mu t})$. Note that both $Q_{c-1,c-1}(t)$ and $Q_{c,c-1}(t)$ can be obtained from Lemma 1.1 (i) by noting that $Q_{i,j}(t) = p(i-1, j; t)$.

(iii.) Let $i \geq c, j = c-1$, then $Q_{i,c-1}(t) = c^{i-c+1}e^{-(c-1)\mu t}\left[1 - \sum_{r=0}^{i-c} \frac{(\mu t)^r}{r!} e^{-\mu t}\right]$. This can be recognized as the distribution function of an Erlang with parameters $i-c+1$ and $c\mu$.

(iv.) Let $i = c, j \leq c-1$, then

$$Q_{c,j}(t) = \sum_{k=1}^{c-j} C_{k,c-j} \frac{(c-k)c}{jk} \left[e^{-(c-k)\mu t} - e^{-c\mu t} \right] .$$

3. The Imbedded Stationary Distribution

In this section we determine the transition probabilities, the system size probabilities at arrival instants, the time average stationary distribution, and measures of performance of the $G/M/c$ model.

A. One Step Transition Probabilities

We develop an embedded Markov chain, then solve for the system size arrival epochs' probabilities. Let $\{X_n, n \geq 1\}$ be a stochastic process that represents the number of customers present right before an arrival (i.e., number of customers as seen by an arrival.) Then, X_{n+1} is related to X_n as follows:

$$X_{n+1} = \begin{cases} X_n + 1 - D_n & D_n \leq X_n + 1, X_n \geq 0 \\ 0 & otherwise \end{cases} ,$$

where D_n is the number of customers served during the n^{th} inter-arrival time.

Since inter-arrival times are $i.i.d.$ and service completions during an inter-arrival, D_n, depend only on X_n, and not on the past history of the system $\{X_k, k \leq n-1\}$, we conclude that that $\{X_n, n \geq 1\}$ is a Markov chain. Let $\{\pi(i), i = 0, \ldots, \}$ be defined as $\pi(i) = \lim_{n\to\infty} P\{X_n = i\}$; that is, $\pi(i) = \lim_{n\to\infty} P\{n^{th}$ arrival sees i customers$\}$.

The transition probabilities of $\{X_n, n \geq 1\}$, are defined as $p(i,j) = P\{X_n = j | X_{n-1} = i\}, i = 0, \ldots, ; j = 0, \ldots,$. Note that $p(i,j)$ is the probability that the pure birth process, $\{Y(t), t \geq 0\}$, representing the number of customers in the system produce $i+1-j$ births (service completions) during an inter-arrival time. Therefore,

$$\begin{aligned} p(i,j) &= \int_0^\infty P\{Y(t) = j \mid Y(0) = i+1\} dA(t) \\ &= \int_0^\infty Q_{i+1,j}(t) dA(t), \end{aligned} \tag{2}$$

where $Q_{ij}(t)$ are the transition probabilities of the birth process $\{Y(t), t \geq 0\}$. Now the transition probability matrix $P = [p(i,j)]$ of the Markov chain $\{X_n, n = 0, 1, \cdots, \}$ is evaluated utilizing equation (2) above.

Let $A^*(s) = \int_0^\infty e^{-st} dA(t)$ be the Laplace-Stieltjes transform (LST) of the inter-arrival times distribution function $A(t)$. Also let $A_n^*(s) = (-1)^n \frac{d^n A^*(s)}{ds^n})$, where $\frac{d^n A^*(s)}{ds^n}$ is the n^{th} derivative of $A^*(s)$. It follows that $A_n^*(s) = \int_0^\infty t^n e^{-st} dA(t)$ for all $n \geq 0$, where $A_0^*(s) = A^*(s)$. Then we have the main computationally useful result.

Theorem 3.1: (i) For $j \leq i+1 \leq c$, the transition probabilities of $p(i,j)$ are given by

$$p(i,j) = \binom{i+1}{i-j+1} \sum_{r=0}^{i-j+1} (-1)^r \binom{i-j+1}{r} A^*((j+r)\mu)) .$$

(ii) For $c \leq j \leq i+1$, the transition probabilities of $p(i,j)$ are given by

$$p(i,j) = \frac{(c\mu)^{i-j+1} A_{i-j+1}^*(c\mu)}{(i-j+1)!} .$$

(iii) for $j > 0, j+1 \leq c \leq i$

$$\begin{aligned} p(i,j) = & \sum_{k=1}^{c-j} \frac{-1)^{c-j-k} C_{k,c-j}^a (c-k)}{j} \left(\frac{c}{k} \right)^{i-c+2} \\ & \left[A^*((c-k)\mu) - \sum_{r=0}^{i-c+1} \frac{(k\mu)^r A_r^*(c\mu)}{r!} \right] ; \end{aligned} \tag{3}$$

where $C_{k,c-j}^a = \prod_{m=1}^{k-1} \frac{c-m}{k-m} \times \prod_{m=k+1}^{c-j} \frac{c-m}{m-k}$ and $p(i,j) = 0$, otherwise.

Proof. Using Lemma 1.1 (i) and (ii) and equation (2) one can easily prove (i) and (ii) of the Corollary respectively. To prove (iii) we appeal to equation (2), Lemma 2.3 and noting that $C_{k,c-j} = (-1)^{c-j-k} C_{k,c-j}^a$. ∎

Note that in contrast to $C_{k,c-j}$, the new term $C_{k,c-j}^a$ is always positive. This will make it possible to to rearrange the terms of $p(i,j)$ to avoid propagation of error due to subtraction and improve numerical stability as we see next.

Remarks on Numerical Stability. Note that the term $(-1)^j$ in Theorem 3.1 (i) and (iii) which causes subtraction in every other step can be a source of numerical instability. To remedy that consider a sum of the form $S(.) = \sum_{k=1}^{J} (-1)^{m-k} f_k(.)$, m is an integer, and rewrite as

$$S(.) = |\sum_{u=1}^{[J/2]} f_{2u}(.) - \sum_{u=1}^{[(J+1)/2]} f_{2u-1}(.)| \tag{4}$$

Now, we have two sums and only one subtraction at the end. We utilize (4) to rewrite (3) as

$$
\begin{aligned}
p(i,j) \quad = \quad & \left| \sum_{u=1}^{[(c-j+1)/2]} \frac{C^a_{2u-1,c-j}(c-2u+1)}{j} \right. \\
& \times \left(\frac{c}{2u-1}\right)^{i-c+2} \Big[A^*((c-2u+1)\mu) \\
& - \sum_{r=0}^{i-c+1} \frac{((2u-1)\mu)^r A^*_r(c\mu)}{r!}\Big] \\
& - \sum_{u=1}^{[(c-j)/2]} \frac{C^a_{2u,c-j}(c-2u)}{j} \left(\frac{c}{2u}\right)^{i-c+2} \\
& \times \left. \left[A^*((c-2u)\mu) - \sum_{r=0}^{i-c+1} \frac{(2u\mu)^r A^*_r(c\mu)}{r!}\right] \right| \qquad (5)
\end{aligned}
$$

In a similar way the formula in (i) of Theorem 3.1 can be written in a more numerically stable form. Ferreira and Pacheco [4] use a uniformization technique that allows for a computation of the transition probabilities $p(i,j)$ of the imbedded Markov chain. Their method, though numerically stable, requires truncation. The convolution method does not require any approximations. Moreover, Grassmann and Tavakoli [5] give a review of several methods to compute the transition probabilities and interestingly introduce equation (31), a version of our $Q_{i,j}(t)$ given in Lemma 2.3 but with factorial coefficients instead of $C_k, c-j$ coefficients that we use. We note that this work was done completely independently. In their paper Grassmann and Tavakoli conclude that their formula (31) is numerically unstable because of subtraction and factorials. We avoid these instabilities by using the more numerically stable form given by (5). ∎

Computing the one-step transition probabilities $p(i,j)$ requires the evaluation of the derivatives of the LST of the inter-arrival time distribution functions. This can be done for several distribution functions where a closed form expressions for the derivatives can obtained. This includes the Erlang and the hyper-exponential distribution functions among others.

We compare the complexity of computing $p(i,j)$ using the convolution method given in Theorem 3.1 (iii) and the method based on the Lemma 1.1 (iii). The method using Lemma 1.1 (iii) requires numerical double integration that involves an infinite integral. This would result in questionable accuracy. The method described in Corollary 3.1 (iii) requires two finite sums and computations on the order of $O((i-c)(c-j))$.

B. Arrival-Time Probabilities

We give a recursive algorithm to solve for the departure-time probabilities $\pi(.)$. Using the fact that probability flow across cuts balances (Kelly [8], Lemma 1.4) one can show that for $j = 1, 2, \cdots$,

$$\pi(j)p(j,j+1) \quad = \quad \sum_{k=j+1}^{\infty} \pi(k) \sum_{i=0}^{j} p(k,i)$$

which implies

$$\pi(j) = \sum_{k=j+1}^{\infty} \pi(k)a(k,j)/p(j,j+1) \qquad (6)$$

where $a(k,j) = \sum_{i=0}^{j} p(k,i)$.

One can compute the $\{\pi(.)\}$ recursively using (6). Since the expression on the $r.h.s.$ of (6) involves an infinite sum, we need to truncate to some value $N \geq c$ using an acceptable level of precision ϵ, so that $|\pi(N+1)-\pi(N)| < \epsilon$. Let $|\sigma| < 1$ be the solution of the equation

$$\sigma = \int_0^{\infty} e^{-c\mu(1-\sigma)t} dA(t) \equiv A^*(c\mu(1-\sigma)) \ , \qquad (7)$$

then it follows from Gross and Harris [6] that $\pi(n) = K\sigma^n$ for $n \geq c$ where K is a constant. Now, we describe the main steps to compute the stationary probabilities $\{\pi(j); j \geq 0\}$.

1) Initialize: Determine ϵ, c, the inter-arrival times distribution function; and compute σ using equation (7) .
2) Compute the one transition probabilities: compute $p(i,j)$; $c \leq j \leq i+1$ using Theorem 3.1 (ii); compute $p(i,j)$; $j \leq i+1 \leq c$ using Theorem 3.1 (i); compute $p(i,j); j+1 \leq c \leq i$ using Theorem 3.1 (iii); and compute $a(k,j) = \sum_{i=0}^{j} p(k,i)$; for $j = 0, \cdots N, k = j+1, \cdots N$.
3) For $j = N, N-1, \cdots, c$, compute $\pi'(j) = \sigma^j$, and for $j = c-1, \cdots, 0$, compute $\pi'(j) = \sum_{k=j+1}^{N} \pi'(k)a(k,j)/p(j,j+1)$.
4) Normalize: set $\pi(j) = \frac{\pi'(j)}{\sum_{k=0}^{N} \pi'(k)}$.

We point out an alternative method would be to use the global balance equations (GBE) and solve recursively similar to the above approach. This method is described by Gross and Harris [6]. However, due to subtractions and division in GBE, our method is more numerically stable than the method based on the global balance equations as shown by Stidham [15].

C. Time-Average Probabilities

Having determined the transition probability matrix for the Markov Chain $\{X_n, n = 0, \ldots\}$, the limiting probabilities $\pi(i), i = 0, \ldots,$ are evaluated from the steady state convergence theorem by solving $\pi P = \pi; \sum_{i=0}^{\infty} \pi(i) = 1$. However, we are interested in time-average stationary distribution defined as $p(i) = \lim_{t \to \infty} P\{X(t) = i\}, i = 0, \ldots, .$ where $\{X(t), t \geq\}$ represents the number of customers present at time t. Moreover we are interested in determining the performance measures of the system.

Lemma 3.2: Consider the $G/M/c/N$ queueing model and let $\rho = \lambda/c\mu$. Then the system size probabilities, $p(i), i = 0, \ldots N$, are given by

$$p(n) = \begin{cases} (1-\rho) + \rho\pi(N) - & \rho \sum_{k=0}^{c-2} \frac{c-k-1}{k+1} \pi(k), \\ & n = 0 \ ; \\ c\rho\pi(n-1)/n \ , & 1 \leq n \leq c \ ; \\ \rho\pi(n-1) \ , & c < n \leq N \ . \end{cases}$$

Proof. We know that (El-Taha and Stidham [3])

$$\lambda\pi(n) = \lambda_n p(n) = \mu_{n+1} p(n+1) \ .$$

Therefore,

$$p(n) = \lambda\pi(n-1)/\mu_n \ ; n \geq 1 \ ;$$

where $\mu_n = \min(n, c)\mu$. Thus, $p(n) = c\rho\pi(n-1)/n$ for $1 \leq n \leq c$, and $p(n) = \rho\pi(n-1)$ for $n > c$. We still need to compute $p(0)$. Normalize the $p(i)'s$ to obtain

$$\begin{aligned} 1 - p(0) &= \sum_{k=1}^{N} p(k) \\ &= c\rho \sum_{k=0}^{c-1} \pi(k)/(k+1) + \rho[1 - \sum_{k=0}^{c-1} \pi(k) - \pi(N)] \ . \end{aligned}$$

Therefore

$$\begin{aligned} p(0) &= (1-\rho) - c\rho \sum_{k=0}^{c-1} \pi(k)/(k+1) \\ &+ \ \rho \sum_{k=0}^{c-1} \pi(k) + \rho\pi(N) \\ &= (1-\rho) - \rho \sum_{k=0}^{c-2} \frac{c-k-1}{k+1} \pi(k) + \rho\pi(N) \ ; \end{aligned}$$

which completes the proof. ∎

For the $G/M/c/\infty$ model, i.e. when N is infinite, $p(0)$ is computed as

$$p(0) = (1-\rho) - \rho \sum_{k=0}^{c-2} \frac{c-k-1}{k+1} \pi(k) \ . \tag{8}$$

With $p(i)$'s determined from Lemma 3.2, other measures of performance are now determined easily. For instance, mean number of customers L is given by $E[L] = \sum_{i=1}^{\infty} ip(i)$. Little's law implies that the mean delay in the system is $W = L/\lambda$.

Numerical results are performed for the $G/M/c/N$ model using high traffic intensities on small and large scale problems. We use service time distribution functions with coefficients of variation that range from 0 (deterministic) to one (exponential) to very large values (hyper-exponential). For the small scale problems with less than 10 servers, our initial results are highly accurate and stable. For the large scale problems we are able to achieve high accuracy for problems with high traffic intensity ($\rho > 0.95$) and $c = 30$ servers, and system capacity N ranging from 400 to 1300 depending on the service time distribution function.

In this article we describe a convolution method to compute stationary transition probabilities using summations to replace integration. A stable method to compute the arrival-time and time-average probabilities of $G/M/c$ model is also proposed. We plan to investigate the computational aspects of this approach and expanding it to the $G^X/M/c/N$ model.

References

[1] M.L. Chaudhry and J.J. Kim. Analytically elegant and computationally efficient results in terms of roots for the $GI^X/M/c$ queueing system. *Queueing Systems*, 82:237–257, 2016.

[2] G. Cosmetatos and S. Godsave. Approximations in the multi-server queue with hyper-exponential inter-arrival times and exponential service times. *Journal of the Operational Research Society*, 31(1):57–62, 1980.

[3] M. El-Taha and S. Stidham Jr. *Sample-Path Analysis of Queueing Systems*. Kluwer Academic Publishing, Boston, 1999.

[4] F. Ferreira and A. Pacheco. Analysis of GI/M/s/c queues using uniformisation. *Computers and Mathematics with Applications*, 51:291–304, 2006.

[5] W. Grassmann and J. Tavakoli. Efficient methods to find the equilibrium distribution of the number of customers in GI/M/c queues. *IFOR*, 52:197–205, 2014.

[6] D. Gross, J.F. Shortle, J.M. Thompson, and C. Harris. *Fundamentals of Queueing Theory*. John Wiley, New Jersey, 4th edition, 2008.

[7] Per Hokstad. The G/M/m queue with finite waiting room. *Journal of Applied Probability*, 12(4):779–792, 1975.

[8] F. Kelly. *Reversibility and Stochastic Networks*. Wiley, New York, 1979.

[9] J.J. Kim and M.L. Chaudhry. A novel way of treating the finite-buffer queue $GI/M/c/N$ using roots. *International Journal of Mathematical Models and Methods in Applied Sciences*, 11, 2017.

[10] L. Kleinrock. *Queueing Systems, vol. I.* Wiley Intersciences, New York, 1975.

[11] P. V. Laxmi and U.C. Gupta. Analysis of finite-buffer multi-server queues with group arrivals: $GI^X/M/c/N$. *Queueing Systems*, 36(1-3):125–140, 2000.

[12] J. Medhi. *Stochastic Models in Queueing Theory*. Academic Press, New York, 2nd edition, 2003.

[13] M.F. Neuts. *Matrix-Geometric Solutions in Stochastic Models-An Algorithmic Approach*. Johns Hopkins University Press, 1981.

[14] S.M. Ross. *Introduction to Probability Models*. Academic Press, San Diego, 9th edition, 2007.

[15] S. Stidham Jr. Stable recursive procedures for numerical computations in Markov models. *Annals of Operations Research*, 8:27–40, 1987.

[16] L. Takács. On a queueing problem concerning telephone traffic. *Acta Math.Acad.Sci.Hungar.*, 8:325–335, 1957.

[17] L. Takács. *Introduction to the theory of queues*. Oxford University Press, New York, 1962.

[18] H.C. Tijms. *A First Course in Stochastic Models*. Wiley, New York, 2003.

[19] D.D. Yao, M.L. Chaudhry, and J.G.C. Templeton. A note on some relations in the queue $GI^X/M/c$. *Operations Research Letters*, 3(1):53–56, 1984.