Applied Probability

School of Mathematics and Statistics, University of Sheffield, Sheffield S3 7RH, UK
Telephone: +44 114 222 3920
Fax: +44 114 222 3926
Http: www.appliedprobability.org

Journal of Applied Probability
Advances in Applied Probability
The Mathematical Scientist
Mathematical Spectrum

J. Appl. Prob. **52,** 941–961 (2015)
Printed in England

TOLLBOOTH TANDEM QUEUES WITH INFINITE HOMOGENEOUS SERVERS

XIULI CHAO,* *University of Michigan*
QI-MING HE,** *University of Waterloo*
SHELDON ROSS,*** *University of Southern California*

Abstract

In this paper we analyze a tollbooth tandem queueing problem with an infinite number of servers. A customer starts service immediately upon arrival but cannot leave the system before all customers who arrived before him/her have left, i.e. customers depart the system in the same order as they arrive. Distributions of the total number of customers in the system, the number of departure-delayed customers in the system, and the number of customers in service at time t are obtained in closed form. Distributions of the sojourn times and departure delays of customers are also obtained explicitly. Both transient and steady state solutions are derived first for Poisson arrivals, and then extended to cases with batch Poisson and nonstationary Poisson arrival processes. Finally, we report several stochastic ordering results on how system performance measures are affected by arrival and service processes.

Keywords: Tollbooth tandem queue; departure delay; departure-delayed customer

2010 Mathematics Subject Classification: Primary 60K25
Secondary 90B22

1. Introduction

In this paper we consider an infinite-sever queueing system where customers must depart in the order they arrive. This model is motivated by the tandem tollbooth in transportation system, where servers are ordered in series and an arriving vehicle needs to be served by one and only one server (see Hall and Daganzo (1983)), but cannot leave until all vehicles ahead of it have departed. Hence, we refer to this queueing system as an infinite-server tollbooth tandem queue.

The tollbooth tandem queue is similar to the classical tandem queue with infinite stages and no waiting space between stages (Papadopoulos and O'Kelly (1993)) as both have an infinite number of servers in series. But it differs in that each customer only requires service in one stage (server). The tollbooth tandem queue is similar to the classical M/G/∞ queue (Ross (2010, Chapter 5), and Eliazar (2007)) in that there are an infinite number of servers available so that no customer has to wait for service, but it differs from the classical M/G/∞ queue in that servers are lined up in serial, not in parallel; hence, a customer cannot leave the system unless all customers who arrived earlier have completed service. The tollbooth tandem queue also differs from some of the studies on M/G/∞ queues with synchronized departures,

Received 17 December 2013; revision received 22 November 2014.
* Postal address: Department of Industrial and Operations Engineering, University of Michigan, Ann Arbor, MI 48109, USA. Email address: xchao@umich.edu
** Postal address: Department of Management Sciences, University of Waterloo, 200 University Avenue West, Waterloo, ON N2L 3G1, Canada. Email address: q7he@uwaterloo.ca
*** Postal address: Daniel J. Epstein Department of Industrial and Systems Engineering, University of Southern California, Los Angeles, CA 90089, USA. Email address: smross@usc.edu

for example, Zazanis (2004), studied an infinite server queue where departures are synchronized by a given point process; in the tollbooth tandem queue, the departure of a customer is affected by departures of the customers ahead of him/her which depends on the state of the system. The tollbooth tandem queue is also different from the asymmetric simple inclusion process (ASIP) studied by, for example, Reuveni *et al.* (2012), (2014), in which each server opens its gate to accept customers from an up-stage server according to a Poisson process. In this paper we shall obtain all the quantities of interest related to this tollbooth tandem queue, both transient and steady state solutions, in closed form, when all servers are homogeneous. The case of finite heterogeneous servers has been studied by He and Chao (2014) using a matrix-geometric computational method.

The tollbooth tandem queue can be applied in the analysis of a one-way single-lane highway with no overtaking. Consider a segment of the single-lane highway of length, say H, and vehicles enter the starting point of this segment according to a Poisson process with rate λ. Each vehicle entering is attached a random value V representing the velocity of this vehicle, with the proviso that whenever the vehicle encounters a slower moving vehicle it must decrease its velocity to that of the slower moving vehicle. Let V_i denote the velocity of the ith entering vehicle and suppose $V_1, V_2, \dots,$ are independent and identically distributed (i.i.d.) with distribution function $V(t)$. Therefore, if there is no vehicle on the road then the time it takes vehicle i to pass this road segment is H/V_i with distribution function $G(t) = 1 - V(L/t)$. It is not hard to see that the distribution of the number of vehicles on this segment of the highway is precisely that of the number of customers in the tollbooth tandem queue described above. Therefore, using the results for the tollbooth tandem queue derived in this paper, we can obtain such quantities as the distribution of the number of vehicles on the segment of interest on the road, the distribution of vehicle traversal time on this road segment, and others.

We assume that the customer service time has an arbitrary distribution. We are concerned with the probability distributions of the following quantities of interest:

(i) $Q(t)$—the number of customers in the system at time t;

(ii) $Q_{\mathrm{d}}(t)$—the number of departure-delayed customers in the system (who finished service but cannot depart) at time t;

(iii) $D(t)$—the number of departures by time t;

(iv) $W(t)$—the time spent in the system by a customer arrived at time t (i.e. the sojourn time);

(v) $W_{\mathrm{d}}(t)$—the departure delay (time waiting to depart after finishing service) of a customer arrived at time t; and

(vi) $L(t)$—the number of customers left behind by a departing customer who arrived at time t.

We first consider the case when the customer arrival process is Poisson, and obtain closed form solutions for (i)–(vi), as well as the corresponding steady state distributions, in Section 2. Then, in Section 3, we extend the results to the cases of batch Poisson and nonhomogeneous Poisson arrival processes. Some stochastic ordering results are reported in Section 4 before we conclude the paper in Section 5.

2. The M/G/∞ tollbooth tandem queue

We consider a tandem queue with an infinite number of servers in series, denoted by $1, 2, \dots$. All servers are identical, and there is no waiting space between servers. Customers arrive according to a Poisson process with rate λ. An arriving customer joins server j if server $j-1$ is occupied, but all servers with indices j or higher are available. Thus, any arriving customer receives service immediately upon its arrival. On the other hand, a customer finishing service on server j can leave the queueing system only if there are no customers occupying servers $1, 2, \ldots, j-1$. That is, a customer may have to wait on server j until all customers ahead of him/her are cleared, leading to a batch departure process.

Denote by $N(t)$ the number of customers who arrived up to time t with $N(0) = 0$. Let $S_0 = 0$, and let S_i be the arrival time of the ith customer, $i = 1, 2, \ldots$. Let the service time distribution for all customers (servers) be denoted by $G(\cdot)$. Let $\bar{G}(t) = 1 - G(t)$. Let X_i be the service time of the ith customer, $i = 1, 2, \ldots$, and X a generic service time, then $\mathbb{E}[X] = \int_0^\infty \bar{G}(s)\,\mathrm{d}s$. Throughout this paper we assume that the queue is empty at time 0.

Theorem 1. *For all $t \geq 0$, we have*

$$\mathbb{P}\{Q(t) = 0\} = \exp\biggl(-\lambda \int_0^t \bar{G}(s)\,\mathrm{d}s\biggr), \tag{1}$$

$$\mathbb{P}\{Q(t) = n\} = \lambda \int_0^t \frac{(\lambda u)^{n-1}}{(n-1)!}\bar{G}(u)\exp\biggl(-\lambda\biggl(u + \int_u^t \bar{G}(s)\,\mathrm{d}s\biggr)\biggr)\,\mathrm{d}u, \qquad n > 0, \tag{2}$$

$$\begin{aligned}\mathbb{P}\{Q_{\mathrm{d}}(t) = 0\} &= \exp\biggl(-\lambda \int_0^t \bar{G}(s)\,\mathrm{d}s\biggr) \\ &\quad + \lambda \int_0^t \bar{G}(u)\exp\biggl(-\lambda\biggl(\int_0^u G(s)\,\mathrm{d}s + \int_u^t \bar{G}(s)\,\mathrm{d}s\biggr)\biggr)\,\mathrm{d}u,\end{aligned} \tag{3}$$

$$\begin{aligned}\mathbb{P}\{Q_{\mathrm{d}}(t) = n\} &= \lambda \int_0^t \frac{(\lambda \int_0^u G(s)\,\mathrm{d}s)^n}{n!}\bar{G}(u) \\ &\qquad \times \exp\biggl(-\lambda\biggl(\int_0^u G(s)\,\mathrm{d}s + \int_u^t \bar{G}(s)\,\mathrm{d}s\biggr)\biggr)\,\mathrm{d}u, \qquad n > 0,\end{aligned} \tag{4}$$

$$\begin{aligned}\mathbb{P}\{D(t) = n\} &= \lambda \int_0^t \frac{(\lambda \int_u^t G(s)\,\mathrm{d}s)^n}{n!}\bar{G}(u)\mathrm{e}^{-\lambda(t-u)}\,\mathrm{d}u \\ &\quad + \frac{(\lambda \int_0^t G(s)\,\mathrm{d}s)^n}{n!}\mathrm{e}^{-\lambda t}, \qquad n \geq 0,\end{aligned} \tag{5}$$

$$\mathbb{P}\{W(t) \leq x\} = G(x)\exp\biggl(-\lambda \int_x^{t+x} \bar{G}(u)\,\mathrm{d}u\biggr), \tag{6}$$

$$\mathbb{P}\{W_{\mathrm{d}}(t) \leq x\} = \int_0^\infty \exp\biggl(-\lambda \int_{u+x}^{u+x+t} \bar{G}(s)\,\mathrm{d}s\biggr)\,\mathrm{d}G(u), \tag{7}$$

$$\begin{aligned}\mathbb{P}\{L(t) = n\} &= \int_0^\infty \frac{(\lambda x)^n}{n!}\exp\biggl(-\lambda x - \lambda \int_x^{x+t} \bar{G}(u)\,\mathrm{d}u\biggr)\,\mathrm{d}G(x) \\ &\quad + \lambda \int_0^\infty \frac{(\lambda x)^n}{n!}G(x)(G(x+t) - G(x)) \\ &\qquad \times \exp\biggl(-\lambda x - \lambda \int_x^{x+t} \bar{G}(u)\,\mathrm{d}u\biggr)\,\mathrm{d}x, \qquad n \geq 0.\end{aligned} \tag{8}$$

In the rest of this section we prove (1)–(8), and present additional results such as the steady state version of the above results and the means and probability generating functions (PGFs) of the quantities of interest.

To prove these results we first note that there is a close relationship between the M/G/∞ tollbooth tandem queue and the classical M/G/∞ queue. First, it is easy to see that busy periods of the two queues are identical. Consequently, the probability that the queue is empty is the same for the two queues. In fact, the queue length in the classical M/G/∞ queue is equal in distribution to the number of working servers in the tollbooth tandem queue. Thus, if we let $Q_\mathrm{b}(t)$ represent the number of servers that are working in the tollbooth tandem queue at time t, then it has the same distribution as that of the queue length in the classical queue at time t. In the tollbooth tandem queue, in addition to $Q_\mathrm{b}(t)$, there are other customers whose services have been completed but cannot leave the system because one or more customers who arrived earlier are still in service; hence, $Q(t)$ is greater than or equal to $Q_{b(t)}$. The distribution of $Q_\mathrm{b}(t)$ is well known (see, e.g. Ross (2010, Chapter 5)), and is given by

$$\mathbb{P}\{Q_\mathrm{b}(t)=n\}=\frac{(\lambda\int_0^t \bar G(s)\,\mathrm{d}s)^n}{n!}\exp\biggl(-\lambda\int_0^t \bar G(s)\,\mathrm{d}s\biggr)\quad \text{for } n=0,1,2,\ldots.$$

Proof of (1). The probability that the tollbooth tandem system is empty at time t is the same as $\mathbb{P}\{Q_\mathrm{b}(t)=0\}$; thus, we obtain (1). Letting t go to ∞, we obtain the steady state probability for the tollbooth tandem queue to be empty given by $\mathrm{e}^{-\rho}$, where $\rho=\lambda\int_0^\infty \bar G(s)\,\mathrm{d}s$.

Proof of (2). To obtain $\mathbb{P}\{Q(t)=n\}$ for $n>0$ and the distributions of other quantities, we make use of a simple result on sampling of Poisson processes. For a fixed time t, we define an arrival as a type 1 event if it enters the system at some time and it would be in the system at time t even if no customers arrived before him/her. In other words, an arrival at time s, $s<t$, is type 1 if its service time is greater than $t-s$. Clearly, an arrival at time s is a type 1 event with probability $\mathbb{P}_1(s)=\bar G(t-s)$ if $s\le t$ and $\mathbb{P}_1(s)=0$ if $s>t$. Letting $N_{1,t}(y)$ denote the number of type 1 events in $[0,y]$, then it follows from Ross (2010, Proposition 5.3) that $N_{1,t}(y)$ has a Poisson distribution with mean

$$\mathbb{E}[N_{1,t}(y)]=\lambda\int_0^y \mathbb{P}_1(s)\,\mathrm{d}s=\begin{cases}\lambda\displaystyle\int_0^y \bar G(t-s)\,\mathrm{d}s, & y\le t,\\ \lambda\displaystyle\int_0^t \bar G(t-s)\,\mathrm{d}s, & y>t.\end{cases}$$

Let Y be the time the first type 1 event takes place, which is defined as ∞ if there are no type 1 events in $[0,t]$. The distribution of Y is

$$F_Y(y)=\mathbb{P}\{Y\le y\}=\mathbb{P}\{N_{1,t}(y)>0\}=\begin{cases}1-\exp\biggl(-\lambda\displaystyle\int_0^y \bar G(t-s)\,\mathrm{d}s\biggr), & 0\le y<t,\\ 1-\exp\biggl(-\lambda\displaystyle\int_0^t \bar G(t-s)\,\mathrm{d}s\biggr), & t\le y<\infty.\end{cases}$$

Differentiating gives the density function of Y as

$$f_Y(y)=\begin{cases}\lambda\bar G(t-y)\exp\biggl(-\lambda\displaystyle\int_0^y \bar G(t-s)\,\mathrm{d}s\biggr), & 0\le y<t,\\ 0, & t\le y<\infty.\end{cases}$$

Note that Y is not a regular random variable and it takes value ∞ with probability

$$\mathbb{P}\{Y=\infty\}=\mathbb{P}\{N_{1,t}(t)=0\}=\exp\biggl(-\lambda\int_0^t \bar G(s)\,\mathrm{d}s\biggr).$$

Clearly, if $Y=\infty$ then the system is empty at time t.

It is easy to see that the system is empty at time t if and only if $N_{1,t}(t) = 0$. Thus, $\mathbb{P}\{Q(t) = 0\} = \mathbb{P}\{N_{1,t}(t) = 0\} = \exp(-\lambda \int_0^t \bar{G}(s)\,\mathrm{d}s)$, which is the same as (1) obtained above using the relationship with the classic M/G/∞ queue. To obtain $\mathbb{P}\{Q(t) = n\}$ when $n > 0$, we condition on Y. If $Y = y \le t$ then the customer who arrives at time y is the earliest arrival among the customers still in the system at time t. This implies that all the arrivals who enter the system between y and t will also be in the system at time t. It follows that conditional on $Y = y$, $Q(t)$ is distributed as 1 plus a Poisson random variable with mean $\lambda(t - y)$. Hence, by conditioning on Y we obtain, for any $n > 0$ and $0 \le y \le t$,

$$\mathbb{P}\{Q(t) = n \mid Y = y\} = \mathrm{e}^{-\lambda(t-y)} \frac{(\lambda(t-y))^{n-1}}{(n-1)!},$$

and $\mathbb{P}\{Q(t) = n \mid Y = \infty\} = 0$ for $n > 0$. Unconditioning, we obtain (2). The PGF and the mean of $Q(t)$ can be obtained as

$$\begin{aligned}
\mathbb{E}[z^{Q(t)}] &= \exp\biggl(-\lambda \int_0^t \bar{G}(s)\,\mathrm{d}s\biggr) \\
&\quad + \lambda z \int_0^t \bar{G}(u) \exp\biggl(-\lambda\biggl(u(1-z) + \int_u^t \bar{G}(s)\,\mathrm{d}s\biggr)\biggr)\,\mathrm{d}u, \qquad 0 \le z \le 1, \\
\mathbb{E}[Q(t)] &= \lambda \int_0^t (1 + \lambda u)\bar{G}(u) \exp\biggl(-\lambda \int_u^t \bar{G}(s)\,\mathrm{d}s\biggr)\,\mathrm{d}u \\
&= \lambda t + 1 - \exp\biggl(-\lambda \int_0^t \bar{G}(s)\,\mathrm{d}s\biggr) - \lambda \int_0^t \exp\biggl(-\lambda \int_u^t \bar{G}(s)\,\mathrm{d}s\biggr)\,\mathrm{d}u.
\end{aligned}$$

Letting $t \to \infty$, we obtain the steady state distribution of the number of customers, denoted by Q, in the system:

$$\mathbb{P}\{Q = n\} = \begin{cases} \exp\biggl(-\lambda \displaystyle\int_0^\infty \bar{G}(s)\,\mathrm{d}s\biggr) & \text{for } n = 0, \\ \lambda \displaystyle\int_0^\infty \frac{(\lambda u)^{n-1}}{(n-1)!} \bar{G}(u) \exp\biggl(-\lambda\biggl(u + \int_u^\infty \bar{G}(s)\,\mathrm{d}s\biggr)\biggr)\,\mathrm{d}u & \text{for } n > 0. \end{cases} \tag{9}$$

The steady state average number of customers in the system is

$$\mathbb{E}[Q] = \lambda \int_0^\infty (1 + \lambda u)\bar{G}(u) \exp\biggl(-\lambda \int_u^\infty \bar{G}(s)\,\mathrm{d}s\biggr)\,\mathrm{d}u.$$

Proofs of (3) and (4). We study the number of departure-delayed customers in the system at time t, that is, those customers who have finished service but cannot depart due to customers ahead of them. To that end, we condition on the arrival time of the first type 1 event, Y, defined earlier. Conditioning on $Y = y$, or the first customer whose service would not finish by time t occurs at y, then all customers who arrived before y would have finished service by time t, implying that they would have all departed the system by time t. This implies that the number of departure-delayed customers at time t is precisely the number of arrivals between y and t that have finished service by time t (thus delayed by the first type 1 customer). In addition, if $Y = y < t$ then the number of customers who arrive between y and t and finish service by time t has the same distribution as the number of departures in M/G/∞ queue at time $t - y$, which has a Poisson distribution with parameter $\lambda\int_0^{t-y} G(u)\,\mathrm{d}u$. If $Y \geq t$ then there is no departure-delayed customers at time t or $Q_{\mathrm{d}}(t) = 0$. Hence, by conditioning, we obtain

$$\begin{aligned}\mathbb{P}\{Q_{\mathrm{d}}(t) = 0\} &= \mathbb{P}\{Y \geq t\} + \int_0^t \mathbb{P}\{Q_{\mathrm{d}}(t) = 0 \mid Y = y\} f_Y(y)\,\mathrm{d}y \\ &= \exp\biggl(-\lambda\int_0^t \bar{G}(s)\,\mathrm{d}s\biggr) \\ &\quad + \lambda\int_0^t \bar{G}(t-y)\exp\biggl(-\lambda\int_0^{t-y} G(s)\,\mathrm{d}s - \lambda\int_0^y \bar{G}(t-s)\,\mathrm{d}s\biggr)\,\mathrm{d}y,\end{aligned}$$

and for $n > 0$, we have

$$\begin{aligned}\mathbb{P}\{Q_{\mathrm{d}}(t) = n\} = \lambda\int_0^t &\frac{(\lambda\int_0^{t-y} G(s)\,\mathrm{d}s)^n}{n!}\bar{G}(t-y) \\ &\times \exp\biggl(-\lambda\int_0^{t-y} G(s)\,\mathrm{d}s - \lambda\int_0^y \bar{G}(t-s)\,\mathrm{d}s\biggr)\,\mathrm{d}y.\end{aligned}$$

Using a change of variable $u = t - y$, we obtain (3) and (4).

It is interesting to compare the expressions for $\mathbb{P}\{Q(t) = k\}$ and $\mathbb{P}\{Q_{\mathrm{d}}(t) = k\}$. The main difference is that x is replaced by $\int_0^x G(s)\,\mathrm{d}s$. Since $x = \int_0^x G(s)\,\mathrm{d}s + \int_0^x \bar{G}(s)\,\mathrm{d}s$, the difference in the two expressions can be explained as follows. For $\mathbb{P}\{Q(t) = n\}$, both customers in service (with probability $\bar{G}(s)$) and completed service (with probability $G(s)$) are included. For $\mathbb{P}\{Q_{\mathrm{d}}(t) = n\}$, only customers that completed service (with probability $G(s)$) are included. Therefore, in the expressions, $1 = G(s) + \bar{G}(s)$ is replaced by $G(s)$, or x is replaced by $\int_0^x G(s)\,\mathrm{d}s$.

The PGF and the mean of $Q_{\mathrm{d}}(t)$ are given by

$$\begin{aligned}\mathbb{E}[z^{Q_{\mathrm{d}}(t)}] &= \exp\biggl(-\lambda\int_0^t \bar{G}(s)\,\mathrm{d}s\biggr) \\ &\quad + \lambda\int_0^t \bar{G}(u)\exp\biggl(-\lambda\biggl((1-z)\int_0^u G(s)\,\mathrm{d}s + \int_u^t \bar{G}(s)\,\mathrm{d}s\biggr)\biggr)\,\mathrm{d}u, \quad 0 \leq z \leq 1,\end{aligned}$$

$$\mathbb{E}[Q_{\mathrm{d}}(t)] = \lambda\int_0^t G(u)\biggl(1 - \exp\biggl(-\lambda\int_u^t \bar{G}(s)\,\mathrm{d}s\biggr)\biggr)\,\mathrm{d}u.$$

Letting $t \to \infty$, we obtain the steady state distribution of the number of departure-delayed customers Q_{d}:

$$\mathbb{P}\{Q_{\mathrm{d}} = 0\} = \mathrm{e}^{-\rho} + \lambda\int_0^\infty \bar{G}(u)\exp\biggl(-\lambda\biggl(\int_0^u G(s)\,\mathrm{d}s + \int_u^\infty \bar{G}(s)\,\mathrm{d}s\biggr)\biggr)\,\mathrm{d}u,$$

and, for $n > 0$,

$$\mathbb{P}\{Q_{\mathrm{d}} = n\} = \lambda\int_0^\infty \frac{(\lambda\int_0^u G(s)\,\mathrm{d}s)^n}{n!}\bar{G}(u)\exp\biggl(-\lambda\biggl(\int_0^u G(s)\,\mathrm{d}s + \int_u^\infty \bar{G}(s)\,\mathrm{d}s\biggr)\biggr)\,\mathrm{d}u.$$

The steady state average number of departure-delayed customers is

$$\mathbb{E}[Q_{\mathrm{d}}] = \lambda\int_0^\infty G(u)\biggl(1 - \exp\biggl(-\lambda\int_u^\infty \bar{G}(s)\,\mathrm{d}s\biggr)\biggr)\,\mathrm{d}u.$$

Proof of (5). We now study the distribution of $D(t)$, i.e. the number of customers who depart the tollbooth tandem queue by time t. To that end, we define, for a given time t, a customer as of type 2 if he/she would finish service by time t (though he/she may or may not be able to depart), and let $N_{2,t}(y)$ denote the number of type 2 customers by time $y \leq t$. Clearly, a customer arriving at time $s \leq t$ is of a type 2 with probability $G(t-s)$, independent of everything else. The number of type 2 events arrived between 0 and y has a Poisson distribution with mean $\lambda \int_0^y G(t-s)\,\mathrm{d}s$. By Ross (2010, Proposition 5.3), we know that $N_{1,t}(y)$ and $N_{2,t}(y)$ are independent Poisson random variables. Conditional on the first type 1 event arriving at time $Y = y < t$, then all arrivals in $[0, y]$ have departed by time t. Since $D(t) = N_{1,t}(y) + N_{2,t}(y)$, given $Y = y < t$, we must have $D(t) = N_{2,t}(y)$. If $Y = \infty$ then all arrivals in $[0, t]$ have departed by time t and $D(t) = N_{2,t}(t)$. Thus, by conditioning on Y, we obtain, for all $n \geq 0$,

$$
\begin{aligned}
&\mathbb{P}\{D(t) = n\} \\
&\quad = \int_0^t \exp\biggl(-\lambda \int_0^y G(t-s)\,\mathrm{d}s\biggr) \frac{(\lambda \int_0^y G(t-s)\,\mathrm{d}s)^n}{n!} f_Y(y)\,\mathrm{d}y \\
&\qquad + \exp\biggl(-\lambda \int_0^t G(s)\,\mathrm{d}s\biggr) \frac{(\lambda \int_0^t G(s)\,\mathrm{d}s)^n}{n!} \mathbb{P}\{Y > t\} \\
&\quad = \lambda \int_0^t \frac{(\lambda \int_0^y G(t-s)\,\mathrm{d}s)^n}{n!} \bar{G}(t-y)\mathrm{e}^{-\lambda y}\,\mathrm{d}y + \frac{(\lambda \int_0^t G(s)\,\mathrm{d}s)^n}{n!} \mathrm{e}^{-\lambda t}.
\end{aligned}
$$

This establishes (5). The PGF and the mean of $D(t)$ are

$$
\begin{aligned}
\mathbb{E}[z^{D(t)}] &= \lambda \int_0^t \bar{G}(u) \exp\biggl(-\lambda \int_u^t (1 - zG(s))\,\mathrm{d}s\biggr)\,\mathrm{d}u \\
&\qquad + \exp\biggl(-\lambda \int_0^t (1 - zG(s))\,\mathrm{d}s\biggr), \qquad 0 \leq z \leq 1, \\
\mathbb{E}[D(t)] &= \lambda \int_0^t G(u) \exp\biggl(-\lambda \int_u^t \bar{G}(s)\,\mathrm{d}s\biggr)\,\mathrm{d}u \\
&= \lambda \int_0^t \exp\biggl(-\lambda \int_u^t \bar{G}(s)\,\mathrm{d}s\biggr)\,\mathrm{d}u + \exp\biggl(-\lambda \int_0^t \bar{G}(s)\,\mathrm{d}s\biggr) - 1.
\end{aligned}
$$

Note that $\mathbb{E}[N(t)] = \mathbb{E}[Q(t)] + \mathbb{E}[D(t)]$, which has to hold since $N(t) = Q(t) + D(t)$.

Proof of (6). We next compute the distribution of the time a customer spends in the system, that is, the customer's sojourn time. As $W(s)$ is the amount of time a customer spends in the system when he/she arrives at time s, if we let X denote this customer's service time, then $W(s) \leq x$ if and only if the following two conditions hold:

(i) $X \leq x$;

(ii) any customer who arrived before s has departed by time $s + x$ or $N_{1,s+x}(s) = 0$.

This argument leads to

$$
\mathbb{P}\{W(s) \leq x\} = \mathbb{P}\{X \leq x,\, N_{1,s+x}(s) = 0\} = G(x) \exp\biggl(-\lambda \int_0^s \bar{G}(s + x - u)\,\mathrm{d}u\biggr).
$$

Changing variables gives (6). Letting $s \to \infty$, we obtain the distribution of the steady state sojourn time, denoted by W, of an arrival as

$$\mathbb{P}\{W \le x\} = G(x) \exp\biggl(-\lambda \int_x^\infty \bar{G}(u)\biggr) \, \mathrm{d}u.$$

Proof of (7). We next analyze the distribution of the departure delay $W_{\mathrm{d}}(t)$ of a customer who arrives at time t. Clearly, $W_{\mathrm{d}}(t)$ is the difference between the departure time and the service completion time. A customer that arrives at time t (to be called a tagged customer) finishes service at time $t + X$, where X is its service time. The tagged customer can leave the system only after all customers who arrive before him/her have completed service. It is clear that the last departure time of the customers who arrive before t is $\max_{1 \le i \le N(t)}\{S_i + X_i\}$. Hence, the tagged customer's departure delay can be expressed as

$$W_{\mathrm{d}}(t) = \Bigl(\max_{1 \le i \le N(t)}\{S_i + X_i\} - t - X\Bigr)^+,$$

where $(x)^+ = \max\{x, 0\}$ for any real number x.

We compute the distribution function of $W_{\mathrm{d}}(t)$ by conditioning on the number of arrivals $N(t)$. Given $N(t) = n$, it is well known that $\{S_1, S_2, \dots, S_n\}$ have the same joint probability distribution as the order statistics of $\{U_1(t), U_2(t), \dots, U_n(t)\}$, which are independent and uniformly distributed on $[0, t]$. Using the fact that $U_i(t)$ and $t - U_i(t)$ have the same distribution, we obtain, for any $x \ge 0$,

$$\begin{aligned}
&\mathbb{P}\{W_{\mathrm{d}}(t) \le x\} \\
&\quad = \mathbb{P}\Bigl\{\max_{1 \le i \le N(t)}\{S_i + X_i\} - t - X \le x\Bigr\} \\
&\quad = \int_0^\infty \mathbb{P}\Bigl\{\max_{1 \le i \le N(t)}\{S_i + X_i\} - t - u \le x\Bigr\} \, \mathrm{d}G(u) \\
&\quad = \int_0^\infty \sum_{n=0}^\infty \mathbb{P}\Bigl\{\max_{1 \le i \le N(t)}\{S_i + X_i\} - t - u \le x \Bigm| N(t) = n\Bigr\} \mathbb{P}\{N(t) = n\} \, \mathrm{d}G(u) \\
&\quad = \int_0^\infty \sum_{n=0}^\infty \mathbb{P}\Bigl\{\max_{1 \le i \le n}\{U_i(t) + X_i\} \le t + u + x\Bigr\} \mathbb{P}\{N(t) = n\} \, \mathrm{d}G(u) \\
&\quad = \int_0^\infty \sum_{n=0}^\infty (\mathbb{P}\{U_1(t) + X_1 \le t + u + x\})^n \frac{(\lambda t)^n}{n!} \mathrm{e}^{-\lambda t} \, \mathrm{d}G(u) \\
&\quad = \int_0^\infty \exp(-\lambda t + \lambda t \mathbb{P}\{U_1(t) + X_1 \le t + u + x\}) \, \mathrm{d}G(u) \\
&\quad = \int_0^\infty \exp\biggl(-\lambda t + \lambda \int_0^t G(t + u + x - s) \, \mathrm{d}s\biggr) \, \mathrm{d}G(u) \\
&\quad = \int_0^\infty \exp\biggl(-\lambda \int_0^t \bar{G}(t + u + x - s) \, \mathrm{d}s\biggr) \, \mathrm{d}G(u) \\
&\quad = \int_0^\infty \exp\biggl(-\lambda \int_0^t \bar{G}(u + x + s) \, \mathrm{d}s\biggr) \, \mathrm{d}G(u) \\
&\quad = \int_0^\infty \exp\biggl(-\lambda \int_{u+x}^{u+x+t} \bar{G}(s) \, \mathrm{d}s\biggr) \, \mathrm{d}G(u),
\end{aligned}$$

where the fifth equality follows from the fact that $\{U_i(t) + X_i, i = 1, 2, \dots, n\}$ are i.i.d., and the last two equalities follow from a change of variable. This proves (7). Letting $t \to \infty$, we obtain the steady state distribution of departure delays

$$\mathbb{P}\{W_{\mathrm{d}} \le x\} = \int_0^\infty \exp\biggl(-\lambda \int_{u+x}^\infty \bar{G}(s)\,\mathrm{d}s\biggr)\,\mathrm{d}G(u).$$

In particular, the probability of having no departure delay is

$$\mathbb{P}\{W_{\mathrm{d}} = 0\} = \int_0^\infty \exp\biggl(-\lambda \int_{u}^\infty \bar{G}(s)\,\mathrm{d}s\biggr)\,\mathrm{d}G(u).$$

Proof of (8). Suppose a tagged customer arrives at time t. The number of customers left behind by the departure of the tagged customer, $L(t)$, is the number of arrivals while the tagged arrival is in the system. That is, $L(t) = N(t + W(t)) - N(t^+)$. Let $L_1(t)$ denote the number of customers in the system immediately after the tagged customer departs, and let $L_2(t)$ denote the number of customers that arrive after but depart at the same moment as the tagged arrival. Then $L(t) = L_1(t) + L_2(t)$.

Note that the sojourn time of the tagged customer $W(t)$ is independent of the service times of all customers arriving after him/her. Therefore, the distribution for the number of customers left behind by this customer is, for $n \ge 0$,

$$\begin{aligned}
\mathbb{P}\{L(t) = n\} &= \int_0^\infty \mathbb{P}\{N(x) = n\}\,\mathrm{d}\mathbb{P}\{W(t) \le x\} \\
&= \int_0^\infty \frac{(\lambda x)^n}{n!} \exp\biggl(-\lambda x - \lambda \int_x^{x+t} \bar{G}(u)\,\mathrm{d}u\biggr)\,\mathrm{d}G(x) \\
&\quad + \lambda \int_0^\infty \frac{(\lambda x)^n}{n!} G(x)(G(x+t) - G(x)) \exp\biggl(-\lambda x - \lambda \int_x^{x+t} \bar{G}(u)\,\mathrm{d}u\biggr)\,\mathrm{d}x.
\end{aligned}$$

This proves (8). By the above relationship, we have $\mathbb{E}[L(t)] = \lambda \mathbb{E}[W(t)]$.

To find the distribution of $L_1(t)$ and $L_2(t)$, we condition on $W(t) = x$. Given $W(t) = x$, each of the customers that arrive during t and $t + x$ can either leave with this customer or stay in the system. A moment of reflection shows that, given $W(t) = x$, the number of customers who will be in the system at time $t + x$ has the same distribution as $Q(x)$, the number of customers at time x in the tollbooth tandem system with infinite servers that starts with 0 customer at time 0, and the number of customers who will leave with this customer has the same distribution as $D(x)$, the number of departures in x units of time in the tollbooth system. Hence, we have

$$\mathbb{P}\{L_1(t) = n\} = \int_0^\infty \mathbb{P}\{Q(x) = n\}\,\mathrm{d}\mathbb{P}\{W(t) \le x\},$$
$$\mathbb{P}\{L_2(t) = n\} = \int_0^\infty \mathbb{P}\{D(x) = n\}\,\mathrm{d}\mathbb{P}\{W(t) \le x\}.$$

Substituting the results from $Q(x)$, $D(x)$, and $W(t)$, we obtain the results for $L_1(t)$ and $L_2(t)$. The details are omitted here.

3. Two extensions

In this section we extend the results obtained in the previous section to two other arrival processes: a batch Poisson arrival process (Section 3.1), and a nonhomogeneous Poisson arrival process (Section 3.2).

3.1. The M^X/G/∞ tollbooth tandem queue

The first extension we study is the case when the arrival process is batch Poisson. Classic infinite server queues with batch Poisson arrival processes have been studied in the literature; see, e.g. Shanbhag (1966) and Holman *et al.* (1983). Suppose that customers arrive in batches at rate λ, and the batch sizes $B_1, B_2, \dots,$ at these arrival epochs are i.i.d. with probability mass function $\mathbb{P}\{B = k\} = b_k,\ k = 1, 2 \dots,$ where B is a generic batch size. Let $B^*(z)$ be the PGF of B, i.e.

$$B^*(z) = \sum_{n=1}^{\infty} b_n z^n.$$

The service times of individual customers are independent random variables with common distribution function $G(\cdot)$, the same as the one defined in Section 2.

Since customers arrive in batches, some new issues arise in the analysis of the quantities defined in Section 1. For instance, to analyze the queue length we need to determine how many customers within a batch have completed their service by a certain time. Thus, the analysis of the quantities of interest is more involved than that in Section 2. For that reason, the distributions of queue length $Q(t)$ and the number of departure-delayed customers $Q_\mathrm{d}(t)$ are given in terms of their PGFs. For waiting time $W(t)$ and departure delay $W_\mathrm{d}(t)$, we have to identify the position of a tagged customer within its batch. In this section we consider the $W(t)$ and $W_\mathrm{d}(t)$ of an arbitrary customer.

Theorem 2. *For all $t \geq 0$, $x \geq 0$, and $0 \leq z \leq 1$, we have*

$$\begin{aligned}\mathbb{E}[z^{Q(t)}] &= \exp\bigg(-\lambda \int_0^t (1 - B^*(G(s))\,\mathrm{d}s\bigg) \\ &\quad + \lambda z \int_0^t \frac{(B^*(G(u)) - B^*(z))}{G(u) - z} \bar{G}(u) \\ &\qquad \times \exp\bigg(-\lambda\bigg(u(1 - B^*(z)) + \int_u^t (1 - B^*(G(s)))\,\mathrm{d}s\bigg)\bigg)\,\mathrm{d}u, \qquad (10)\end{aligned}$$

$$\begin{aligned}\mathbb{E}[z^{Q_\mathrm{d}(t)}] &= \exp\bigg(-\lambda \int_0^t (1 - B^*(G(s))\,\mathrm{d}s\bigg) \\ &\quad + \lambda \int_0^t \frac{(B^*(G(u)) - B^*(zG(u) + \bar{G}(u)))}{G(u) - (G(u)z + \bar{G}(u))} \bar{G}(u) \\ &\qquad \times \exp\bigg(-\lambda\bigg(\int_0^u (1 - B^*(zG(s) + \bar{G}(s)))\,\mathrm{d}s \\ &\qquad\qquad + \int_u^t (1 - B^*(G(s)))\,\mathrm{d}s\bigg)\bigg)\,\mathrm{d}u \qquad (11)\end{aligned}$$

$$\begin{aligned}\mathbb{E}[z^{D(t)}] &= \lambda \int_0^t \frac{(1 - B^*(zG(u)))}{1 - zG(u)} \bar{G}(u) \\ &\qquad \times \exp\bigg(-\lambda \int_u^t (2 - B^*(zG(s) + \bar{G}(s)) - B^*(G(s)))\,\mathrm{d}s\bigg)\,\mathrm{d}u \\ &\quad + \exp\bigg(-\lambda \int_0^t (2 - B^*(G(s)z + \bar{G}(s)) - B^*(G(s)))\,\mathrm{d}s\bigg) \qquad (12)\end{aligned}$$

$$\mathbb{P}\{W(t) < x\} = \frac{G(x)(1 - B^*(G(x)))}{\bar{G}(x)\mathbb{E}[B]} \exp\biggl(-\lambda \int_x^{x+t} (1 - \bar{B}^*(G(s)))\,\mathrm{d}s\biggr) \tag{13}$$

$$\mathbb{P}\{W_\mathrm{d}(t) < x\} = \int_0^\infty \frac{(1 - B^*(G(x+u)))}{\bar{G}(x+u)\mathbb{E}[B]} \exp\biggl(-\lambda \int_{x+u}^{x+u+t} (1 - B^*(G(s)))\,\mathrm{d}s\biggr)\,\mathrm{d}G(u). \tag{14}$$

Proof of (10). In the last section, the computation of distribution of the number of customers in the system and the number of departure-delayed customers in the system depend critically on the sampling of a Poisson process. To use a similar idea for batch Poisson processes, we need to carefully define the classification of events. We define an arrival epoch of a batch as of type 1 if at least one customer from this batch has not finished service at time t. Let X_{ij} be the service time of customer j in batch i, and we introduce the notation

$$\bar{X}_i = \max_{1 \le j \le B_i} X_{ij}.$$

Then, a batch arriving at time s is of type 1 with probability, if $s \le t$,

$$\begin{aligned}
\mathbb{P}_1(s) &= \mathbb{P}\{\bar{X}_i > t - s\} \\
&= \sum_{n=1}^{\infty} \mathbb{P}\{\bar{X}_i > t - s \mid B_i = n\} b_n \\
&= \sum_{n=1}^{\infty} (1 - \mathbb{P}\{\bar{X}_i \le t - s \mid B_i = n\}) b_n \\
&= \sum_{n=1}^{\infty} (1 - (G(t-s))^n) b_n \\
&= 1 - B^*(G(t-s)),
\end{aligned}$$

and $\mathbb{P}_1(s) = 0$ if $s > t$. The number of type 1 events by time $y \le t$, $N_{1,t}(y)$, has a Poisson distribution with mean

$$\mathbb{E}[N_{1,t}(y)] = \lambda \int_0^y \mathbb{P}_1(s)\,\mathrm{d}s = \begin{cases} \lambda \displaystyle\int_0^y (1 - B^*(G(t-s)))\,\mathrm{d}s & \text{if } y \le t, \\ \lambda \displaystyle\int_0^t (1 - B^*(G(t-s)))\,\mathrm{d}s & \text{if } y > t. \end{cases}$$

Similar to the last section, let Y be the time the first type 1 event takes place, defined as ∞ if it never occurs. Then

$$\mathbb{P}\{Y \le y\} = \mathbb{P}\{N_{1,t}(y) > 0\} = \begin{cases} 1 - \exp\biggl(-\lambda \displaystyle\int_0^y (1 - B^*(G(t-s)))\,\mathrm{d}s\biggr) & \text{if } y \le t, \\ 1 - \exp\biggl(-\lambda \displaystyle\int_0^t (1 - B^*(G(t-s)))\,\mathrm{d}s\biggr) & \text{if } y > t. \end{cases}$$

The density function of Y is

$$f_Y(y) = \begin{cases} \lambda(1 - B^*(G(t-y)) \exp\biggl(-\lambda \displaystyle\int_0^y (1 - B^*(G(t-s)))\,\mathrm{d}s\biggr) & \text{if } 0 \le y \le t, \\ 0 & \text{if } y > t. \end{cases}$$

To compute the distribution of the number of customers in the system, the following is needed. Given $\max_{1\le j\le B} X_j > x$, what is the distribution of B? This is obtained as follows:

$$\tilde{b}_n(x) = \mathbb{P}\Big\{B = n \;\Big|\; \max_{1\le j\le B} X_j > x\Big\} = \frac{1-(G(x))^n}{1-B^*(G(x))} b_n, \qquad n = 1, 2, \dots.$$

We also need to compute, given $\max_{1\le j\le B} X_j > x$, the distribution of the number of customers in batch B whose indices are at least as large as the index of the first customer whose service time is longer than x. Given $\max_{1\le j\le B} X_j > x$, the index of the first customer whose service time is longer than x is

$$J(x) = \min\Big\{j;\, X_j > x \;\Big|\; \max_{1\le i\le B} X_i > x\Big\}$$

and it takes value in $\{1, \dots, B\}$. The number of customers in B whose index is at least $J(x)$ can be written as $K(x) = B - J(x) + 1$. It is easy to verify that

$$\mathbb{P}\{J(x) = j \mid B = n\} = \frac{(G(x))^{j-1}\bar{G}(x)}{1-(G(x))^n}, \qquad 1 \le j \le n.$$

Hence,

$$\begin{aligned}
&\mathbb{P}\Big\{K(x) = k \;\Big|\; \max_{1\le i\le B} X_i > x\Big\} \\
&\quad = \mathbb{P}\Big\{J(x) = B + 1 - k \;\Big|\; \max_{1\le i\le B} X_i > x\Big\} \\
&\quad = \sum_{n=1}^{\infty} \mathbb{P}\{J(x) = B + 1 - k \mid B = n\}\mathbb{P}\Big\{B = n \;\Big|\; \max_{1\le i\le B} X_i > x\Big\} \\
&\quad = \sum_{n=k}^{\infty} \mathbb{P}\{J(x) = n + 1 - k \mid B = n\}\frac{1-(G(x))^n}{1-B^*(G(x))} b_n \\
&\quad = \sum_{n=k}^{\infty} \frac{(G(x))^{n-k}\bar{G}(x)}{1-B^*(G(x))} b_n.
\end{aligned}$$

The PGF of $K(x)$ is

$$\begin{aligned}
K^*(z \mid x) &= \sum_{k=1}^{\infty} \mathbb{P}\Big\{K(x) = k \;\Big|\; \max_{1\le i\le B} X_i > x\Big\} z^k \\
&= \frac{z\bar{G}(x)}{(G(x)-z)(1-B^*(G(x)))}(B^*(G(x)) - B^*(z)).
\end{aligned}$$

To compute the PGF of $Q(t)$, we condition on $Y = y$. Given $Y = y$, $Q(t)$ is equal in distribution to $K(t-y)$ plus the number of arrivals between y and t. Hence,

$$\{Q(t) \mid Y = y\} = K(t-y) + \sum_{i=1}^{N(t-y)} B_i.$$

Since all the terms are independent, conditioning on Y, the PGF of $Q(t)$ is

$$\mathbb{E}[z^{Q(t)} \mid Y = y] = K^*(z \mid t-y)\mathrm{e}^{-\lambda(t-y)(1-B^*(z))}$$

and unconditioning we obtain

$$\begin{aligned}\mathbb{E}[z^{Q(t)}] = \lambda z \int_0^t & \frac{\bar{G}(t-y)}{(G(t-y)-z)}(B^*(G(t-y)) - B^*(z)) \\ & \times \exp\left(-\lambda(t-y)(1-B^*(z)) - \lambda \int_0^y (1 - B^*(G(t-s)))\,\mathrm{d}s\right)\mathrm{d}y \\ & + \exp\left(-\lambda \int_0^t (1 - B^*(G(t-s)))\,\mathrm{d}s\right),\end{aligned}$$

which leads to (10) by using the change of variable $u = t - y$. Letting $t \to \infty$, we obtain the steady state distribution of the number of customers in the system. The details are omitted.

Proof of (11). To compute $Q_\mathrm{d}(t)$, we also condition on $Y = y$. Given $Y = y$, $Q_\mathrm{d}(t)$ contains two parts. First, those in the same batch that finish service but cannot depart, and those customers who arrive between y and t and finish their service by time t. From the analysis above it is seen that there are $K(t-y)-1$ customers who are behind customer $J \equiv J(t-y)$ in batch B. Given J, the service times of these customers have distribution G; hence, each finishes service with probability $G(t-y)$. Hence, $\sum_{i=1}^{K(t-y)-1} \mathbf{1}_{\{X_i \le t-y\}}$ is the number of customers from the same batch B who have finished service but are blocked by customer J. To determine the number of customers who arrive between y and t who are departure-delayed at time t, we again use the fact that, given $N(t) - N(y) = n$, the arrival times of these n batches, denoted by $S_1, S_2, \dots, S_n$ are the order statistics of uniform random variables $U_1(t-y), \dots, U_n(t-y)$. Hence, the number of these customers who are departure-delayed at t can be written as

$$\sum_{i=1}^{N(t-y)-1} \sum_{j=1}^{B_i} \mathbf{1}_{\{U_i(t-y)+X_{ij} \le t-y\}} = \sum_{i=1}^{N(t-y)} \sum_{j=1}^{B_i} \mathbf{1}_{\{X_{ij} \le U_i(t-y)\}}.$$

Hence,

$$\{Q_\mathrm{d}(t) \mid Y = y\} = \sum_{i=1}^{K(t-y)-1} \mathbf{1}_{\{X_i \le t-y\}} + \sum_{i=1}^{N(t-y)} \sum_{j=1}^{B_i} \mathbf{1}_{\{X_{ij} \le U_i(t-y)\}}.$$

Recall the following facts.

- If a batch with n customers arrives at time 0, the PGF of the number of customers that have completed their service at time x is $(G(x)z + \bar{G}(x))^n$.
- If a batch of customers arrives at time 0, the PGF of the number of customers that have completed their service at time x is $B^*(G(x)z + \bar{G}(x))$.
- If a batch of customers arrives at time U that is uniformly distributed over $[0, t]$, the PGF of the number of customers that have completed their service at time t is
$$\frac{1}{t}\int_0^t B^*(G(t-s)z + \bar{G}(t-s))\,\mathrm{d}s = \frac{1}{t}\int_0^t B^*(G(s)z + \bar{G}(s))\,\mathrm{d}s.$$
- The PGF of the number of customers who arrive in $[0, t]$ and have completed their service at time t is
$$\exp\left(-\lambda \int_0^t (1 - B^*(G(s)z + \bar{G}(s)))\,\mathrm{d}s\right).$$

Applying these results, the PGF of $Q_{\mathrm{d}}(t)$, given $Y = y \le t$, can be obtained as

$$\mathbb{E}[z^{Q_{\mathrm{d}}(t)} \mid Y = y] = \frac{(B^*(G(t-y)) - B^*(zG(t-y) + \bar{G}(t-y)))}{G(t-y) - (G(t-y)z + \bar{G}(t-y))} \times \exp\biggl(-\lambda \int_0^{t-y} (1 - B^*(zG(s) + \bar{G}(s)))\,\mathrm{d}s\biggr).$$

Unconditioning and using a change of variable $u = t - y$, we obtain (11).

Intuitively, (11) can be obtained from (10) as follows. Every customer counted in $Q(t)$ except for the oldest one, which blocks all other customers following him/her and is not counted in $Q_{\mathrm{d}}(t)$, is counted in $Q_{\mathrm{d}}(t)$ with probability $G(t-s)$ and not counted in $Q_{\mathrm{d}}(t)$ with probability $\bar{G}(t-s)$, if the customer arrives at s. Thus, (11) can be obtained by dropping the 'z' terms after λ and in front of the integral in the second part of (10) and replacing all other 'z' terms in (10) with $G(t-s)z + \bar{G}(t-s)$. Letting $t \to \infty$, we obtain the steady state distribution of the number of departure-delayed customers in the system. The details are omitted.

Proof of (12). First note that the quantity $N_{2,t}(y)$ has the PGF

$$\mathbb{E}[z^{N_{2,t}(y)}] = \exp\biggl(-\lambda \int_0^{y} (1 - B^*(G(t-s)z + \bar{G}(t-s)))\,\mathrm{d}s\biggr).$$

Also note that, conditioning on $Y = y$, $D(t)$ is equal to $N_{2,t}(y)$ plus $J(y) - 1$ if $y < t$; and is equal to $N_{2,t}(t)$ if $y = \infty$. For $y < t$, we have

$$\begin{aligned}
\mathbb{E}[z^{J(t-y)-1}] &= \sum_{j=0}^{\infty} z^j \mathbb{P}\{J(t-y) = j+1\} \\
&= \sum_{j=0}^{\infty} z^j \sum_{n=j+1}^{\infty} \mathbb{P}\{J(t-y) = j+1 \mid B = n\} \mathbb{P}\Bigl\{B = n \mid \max_{1 \le i \le B} X_i > t - y\Bigr\} \\
&= \sum_{j=0}^{\infty} z^j \sum_{n=j+1}^{\infty} \frac{(G(t-y))^{j-1}\bar{G}(t-y)}{1 - (G(t-y))^n} \frac{(1 - (G(t-y))^n)}{(1 - B^*(G(t-y)))} b_n \\
&= \frac{\bar{G}(t-y)(1 - B^*(zG(t-y)))}{(1 - B^*(G(t-y)))(1 - zG(t-y))}.
\end{aligned}$$

Conditioning on Y, the PGF of $D(t)$ is given by

$$\begin{aligned}
&\mathbb{E}[z^{D(t)}] \\
&\quad = \lambda \int_0^t \exp\biggl(-\lambda \int_0^y (1 - B^*(zG(t-s) + \bar{G}(t-s)))\,\mathrm{d}s\biggr) \\
&\qquad \times \frac{(1 - B^*(zG(t-y)))}{1 - zG(t-y)} \bar{G}(t-y) \exp\biggl(-\lambda \int_0^y (1 - B^*(G(t-s)))\,\mathrm{d}s\biggr)\,\mathrm{d}y \\
&\qquad + \exp\biggl(-\lambda \int_0^t (1 - B^*(G(t-s)z + \bar{G}(t-s)))\,\mathrm{d}s\biggr) \mathbb{P}\{Y > t\},
\end{aligned}$$

which leads to (12).

Proof of (13). Suppose that a tagged customer arrives at time t and we want to calculate his/her sojourn time distribution. Clearly, this customer is selected from a batch of size n with probability $nb_n/\mathbb{E}[B]$. Let $\hat{B}$ denote the size of the selected batch.

Let $N(t)$ be the number of batches that arrive by time t. Let X be the service time of the tagged customer, then the tagged customer arriving at time t finishes his/her service at time $t+X$, and leaves the system if and only if all customers who arrived before t, as well as those customers who are in the same batch but are ahead of him/her, have finished service. Let I be the index of the tagged customer in his/her batch. Given the batch size is n, I is equal to k with probability $1/n$, $k = 1, \dots, n$; hence,

$$\mathbb{P}\{I = k\} = \sum_{n=k}^{\infty} \mathbb{P}\{I = k \mid \hat{B} = n\} \frac{nb_n}{\mathbb{E}[B]} = \frac{\bar{b}_k}{\mathbb{E}[B]}, \qquad k = 1, 2, \dots,$$

where $\bar{b}_k = \sum_{n=k}^{\infty} b_n$. Then his/her departure time can be expressed as

$$\max\Big\{\max_{1\le i\le N(t), 1\le j\le B_i} (S_i + X_{ij}), t + \max_{1\le k\le I} X_k\Big\},$$

where X_{ij} is the service time of customer j in batch i. Thus, the sojourn time of the tagged customer is

$$\begin{aligned} W(t) &= \max\Big\{\max_{1\le i\le N(t), 1\le j\le B_i} (S_i + X_{ij}), t + \max_{1\le k\le I} X_k\Big\} - t \\ &= \max\Big\{\max_{1\le i\le N(t), 1\le j\le B_i} (S_i + X_{ij} - t), \max_{1\le k\le I} X_k\Big\}. \end{aligned}$$

Since it is known that, given $N(t) = n$, $S_1, S_2, \dots, S_n$ are order statistic of n uniform random variables on [0, 1] (see, e.g. Ross (2010, Theorem 5.2)); hence, $W(t)$ has the same distribution as

$$\begin{aligned} W(t) &= \max\Big\{\max_{1\le i\le N(t), 1\le j\le B_i} (U_i(t) + X_{ij} - t), \max_{1\le k\le I} X_k\Big\} \\ &= \max\Big\{\max_{1\le i\le N(t), 1\le j\le B_i} (X_{ij} - U_i(t)), \max_{1\le k\le I} X_k\Big\}, \end{aligned}$$

where $U_i(t)$, $i = 1, \dots, n$ are i.i.d. uniform random variables on $[0, t]$, and in the second equality we have used that fact that $U_i(t)$ and $t - U_i(t)$ have the same distribution. It follows that

$$\begin{aligned} &\mathbb{P}\{W(t) \le a\} \\ &\quad = \mathbb{P}\Big\{\max\Big\{\max_{1\le i\le N(t), 1\le j\le B_i} (X_{ij} - U_i(t)), \max_{1\le k\le I} X_k\Big\} \le a\Big\} \\ &\quad = \mathbb{P}\Big\{\max_{1\le i\le N(t)} \Big(\max_{1\le j\le B_i} X_{ij} - U_i(t)\Big) \le a\Big\} \mathbb{P}\Big\{\max_{1\le k\le I} X_k \le a\Big\} \\ &\quad = \Big(\sum_{n=0}^{\infty} \mathbb{P}\Big\{\max_{1\le i\le N(t)} \Big(\max_{1\le j\le B_i} X_{ij} - U_i(t)\Big) \le a \;\Big|\; N(t) = n\Big\} \mathbb{P}\{N(t) = n\}\Big) \\ &\qquad \times \Big(\sum_{k=1}^{\infty} \mathbb{P}\Big\{\max_{1\le k\le I} X_k \le a \;\Big|\; I = k\Big\} \mathbb{P}\{I = k\}\Big) \end{aligned}$$

$$
\begin{aligned}
&= \exp\Big(-\lambda t + \lambda t\mathbb{P}\Big\{\max_{1\le j\le B_i} X_{ij} - U_i(t) \le a\Big\}\Big)\frac{\sum_{n=1}^{\infty}(G(a))^n\bar{b}_n}{\mathbb{E}[B]} \\
&= \exp\Big(-\lambda t + \lambda\int_0^t \mathbb{P}\Big\{\max_{1\le j\le B_i} X_{ij} \le a+u\Big\}\,\mathrm{d}u\Big)\frac{G(a)}{(1-G(a))\mathbb{E}[B]}(1-B^*(G(a))) \\
&= \frac{G(a)}{(1-G(a))\mathbb{E}[B]}(1-B^*(G(a))) \\
&\quad\times \exp\Big(-\lambda t + \lambda\int_0^t \sum_{n=1}^{\infty} b_n (G(a+u))^n\,\mathrm{d}u\Big) \\
&= \frac{G(a)}{(1-G(a))\mathbb{E}[B]}(1-B^*(G(a)))\exp\Big(-\lambda\int_0^t (1-B^*(G(a+u)))\,\mathrm{d}u\Big),
\end{aligned}
$$

which leads to (13).

Proof of (14). Then, we calculate the distribution of departure delay of the tagged customer. The departure delay of a customer who arrives at time t is

$$
W_{\mathrm{d}}(t) = \Big(\max\Big\{\max_{1\le i\le N(t),1\le j\le B_i}(X_{ij}-U_i(t)),\ \max_{1\le k\le I} X_k\Big\} - X_I\Big)^+.
$$

For $a \ge 0$, following a similar computation as that of the previous case, we obtain

$$
\begin{aligned}
&\mathbb{P}\{W_{\mathrm{d}}(t) \le a\} \\
&\quad= \mathbb{P}\Big\{\max\Big\{\max_{1\le i\le N(t),1\le j\le B_i}(X_{ij}-U_i(t)),\ \max_{1\le k\le I} X_k\Big\} - X_I \le a\Big\} \\
&\quad= \int_0^{\infty} \mathbb{P}\Big\{\max_{1\le i\le N(t),1\le j\le B_i}(X_{ij}-U_i(t)) \le a+x\Big\}\mathbb{P}\Big\{\max_{1\le k\le I-1} X_k \le a+x\Big\}\,\mathrm{d}G(x) \\
&\quad= \int_0^{\infty} \exp\Big(-\lambda\int_0^t (1-B^*(G(a+x+u)))\,\mathrm{d}u\Big)\frac{1-B^*(G(a+x))}{\bar{G}(a+x)\mathbb{E}[B]}\,\mathrm{d}G(x),
\end{aligned}
$$

which leads to (14).

Note that for a simple Poisson arrival process, we have $\tilde{B}(z) = z$ and $\mathbb{E}[B] = 1$, and so the results above reduce to those in Section 2.

We remark that the distribution of I is the equilibrium distribution of the batch size. Thus, the above result actually gives the sojourn time distribution and departure delay distribution of an arbitrarily chosen customer in steady state. If we consider a customer in a batch arrives at time t, the distribution of I, the position of the tagged customer in its batch, is given by

$$
\mathbb{P}\{I = k\} = \sum_{n=k}^{\infty} b_n \frac{1}{n} \quad \text{for } k = 1, 2, \ldots.
$$

By using the above method, the distributions of the sojourn time and the departure delay of the tagged customer can be obtained but the results are much more complicated. The details are omitted.

Using the above results, the waiting time of the first/last customer in an arbitrary batch or a batch arriving at time t can also be found. The details are omitted.

3.2. The M(t)/G/∞ tollbooth tandem queue

The second extension we consider is the case where customers arrive according to a nonhomogeneous Poisson process $\{N(t); t \geq 0\}$ with arrival intensity $\{\lambda(t); t \geq 0\}$, and we compute the sojourn time of a customer who arrives at time s. The classic infinite-server queue with nonhomogeneous Poisson arrival processes have important applications in transportation systems; see, e.g. Eick *et al.* (1993) and Massey and Whitt (1993).

Recall that, for a nonhomogeneous Poisson process, given $N(t) = n$, the arrival times of the n customers are order statistics of $V_1(t), V_2(t), \ldots, V_n(t)$ that have density function $\lambda(s)/m(t)$, $0 \leq s \leq t$, where $m(t) = \int_0^t \lambda(s)\,\mathrm{d}s$.

To find the distribution for the number of customers in the system at time t, we again define type 1 events as we did in the previous section. Then, the number of type 1 events up to time y, $N_{1,t}(y)$, has a Poisson distribution with parameter $\int_0^y \lambda(s)\bar{G}(t-s)\,\mathrm{d}s$. Let Y be the time at which the first type 1 event takes place, which is equal to ∞ if it never occurs, then

$$\mathbb{P}\{Y \leq y\} = \mathbb{P}\{N_{1,t}(y) > 0\} = 1 - \exp\biggl(-\int_0^y \lambda(s)\bar{G}(t-s)\,\mathrm{d}s\biggr),$$

and

$$f_Y(y) = \lambda(y)\bar{G}(t-y)\exp\biggl(-\int_0^y \lambda(s)\bar{G}(t-s)\,\mathrm{d}s\biggr).$$

Given $Y = y \leq t$, the number of customers in the system at time t is 1 plus the number of arrivals between y and t; thus,

$$\mathbb{P}\{Q(t) = 0\} = \exp\biggl(-\int_0^t \lambda(s)\bar{G}(t-s)\,\mathrm{d}s\biggr),$$

and, for $n > 0$,

$$\begin{aligned}\mathbb{P}\{Q(t) = n\} = \int_0^t & \frac{(\int_y^t \lambda(s)\,\mathrm{d}s)^{n-1}}{(n-1)!}\lambda(y)\bar{G}(t-y) \\ & \times \exp\biggl(-\int_y^t \lambda(s)\,\mathrm{d}s - \int_0^y \lambda(s)\bar{G}(t-s)\,\mathrm{d}s\biggr)\,\mathrm{d}y.\end{aligned}$$

Next, we find the distribution of $Q_{\mathrm{d}}(t)$ as

$$\begin{aligned}&\mathbb{P}\{Q_{\mathrm{d}}(t) = 0\} \\ &\quad = \exp\biggl(-\int_0^t \lambda(s)\bar{G}(t-s)\,\mathrm{d}s\biggr) \\ &\qquad + \int_0^t \lambda(y)\bar{G}(t-y)\exp\biggl(-\int_y^t \lambda(s)G(t-s)\,\mathrm{d}s - \int_0^y \lambda(s)\bar{G}(t-s)\,\mathrm{d}s\biggr)\,\mathrm{d}y,\end{aligned}$$

and, for $n > 0$,

$$\begin{aligned}\mathbb{P}\{Q_{\mathrm{d}}(t) = n\} = \int_0^t & \frac{(\int_y^t \lambda(s)G(t-s)\,\mathrm{d}s)^{n-1}}{(n-1)!}\lambda(y)\bar{G}(t-y) \\ & \times \exp\biggl(-\int_y^t \lambda(s)G(t-s)\,\mathrm{d}s - \int_0^y \lambda(s)\bar{G}(t-s)\,\mathrm{d}s\biggr)\,\mathrm{d}y.\end{aligned}$$

Using a similar argument as used in Sections 2 and 3.1, we obtain the distribution of the number of departures by any time t. For all $n \geq 0$,

$$\mathbb{P}\{D(t)=n\} = \int_0^t \frac{(\int_0^y \lambda(s)G(t-s)\,\mathrm{d}s)^n}{n!}\lambda(y)\bar{G}(t-y)\exp\biggl(-\int_0^y \lambda(s)\,\mathrm{d}s\biggr)\,\mathrm{d}y + \frac{(\int_0^t \lambda(s)G(s)\,\mathrm{d}s)^n}{n!}\exp\biggl(-\int_0^t \lambda(s)\,\mathrm{d}s\biggr).$$

The same argument as in the previous section yields the distribution function of the sojourn time $W(t)$ of a customer that arrives at time s:

$$\begin{aligned}\mathbb{P}\{W(s)\leq a\} &= G(a)\sum_{n=0}^{\infty}(\mathbb{P}\{V_1(s)+X_1-s\leq a\})^n\mathbb{P}\{N(s)=n\}\\ &= G(a)\sum_{n=0}^{\infty}(\mathbb{P}\{V_1(s)+X_1-s\leq a\})^n\mathrm{e}^{-m(s)}\frac{(m(s))^n}{n!}\\ &= G(a)\exp(-m(t)(1-\mathbb{P}\{V_1(s)+X_1-s\leq a\})\\ &= G(a)\exp\biggl(-\int_0^s \lambda(u)\bar{G}(a+s-u)\,\mathrm{d}u\biggr).\end{aligned}$$

Note that, for a nonhomogeneous Poisson arrival process, we will not be able to take limit $s\to\infty$.

Similarly, the departure delay distribution can be computed as, for $a\geq 0$,

$$\begin{aligned}&\mathbb{P}\{W_{\mathrm{d}}(s)\leq a\}\\ &\quad= \int_0^\infty \sum_{n=0}^{\infty}(\mathbb{P}\{V_1(s)+X_1\leq s+x+a\})^n\mathbb{P}\{N(s)=n\}\,\mathrm{d}G(x)\\ &\quad= \int_0^\infty \exp(-m(s)(1-\mathbb{P}\{U_1(s)+X_1\leq s+x+a\}))\,\mathrm{d}G(x)\\ &\quad= \int_0^\infty \exp\biggl(-\int_0^s \lambda(u)\bar{G}(s+x+a-u)\,\mathrm{d}u\biggr)\,\mathrm{d}G(x).\end{aligned}$$

4. Some stochastic comparison results

In this section we present some qualitative results on the base model introduced in Section 2. We answer the following basic question: how are the number of customers in the system and customer delays affected by arrival process (i.e. the arrival rate) and customer service times? This is answered using stochastic ordering. Recall that a random variable X is said to be stochastically larger than another random variable Y if $\mathbb{P}\{X\geq t\}\geq\mathbb{P}\{Y\geq t\}$ for all t; see Ross (1996, Chapter 8) or Shaked and Shanthikumar (2006).

Theorem 3. *For the M/G/∞ tollbooth tandem queue defined in Section 2, we have the following results.*

(i) *For any $t\geq 0$, $Q(t)$, $Q_{\mathrm{d}}(t)$, $D(t)$, $W(t)$, and $W_{\mathrm{d}}(t)$ are stochastically increasing in the arrival rate, λ.*

(ii) *If the service time distribution becomes stochastically longer, then $Q(t)$ and $W(t)$ become stochastically larger for any $t\geq 0$.*

Proof of Theorem 3(i). First, the results for $W(t)$ and $W_{\mathrm{d}}(t)$ can easily be obtained from (6) and (7). A sample path, or coupling, approach is utilized to prove the rest of Theorem 3(i). To that end, we construct two tollbooth queueing systems with arrival rates λ and $\lambda+\delta$, respectively. For convenience in this proof, we use $\mathrm{M}^{\lambda}/\mathrm{G}/\infty$ and $\mathrm{M}^{\lambda+\delta}/\mathrm{G}/\infty$ to denote these two tollbooth tandem queues. In the $\mathrm{M}^{\lambda+\delta}/\mathrm{G}/\infty$ queue, we mark an arrival as a type I customer with probability $\lambda/(\lambda+\delta)$ and as a type II customer with probability $\delta/(\lambda+\delta)$, independently of everything else. Then, type I customers form a Poisson process with arrival rate λ. Thus, the arrival and service of all type I customers are stochastically equivalent to that of the customers in the $\mathrm{M}^{\lambda}/\mathrm{G}/\infty$ queue. We couple the service times of the type I customers in the $\mathrm{M}^{\lambda+\delta}/\mathrm{G}/\infty$ queue with the service times of customers in the $\mathrm{M}^{\lambda}/\mathrm{G}/\infty$ queue. However, the departure process of type I customers in the $\mathrm{M}^{\lambda+\delta}/\mathrm{G}/\infty$ queue is different from that in the $\mathrm{M}^{\lambda}/\mathrm{G}/\infty$ queue due to the existence of type II customers. If a type I customer has departed in the $\mathrm{M}^{\lambda+\delta}/\mathrm{G}/\infty$ queue, the corresponding customer in the $\mathrm{M}^{\lambda}/\mathrm{G}/\infty$ queue must have departed too. Let $Q^{\lambda}(t)$ be the number of customers in the $\mathrm{M}^{\lambda}/\mathrm{G}/\infty$ queue, and $Q^{\lambda+\delta,I}(t)$ be the number of type I customers in the $\mathrm{M}^{\lambda+\delta}/\mathrm{G}/\infty$ queue at time t. The argument above shows that

$$Q^{\lambda}(t) \le Q^{\lambda+\delta,I}(t) \le Q^{\lambda+\delta}(t).$$

This proves that $Q(t)$ is stochastically increasing in the arrival rate λ.

To construct $D^{\lambda}(t)$ and $D^{\lambda+\delta}(t)$, let $N(t)$ be the number of customers who arrived in $[0,t]$ in the $\mathrm{M}^{\lambda}/\mathrm{G}/\infty$ queue. These $N(t)$ customers are generated as the type I customers who arrived in $[0,t]$ in the $\mathrm{M}^{\lambda+\delta}/\mathrm{G}/\infty$ queue. We assign the service times of the $N(t)$ customers in the $\mathrm{M}^{\lambda}/\mathrm{G}/\infty$ queue to the first $N(t)$ arrivals in the $\mathrm{M}^{\lambda+\delta}/\mathrm{G}/\infty$ queue. As service times of all customers have the same distribution, probabilistically, this assignment of service times does not change the number of departures in the $\mathrm{M}^{\lambda+\delta}/\mathrm{G}/\infty$ queue. Since the first $N(t)$ customers in the $\mathrm{M}^{\lambda+\delta}/\mathrm{G}/\infty$ queue arrive earlier than the first $N(t)$ customers in the $\mathrm{M}^{\lambda}/\mathrm{G}/\infty$ queue (and yet they have the same service times) and their departures are not affected by arrivals after them, it is clear that in the processes thus constructed, $D^{\lambda}(t)$ is less than or equal to $D^{\lambda+\delta}(t)$. This shows that $D(t)$ is stochastically increasing in the arrival rate λ as well.

Proof of Theorem 3(ii). We first write (2) as follows. For $n=1$,

$$\begin{aligned}
\mathbb{P}\{Q(t)=1\} &= \lambda\int_0^t \bar{G}(u)\exp\biggl(-\lambda\biggl(u+\int_u^t \bar{G}(s)\,\mathrm{d}s\biggr)\biggr)\,\mathrm{d}u \\
&= \int_0^t \mathrm{e}^{-\lambda u}\,\mathrm{d}\biggl(\exp\biggl(-\lambda\int_u^t \bar{G}(s)\,\mathrm{d}s\biggr)\biggr) \\
&= \mathrm{e}^{\lambda t} - \exp\biggl(-\lambda\int_0^t \bar{G}(s)\,\mathrm{d}s\biggr) + \lambda\int_0^t \exp\biggl(-\lambda\biggl(u+\int_u^t \bar{G}(s)\,\mathrm{d}s\biggr)\biggr)\,\mathrm{d}u,
\end{aligned}$$

and, for $n\ge 2$,

$$\begin{aligned}
\mathbb{P}\{Q(t)=n\} &= \int_0^t \frac{(\lambda u)^{n-1}}{(n-1)!}\mathrm{e}^{-\lambda u}\,\mathrm{d}\biggl(\exp\biggl(-\lambda\int_u^t \bar{G}(s)\,\mathrm{d}s\biggr)\biggr) \\
&= \frac{(\lambda t)^{n-1}}{(n-1)!}\mathrm{e}^{-\lambda t} - \lambda\int_0^t \biggl(\frac{(\lambda u)^{n-2}}{(n-2)!} - \frac{(\lambda u)^{n-1}}{(n-1)!}\biggr) \\
&\qquad\qquad \times \exp\biggl(-\lambda\biggl(u+\int_u^t \bar{G}(s)\,\mathrm{d}s\biggr)\biggr)\,\mathrm{d}u
\end{aligned}$$

$$= \frac{(\lambda t)^{n-1}}{(n-1)!}\mathrm{e}^{-\lambda t} - \lambda \int_0^t \frac{(\lambda u)^{n-2}}{(n-2)!} \exp\bigg(-\lambda\bigg(u + \int_u^t \bar{G}(s)\,\mathrm{d}s\bigg)\bigg)\,\mathrm{d}u$$
$$+ \lambda \int_0^t \frac{(\lambda u)^{n-1}}{(n-1)!} \exp\bigg(-\lambda\bigg(u + \int_u^t \bar{G}(s)\,\mathrm{d}s\bigg)\bigg)\,\mathrm{d}u.$$

Using these expressions and (1), it is readily verified that, for $k \geq 1$,

$$\begin{aligned}\mathbb{P}\{Q(t) \leq k\} &= \sum_{n=0}^{k} \mathbb{P}\{Q(t) = n\} \\ &= \sum_{n=1}^{k} \frac{(\lambda t)^{n-1}}{(n-1)!}\mathrm{e}^{-\lambda t} \\ &\quad + \lambda \int_0^t \frac{(\lambda u)^{k-1}}{(k-1)!} \exp\bigg(-\lambda\bigg(u + \int_u^t \bar{G}(s)\,\mathrm{d}s\bigg)\bigg)\,\mathrm{d}u.\end{aligned}$$

Therefore, if $\bar{G}(s)$ becomes larger (for any $s > 0$), the above sum gets smaller. This shows that, if the service time X becomes stochastically larger then $Q(t)$ will be stochastically larger too.

The result for $W(t)$ is obvious by (6). This completes the proof of Theorem 3(ii).

5. Conclusions

In this paper we have studied a tollbooth tandem queue with an infinite number of homogeneous servers. We obtained closed-form solutions for the distributions for the number of customers in the system, number of departure-delayed customers in the system, customer sojourn times, and customer departure delays. We first analyzed the case with Poisson arrival processes and then extended the results to batch Poisson arrival processes and nonhomogeneous Poisson arrival processes. These results offer useful insights for infinite-server queues when customers have to depart in the order they arrive, as well as other real world models such as tollbooth tandem queues and gas stations, whose studies have been mainly relying on numerical methods; see, e.g. Hall and Daganzo (1983), Hong *et al.* (2009), Daskin *et al.* (1976), and Teimoury and Yazdi (2011).

The results can be extended to the case when the arrivals follow a batch Markov arrival process. In that case, the fundamental differential equations can be derived for the probability distributions and the moments of the quantities of interest can be computed using standard methods; see, e.g. Ramaswami and Neuts (1980).

Acknowledgements

The authors thank two anonymous referees for their helpful comments and suggestions to improve the exposition of this paper. The research of Xiuli Chao is partially supported by the NSF (grants CMMI-1131249 and CMMI-1362619), and the research of Sheldon Ross is supported by U.S. Army Research Laboratory and the U.S. Army Research Office (grant no. W911NF-11-1-0115).

References

DASKIN, M. S., SHLADOVER, S. AND SOBEL, K. (1976). An analysis of service station queues under gasoline shortage conditions. *Comput. Operat. Res.* **3,** 83–93.

EICK, S. G., MASSEY, W. A. AND WHITT, W. (1993). $M(t)/G/\infty$ queues with sinusoidal arrival rates. *Manag. Sci.* **39,** 241–252.

ELIAZAR, I. (2007). The M/G/∞ system revisited: finiteness, summability, long range dependence, and reverse engineering. *Queueing Systems* **55,** 71–82.

HALL, R. W. AND DAGANZO, C. F. (1983). Tandem tollbooths for the Golden Gate Bridge. *Transportation Res. Rec.* **905,** 7–14.

HE, Q-M. AND CHAO, X. (2014). A tollbooth tandem queue with heterogeneous servers. *Europ. J. Operat. Res.* **236,** 177–189.

HOLMAN, D. F., CHAUDHRY, M. L. AND KASHYAP, B. R. K. (1983). On the service system $M^X/G/\infty$. *Europ. J. Operat. Res.* **13,** 142-145.

HONG, Y-C. *et al.* (2009). Modeling and simulation of tandem tollbooth operations with max-algebra approach. In *Future Generation Information Technology*, Springer, Berlin, pp. 138–150.

MASSEY, W. A. AND WHITT, W. (1993). Networks of infinite-server queues with nonstationary Poisson input. *Queueing Systems* **13,** 183–250.

PAPADOPOULOS, H. T. AND O'KELLY, M. E. J. (1993). Exact analysis of production lines with no intermediate buffers. *Europ. J. Operat. Res.* **65,** 118–137.

RAMASWAMI, V. AND NEUTS, M. F. (1980). Some explicit formulas and computational methods for infinite-server queues with phase-type arrivals. *J. Appl. Prob.* **17,** 498–514.

REUVENI, S., ELIAZAR, I. AND YECHIALI, U. (2012). Asymmetric inclusion process as a showcase of complexity. *Phys. Rev. Lett.* **109,** 020603.

REUVENI, S., HIRSCHBERG, O., ELIAZAR, I. AND YECHIALI, U. (2014). Occupation probabilities and fluctuations in the asymmetric simple inclusion process. *Phys. Rev. E* **89,** 042109.

ROSS, S. M. (1996). *Stochastic Processes*, 2nd edn. John Wiley, New York.

ROSS, S. M. (2010). *Introduction to Probability Models*, 10th edn. Academic Press, Boston, MA.

SHAKED, M. AND SHANTHIKUMAR, J. G. (2007). *Stochastic Orders*. Springer, New York.

SHANBHAG, D. N. (1966). On infinite server queues with batch arrivals. *J. Appl. Prob.* **3,** 274–279.

TEIMOURY, E., YAZDI, M. M., HADDADI, M. AND FATHI, M. (2011). Modelling and improvement of non-standard queuing systems: a gas station case study. *Internat. J. Appl. Decision Sci.* **4,** 324–340.

ZAZANIS, M. A. (2004). Infinite server queues with synchronized departures driven by a single point process. *Queueing Systems* **48,** 309–338.