Probability in the Engineering and Informational Sciences (2023), **37**:4 997–1019
doi:10.1017/S0269964822000365

CAMBRIDGE UNIVERSITY PRESS

RESEARCH ARTICLE

Analyzing a single hyper-exponential working vacation queue from its governing difference equation

Miaomiao Yu and Yinghui Tang

School of Mathematical Science, Sichuan Normal University, Chengdu, Sichuan 610066, P. R. China. E-mail: mmyu75@163.com; tangyh@sicnu.edu.cn.

Keywords: Difference equation, Hyper-exponential working vacation, Queueing system, Rouché's theorem, Shift operator

Abstract
As the queue becomes exhausted, different maintenance tasks can be performed according to the fatigue load and wear degree of the service equipment. At the same time, considering the customer's sensitivity to time delay, the service facility will not completely remain inactive during the maintenance period. To describe this objectively existing phenomenon arising in the waiting line system, we consider a hyper-exponential working vacation queue with a batch renewal arrival process. Through the calculation of the well-structured roots of the associated characteristic equation, the shift operator method in the theory of difference equations and the supplementary variable technique for stochastic modeling plays a central role in the queue-length distribution analysis. Comparison with other ways to analyze queueing models, the advantage of our approach is that we can avoid deriving the complex transition probability matrix of the queue-length process embedded at input points. The feasibility of this approach is verified by extensive numerical examples.

1. Introduction

The queueing system with server vacations is useful to model a system in which the server has an additional task during its idle period. The additional task may represent the server's working on some supplementary jobs, performing service equipment maintenance inspections and repairs, or server's rest after queue exhaustion. Since such a queue has broad applicability in analyzing the performance of industrial production systems and data communication networks, it has attracted many researchers' attention over the past several decades. Various vacation policies provide more flexibility for optimal design and operating control of the systems. The working vacation policy introduced by Servi and Finn [23] is a kind of semi-vacation policy. It is characterized by the feature that the service facility works at a lower service rate rather than completely stopping service during the vacation period. At first glance, the concept of working vacation is a bit of an oxymoron because work and vacation are essentially two different things. But when we combine them into a waiting line system arising from the manufacturing environment, it might be a perfect way to reduce customer service response time because we can maintain the production equipment during the vacation period without suspending production completely. In the last twenty years, numerous researchers, including Wu and Takagi [26], Liu *et al.* [19], Li *et al.* [18], Zhang and Hou [28,29] , Selvaraju and Goswami [22], Gao and Yao [11], Lee and Kim [16], Ma *et al.* [20], used different methods to study this kind of queue under the assumption that customers arrive at a service facility according to a Poisson stream. Since our research is mainly concerned with the general renewal input working vacation queue, we only give a brief literature review in this area and point out the limitations of the current study to clarify our work's motivation.

So far, the matrix-geometric approach and the embedded Markov Chain technique are the mainstream methods to study the GI/M/1 type queue, and there is no exception for the working vacation queue. Baba [1] applied the matrix-geometric method to investigate a GI/M/1 queue with multiple exponential working vacations. Li *et al.* [17] used the same way to study the discrete-time version of the above model. Furthermore, both in continuous and discrete-time cases, Chae *et al.* [7] discussed general input queues with a single working vacation using the embedded Markov chain. Considering the fact that customers arrive in batches instead of individually, Guha and Banik [14] and Guha *et al.* [15] also analyzed the renewal input batch arrival queue under single and multiple exponential working vacation policies. Recently, Yu [27] extended the above work so that the bulk service rule is included. After carefully reading these papers mentioned above, we may note that each one contains a more complex transition probability analysis for the queue-length process embedded at the pre-arrival epoch. The corresponding probabilistic arguments involved in these studies are cumbersome to implement and error-prone. Additionally, for the general input working vacation queue, except for the work done by Chen *et al.* [9], nearly all other authors assume that the server's vacation time is exponentially distributed. We think that this is not a very reasonable assumption for many cases in the industrial environment. Under such a situation, when the waiting line system becomes exhausted, we usually decide what maintenance needs to be done according to the fatigue state of service equipment. Maintenance may represent the actions of replacement, repair, adjustment, overhaul, improvement, and checking of the service equipment. Since each action corresponds to a degree of complexity, the maintenance time cannot be assumed to follow a common exponential distribution. Thus, to give a more realistic model assumption, we intend to extend the exponential vacation time to hyper-exponential vacation time in this research. Moreover, to provide an alternative yet simple problem-solving methodology, we try to use an approach based on the theory of difference equations to carry out a comprehensive analysis of GI^X/M/1 single working vacation queue. Here, we must admit that the recent work done by Barbhuiya and Gupta [2–5] gives us some basic ideas to complete the model analysis. We will see that the method adopted in this paper can make us get rid of the chain matrix analysis. So for most people, this is an easy to accept method.

The rest of this paper is organized as follows. In Section 2, we describe the mathematical queueing model in detail. A set of differential-difference equations that represent the dynamics of the queue-length process is developed in Section 3. In Section 4, queue-length probabilities at different epochs are derived explicitly by solving simultaneous nonhomogeneous difference equations. Section 5 is devoted to the sojourn time of a randomly selected customer in an arriving batch. A three-step algorithm for computing the queue-length distribution is summarized in Section 6. To validate our computational algorithm, we further provide several typical numerical examples in this section. Finally, conclusions and future scopes are presented in Section 7.

2. Model formulation

The model is defined by making the following assumptions.

(1) Consider a single-station queueing system where customers arrive in batches according to a renewal process with independent identically distributed (i.i.d.) inter-batch arrival times having a common cumulative distribution function $A(t)$, and probability density function $a(t)$. Let $a^*(s) = \int_0^\infty e^{-st}\,\mathrm{d}A(t)$ be the Laplace–Stieltjes transform (L.S.T.) of $A(t)$ and let the mean inter-batch arrival time be denoted by $1/\lambda$. Differentiation of the L.S.T. with respect to s is justified and yields $1/\lambda = -(\mathrm{d}/\mathrm{d}s)a^*(s)|_{s=0} < \infty$.

(2) At every arrival epoch, a batch of k customers arrives with probability g_k. For mathematical convenience and from a more realistic point of view, we assume that the maximum batch size is equal to b. Consequently, the probability generating function of the sequence $\{g_k, k = 1, 2, \ldots, b\}$ is $G(z) = \sum_{k=1}^{b} g_k z^k$ with the first moment $\overline{g} = \sum_{k=1}^{b} k g_k$.

https://doi.org/10.1017/S0269964822000365 Published online by Cambridge University Press

(3) In a normal busy period, the service times are i.i.d. exponential random variables with mean $1/\mu_0$. When the system becomes empty, the server goes on vacation for a random duration, where the vacation time has an h-phase hyper-exponential distribution, which can be properly represented by a probabilistic mixture of exponential distributions. The distribution function of the vacation time V is defined as $\Pr\{V \le t\} = V(t) = \sum_{j=1}^{h} \alpha_j(1 - e^{-\theta_j t})$. In other words, at the end of a service, if no customer is left in the system, the server takes type j vacation with probability α_j, then the vacation time distribution is $\mathrm{Exp}(\theta_j)$, $j = 1, 2, \ldots, h$.
(4) Customers arriving during a type j vacation will be served at a lower service rate, where the service time obeys the exponential distribution with rate μ_j ($\mu_j < \mu_0$, $j = 1, 2, \ldots, h$). Upon completion of service at a lower rate, if the vacation is not over, and no customer is waiting for service, the server will continue the current vacation. On the contrary, if there is at least one customer in the system, the server will keep this service mode.
(5) As this type of j vacation gets over, the server turns to the normal working mode immediately. The service of a customer being served will be interrupted and restarted from the beginning in the normal busy period. Alternatively, if no customers are found in the queue at the end of a vacation, the server remains idle and is ready to serve new arrivals at a normal service rate μ_0.
(6) We further assume that the inter-batch arrival times, service times, and vacation times are mutually independent. A necessary and sufficient condition for model stability is $\rho = \lambda\overline{g}/\mu_0 < 1$ (see the proofs of Lemma 1 and the analysis of Eq. (35)).

3. Governing difference equation

Queueing systems with general inter-batch arrival time distribution and hyper-exponential vacation time are difficult to analyze mathematically due to the queueing process being non-Markovian. To enable the system to be characterized as a Markov system, the following random variables will be used for the development of our model. Let $N(t)$ and $\xi(t)$ denote the number of customers in the system (including the one in service) and the state of the server at time t, respectively. Here

$$\xi(t) = \begin{cases} 0, & \text{if the server is in normal busy period at time } t, \\ j, & \text{if the server is on type } j \text{ vacation at time } t, \ \ j = 1, 2, \ldots, h. \end{cases}$$

With the inclusion of the supplementary variable technique, and together with the remaining inter-batch arrival time $\tilde{A}(t)$ at time t, we may obtain a tri-variate Markov process $\{N(t), \xi(t), \tilde{A}(t)\}$. Furthermore, to establish the dynamic model of the above Markov process, let us define some probabilities as below

$$P_i(x,t)\,\mathrm{d}x = \Pr\{N(t) = i, \xi(t) = 0, x < \tilde{A}(t) \le x + \mathrm{d}x\}, \quad i = 0, 1, 2\ldots,$$
$$Q_{i,j}(x,t)\,\mathrm{d}x = \Pr\{N(t) = i, \xi(t) = j, x < \tilde{A}(t) \le x + \mathrm{d}x\}, \quad i = 0, 1, 2\ldots, \ j = 1, \ldots, h.$$

Employing the above-stated probabilities, and considering the state transitions between time t and $t + \Delta t$ like the usual arguments as in the birth and death model, we can have the following partial differential equations (1) to (6) for the tri-variate Markov process.

$$\left(\frac{\partial}{\partial t} - \frac{\partial}{\partial x}\right) P_0(x,t) = \sum_{j=1}^{h} \theta_j Q_{0,j}(x,t), \tag{1}$$

$$\left(\frac{\partial}{\partial t} - \frac{\partial}{\partial x}\right) P_n(x,t) = -\mu_0 P_n(x,t) + \mu_0 P_{n+1}(x,t) + a(x)\sum_{k=1}^{n} g_k P_{n-k}(0,t) + \sum_{j=1}^{h} \theta_j Q_{n,j}(x,t), \quad n = 1, 2, \ldots, b-1, \tag{2}$$

https://doi.org/10.1017/S0269964822000365 Published online by Cambridge University Press

$$\left(\frac{\partial}{\partial t}-\frac{\partial}{\partial x}\right)P_n(x,t)=-\mu_0P_n(x,t)+\mu_0P_{n+1}(x,t)+a(x)\sum_{k=1}^{b}g_kP_{n-k}(0,t)$$
$$+\sum_{j=1}^{h}\theta_jQ_{n,j}(x,t),\quad n\geq b, \tag{3}$$
$$\left(\frac{\partial}{\partial t}-\frac{\partial}{\partial x}\right)Q_{0,j}(x,t)=-\theta_jQ_{0,j}(x,t)+\mu_jQ_{1,j}(x,t)+\mu_0\alpha_jP_1(x,t),\quad j=1,2,\ldots,h, \tag{4}$$
$$\left(\frac{\partial}{\partial t}-\frac{\partial}{\partial x}\right)Q_{n,j}(x,t)=-(\theta_j+\mu_j)Q_{n,j}(x,t)+\mu_jQ_{n+1,j}(x,t)$$
$$+a(x)\sum_{k=1}^{n}g_kQ_{n-k,j}(0,t),\quad n=1,\ldots,b-1,\ j=1,\ldots,h, \tag{5}$$
$$\left(\frac{\partial}{\partial t}-\frac{\partial}{\partial x}\right)Q_{n,j}(x,t)=-(\theta_j+\mu_j)Q_{n,j}(x,t)+\mu_jQ_{n+1,j}(x,t)$$
$$+a(x)\sum_{k=1}^{b}g_kQ_{n-k,j}(0,t),\quad n\geq b,\ j=1,\ldots,h. \tag{6}$$

As many articles have pointed out, some probabilistic interpretations will be helpful to understand how these equations are derived. For example, Eq. (6) is obtained at time $t+\Delta t$ considering all possibilities. Note that when time t is increased by Δt, the remaining inter-batch arrival time will be reduced by $x-\Delta t$. Thus, we have

$$Q_{n,j}(x-\Delta t,t+\Delta t)=Q_{n,j}(x,t)(1-\theta_j\Delta t+o(\Delta t))(1-\mu_j\Delta t+o(\Delta t))+\mu_j\Delta tQ_{n+1,j}(x,t)$$
$$+a(x)\Delta t\sum_{k=1}^{b}g_kQ_{n-k,j}(0,t),\quad n\geq b,\ j=1,\ldots,h.$$

The above equation explains possible cases for the probability that there are $n(n\geq b)$ customers in the system and the server is on type j vacation when the remaining inter-batch arrival time is $x-\Delta t$ at time $t+\Delta t$. In a similar way, Eqs. (1)–(5) can be also obtained.

Let $\lim_{t\to\infty}P_n(x,t)=P_n(x)$ and $\lim_{t\to\infty}Q_{n,j}(x,t)=Q_{n,j}(x)$, the Kolmogorov forward equations governing the system in steady-state for the proposed model is:

$$-\frac{\mathrm{d}}{\mathrm{d}x}P_0(x)=\sum_{j=1}^{h}\theta_jQ_{0,j}(x), \tag{7}$$
$$-\frac{\mathrm{d}}{\mathrm{d}x}P_n(x)=-\mu_0P_n(x)+\mu_0P_{n+1}(x)+a(x)\sum_{k=1}^{n}g_kP_{n-k}(0)$$
$$+\sum_{j=1}^{h}\theta_jQ_{n,j}(x),\quad n=1,\ldots,b-1, \tag{8}$$
$$-\frac{\mathrm{d}}{\mathrm{d}x}P_n(x)=-\mu_0P_n(x)+\mu_0P_{n+1}(x)+a(x)\sum_{k=1}^{b}g_kP_{n-k}(0)$$
$$+\sum_{j=1}^{h}\theta_jQ_{n,j}(x),\quad n\geq b, \tag{9}$$

https://doi.org/10.1017/S0269964822000365 Published online by Cambridge University Press

$$-\frac{\mathrm{d}}{\mathrm{d}x}Q_{0,j}(x) = -\theta_j Q_{0,j}(x) + \mu_j Q_{1,j}(x) + \mu_0\alpha_j P_1(x), \quad j = 1, 2, \ldots, h, \tag{10}$$

$$-\frac{\mathrm{d}}{\mathrm{d}x}Q_{n,j}(x) = -(\theta_j + \mu_j)Q_{n,j}(x) + \mu_j Q_{n+1,j}(x) + a(x)\sum_{k=1}^{n} g_k Q_{n-k,j}(0), \quad n = 1, \ldots, b-1,\ j = 1, \ldots, h, \tag{11}$$

$$-\frac{\mathrm{d}}{\mathrm{d}x}Q_{n,j}(x) = -(\theta_j + \mu_j)Q_{n,j}(x) + \mu_j Q_{n+1,j}(x) + a(x)\sum_{k=1}^{b} g_k Q_{n-k,j}(0), \quad n \geq b,\ j = 1, \ldots, h. \tag{12}$$

Since the Laplace transform turns the operation of differentiation into the algebraic operation in s-domain, and can greatly simplify the solution of problems involving differential equations, we further define the Laplace transforms of $P_n(x)$ and $Q_{n,j}(x)$ as follows

$$P_n^*(s) = \int_0^\infty e^{-sx} P_n(x)\,\mathrm{d}x, \quad n \geq 0, \quad Q_{n,j}^*(s) = \int_0^\infty e^{-sx} Q_{n,j}(x)\,\mathrm{d}x, \quad n \geq 0,\ j = 1, \ldots, h.$$

Additionally, for $s = 0$, we define

$$P_n^*(0) \equiv P_n = \int_0^\infty P_n(x)\,\mathrm{d}x, \quad n \geq 0, \quad Q_{n,j}^*(0) \equiv Q_{n,j} = \int_0^\infty Q_{n,j}(x)\,\mathrm{d}x, \quad n \geq 0,\ j = 1, \ldots, h.$$

Thus, P_n is the probability that there are n customers in the system when the server is in a normal busy period. Similarly, $Q_{n,j}$ is the probability that there are n customers in the system while the server is on type j vacation. Taking the Laplace transform on both sides of Eqs. (7)–(12), the corresponding algebraic equations are given by

$$sP_0^*(s) = P_0(0) - \sum_{j=1}^{h} \theta_j Q_{0,j}^*(s), \tag{13}$$

$$(s - \mu_0)P_n^*(s) + \mu_0 P_{n+1}^*(s) = P_n(0) - a^*(s)\sum_{k=1}^{n} g_k P_{n-k}(0) - \sum_{j=1}^{h} \theta_j Q_{n,j}^*(s), \quad n = 1, \ldots, b-1, \tag{14}$$

$$(s - \mu_0)P_n^*(s) + \mu_0 P_{n+1}^*(s) = P_n(0) - a^*(s)\sum_{k=1}^{b} g_k P_{n-k}(0) - \sum_{j=1}^{h} \theta_j Q_{n,j}^*(s), \quad n \geq b, \tag{15}$$

$$(s - \theta_j)Q_{0,j}^*(s) + \mu_j Q_{1,j}^*(s) = Q_{0,j}(0) - \mu_0\alpha_j P_1^*(s), \quad j = 1, \ldots, h, \tag{16}$$

$$[s - (\theta_j + \mu_j)]Q_{n,j}^*(s) + \mu_j Q_{n+1,j}^*(s) = Q_{n,j}(0) - a^*(s)\sum_{k=1}^{n} g_k Q_{n-k,j}(0), \quad n = 1, \ldots, b-1,\ j = 1, \ldots, h, \tag{17}$$

$$[s - (\theta_j + \mu_j)]Q_{n,j}^*(s) + \mu_j Q_{n+1,j}^*(s) = Q_{n,j}(0) - a^*(s)\sum_{k=1}^{b} g_k Q_{n-k,j}(0), \quad n \geq b, \quad j = 1, \ldots, h. \tag{18}$$

https://doi.org/10.1017/S0269964822000365 Published online by Cambridge University Press

Adding Eqs. (13) to (18), term by term on both sides, it yields

$$\sum_{n=0}^{\infty} P_n^*(s) + \sum_{n=0}^{\infty}\sum_{j=1}^{h} Q_{n,j}^*(s) = \frac{1-a^*(s)}{s}\left(\sum_{n=0}^{\infty} P_n(0) + \sum_{n=0}^{\infty}\sum_{j=1}^{h} Q_{n,j}(0)\right).$$

Notice that $P_n^*(s) \to P_n$, $Q_{n,j}^*(s) \to Q_{n,j}$ and $1 - a^*(s) \to 0$ as $s \to 0$, so L'Hôspital's Rule and the normalization condition give

$$1 = \sum_{n=0}^{\infty} P_n + \sum_{n=0}^{\infty}\sum_{j=1}^{h} Q_{n,j} = \frac{1}{\lambda}\left(\sum_{n=0}^{\infty} P_n(0) + \sum_{n=0}^{\infty}\sum_{j=1}^{h} Q_{n,j}(0)\right).$$

Let P_n^- denote the probability that an arriving batch finds the server in a normal working mode and sees n customers in the system, and $Q_{n,j}^-$ is defined as the probability that an incoming batch of the customers also finds n customers in the system and the server is on type j vacation. According to Bayes' theorem, the following formulas are obtained:

$$\begin{aligned} P_n^- &= \Pr\{n \text{ customers in the system prior to an arrival of a batch when the server} \\ &\qquad \text{is in normal working mode} \mid \text{a group of customers will arrive soon}\} \\ &= \frac{P_n(0)}{\sum_{n=0}^{\infty} P_n(0) + \sum_{n=0}^{\infty}\sum_{j=1}^{h} Q_{n,j}(0)} = \frac{1}{\lambda} P_n(0), \quad n = 0, 1, \ldots, \end{aligned} \tag{19}$$

$$\begin{aligned} Q_{n,j}^- &= \Pr\{n \text{ customers in the system prior to an arrival of a batch when the server} \\ &\qquad \text{is on type } j \text{ vacation} \mid \text{a group of customers will arrive soon}\} \\ &= \frac{Q_{n,j}(0)}{\sum_{n=0}^{\infty} P_n(0) + \sum_{n=0}^{\infty}\sum_{j=1}^{h} Q_{n,j}(0)} = \frac{1}{\lambda} Q_{n,j}(0), \quad n = 0, 1, \ldots, \ j = 1, \ldots, h. \end{aligned} \tag{20}$$

Clearly, it may be seen that once the expressions of $P_n(0)$, $P_n^*(s)$, $Q_{n,j}(0)$, and $Q_{n,j}^*(s)$ are given, the queue-length distribution both at pre-arrival (P_n^- and $Q_{n,j}^-$) and arbitrary epochs (P_n and $Q_{n,j}$) can be determined from these quantities. We will address this topic in the following section by using the shift operator method for solving the sets of difference equations (see (13)–(18)) that arise in analyzing such a queue.

4. Stationary queue-length distributions at two different epochs

For the purpose of analysis, the discrete variable n which takes on values in nonnegative integers will be viewed as the independent variable while $P_n(0)$, $P_n^*(s)$, $Q_{n,j}(0)$, and $Q_{n,j}^*(s)$ (whose value "depends" on the value of n) can be viewed as functions of n. The forward shift operator $\mathcal{D}$ acting on the sequences $\{P_n(0), n \geq 0\}$, $\{P_n^*(s), n \geq 0\}$, $\{Q_{n,j}(0), n \geq 0\}$, and $\{Q_{n,j}^*(s), n \geq 0\}$ is defined by

$$\begin{aligned} &\mathcal{D}^l P_n(0) = P_{n+l}(0), \quad \mathcal{D}^l P_n^*(s) = P_{n+l}^*(s), \quad l \geq 1, \\ &\mathcal{D}^l Q_{n,j}(0) = Q_{n+l,j}(0), \quad \mathcal{D}^l Q_{n,j}^*(s) = Q_{n+l,j}^*(s), \quad l \geq 1, \ j = 1, \ldots, h. \end{aligned}$$

With the aid of the notation of the forward shift operator $\mathcal{D}$, the difference equation (18) can be written as

$$[s - (\theta_j + \mu_j) + \mu_j \mathcal{D}] Q_{n,j}^*(s) = \left(\mathcal{D}^b - a^*(s)\sum_{k=1}^{b} g_k \mathcal{D}^{b-k}\right) Q_{n-b,j}(0), \quad n \geq b, \ j = 1, \ldots, h. \tag{21}$$

https://doi.org/10.1017/S0269964822000365 Published online by Cambridge University Press

Re-indexing the variable of $Q_{n-b,j}(0)$ as $n-b \longrightarrow n$, and setting $s = \theta_j + \mu_j - \mu_j \mathcal{D}$ for $j = 1, \ldots, h$, Eq. (21) can be reduced to a homogeneous difference equation with constant coefficients

$$\left(\mathcal{D}^b - a^*(\theta_j + \mu_j - \mu_j \mathcal{D}) \sum_{k=1}^{b} g_k \mathcal{D}^{b-k}\right) Q_{n,j}(0) = 0, \quad n \geq 0,\ j = 1, \ldots, h. \tag{22}$$

According to the basic theory of difference equation, for a fixed j, the characteristic equation associated with Eq. (22) is

$$z^b - a^*(\theta_j + \mu_j - \mu_j z) \sum_{k=1}^{b} g_k z^{b-k} = 0, \quad j = 1, \ldots, h. \tag{23}$$

Next, we use Rouché's theorem to find the number of zeros of the function $p(z) = z^b - a^*(\theta_j + \mu_j - \mu_j z) \sum_{k=1}^{b} g_k z^{b-k}$ that lies inside the unit circle. Our results will be presented in the form of the following lemma.

Lemma 1. *If $\lambda \bar{g}/\mu_j < 1$, then $p(z)$ has b zeros (counted with multiplicity) in the disk $\{|z| < 1\}$.*

Proof. To apply Rouché's theorem, we seek to express $p(z)$ in the form $p(z) = f(z) + h(z)$, where the function $f(z)$ dominates $h(z)$ on the unit circle, and where it is apparent how many zeros has inside the unit circle. Thus, our choice for $f(z)$ in this case is $f(z) = z^b$, which has b zeros inside the unit circle, all at the origin. Furthermore, we take $h(z) = -a^*(\theta_j + \mu_j - \mu_j z) \sum_{k=1}^{b} g_k z^{b-k}$. For notational convenience, let $U(z) = a^*(\theta_j + \mu_j - \mu_j z)$. When $\epsilon > 0$ is sufficiently small, $U(z)$ is analytic in $|z| \leq 1 + \epsilon$. Thus, according to the Taylor's theorem for analytic function, $U(z)$ can be represented as a power series $\sum_{l=0}^{\infty} u_l (z-1)^l$. Also, we have formulas for the coefficients $u_l = U^{(l)}(1)/l! = (1/2\pi \mathbf{i}) \oint_\Gamma (U(z)/(z-1)^{l+1})\, \mathrm{d}z$, where $\mathbf{i}$ is a square root of -1, and Γ is any simple closed curve in $|z| \leq 1 + \epsilon$ around 1. Now, employing the Taylor expansion for $U(z)$, we may estimate $|h(z)|$ and $|f(z)|$ on the simple closed curve $|z| = 1 - \delta$, where δ is a small positive real number. Since

$$\begin{aligned}
|h(z)|_{z=1-\delta} &= |U(z)| \left| \sum_{k=1}^{b} g_k z^{b-k} \right|_{z=1-\delta} \leq U(|z|) \sum_{k=1}^{b} g_k |z|_{z=1-\delta}^{b-k} = U(1-\delta) \sum_{k=1}^{b} g_k (1-\delta)^{b-k} \\
&= \left[U(1) + \frac{U'(1)}{1!}(1-\delta-1) + \sum_{l=2}^{\infty} \frac{U^{(l)}(1)}{l!}(1-\delta-1)^l \right] \sum_{k=1}^{b} g_k [1 - (b-k)\delta + \mathrm{o}(\delta)] \\
&= \left[\int_0^\infty e^{-\theta_j t}\, \mathrm{d}A(t) - \delta \int_0^\infty e^{-\theta_j t} \mu_j t\, \mathrm{d}A(t) + \mathrm{o}(\delta) \right] \sum_{k=1}^{b} g_k [1 - (b-k)\delta + \mathrm{o}(\delta)] \\
&= \left[\int_0^\infty e^{-\theta_j t} (1 - \delta \mu_j t)\, \mathrm{d}A(t) + \mathrm{o}(\delta) \right] \sum_{k=1}^{b} g_k [1 - (b-k)\delta + \mathrm{o}(\delta)] \\
&\leq \left[\int_0^\infty (1 - \delta \mu_j t)\, \mathrm{d}A(t) + \mathrm{o}(\delta) \right] \sum_{k=1}^{b} g_k [1 - (b-k)\delta + \mathrm{o}(\delta)] \\
&= \left[1 - \delta \frac{\mu_j}{\lambda} + \mathrm{o}(\delta) \right] [1 - b\delta + \bar{g}\delta + \mathrm{o}(\delta)] = 1 - b\delta - \left(\frac{\mu_j}{\lambda} - \bar{g} \right) \delta + \mathrm{o}(\delta) \\
&< |f(z)|_{z=1-\delta} = (1-\delta)^b = 1 - b\delta + \mathrm{o}(\delta),
\end{aligned}$$

and both $f(z)$ and $h(z)$ are analytic for $|z| \leq 1 - \delta$, letting δ tend to zero, Rouché's theorem tells us that $f(z)$ and $p(z) = f(z) + h(z)$ have the same number of zeros in $|z| < 1$, which is b. □

On the other hand, in the existing literature, authors like Chaudhry [8] and Tijms [25] have emphasized that root-finding in queueing theory is well structured, in the sense that the roots of the characteristic

https://doi.org/10.1017/S0269964822000365 Published online by Cambridge University Press

equation are distinct for most queueing models. Therefore, based on the above point of view, we assume that the b roots of the characteristic equation (23) are distinct, denoted by $r_{m,j}$, $m = 1, \ldots, b$, $j = 1, \ldots, h$. Then, an immediate consequence for this particular case is that the general solution of the homogeneous equation (22) is as follows

$$Q_{n,j}(0) = \sum_{m=1}^{b} c_{m,j} r_{m,j}^{n}, \quad n \geq 0, \; j = 1, \ldots, h, \tag{24}$$

where for fixed j, $c_{1,j}, c_{2,j}, \ldots$, and $c_{b,j}$ are real or complex constants whose values can be determined by a system of linear equations (see subsequent discussions). Substitution of Eq. (24) into Eq. (21) gives

$$\begin{aligned}
[s - (\theta_j + \mu_j) + \mu_j \mathcal{D}] Q_{n,j}^{*}(s) &= \sum_{m=1}^{b} c_{m,j} r_{m,j}^{n} - a^{*}(s) \sum_{k=1}^{b} g_k \left(\sum_{m=1}^{b} c_{m,j} r_{m,j}^{n-k} \right) \\
&= \sum_{m=1}^{b} c_{m,j} r_{m,j}^{n} - a^{*}(s) \sum_{m=1}^{b} c_{m,j} \sum_{k=1}^{b} g_k r_{m,j}^{n-k}, \quad n \geq b, \; j = 1, \ldots, h.
\end{aligned} \tag{25}$$

Any solution $Q_{n,j}^{*}(s)$ of Eq. (25) may be written as

$$Q_{n,j}^{*}(s) = Q_{n,j}^{*(\text{par})}(s) + Q_{n,j}^{*(\text{hom})}(s) = Q_{n,j}^{*(\text{par})}(s) + \mathbb{C}_1 \left(1 + \frac{\theta_j - s}{\mu_j} \right)^{n}, \quad n \geq b,$$

where $Q_{n,j}^{*(\text{par})}(s)$ is a particular solution of the nonhomogeneous equation (25), and $Q_{n,j}^{*(\text{hom})}(s) = \mathbb{C}_1 (1 + (\theta_j - s)/\mu_j)^n$ is the general solution of the corresponding homogenous equation $[s - (\theta_j + \mu_j) + \mu_j \mathcal{D}] Q_{n,j}^{*}(s) = 0$. It is easy to verify that the undetermined constant $\mathbb{C}_1 = 0$ because of the condition $\sum_{n=b}^{\infty} Q_{n,j}^{*}(0) = \sum_{n=b}^{\infty} Q_{n,j} < \infty$. Thus, for $n \geq b$, we finally have $Q_{n,j}^{*}(s) = Q_{n,j}^{*(\text{par})}(s)$.

Now, we focus our attention on finding a particular solution of Eq. (25). Since the nonhomogeneous term is a finite linear combination of the function $r_{m,j}^{n}$, the method of undetermined coefficient is used to find $Q_{n,j}^{*(\text{par})}(s)$. Here, we can guess the trial solutions with undetermined coefficients, plug them into the difference equation (25), and then solve for the unknown coefficients to obtain the particular solution as below

$$Q_{n,j}^{*}(s) = Q_{n,j}^{*(\text{par})}(s) = \sum_{m=1}^{b} c_{m,j} \frac{r_{m,j}^{n} - a^{*}(s) \sum_{k=1}^{b} g_k r_{m,j}^{n-k}}{s - \theta_j - \mu_j (1 - r_{m,j})}, \quad n \geq b, \; j = 1, \ldots, h. \tag{26}$$

Since Eqs. (17) and (18) have almost exactly the same form, we now seek the suitable conditions under which $Q_{n,j}^{*}(s)|_{s=0}$ has the same expression as in Eq. (26) for $1 \leq n \leq b-1$, when $s = 0$. This point can be justified by the matrix geometric algorithm (see Neuts [21]), and is not our intuitive guess. Actually, by assuming that the size of an arriving batch is bounded, the jumps to the right of the embedded Markov chain are bounded (although our approach can avoid discussion of the embedded Markov chain). This implies that the Markov chain fits into the standard matrix-geometric framework by appropriately defining the levels, and its stationary distribution is readily available in matrix-geometric form. In other words, for $n = 0, 1, 2, \ldots$, if we let $\boldsymbol{\pi}_n = (P_n^{-}, Q_{n,1}^{-}, Q_{n,2}^{-}, \ldots, Q_{n,h}^{-})$, Neuts has shown that there exists a positive matrix $\boldsymbol{R}$ such that $\boldsymbol{\pi}_n = \boldsymbol{\pi}_0 \boldsymbol{R}^n$. Furthermore, from the results regarding the steady-state queue-length distributions immediately before batch arrival, we can derive the stationary queue-length distribution at arbitrary epochs by employing the classical arguments based on renewal theory and semi-Markov process. For the above reasons, and noting that $P_n^{*}(s)|_{s=0} = P_n$ and $Q_{n,j}^{*}(s)|_{s=0} = Q_{n,j}$, we think the expressions of $Q_{n,j}^{*}(0)$ and $P_n^{*}(0)$ can necessarily be unified regardless of $n \geq b$ or not. Hence, by putting Eq. (24) into Eqs. (18) and (17), and comparing the last term of the right-hand side of Eqs. (18) and (17) gives the following relation, which the $c_{m,j}$ must satisfy:

$$\sum_{k=1}^{n} g_k \sum_{m=1}^{b} c_{m,j} r_{m,j}^{n-k} = \sum_{k=1}^{b} g_k \sum_{m=1}^{b} c_{m,j} r_{m,j}^{n-k} \Rightarrow \sum_{m=1}^{b} c_{m,j} \sum_{k=1}^{n} g_k r_{m,j}^{n-k} = \sum_{m=1}^{b} c_{m,j} \sum_{k=1}^{b} g_k r_{m,j}^{n-k}$$
$$\Rightarrow \sum_{m=1}^{b} c_{m,j} \sum_{k=n+1}^{b} g_k r_{m,j}^{n-k} = 0, \quad n = b-1, b-2, \ldots, 1, \ j = 1, 2, \ldots, h. \tag{27}$$

Notice that for $g_b \neq 0$, setting $n = b-1, b-2, \ldots, 1$ in Eq. (27), respectively, allows us to conclude that

$$\begin{cases} \sum_{m=1}^{b} \frac{c_{m,1}}{r_{m,1}} = 0 \\ \sum_{m=1}^{b} \frac{c_{m,1}}{r_{m,1}^2} = 0 \\ \vdots \\ \sum_{m=1}^{b} \frac{c_{m,1}}{r_{m,1}^{b-2}} = 0 \\ \sum_{m=1}^{b} \frac{c_{m,1}}{r_{m,1}^{b-1}} = 0 \end{cases}, \quad \begin{cases} \sum_{m=1}^{b} \frac{c_{m,2}}{r_{m,2}} = 0 \\ \sum_{m=1}^{b} \frac{c_{m,2}}{r_{m,2}^2} = 0 \\ \vdots \\ \sum_{m=1}^{b} \frac{c_{m,2}}{r_{m,2}^{b-2}} = 0 \\ \sum_{m=1}^{b} \frac{c_{m,2}}{r_{m,2}^{b-1}} = 0 \end{cases}, \ldots\ldots, \begin{cases} \sum_{m=1}^{b} \frac{c_{m,h-1}}{r_{m,h-1}} = 0 \\ \sum_{m=1}^{b} \frac{c_{m,h-1}}{r_{m,h-1}^2} = 0 \\ \vdots \\ \sum_{m=1}^{b} \frac{c_{m,h-1}}{r_{m,h-1}^{b-2}} = 0 \\ \sum_{m=1}^{b} \frac{c_{m,h-1}}{r_{m,h-1}^{b-1}} = 0 \end{cases}, \quad \begin{cases} \sum_{m=1}^{b} \frac{c_{m,h}}{r_{m,h}} = 0 \\ \sum_{m=1}^{b} \frac{c_{m,h}}{r_{m,h}^2} = 0 \\ \vdots \\ \sum_{m=1}^{b} \frac{c_{m,h}}{r_{m,h}^{b-2}} = 0 \\ \sum_{m=1}^{b} \frac{c_{m,h}}{r_{m,h}^{b-1}} = 0 \end{cases}. \tag{28}$$

For $j = 1, 2, \ldots, h$, let $\boldsymbol{c}_j = (c_{1,j}, c_{2,j}, \ldots, c_{b,j})^{\mathsf{T}}$ and

$$\boldsymbol{R}_j = \begin{pmatrix} r_{1,j}^{-1} & r_{2,j}^{-1} & \cdots & r_{b-1,j}^{-1} & r_{b,j}^{-1} \\ r_{1,j}^{-2} & r_{2,j}^{-2} & \cdots & r_{b-1,j}^{-2} & r_{b,j}^{-2} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ r_{1,j}^{-(b-2)} & r_{2,j}^{-(b-2)} & \cdots & r_{b-1,j}^{-(b-2)} & r_{b,j}^{-(b-2)} \\ r_{1,j}^{-(b-1)} & r_{2,j}^{-(b-1)} & \cdots & r_{b-1,j}^{-(b-1)} & r_{b,j}^{-(b-1)} \end{pmatrix},$$

where "T" represents the transpose operation on a vector or a matrix. With these notations, Eq. (27) may be written in matrix form as

$$\begin{pmatrix} \boldsymbol{R}_1 & \boldsymbol{0} & \boldsymbol{0} & \cdots & \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{R}_2 & \boldsymbol{0} & \cdots & \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{R}_3 & \cdots & \boldsymbol{0} & \boldsymbol{0} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} & \cdots & \boldsymbol{R}_{h-1} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} & \cdots & \boldsymbol{0} & \boldsymbol{R}_h \end{pmatrix} \begin{pmatrix} \boldsymbol{c}_1 \\ \boldsymbol{c}_2 \\ \boldsymbol{c}_3 \\ \vdots \\ \boldsymbol{c}_{h-1} \\ \boldsymbol{c}_h \end{pmatrix} = \begin{pmatrix} \boldsymbol{0} \\ \boldsymbol{0} \\ \boldsymbol{0} \\ \vdots \\ \boldsymbol{0} \\ \boldsymbol{0} \end{pmatrix}, \tag{29}$$

where $\boldsymbol{0}$ denotes a zero matrix or a zero column vector of appropriate dimension. If the constants $c_{1,1}$, …, $c_{b,1}$, …, $c_{1,h}$, …, $c_{b,h}$ satisfy Eq. (28), the same expression for $Q_{n,j}^*(s)$ can be derived from Eqs. (17) and (18). This also means that for any $n \geq 1$, we have

$$Q_{n,j}^*(s) = \sum_{m=1}^{b} c_{m,j} \frac{r_{m,j}^n - a^*(s) \sum_{k=1}^{b} g_k r_{m,j}^{n-k}}{s - \theta_j - \mu_j (1 - r_{m,j})}, \quad j = 1, \ldots, h. \tag{30}$$

https://doi.org/10.1017/S0269964822000365 Published online by Cambridge University Press

However, since the number of unknowns is greater than the number of equations, the constants $c_{m,j}$, $m = 1, \dots, b$, $j = 1, \dots, h$ can not be uniquely determined by Eq. (29). So we need to find more conditions to get $c_{m,j}$.

Again, as in Eq. (18) treated above, Eq. (15) can also be expressed in terms of the forward shift operator $\mathcal{D}$ as below

$$(s - \mu_0 + \mu_0 \mathcal{D}) P_n^*(s) = \left(\mathcal{D}^b - a^*(s) \sum_{k=1}^{b} g_k \mathcal{D}^{b-k} \right) P_{n-b}(0) - \sum_{j=1}^{h} \theta_j Q_{n,j}^*(s), \quad n \geq b. \tag{31}$$

Letting $s = \mu_0 - \mu_0 \mathcal{D}$ on both sides of Eq. (30) gives

$$\left(\mathcal{D}^b - a^*(\mu_0 - \mu_0 \mathcal{D}) \sum_{k=1}^{b} g_k \mathcal{D}^{b-k} \right) P_{n-b}(0) = \sum_{j=1}^{h} \theta_j Q_{n,j}^*(\mu_0 - \mu_0 \mathcal{D}), \quad n \geq b. \tag{32}$$

Similarly, by setting $n \geq 0$ instead of $n \geq b$ and substituting Eq. (30) into Eq. (32) yields

$$\begin{aligned} \left(\mathcal{D}^b - a^*(\mu_0 - \mu_0 \mathcal{D}) \sum_{k=1}^{b} g_k \mathcal{D}^{b-k} \right) P_n(0) &= \sum_{j=1}^{h} \theta_j Q_{n+b,j}^*(\mu_0 - \mu_0 \mathcal{D}) \\ &= \sum_{j=1}^{h} \theta_j \sum_{m=1}^{b} c_{m,j} \left\{ \frac{r_{m,j}^b - a^*(\mu_0 - \mu_0 \mathcal{D}) \sum_{k=1}^{b} g_k r_{m,j}^{b-k}}{\mu_0 - \mu_0 \mathcal{D} - \theta_j - \mu_j (1 - r_{m,j})} \right\} r_{m,j}^n, \quad n \geq 0. \end{aligned} \tag{33}$$

Eq. (33) is a nonhomogeneous equation for $P_n(0)$ as a function of n. The general solution of Eq. (33) consists of the sum of the solution to the homogeneous equation

$$\left(\mathcal{D}^b - a^*(\mu_0 - \mu_0 \mathcal{D}) \sum_{k=1}^{b} g_k \mathcal{D}^{b-k} \right) P_n(0) = 0, \quad n \geq 0, \tag{34}$$

and any particular solution of Eq. (33). Notice that the characteristic equation associated with Eq. (34) is

$$z^b - a^*(\mu_0 - \mu_0 z) \sum_{k=1}^{b} g_k z^{b-k} = 0. \tag{35}$$

If we start with Eq. (35) and still choose $f(z) = z^b$ and $q(z) = -a^*(\mu_0 - \mu_0 z) \sum_{k=1}^{b} g_k z^{b-k}$, analogous to the proof of Lemma 1, we can prove the characteristic equation (35) has exactly b roots inside the unit disk under the assumption that $\bar{g}\lambda < \mu_0$. Let these roots be denoted by $\omega_1, \omega_2, \dots, \omega_b$, and also assume that they are distinct. Then the general solution to the homogeneous equation (34) is $\sum_{m=1}^{b} f_m \omega_m^n$, where $f_1, f_2, \dots,$ and f_b are b arbitrary constants that need to be determined. Next, to find a particular solution of Eq. (33), we observe the structure of the nonhomogeneous term, and guess that it has to have the form

$$P_n^{(\text{par})}(0) = \sum_{j=1}^{h} \sum_{m=1}^{b} d_{m,j} r_{m,j}^n, \quad n \geq 0, \tag{36}$$

where $d_{m,j}$ can be expressed in terms of $c_{m,j}$ and $r_{m,j}$ by substituting Eq. (36) into Eq. (33). Doing this gives

$$d_{m,j} = \frac{c_{m,j} \theta_j}{(\mu_0 - \mu_j)(1 - r_{m,j}) - \theta_j}, \quad m = 1, \dots, b, \ j = 1, \dots, h.$$

https://doi.org/10.1017/S0269964822000365 Published online by Cambridge University Press

Hence final general solution to the nonhomogeneous equation (32) is given by

$$P_n(0) = \sum_{m=1}^{b} f_m \omega_m^n + \sum_{j=1}^{h} \sum_{m=1}^{b} \frac{c_{m,j}\theta_j}{(\mu_0 - \mu_j)(1 - r_{m,j}) - \theta_j} r_{m,j}^n, \quad n \geq 0. \tag{37}$$

The task now is to find the expression of $P_n^*(s)$. Plugging Eqs. (30) and (37) into Eq. (31), and after some algebraic manipulation, we get

$$\begin{aligned}
&(s - \mu_0 + \mu_0 \mathcal{D}) P_n^*(s) \\
&= \left(\mathcal{D}^b - a^*(s) \sum_{k=1}^{b} g_k \mathcal{D}^{b-k}\right) \left(\sum_{m=1}^{b} f_m \omega_m^{n-b} + \sum_{j=1}^{h} \sum_{m=1}^{b} \frac{c_{m,j}\theta_j r_{m,j}^{n-b}}{(\mu_0 - \mu_j)(1 - r_{m,j}) - \theta_j}\right) \\
&\quad - \sum_{j=1}^{h} \theta_j \sum_{m=1}^{b} c_{m,j} \frac{r_{m,j}^n - a^*(s) \sum_{k=1}^{b} g_k r_{m,j}^{n-k}}{s - \theta_j - \mu_j(1 - r_{m,j})} \\
&= \sum_{m=1}^{b} f_m \left(\omega_m^n - a^*(s) \sum_{k=1}^{b} g_k \omega_m^{n-k}\right) + \sum_{j=1}^{h} \sum_{m=1}^{b} \frac{c_{m,j}\theta_j (r_{m,j}^n - a^*(s) \sum_{k=1}^{b} g_k r_{m,j}^{n-k})}{(\mu_0 - \mu_j)(1 - r_{m,j}) - \theta_j} \\
&\quad - \sum_{j=1}^{h} \sum_{m=1}^{b} c_{m,j}\theta_j \left\{\frac{r_{m,j}^n - a^*(s) \sum_{k=1}^{b} g_k r_{m,j}^{n-k}}{s - \theta_j - \mu_j(1 - r_{m,j})}\right\} \\
&= \sum_{m=1}^{b} f_m \left(1 - a^*(s) \sum_{k=1}^{b} g_k \omega_m^{-k}\right) \omega_m^n \\
&\quad + \sum_{j=1}^{h} \sum_{m=1}^{b} \frac{c_{m,j}\theta_j (1 - a^*(s) \sum_{k=1}^{b} g_k r_{m,j}^{-k})[s - \mu_0(1 - r_{m,j})]}{[(\mu_0 - \mu_j)(1 - r_{m,j}) - \theta_j][s - \theta_j - \mu_j(1 - r_{m,j})]} r_{m,j}^n, \quad n \geq b.
\end{aligned} \tag{38}$$

This indicates that the sequence $\{P_n^*(s), n \geq b\}$ satisfies the first-order nonhomogeneous difference equation (38). Similar to our previous discussion, we of course first find the general solution to the corresponding homogeneous equation of Eq. (38). Here, the general solution is given as $P_n^{*(\mathrm{hom})}(s) = \mathbb{C}_2(1 - s/\mu_0)^n$, and $\mathbb{C}_2$ is an undetermined constant. Notice also that the nonhomogeneous term is a linear combination of ω_m^n and $r_{m,j}^n$, a particular solution of Eq. (38) can be given by

$$\begin{aligned}
P_n^{*(\mathrm{par})}(s) &= \sum_{m=1}^{b} \frac{f_m[1 - a^*(s)G(\omega_m^{-1})]}{s - \mu_0(1 - \omega_m)} \omega_m^n \\
&\quad + \sum_{j=1}^{h} \sum_{m=1}^{b} \frac{c_{m,j}\theta_j [1 - a^*(s)G(r_{m,j}^{-1})] r_{m,j}^n}{[(\mu_0 - \mu_j)(1 - r_{m,j}) - \theta_j][s - \theta_j - \mu_j(1 - r_{m,j})]}, \quad n \geq b.
\end{aligned} \tag{39}$$

Thus, for $n \geq b$, the general solution to Eq. (38) is the sum of the homogeneous solution and the particular solution (see Eq. (39)),

$$\begin{aligned}
P_n^*(s) &= \mathbb{C}_2 \left(1 - \frac{s}{\mu}\right)^n + \sum_{m=1}^{b} \frac{f_m[1 - a^*(s)G(\omega_m^{-1})]}{s - \mu_0(1 - \omega_m)} \omega_m^n \\
&\quad + \sum_{j=1}^{h} \sum_{m=1}^{b} \frac{c_{m,j}\theta_j [1 - a^*(s)G(r_{m,j}^{-1})]}{[(\mu_0 - \mu_j)(1 - r_{m,j}) - \theta_j][s - \theta_j - \mu_j(1 - r_{m,j})]} r_{m,j}^n, \quad n \geq b.
\end{aligned} \tag{40}$$

https://doi.org/10.1017/S0269964822000365 Published online by Cambridge University Press

Since $\sum_{n=b}^{\infty} P_n^*(0) = \sum_{n=b}^{\infty} P_n < \infty$, we must have $\mathbb{C}_2 = 0$. Hence, Eq. (40) reduces to

$$P_n^*(s) = \sum_{m=1}^{b} \frac{f_m[1 - a^*(s)G(\omega_m^{-1})]}{s - \mu_0(1-\omega_m)} \omega_m^n + \sum_{j=1}^{h} \sum_{m=1}^{b} \frac{c_{m,j}\theta_j[1 - a^*(s)G(r_{m,j}^{-1})] r_{m,j}^n}{[(\mu_0 - \mu_j)(1 - r_{m,j}) - \theta_j][s - \theta_j - \mu_j(1 - r_{m,j})]}, \quad n \geq b. \tag{41}$$

Then, we do wish to find the right conditions such that the expression for $P_n^*(s)|_{s=0}$ presented in Eq. (41) still holds when $1 \leq n \leq b-1$ and $s = 0$. We compare the second term on the right-hand side of Eqs. (14) and (15) and conclude that f_m and $c_{m,j}$ satisfy the relation

$$\sum_{m=1}^{b} f_m \sum_{k=n+1}^{b} g_k \omega_m^{n-k} + \sum_{j=1}^{h} \sum_{m=1}^{b} c_{m,j} \sum_{k=n+1}^{b} g_k \times \frac{\theta_j r_{m,j}^{n-k}}{(\mu_0 - \mu_j)(1 - r_{m,j}) - \theta_j} = 0, \quad n = b-1, b-2, \ldots, 1. \tag{42}$$

Let us take $n = b-1, b-2, \ldots$, and 1 in Eq. (42), respectively, and note that for $g_b \neq 0$, a system of linear equations in $b + hb$ variables can be derived from Eq. (42)

$$\begin{cases} \sum_{m=1}^{b} \frac{f_m}{\omega_m} + \sum_{j=1}^{h} \sum_{m=1}^{b} \frac{c_{m,j}\theta_j}{[(\mu_0 - \mu_j)(1 - r_{m,j}) - \theta_j] r_{m,j}} = 0 \\ \sum_{m=1}^{b} \frac{f_m}{\omega_m^2} + \sum_{j=1}^{h} \sum_{m=1}^{b} \frac{c_{m,j}\theta_j}{[(\mu_0 - \mu_j)(1 - r_{m,j}) - \theta_j] r_{m,j}^2} = 0 \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ \sum_{m=1}^{b} \frac{f_m}{\omega_m^{b-1}} + \sum_{j=1}^{h} \sum_{m=1}^{b} \frac{c_{m,j}\theta_j}{[(\mu_0 - \mu_j)(1 - r_{m,j}) - \theta_j] r_{m,j}^{b-1}} = 0 \end{cases}. \tag{43}$$

In other words, when $1 \leq n \leq b-1$, if the above Eq. (43) holds, then $P_n^*(s)$ has the following uniform expression

$$P_n^*(s) = \sum_{m=1}^{b} \frac{f_m[1 - a^*(s)G(\omega_m^{-1})]}{s - \mu_0(1-\omega_m)} \omega_m^n + \sum_{j=1}^{h} \sum_{m=1}^{b} \frac{c_{m,j}\theta_j[1 - a^*(s)G(r_{m,j}^{-1})] r_{m,j}^n}{[(\mu_0 - \mu_j)(1 - r_{m,j}) - \theta_j][s - \theta_j - \mu_j(1 - r_{m,j})]}, \quad n \geq 1. \tag{44}$$

https://doi.org/10.1017/S0269964822000365 Published online by Cambridge University Press

Furthermore, letting $s = \theta_j$ in Eq. (16) and substituting Eqs. (24), (30), and (44) into Eq. (16) gives another linear system with h equations and $hb + b$ variables.

$$\left\{\begin{array}{l}
\sum_{m=1}^{b} c_{m,1}\left(\frac{1-a^*(\theta_1)G(r_{m,1}^{-1})r_{m,1}}{r_{m,1}-1}+\frac{\alpha_1\mu_0\theta_1[1-a^*(\theta_1)G(r_{m,1}^{-1})]r_{m,1}}{[(\mu_0-\mu_1)(1-r_{m,1})-\theta_1]\mu_1(r_{m,1}-1)}\right)\\
\quad+\sum_{l=2}^{h}\sum_{m=1}^{b}\frac{\alpha_1\mu_0 c_{m,l}\theta_l[1-a^*(\theta_1)G(r_{m,l}^{-1})]r_{m,l}}{[(\mu_0-\mu_l)(1-r_{m,l})-\theta_l][\theta_1-\theta_l-\mu_l(1-r_{m,l})]}\\
\quad+\sum_{m=1}^{b}\frac{\alpha_1\mu_0 f_m[1-a^*(\theta_1)G(\omega_m^{-1})]\omega_m}{\theta_1-\mu_0(1-\omega_m)}=0\\
\sum_{m=1}^{b} c_{m,2}\left(\frac{1-a^*(\theta_2)G(r_{m,2}^{-1})r_{m,2}}{r_{m,2}-1}+\frac{\alpha_2\mu_0\theta_2[1-a^*(\theta_2)G(r_{m,2}^{-1})]r_{m,2}}{[(\mu_0-\mu_2)(1-r_{m,2})-\theta_2]\mu_2(r_{m,2}-1)}\right)\\
\quad+\sum_{\substack{l=1\\ l\neq 2}}^{h}\sum_{m=1}^{b}\frac{\alpha_2\mu_0 c_{m,l}\theta_l[1-a^*(\theta_2)G(r_{m,l}^{-1})]r_{m,l}}{[(\mu_0-\mu_l)(1-r_{m,l})-\theta_l][\theta_2-\theta_l-\mu_l(1-r_{m,l})]}\\
\quad+\sum_{m=1}^{b}\frac{\alpha_2\mu_0 f_m[1-a^*(\theta_2)G(\omega_m^{-1})]\omega_m}{\theta_2-\mu_0(1-\omega_m)}=0\\
\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\\
\sum_{m=1}^{b} c_{m,h}\left(\frac{1-a^*(\theta_h)G(r_{m,h}^{-1})r_{m,h}}{r_{m,h}-1}+\frac{\alpha_h\mu_0\theta_h[1-a^*(\theta_h)G(r_{m,h}^{-1})]r_{m,h}}{[(\mu_0-\mu_h)(1-r_{m,h})-\theta_h]\mu_h(r_{m,h}-1)}\right)\\
\quad+\sum_{l=1}^{h-1}\sum_{m=1}^{b}\frac{\alpha_h\mu_0 c_{m,l}\theta_l[1-a^*(\theta_h)G(r_{m,l}^{-1})]r_{m,l}}{[(\mu_0-\mu_l)(1-r_{m,l})-\theta_l][\theta_h-\theta_l-\mu_l(1-r_{m,l})]}\\
\quad+\sum_{m=1}^{b}\frac{\alpha_h\mu_0 f_m[1-a^*(\theta_h)G(\omega_m^{-1})]\omega_m}{\theta_h-\mu_0(1-\omega_m)}=0
\end{array}\right. . \tag{45}$$

To enhance the calculation and simplify the programs coded in Matlab, we rewrite Eqs. (43) and (45) in the matrix form as below

$$\begin{pmatrix} \boldsymbol{H}_1 & \boldsymbol{H}_2 & \cdots & \boldsymbol{H}_h & \boldsymbol{W} \\ \boldsymbol{\Lambda}_1 & \boldsymbol{\Lambda}_2 & \cdots & \boldsymbol{\Lambda}_h & \boldsymbol{B} \end{pmatrix}\begin{pmatrix} \boldsymbol{c}_1 \\ \vdots \\ \boldsymbol{c}_h \\ \boldsymbol{f} \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 0 \end{pmatrix}_{(b-1+h)\times 1}, \tag{46}$$

where $\boldsymbol{f} = (f_1, f_2, \ldots, f_b)^{\mathsf{T}}$,

$$\boldsymbol{H}_j = \begin{pmatrix} \zeta_{1,j}r_{1,j}^{-1} & \zeta_{2,j}r_{2,j}^{-1} & \cdots & \zeta_{b,j}r_{b,j}^{-1} \\ \zeta_{1,j}r_{1,j}^{-2} & \zeta_{2,j}r_{2,j}^{-2} & \cdots & \zeta_{b,j}r_{b,j}^{-2} \\ \vdots & \vdots & \vdots & \vdots \\ \zeta_{1,j}r_{1,j}^{-(b-1)} & \zeta_{2,j}r_{2,j}^{-(b-1)} & \cdots & \zeta_{b,j}r_{b,j}^{-(b-1)} \end{pmatrix}, \quad j = 1, 2, \ldots, h,$$

$$\boldsymbol{\Lambda}_j = j\text{th row}\begin{pmatrix} \eta_{1,j}^{(1)} & \eta_{2,j}^{(1)} & \cdots & \eta_{b-1,j}^{(1)} & \eta_{b,j}^{(1)} \\ \eta_{1,j}^{(2)} & \eta_{2,j}^{(2)} & \cdots & \eta_{b-1,j}^{(2)} & \eta_{b,j}^{(2)} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \psi_{1,j} & \psi_{2,j} & \cdots & \psi_{b-1,j} & \psi_{b,j} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \eta_{1,j}^{(h)} & \eta_{2,j}^{(h)} & \cdots & \eta_{b-1,j}^{(h)} & \eta_{b,j}^{(h)} \end{pmatrix}, \quad j = 1, 2, \ldots, h,$$

https://doi.org/10.1017/S0269964822000365 Published online by Cambridge University Press

$$\boldsymbol{W} = \begin{pmatrix} \omega_1^{-1} & \omega_2^{-1} & \cdots & \omega_b^{-1} \\ \omega_1^{-2} & \omega_2^{-2} & \cdots & \omega_b^{-2} \\ \vdots & \vdots & \vdots & \vdots \\ \omega_1^{-(b-1)} & \omega_2^{-(b-1)} & \cdots & \omega_b^{-(b-1)} \end{pmatrix}, \quad \boldsymbol{B} = \begin{pmatrix} \beta_{1,1} & \beta_{2,1} & \cdots & \beta_{b,1} \\ \beta_{1,2} & \beta_{2,2} & \cdots & \beta_{b,2} \\ \vdots & \vdots & \vdots & \vdots \\ \beta_{1,h} & \beta_{2,h} & \cdots & \beta_{b,h} \end{pmatrix}.$$

The elements in matrices $\boldsymbol{H}_j$, $\boldsymbol{\Lambda}_j$, and $\boldsymbol{B}$ are given, respectively, by

$$\zeta_{m,j} = \frac{\theta_j}{(\mu_0 - \mu_j)(1 - r_{m,j}) - \theta_j}, \quad m = 1, \ldots, b, \ j = 1, \ldots, h,$$

$$\eta_{m,j}^{(i)} = \frac{\alpha_i \mu_0 \theta_j [1 - a^*(\theta_i) G(r_{m,j}^{-1})] r_{m,j}}{[(\mu_0 - \mu_j)(1 - r_{m,j}) - \theta_j][\theta_i - \theta_j - \mu_j(1 - r_{m,j})]}, \quad m = 1, \ldots, b, \ i, j = 1, \ldots, h, \ i \neq j,$$

$$\psi_{m,j} = \frac{1 - a^*(\theta_j) G(r_{m,j}^{-1}) r_{m,j}}{r_{m,j} - 1} + \frac{\alpha_j \mu_0 \theta_j [1 - a^*(\theta_j) G(r_{m,j}^{-1})] r_{m,j}}{[(\mu_0 - \mu_j)(1 - r_{m,j}) - \theta_j] \mu_j (r_{m,j} - 1)}, \quad m = 1, \ldots, b, \ j = 1, \ldots, h,$$

$$\beta_{m,j} = \frac{\alpha_j \mu_0 [1 - a^*(\theta_j) G(\omega_m^{-1})] \omega_m}{\theta_j - \mu_0 (1 - \omega_m)}, \quad m = 1, \ldots, b, \ j = 1, \ldots, h.$$

Since $\sum_{n=0}^{\infty} P_n(0) + \sum_{n=0}^{\infty} \sum_{j=1}^{h} Q_{n,j}(0) = \lambda$, for $m = 1, \ldots, b$ and $j = 1, \ldots, h$, we can further derive a relationship for $c_{m,j}$ and f_m

$$\sum_{m=1}^{b} \frac{f_m}{1 - \omega_m} + \sum_{j=1}^{h} \sum_{m=1}^{b} \frac{c_{m,j}(\mu_0 - \mu_j)}{(\mu_0 - \mu_j)(1 - r_{m,j}) - \theta_j} = \lambda. \tag{47}$$

Obviously, Eqs. (29), (46) and (47) constitute a system of $hb+b$ linear equations with $hb+b$ unknowns to compute the undetermined constants. Having found $c_{m,j}$ and f_m, we are now able to give the stationary queue-length distributions at pre-arrival and arbitrary epochs. According to Eqs. (19) and (20), we have

$$P_n^- = \frac{1}{\lambda} P_n(0) = \frac{1}{\lambda} \left(\sum_{m=1}^{b} f_m \omega_m^n + \sum_{j=1}^{h} \sum_{m=1}^{b} \frac{c_{m,j} \theta_j}{(\mu_0 - \mu_j)(1 - r_{m,j}) - \theta_j} r_{m,j}^n \right), \quad n \geq 0, \tag{48}$$

$$Q_{n,j}^- = \frac{1}{\lambda} Q_{n,j}(0) = \frac{1}{\lambda} \sum_{m=1}^{b} c_{m,j} r_{m,j}^n, \quad n \geq 0, \ j = 1, \ldots, h. \tag{49}$$

Taking $s = 0$ in Eqs. (30) and (44), we further have

$$\begin{aligned} P_n = P_n^*(0) &= \sum_{m=1}^{b} \frac{f_m [G(\omega_m^{-1}) - 1] \omega_m^n}{\mu_0 (1 - \omega_m)} \\ &+ \sum_{j=1}^{h} \sum_{m=1}^{b} \frac{c_{m,j} \theta_j [G(r_{m,j}^{-1}) - 1] r_{m,j}^n}{[(\mu_0 - \mu_j)(1 - r_{m,j}) - \theta_j][\theta_j + \mu_j (1 - r_{m,j})]}, \quad n \geq 1 \end{aligned} \tag{50}$$

$$Q_{n,j} = Q_{n,j}^*(0) = \sum_{m=1}^{b} \frac{c_{m,j} [G(r_{m,j}^{-1}) - 1] r_{m,j}^n}{\theta_j + \mu_j (1 - r_{m,j})}, \quad n \geq 1, \ j = 1, \ldots, h. \tag{51}$$

https://doi.org/10.1017/S0269964822000365 Published online by Cambridge University Press

For $n = 0$, $Q_{0,j}$ can be derived from Eq. (16) by making the substitution $s = 0$. With some algebraic manipulation, it can be shown that

$$Q_{0,j} = Q^*_{0,j}(0) = \frac{1}{\theta_j}\left[\mu_j \sum_{m=1}^{b} \frac{c_{m,j}[G(r^{-1}_{m,j}) - 1]r_{m,j}}{\theta_j + \mu_j(1 - r_{m,j})} - \sum_{m=1}^{b} c_{m,j} + \mu_0\alpha_j\left(\sum_{m=1}^{b} \frac{f_m[G(\omega^{-1}_m) - 1]\omega_m}{\mu_0(1 - \omega_m)} + \sum_{l=1}^{h}\sum_{m=1}^{b} \frac{c_{m,l}\theta_l[G(r^{-1}_{m,l}) - 1]r_{m,l}}{[(\mu_0 - \mu_l)(1 - r_{m,l}) - \theta_l][\theta_l + \mu_l(1 - r_{m,l})]}\right)\right], \quad j = 1, \ldots, h. \tag{52}$$

As for the stationary probability P_0, it may be determined from the normalization equation

$$\begin{aligned}P_0 &= 1 - \sum_{n=1}^{\infty} P_n - \sum_{n=1}^{\infty}\sum_{j=1}^{h} Q_{n,j} - \sum_{j=1}^{h} Q_{0,j} \\ &= 1 - \sum_{j=1}^{h}\sum_{m=1}^{b} \frac{c_{m,j}[G(r^{-1}_{m,j}) - 1](\mu_0 - \mu_j)r_{m,j}}{[(\mu_0 - \mu_j)(1 - r_{m,j}) - \theta_j][\theta_j + \mu_j(1 - r_{m,j})]} - \sum_{m=1}^{b} \frac{f_m[G(\omega^{-1}_m) - 1]\omega_m}{\mu_0(1 - \omega_m)^2} \\ &\quad - \sum_{j=1}^{h} \frac{1}{\theta_j}\left[\mu_j \sum_{m=1}^{b} \frac{c_{m,j}[G(r^{-1}_{m,j}) - 1]r_{m,j}}{\theta_j + \mu_j(1 - r_{m,j})} - \sum_{m=1}^{b} c_{m,j} + \mu_0\alpha_j\left(\sum_{m=1}^{b} \frac{f_m[G(\omega^{-1}_m) - 1]\omega_m}{\mu_0(1 - \omega_m)}\right.\right. \\ &\quad \left.\left. + \sum_{j=1}^{h}\sum_{m=1}^{b} \frac{c_{m,j}\theta_j[G(r^{-1}_{m,j}) - 1]r_{m,j}}{[(\mu_0 - \mu_j)(1 - r_{m,j}) - \theta_j][\theta_j + \mu_j(1 - r_{m,j})]}\right)\right].\end{aligned} \tag{53}$$

It is not difficult to see that the most critical step in the above queueing analysis is to construct a set of linear algebraic equations based on a discrete dynamic system and its roots of the characteristic equation. After solving the system of linear equations, we may combine Eqs. (48) and (49) to compute the sojourn time of an arbitrary customer. Moreover, the expected queue length can also be numerically obtained to verify that whether the expected queue length and the average sojourn time satisfy Little's Law.

5. Sojourn time distribution in the system

Here, the main purpose of studying the sojourn time of an arbitrary customer is to provide a way to verify the correctness of the algorithm for computing the queue-length distribution.

The time that a customer spends in the system, from the instant of its arrival to the queue to the instant of its departure from the server, is called the sojourn time. Denote the random variable that describes the quantity mentioned above by W_R, and its mean value by $\mathrm{E}[W_R]$. At the time of an arbitrary test customer's arrival, it sees all the customers that were already in the system plus all other customers in front of it arriving in the same batch. Let g^-_k $(k = 0, 1, \ldots, b-1)$ be the probability of k number of customers ahead of a randomly selected test customer within the batch. Following Burke [6], we have $g^-_k = (1/\bar{g}) \sum_{\tau=k+1}^{\infty} g_\tau$. Next, we obtain the probability distribution function of W_R by conditioning on the fact that at a batch arrival epoch, the server is serving in vacation mode or in normal mode. Specifically, we need to consider two cases:

(1) Suppose the number of customers that arrive in the same bulk as the test customer, but enter service before the test customer is k $(k = 0, 1, \ldots, b-1)$. Additionally, we further assume that the test customer finds n $(n \geq 0)$ customers already in the system, and the server is on type j vacation upon its arrival. Thus, if $n + k + 1$ customers are served before the single ongoing vacation ends, then the test customer's sojourn time is the sum of $n + k + 1$ independent exponential service times with common mean $1/\mu_j$. On the other hand, if i $(i = 0, 1, \ldots, n+k)$ customers are served before the single ongoing

https://doi.org/10.1017/S0269964822000365 Published online by Cambridge University Press

vacation ends, and the rest are served in a normal busy period, then the sojourn time of the test customer is the sum of the remaining vacation time and $n+k+1-i$ exponential service times with intensity μ_0.

(2) Suppose a randomly selected test customer's position in an arrival group is $k+1$ ($k=0,1,\dots,b-1$), and it finds the server has got back to normal service mode. Further, if there are n customers already in the system when the test customer arrives, then the test customer's sojourn time is equal to the sum of $n+k+1$ regular service times. Let $S_{v,j}$ denote the service time of the vth customer in the type j vacation, and $\overline{V}_j$ represents the remaining vacation time of type j. Taking into account every scenario mentioned above, we can conclude

$$\begin{aligned}
W_R(t) &= \Pr\{W_R \le t\} \\
&= \sum_{j=1}^{h}\sum_{n=0}^{\infty} Q_{n,j}^{-} \sum_{k=0}^{b-1} g_k^{-} \Bigg[\Pr\left\{ \sum_{v=1}^{n+k+1} S_{v,j} \le t \,\middle|\, \overline{V}_j \ge \sum_{v=1}^{n+k+1} S_{v,j} \right\} \Pr\left\{ \overline{V}_j \ge \sum_{v=1}^{n+k+1} S_{v,j} \right\} \\
&\quad + \sum_{i=0}^{n+k} \Pr\left\{ \overline{V}_j + \sum_{v=i+1}^{n+k+1} S_{v,0} \le t \,\middle|\, \sum_{v=1}^{i} S_{v,j} \le \overline{V}_j < \sum_{v=1}^{i+1} S_{v,j} \right\} \Pr\left\{ \sum_{v=1}^{i} S_{v,j} \le \overline{V}_j < \sum_{v=1}^{i+1} S_{v,j} \right\} \Bigg] \\
&\quad + \sum_{n=0}^{\infty} P_n^{-} \sum_{k=0}^{b-1} g_k^{-} \Pr\left\{ \sum_{v=1}^{n+k+1} S_{v,0} \le t \right\} \\
&= \sum_{j=1}^{h}\sum_{n=0}^{\infty} Q_{n,j}^{-} \sum_{k=0}^{b-1} g_k^{-} \Bigg[\int_0^t \frac{\mu_j(\mu_j x)^{n+k}}{(n+k)!} e^{-(\mu_j+\theta)x}\, \mathrm{d}x \\
&\quad + \sum_{i=0}^{n+k} \int_0^t \theta_j\, e^{-(\theta_j+\mu_j)x} \frac{(\mu_j x)^i}{i!} \left[1 - e^{-\mu_0(t-x)} \sum_{v=0}^{n+k-i} \frac{(\mu_0(t-x))^v}{v!} \right] \mathrm{d}x \Bigg] \\
&\quad + \sum_{n=0}^{\infty} P_n^{-} \sum_{k=0}^{b-1} g_k^{-} \int_0^t \frac{\mu_0(\mu_0 x)^{n+k}}{(n+k)!} e^{-\mu_0 x}\, \mathrm{d}x.
\end{aligned} \tag{54}$$

Let $W_R^*(s) = \int_0^\infty e^{-st}\mathrm{d}W_R(t)$ be the L.S.T. of $W_R(t)$. From the convolution property of the Laplace transform, we have

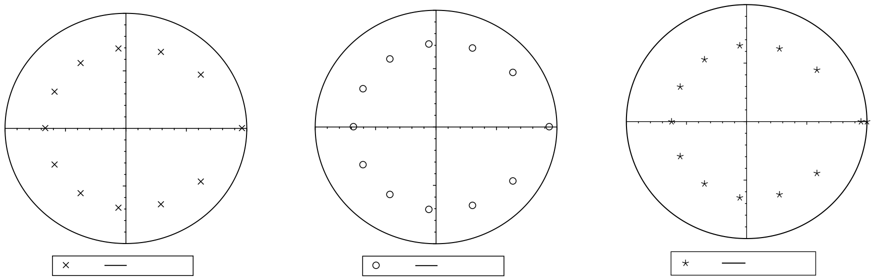
$$\begin{aligned}
W_R^*(s) &= \sum_{j=1}^{h}\sum_{n=0}^{\infty} Q_{n,j}^{-} \sum_{k=0}^{b-1} g_k^{-} \left[\left(\frac{\mu_j}{s+\mu_j+\theta_j} \right)^{n+k+1} + \sum_{i=0}^{n+k} \frac{\theta_j \mu_j^i}{(s+\mu_j+\theta_j)^{i+1}} \left(\frac{\mu_0}{s+\mu_0} \right)^{n+k+1-i} \right] \\
&\quad + \sum_{n=0}^{\infty} P_n^{-} \sum_{k=0}^{b-1} g_k^{-} \left(\frac{\mu_0}{s+\mu_0} \right)^{n+k+1}.
\end{aligned} \tag{55}$$

By differentiation Eq. (55) with respect to s, and setting $s=0$, the expectation of W_R is given by

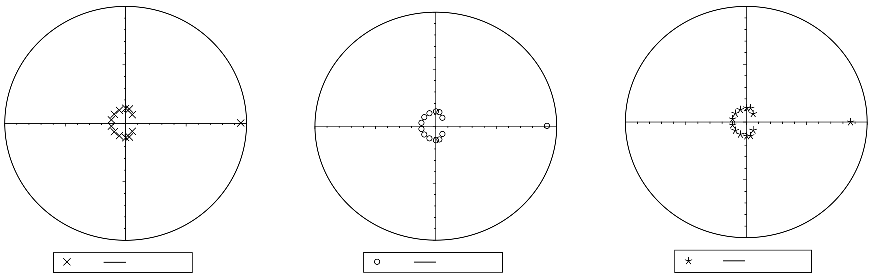
$$\begin{aligned}
\mathrm{E}[W_R] &= \sum_{j=1}^{h}\sum_{n=0}^{\infty} Q_{n,j}^{-} \sum_{k=0}^{b-1} g_k^{-} \left[\frac{(n+k+1)\mu_j^{n+k+1}}{(\mu_j+\theta_j)^{n+k+2}} + \sum_{i=0}^{n+k} \left(\frac{(i+1)\mu_j^i\theta_j}{(\mu_j+\theta_j)^{i+2}} + \frac{\mu_j^i\theta_j(n+k+1-i)}{\mu_0(\mu_j+\theta_j)^{i+1}} \right) \right] \\
&\quad + \sum_{n=0}^{\infty} P_n^{-} \sum_{k=0}^{b-1} g_k^{-} \frac{n+k+1}{\mu_0}.
\end{aligned} \tag{56}$$

It is well known that after obtaining the average queue length L_s, the mean sojourn time of an arbitrary customer can be evaluated by Little's Law, $\mathrm{E}[W_R] = L_s/\lambda\overline{g}$. However, Eq. (56) provides us with another semi-analytical way to get this performance measure. Therefore, we say that comparing the consistency of the calculation results of Eq. (56) and Little's Law is an effective means to test the feasibility of this method.

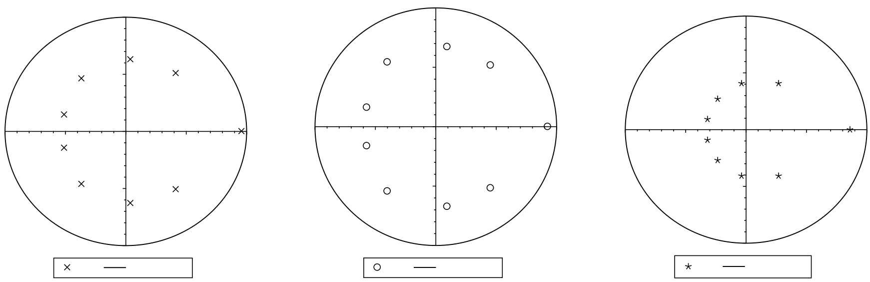
https://doi.org/10.1017/S0269964822000365 Published online by Cambridge University Press



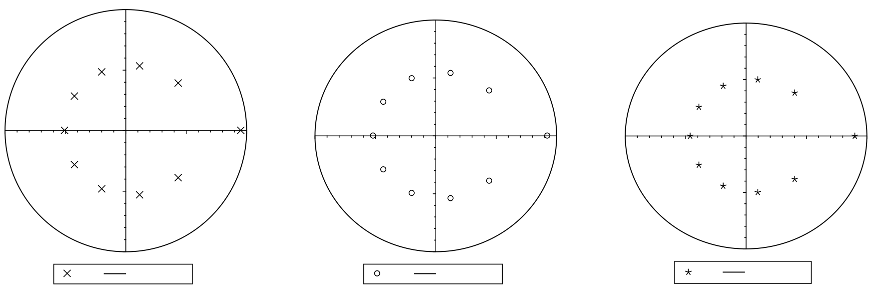
(a) Exponential inter-batch arrival time and uniformly distributed batch size



(b) PH inter-batch arrival time and normalized Poisson distributed batch size



(c) Deterministic inter-batch arrival time and 1-3-6-9 distributed batch size



(d) Inverse Gaussian inter-batch arrival time and normalized geometrically distributed batch size

Figure 1. *All the roots of the characteristic equations (23) and (35) lie inside the unit circle for different cases. (**a**) Exponential inter-batch arrival time and uniformly distributed batch size. (**b**) PH inter-batch arrival time and normalized Poisson distributed batch size. (**c**) Deterministic inter-batch arrival time and 1–3–6–9 distributed batch size. (**d**) Inverse Gaussian inter-batch arrival time and normalized geometrically distributed batch size.*

https://doi.org/10.1017/S0269964822000365 Published online by Cambridge University Press

Table 1. *Queue-length distributions at two different epochs for $M^X/M/1$ queue with single hyper-exponential working vacation.*

	Normal busy period		Working vacation			
n	P_n	P_n^-	$Q_{n,1}$	$Q_{n,1}^-$	$Q_{n,2}$	$Q_{n,2}^-$
0	0.012457	0.012457	0.052550	0.052550	0.008877	0.008877
1	0.006826	0.006826	0.004741	0.004741	0.000575	0.000575
2	0.007397	0.007397	0.005086	0.005086	0.000612	0.000612
3	0.007960	0.007960	0.005434	0.005434	0.000649	0.000649
4	0.008507	0.008507	0.005776	0.005776	0.000688	0.000688
5	0.009031	0.009031	0.006103	0.006103	0.000728	0.000728
6	0.009525	0.009525	0.006399	0.006399	0.000766	0.000766
7	0.009980	0.009980	0.006646	0.006646	0.000802	0.000802
8	0.010388	0.010388	0.006820	0.006820	0.000832	0.000832
9	0.010739	0.010739	0.006889	0.006889	0.000850	0.000850
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
21	0.011497	0.011497	0.004095	0.004095	0.000367	0.000367
22	0.011404	0.011404	0.003922	0.003922	0.000340	0.000340
23	0.011297	0.011297	0.003759	0.003759	0.000315	0.000315
24	0.011178	0.011178	0.003610	0.003610	0.000292	0.000292
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
60	0.004932	0.004932	0.000926	0.000926	0.000031	0.000031
61	0.004789	0.004789	0.000892	0.000892	0.000029	0.000029
62	0.004649	0.004649	0.000858	0.000858	0.000027	0.000027
63	0.004513	0.004513	0.000826	0.000826	0.000025	0.000025
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
Sum	0.699653	0.699653	0.273042	0.273042	0.027304	0.027304

$\mathrm{E}[W_R] = 8.897022$, $\mathrm{E}[W_R]_{\text{Little}} = 8.897022$.

6. Numerical illustrations

To summarize, the procedure for the calculation of the queue-length distribution adopted here consists of the following three-step algorithm.

Step 1. Set the values for the parameters of the system based on the stability condition, and find the roots of the characteristic equations (23) and (35), respectively.

Step 2. Combining Eqs. (29), (46), and (47) gives a system of $bh + b$ linear equations in $bh + b$ unknowns. Since Eq. (29) does not contain the unknown vector $\boldsymbol{f}$, we have to expand the coefficient matrix of Eq. (29) as follows

$$\begin{pmatrix} \boldsymbol{R}_1 & \boldsymbol{0} & \boldsymbol{0} & \cdots & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{R}_2 & \boldsymbol{0} & \cdots & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{R}_3 & \cdots & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} & \cdots & \boldsymbol{R}_{h-1} & \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} & \cdots & \boldsymbol{0} & \boldsymbol{R}_h & \boldsymbol{0} \end{pmatrix}.$$

Step 3. Solve the above system of linear equations and obtain the stationary queue-length distributions at different epochs by using Eqs. (48), (49), (50), (51), (52), and (53).

https://doi.org/10.1017/S0269964822000365 Published online by Cambridge University Press

Table 2. *Queue-length distributions at two different epochs for PH^X/M/1 queue with single hyper-exponential working vacation.*

	Normal busy period		Working vacation			
n	P_n	P_n^-	$Q_{n,1}$	$Q_{n,1}^-$	$Q_{n,2}$	$Q_{n,2}^-$
0	0.001343	0.001173	0.022091	0.019392	0.003824	0.003361
1	0.007579	0.006822	0.012116	0.012182	0.001958	0.001985
2	0.010509	0.009875	0.012079	0.012187	0.001901	0.001932
3	0.012934	0.012426	0.011564	0.011669	0.001758	0.001787
4	0.014836	0.014437	0.011040	0.011140	0.001621	0.001648
5	0.016292	0.015983	0.010542	0.010638	0.001495	0.001520
6	0.017372	0.017140	0.010068	0.010160	0.001379	0.001402
7	0.018137	0.017969	0.009615	0.009703	0.001272	0.001293
8	0.018640	0.018525	0.009182	0.009266	0.001173	0.001192
9	0.018924	0.018854	0.008769	0.008849	0.001082	0.001100
⋮	⋮	⋮	⋮	⋮	⋮	⋮
21	0.014747	0.014851	0.005047	0.005093	0.000410	0.000417
22	0.014202	0.014306	0.004820	0.004864	0.000378	0.000385
23	0.013660	0.013763	0.004603	0.004645	0.000349	0.000355
24	0.013125	0.013227	0.004396	0.004436	0.000322	0.000327
⋮	⋮	⋮	⋮	⋮	⋮	⋮
60	0.002494	0.002518	0.000838	0.000846	0.000018	0.000018
61	0.002379	0.002401	0.000800	0.000807	0.000016	0.000016
62	0.002269	0.002291	0.000764	0.000771	0.000015	0.000015
63	0.002165	0.002185	0.000730	0.000736	0.000014	0.000014
⋮	⋮	⋮	⋮	⋮	⋮	⋮
Sum	0.666505	0.666719	0.303177	0.302995	0.030318	0.030286

$\mathrm{E}[W_R] = 6.215603$, $\mathrm{E}[W_R]_{\text{Little}} = 6.215603$.

The following numerical examples are presented to illustrate the application of our analysis. Consider a GI^X/M/1 queue in which the exponential service distribution in a normal busy period has mean service time equal to 0.2 (i.e., $\mu_0 = 5$), and in which the server's vacation time has a hyper-exponential distribution function $V(t) = 0.8(1 - e^{-0.1t}) + 0.2(1 - e^{-0.25t})$. If such a distribution is used to model a server's vacation, then a server entering a maintenance phase will, with probability 80%, receive imperfect preventive repair that is exponentially distributed with parameter $\theta_1 = 0.1$ and then exit the type 1 vacation mode, or else, with probability 20% receive a preventive replacement of a specific component that is exponentially distributed with parameter $\theta_2 = 0.25$ and then exit the type 2 vacation mode. Furthermore, the exponential service time distributions have an average service time of 0.5 in type 1 vacation, and 0.8 in type 2 vacation, i.e., $\mu_1 = 2$ and $\mu_2 = 1.25$. For the batch arrival process, we consider four different cases listed below:

Case 1. Investigate a batch arrival Poisson queue. We fix the value of arrival rate $\lambda = 0.6$, and illustrate a numerical example by assuming that the p.m.f. of batch size X has the constant value $1/12$ for all possible values k of X, $k = 1, 2, \ldots, 12$. This leads to $\overline{g} = 6.5$, $\rho = 0.78$ and $a^*(s) = 0.6/(s + 0.6)$.

Case 2. The numerical results in this case were obtained by assuming that the group size distribution is normalized Poisson with p.m.f. $g_k = e^{-0.8}(0.8)^k / \left(k! \sum_{n=1}^{13} e^{-0.8}(0.8)^n/n!\right)$, $k = 1, 2, \ldots, 13$, and

https://doi.org/10.1017/S0269964822000365 Published online by Cambridge University Press

Table 3. *Queue-length distributions at two different epochs for D^X/M/1 queue with single hyper-exponential working vacation.*

n	Normal busy period P_n	P_n^-	Working vacation $Q_{n,1}$	$Q_{n,1}^-$	$Q_{n,2}$	$Q_{n,2}^-$
0	0.000544	0.001518	0.008405	0.032911	0.001496	0.006830
1	0.005175	0.006069	0.003235	0.003960	0.000531	0.000520
2	0.006030	0.010547	0.003665	0.004761	0.000533	0.000631
3	0.007521	0.013397	0.005593	0.005601	0.001010	0.000728
4	0.009272	0.015483	0.004632	0.006435	0.000545	0.000823
5	0.011173	0.017082	0.006818	0.006835	0.001051	0.001095
6	0.013047	0.018284	0.009159	0.006341	0.001808	0.001003
7	0.014781	0.019182	0.005396	0.006010	0.000658	0.000733
8	0.016360	0.019816	0.006460	0.006051	0.000913	0.000851
9	0.017606	0.020185	0.007422	0.005554	0.001237	0.000776
⋮	⋮	⋮	⋮	⋮	⋮	⋮
21	0.017748	0.016373	0.003773	0.003362	0.000351	0.000282
22	0.017239	0.015822	0.003621	0.003225	0.000327	0.000262
23	0.016715	0.015273	0.003473	0.003094	0.000304	0.000243
24	0.016181	0.014727	0.003331	0.002969	0.000280	0.000224
⋮	⋮	⋮	⋮	⋮	⋮	⋮
60	0.003684	0.003271	0.000753	0.000671	0.000018	0.000014
61	0.003530	0.003134	0.000723	0.000644	0.000016	0.000013
62	0.003383	0.003004	0.000694	0.000618	0.000015	0.000012
63	0.003242	0.002879	0.000666	0.000593	0.000014	0.000011
⋮	⋮	⋮	⋮	⋮	⋮	⋮
Sum	0.772308	0.762065	0.206993	0.215128	0.020699	0.022807

$\mathrm{E}[W_R] = 6.941437$, $\mathrm{E}[W_R]_{\text{Little}} = 6.941437$.

the inter-batch arrival time follows a phase-type distribution with an irreducible representation $(\boldsymbol{\sigma}, \boldsymbol{L})$ of order two, where $\boldsymbol{\sigma} = (0.7, 0.3)$, $\boldsymbol{L} = \left(\begin{smallmatrix} -3 & 2.7 \\ 4.5 & -21 \end{smallmatrix}\right)$. This leads to $\lambda = 2.699045$, $\overline{g} = 1.452773$, $\rho = 0.784220$, and $a^*(s) = (516s + 5085)/5(20s^2 + 480s + 1017)$.

Case 3. Suppose that the inter-batch arrival time distribution obeys a deterministic distribution with mean 1.25, and the arriving batch size follows a 1–3–6–9 distribution with p.m.f. $g_1 = 0.1$, $g_3 = 0.25$, $g_6 = 0.45$, $g_9 = 0.2$. This leads to $\lambda = 0.8$, $\overline{g} = 5.35$, $\rho = 0.856$, and $a^*(s) = e^{-1.25s}$.

Case 4. Model the batch arrival process by inverse Gaussian distribution with p.d.f. $a(t) = (0.75/\sqrt{2\pi t^3})e^{-(t-0.75)^2/2t}$, $t > 0$, and for which the p.m.f. of bulk size is $g_k = 0.35(1-0.35)^{k-1}/(1-(1-0.35)^{10})$, $k = 1, \dots, 10$. This leads to $\lambda = 4/3$, $\overline{g} = 2.720678$, $\rho = 0.725514$, and $a^*(s) = e^{0.75-0.75\sqrt{1+2s}}$.

We have found that in the above cases not all of the distribution functions for the inter-batch arrival time have rational Laplace–Stieltjes transforms. This makes it hard or even impossible to use mathematical software to get the roots of the characteristic equations. To overcome this difficulty, the classical technique of Padé approximation is employed to approximate a given Laplace–Stieltjes transform by a suitable rational function $\mathcal{R}(s)$. In the numerical experiments, the MATHEMATICA build-in command “PadeApproximant” can be used to easily generate the Padé approximants. Thus, the Laplace–Stieltjes

https://doi.org/10.1017/S0269964822000365 Published online by Cambridge University Press

Table 4. *Queue-length distributions at two different epochs for Inverse Gaussian^X/M/1 queue with single hyper-exponential working vacation.*

	Normal busy period		Working vacation			
n	P_n	P_n^-	$Q_{n,1}$	$Q_{n,1}^-$	$Q_{n,2}$	$Q_{n,2}^-$
0	0.008249	0.005487	0.052040	0.043229	0.008448	0.007005
1	0.009011	0.008744	0.013399	0.013621	0.001953	0.002001
2	0.010458	0.010262	0.012925	0.013141	0.001827	0.001873
3	0.011680	0.011549	0.012464	0.012675	0.001709	0.001752
4	0.012697	0.012623	0.012016	0.012221	0.001599	0.001639
5	0.013528	0.013504	0.011576	0.011777	0.001496	0.001533
6	0.014188	0.014208	0.011145	0.011341	0.001399	0.001434
7	0.014694	0.014751	0.010717	0.010909	0.001307	0.001341
8	0.015060	0.015148	0.010288	0.010476	0.001221	0.001252
9	0.015300	0.015413	0.009852	0.010031	0.001137	0.001166
⋮	⋮	⋮	⋮	⋮	⋮	⋮
21	0.012871	0.013062	0.005579	0.005679	0.000443	0.000455
22	0.012484	0.012673	0.005324	0.005419	0.000410	0.000421
23	0.012092	0.012278	0.005080	0.005170	0.000380	0.000389
24	0.011697	0.011879	0.004847	0.004934	0.000351	0.000360
⋮	⋮	⋮	⋮	⋮	⋮	⋮
60	0.002477	0.002521	0.000896	0.000912	0.000021	0.000022
61	0.002364	0.002406	0.000855	0.000870	0.000020	0.000020
62	0.002255	0.002295	0.000816	0.000830	0.000018	0.000019
63	0.002152	0.002190	0.000778	0.000792	0.000017	0.000017
⋮	⋮	⋮	⋮	⋮	⋮	⋮
Sum	0.603497	0.607617	0.360458	0.357082	0.036045	0.035301

$\mathrm{E}[W_R] = 6.538318$, $\mathrm{E}[W_R]_{\text{Little}} = 6.538318$.

transforms of the inter-batch arrival time distributions in Case 3 and Case 4 can be approximated as

$$e^{-1.25s} \approx \frac{\begin{array}{c}1 - 0.588235s + 0.160846s^2 - 0.026808s^3 + 0.002992s^4 - 0.000230s^5 \\ +0.000012s^6 - 3.89184 \times 10^{-7}s^7 + 6.08099 \times 10^{-9}s^8\end{array}}{\begin{array}{c}1 + 0.661765s + 0.206801s^2 + 0.040211s^3 + 0.005385s^4 + 0.000518s^5 \\ +0.000036s^6 + 1.75133 \times 10^{-6}s^7 + 5.4729 \times 10^{-8}s^8 + 8.44583 \times 10^{-10}s^9\end{array}}, \tag{57}$$

$$e^{0.75-0.75\sqrt{1+2s}} \approx \frac{1 + 4.95748s + 8.70859s^2 + 6.32864s^3 + 1.57069s^4}{\begin{array}{c}1 + 5.70748s + 12.333s^2 + 12.5594s^3 + 6.1046s^4 + 1.27288s^5 \\ +0.092802s^6 + 0.003849s^7\end{array}}. \tag{58}$$

As for the issues on choosing the best Padé approximant, readers may refer to the work done by Singh *et al.* [24]. Substituting the values of parameters into Eqs. (23) and (35), and using the Padé approximants for $a^*(s)$, we may successfully obtain all roots of the characteristic equations inside the unit circle. For a visual illustration, we plot these roots in Figure 1(a–d). We may observe that the imaginary roots of the characteristic equations always exist in complex conjugate pairs. If b is an odd number, there are $(b-1)/2$ pairs of complex conjugate roots for each characteristic equation. In contrast, if b is an even number, there are $(b-2)/2$ pairs of complex conjugate roots for each characteristic equation. Using

https://doi.org/10.1017/S0269964822000365 Published online by Cambridge University Press

these roots, we establish the coefficient matrix of the system of $bh + b$ linear equations that determines the unknown constants $c_{m,j}$ and f_m.

Once $c_{m,j}$ and f_m are obtained, the queue-length distributions can be computed by some basic algebraic operations. Due to lack of space, the queue-length distributions at different epochs presented in Tables 1–4 are reported to six decimal places. The main objective of presenting such results in tabular form is to show that the semi-analytical method has good numerical tractability. Furthermore, we know an important property of the batch Poisson arrival process is that the distribution of customers seen by a batch arrival to a queueing facility is, stochastically, the same as the limiting distribution of customers at that facility (i.e., PASTA property, see Gross and Harris [13]). In other words, once the queueing system has reached a steady-state, the distribution of customers at batch arrival instants is also the same as the distribution of customers at any time instant. There is no doubt that the data in Table 1 confirm this fact. We may also use this property as one of the effective ways to check the accuracy of numerical solutions. Additionally, at the bottom of each table, we show the mean sojourn time of an arbitrary customer calculated by Eq. (56) and Little's Law, respectively. We may find that $\mathrm{E}[W_R]$ evaluated through Eq. (56) exactly matches the one obtained from Little's Law, namely $\mathrm{E}[W_R] = \mathrm{E}[W_R]_{\text{Little}}$.

7. Conclusions

This paper has studied a more realistic working vacation queue that has never been considered in queueing literature. Using the supplementary variable technique and the shift operator method in the theory of difference equations, we obtain the queue-length distributions at the pre-arrival and arbitrary epochs simultaneously in terms of roots of the associated characteristic equation. Since the root-finding procedure is no more difficult with the help of MATHEMATICA software package, the core of computing the queue-length distribution is to solve a nonhomogeneous system of linear equations in our algorithm. We may see that characteristic roots serve as a bridge in our analysis. During the development of queueing theory, several authors have positively evaluated the advantages of the methods based on the characteristic roots. In comparison with the matrix-geometric method, Daigle and Lucantoni [10] stated that whenever the roots method works, it works blindingly fast and is insensitive to traffic intensity ρ, whereas the matrix-geometric method is not. Gouweleeuw [12] also stated that using the roots method to obtain the numerical results for the stationary queue-length probabilities will be more efficient. In fact, the method presented in this paper is straightforward in terms of analysis, notation, and computation. Additionally, this method can effectively avoid discussing the transition probability matrix for the embedded Markov chain, so we do not need to focus our attention on the minimal nonnegative solution to a nonlinear matrix equation arising in many queueing problems. Further, through the extensive numerical experiments, we also see that this algorithm can address many different arrival patterns and is no longer limited to the class of phase-type distributions that have rational Laplace–Stieltjes transform. Finally, incorporating a randomize working vacation policy into a renewal input batch arrival queue is worthy of investigation in our future research.

Acknowledgments. Authors would like to thank the editor and two reviewers for their valuable comments and suggestions. This research was supported by the National Natural Science Foundation of China (Grant No. 71571127), the Specialized Project for Subject Construction of Sichuan Normal University (Grant No. XKZX2021-04), and the National Office for Philosophy and Social Sciences of P. R. China (Grant No. 20BGL109).

References

[1] Baba, Y. (2005). Analysis of a GI/M/1 queue with multiple working vacations. *Operations Research Letters* 33: 201–209.

[2] Barbhuiya, F.P. & Gupta, U.C. (2019). A difference equation approach for analysing a batch service queue with the batch renewal arrival process. *Journal of Difference Equations and Applications* 25: 233–242.

[3] Barbhuiya, F.P. & Gupta, U.C. (2019). Discrete-time queue with batch renewal input and random serving capacity rule: $\mathrm{GI}^X/\mathrm{Geo}^Y/1$. *Queueing Systems* 91: 347–365.

[4] Barbhuiya, F.P. & Gupta, U.C. (2020). Analytical and computational aspects of the infinite buffer single server N policy queue with batch renewal input. *Computers & Operations Research* 118: 104916.

https://doi.org/10.1017/S0269964822000365 Published online by Cambridge University Press

[5] Barbhuiya, F.P. & Gupta, U.C. (2020). A discrete-time GI^X/Geo/1 queue with multiple working vacations under late and early arrival system. *Methodology and Computing in Applied Probability* 22: 599–624.
[6] Burke, P.J. (1975). Delays in single-server queues with batch input. *Operations Research* 23: 830–832.
[7] Chae, K.C., Lim, D.E., & Yang, W.S. (2009). The GI/M/1 queue and the GI/Geo/1 queue both with single working vacation. *Performance Evaluation* 66: 356–367.
[8] Chaudhry, M.L., Harris, C.M., & Marchal, W.G. (1990). Robustness of rootfinding in single server queueing models. *INFORMS Journal of Computing* 2: 273–286.
[9] Chen, H., Li, J., & Tian, N. (2009). The GI/M/1 queue with phase-type working vacations and vacation interruption. *Journal of Applied Mathematics and Computing* 30: 121–141.
[10] Daigle, J.N. & Lucantoni, D.M. (1991). Queueing systems having phase-dependent arrival and service rates. In W.J. Stewart (ed.), *Numerical solutions of Markov chains*. New York: Marcel Dekker Inc.
[11] Gao, S. & Yao, Y. (2014). An M^X/G/1 queue with randomized working vacations and at most J vacations. *International Journal of Computer Mathematics* 91: 368–383.
[12] Gouweleeuw, F.N. (1996). A general approach to computing loss probabilities in finite-buffer queues. Ph.D. thesis, Vrije Universiteit, Amsterdam.
[13] Gross, D. & Harris, C.M (1985). *Fundamentals of queueing theory*, 2nd ed. New York: Wiley.
[14] Guha, D. & Banik, A.D. (2013). On the renewal input batch-arrival queue under single and multiple working vacation policy with application to EPON. *INFOR* 51: 175–191.
[15] Guha, D., Goswami, V., & Banik, A.D. (2015). Equilibrium balking strategies in renewal input batch arrival queues with multiple and single working vacation. *Performance Evaluation* 94: 1–24.
[16] Lee, D.H. & Kim, B.K. (2015). A note on the sojourn time distribution of an M/G/1 queue with a single working vacation and vacation interruption. *Operations Research Perspectives* 2: 57–61.
[17] Li, J., Tian, N., & Liu, W. (2007). The discrete-time GI/Geom/1 queue with multiple working vacations. *Queueing Systems* 56: 53–63.
[18] Li, J., Tian, N., Zhang, Z.G., & Luh, H.P. (2009). Analysis of the M/G/1 queue with exponentially working vacations – a matrix analytic approach. *Queueing Systems* 61: 139–166.
[19] Liu, W., Xu, X., & Tian, N. (2007). Stochastic decompositions in the M/M/1 queue with working vacations. *Operations Research Letters* 35: 595–600.
[20] Ma, Z., Chen, L., & Wang, P. (2020). Analysis of G-queue with pseudo-fault and multiple working vacations. *Journal of Systems Science and Complexity* 33: 1144–1162.
[21] Neuts, M.F (1981). *Matrix-geometric solutions in stochastic models: An algorithmic approach*. Baltimore: The Johns Hopkins University Press.
[22] Selvaraju, N. & Goswami, C. (2013). Impatient customers in an M/M/1 queue with single and multiple working vacations. *Computers & Industrial Engineering* 65: 207–215.
[23] Servi, L.D. & Finn, S.G. (2002). M/M/1 queues with working vacations. *Performance Evaluation* 50: 41–52.
[24] Singh, G., Gupta, U.C., & Chaudhry, M.L. (2014). Analysis of queueing-time distributions for $\mathrm{MAP/D}_N/1$ queue. *International Journal of Computer Mathematics* 94: 1911–1930.
[25] Tijms, H.C (2003). *A first course in stochastic models*. Chichester: John Wiley & Sons.
[26] Wu, D. & Takagi, H. (2006). M/G/1 queue with multiple working vacations. *Performance Evaluation* 63: 654–681.
[27] Yu, M. (2021). Alternative approach based on roots for computing the stationary queue-length distributions in $\mathrm{GI}^X/\mathrm{M}^{(1,b)}/1$ single working vacation queue. *RAIRO-Operations Research* 55: S2259–S2290.
[28] Zhang, M. & Hou, Z. (2010). Performance analysis of M/G/1 queue with working vacations and vacation interruption. *Journal of Computational and Applied Mathematics* 234: 2977–2985.
[29] Zhang, M. & Hou, Z. (2012). M/G/1 queue with single working vacation. *Journal of Applied Mathematics and Computing* 39: 221–234.

Cite this article: Yu M and Tang Y (2023). Analyzing a single hyper-exponential working vacation queue from its governing difference equation. *Probability in the Engineering and Informational Sciences* **37**, 997–1019. https://doi.org/10.1017/S0269964822000365