On the time-dependent behavior of preemptive single-server queueing systems with Poisson arrivals

Brian Fralix
Clemson University
School of Mathematical and Statistical Sciences
Clemson, SC, USA

March 19, 2023

Abstract

We present a detailed study of the time-dependent behavior of both the queue-length process and the workload process of various types of preemptive Last-Come-First-Served queueing systems, such as the preemptive-repeat-different and preemptive-repeat-identical models recently studied in (Asmussen and Glynn, 2017) and (Bergquist and Sigman, 2022). Our main results show various quantities that provide information about the time-dependent behavior of these processes can be expressed in terms of the Laplace-Stieltjes transform of the busy period, and we show how a natural coupling procedure can be used to establish, for each preemptive queue we consider, a recursive procedure for calculating these busy period transforms on the set of all complex numbers having positive real part.

Keywords: LIFO, Palm distribution, preemptive-resume, preemptive-repeat, queue length, workload

2020 MSC: Primary 60K25, Secondary 60G55.

1 Introduction

A handful of recently-published articles have focused on studying single-server queueing systems operating under various types of preemptive-repeat Last-Come-First-Served (LCFS) queueing disciplines. In the work of Asmussen and Glynn [8], the authors establish stability conditions for single-server queues operating under either the preemptive-repeat-identical and preemptive-repeat-different LCFS disciplines (these will be defined shortly) when customers arrive in accordance to a Poisson process, as well as the more general setting where customers arrive in accordance to a Markovian Arrival process (MAP). In Bergquist and Sigman [9] the authors study properties of the limiting distribution of the workload process associated with both the preemptive-repeat-identical and preemptive-repeat-different LCFS disciplines, when customers arrive in accordance to a Poisson process. Another recent work on this topic is that of Horváth et al. [15], where the authors study a single-server queue with Poisson arrivals that operates under a nonpreemptive LCFS with resampling discipline that is similar in some (but not all) respects to the single-server queue operating under the preemptive-repeat-different LCFS discipline.

Our main objective is to build further on [9] by studying the time-dependent behavior of both the queue-length process $\{Q(t); t \geq 0\}$ and the workload process $\{W(t); t \geq 0\}$ associated with both of the preemptive-repeat LCFS disciplines mentioned above, where for each $t \geq 0$, $Q(t)$ and $W(t)$ denote, respectively, the number of customers found in the system and the remaining amount of work present in the system at time t. In each of these models, we assume customers arrive one-at-a-time to a single-server queueing system in accordance to a Poisson arrival process $\{A(t); t \geq 0\}$ having rate λ and points $\{T_n\}_{n\geq 1}$, so that for each integer $n \geq 1$, T_n denotes the arrival time of the

nth customer to arrive to the system strictly after time zero. Here $A(t)$ denotes the number of points found in the set $(0, t]$ (we set $A(0) := 0$), and the one-to-one correspondence between the points $\{T_n\}_{n\geq 1}$ and the counting process $\{A(t); t \geq 0\}$ is reinforced mathematically by noting that for each real number $t \geq 0$, and each integer $n \geq 1$,

$$A(t) := \sum_{k=1}^{\infty} \mathbf{1}_{\{T_k \leq t\}}, \quad T_n := \inf\{t \geq 0 : A(t) \geq n\}$$

where for each event C, $\mathbf{1}_C$ is the indicator function associated with C, in that $\mathbf{1}_C$ is equal to 1 if C occurs, and is equal to 0 if C does not occur.

In each queueing model we study, the server always operates in a preemptive LCFS manner, meaning whenever a customer arrives to the system, the server immediately suspends working on the current job and devotes its full attention to the new arrival, only to return servicing the newly-neglected job when the new arrival leaves the system. One way to picture how such a model behaves is to imagine the jobs physically stacked on top of one another, where new arrivals land on top of the stack, and at any time instant, the server can only process (at unit rate) the work possessed by the job located on top of the stack. It also helps to alternatively refer to each position in the stack as a *slot*, where an arriving customer always occupies the lowest-numbered unoccupied slot upon arrival, and the server always provides its full attention to the customer having, among all present customers, the highest slot position in the stack.

There are multiple ways in which the server can behave once it returns to a customer it tried serving previously:

- *Preemptive-Resume*: The customer arriving at time T_n brings with it to the system an amount of work B_n for processing. Whenever a customer's processing experience with the server is interrupted, upon revisiting the server it continues from where it left off. In this model, we assume the sequence $\{B_n\}_{n\geq 1}$ is i.i.d. with cumulative distribution function (CDF) F and Laplace-Stieltjes transform (LST) β, with $\{B_n\}_{n\geq 1}$ being independent of the sequence $\{T_n\}_{n\geq 1}$ (and $\{A(t); t \geq 0\}$).

- *Preemptive-Repeat-Different*: The (tagged) customer arriving at time T_n brings with it to the system a sequence of work $\{B_{n,k}\}_{k\geq 1}$ for (potential) processing, where the doubly-indexed sequence $\{B_{n,k}\}_{n\geq 1,k\geq 1}$ is assumed to be i.i.d. with CDF F and LST β, and independent of $\{T_n\}_{n\geq 1}$. When this customer first arrives to the system, it immediately begins having its amount of work $B_{n,1}$ processed by the server. If a new customer arrives before this amount of work has been processed, the server moves on to the new arrival: when the server returns to the tagged customer, it then begins processing the amount of work $B_{n,2}$. Again, if a new arrival occurs before the server finishes this amount of work, it moves on to the new arrival, and when it returns to the tagged customer, it begins processing the amount of work $B_{n,3}$. This behavior continues throughout until the instant where the server finally finishes processing one of these amounts of work possessed by the tagged customer, at which time the tagged customer departs from the system.

- *Preemptive-Repeat-Identical*: Similar to the preemptive-resume case, here the customer arriving at time T_n brings with it to the system an amount of work B_n for processing, where we assume the sequence $\{B_n\}_{n\geq 1}$ is i.i.d. with CDF F and LST β, with $\{B_n\}_{n\geq 1}$ being independent of the sequence $\{T_n\}_{n\geq 1}$ (and $\{A(t); t \geq 0\}$). Unlike the preemptive-resume case, whenever the server revisits the customer that arrived at time T_n, it completely restarts processing the job of this customer, meaning all of the previous work performed by the server on that customer is discarded entirely.

In the interest of improving readability, any reference made to a preemptive-resume queue will correspond to a single-server queueing system operating under the Last-Come-First-Served preemptive-resume discipline. Likewise, any reference to a preemptive-repeat-different queue or a preemptive-repeat-identical queue should be interpreted in a similar manner.

Throughout our study, we use arguably the most natural definition for the workload at time t. For each $t \geq 0$, $W(t)$ represents the sum of the amounts of work possessed by each customer present in the system at time t, where the amount of work possessed by the customer within position j of the stack corresponds to either (a) for the customer currently being served at time t, the remaining amount of work present associated with the specific job being processed by the server at that time, or (b) for each customer waiting in the stack to be revisited by the server, the amount of work the server will encounter once it returns to that customer. This is how the workload process is defined in [9].

We associate with each slot $k \geq 1$, the random variable $W_k(t)$ that denotes the remaining amount of work possessed by the customer occupying slot k at time t. Our main objective is to study the joint distribution of $Q(t)$ and $\{W_k(t)\}_{k\geq 1}$, from which we can study the joint distribution of $Q(t)$ and $W(t)$ since for each $t \geq 0$,

$$W(t) = \sum_{k=1}^{\infty} W_k(t).$$

Define the set $\mathbb{C}_+ := \{\alpha \in \mathbb{C} : Re(\alpha) > 0\} \cup \{0\}$ as the union of the singleton $\{0\}$ with the complex open half-plane consisting of complex numbers having a positive real part, define $\mathbb{C}_+^\infty := \{\{\alpha_n\}_{n\geq 1} : \alpha \in \mathbb{C}_+\}$ as the set of all sequences having elements belonging to $\mathbb{C}_+$, and define, for each integer $n_0 \geq 0$ and each integer $n \geq 0$, the function $\phi_{n_0;n} : \mathbb{C}_+ \times \mathbb{C}_+^\infty \to \mathbb{C}$ as

$$\phi_{n_0;n}(\alpha, \boldsymbol{\gamma}) := \int_0^\infty e^{-\alpha t} \mathbb{E}_{n_0}\left[\mathbf{1}(Q(t) = n) e^{-\sum_{k=1}^n \gamma_k W_k(t)}\right] dt, \quad \alpha \in \mathbb{C}_+,\ \boldsymbol{\gamma} \in \mathbb{C}_+^\infty$$

where $\mathbb{E}_{n_0}$ represents conditional expectation, conditional on $Q(0) = n_0$, where each of the n_0 customers present in the system at time 0 begin with an amount of work having CDF F, independent of everything else.

The method we use to derive the $\phi_{n_0;n}$ functions is a simple modification of the approach used in [14], where various factorization results were established for what was referred to in [14] as Preemptive-Resume-Production systems, or PRP systems. It is notable that the overall idea behind the approach found in [14] is similar to ideas used by the matrix-analytic community (see e.g. [17]), and so we suspect many of our results will also carry over to the case where customers arrive in accordance to a Markovian Arrival Process, at the cost of more tedious proofs/derivations, in particular having to establish the invertibility of various matrices that will inevitably be encountered in this more general setting. We plan to address this more general case in a future study.

For each of the three preemptive queueing systems we study in this manuscript, the formulas we derive for each $\phi_{n_0;n}$ function are in terms of the LST of the busy period associated with this queue. This is not surprising, as this fact is in-line with what was observed in a series of papers written by Abate and Whitt in the late 1980s [1, 2, 3, 4, 6] that address the transient (i.e. time-dependent) behavior of the M/M/1 queue, the M/G/1 workload, and of regulated Brownian motion, as well as in related work of the author [10, 11, 12]. Hence, we also thoroughly discuss how, for each of the three preemptive queueing systems, the LST of the busy period can be calculated using a natural iterative scheme, at all complex numbers having positive real part. Through a single coupling arguent, we will derive not only the iteration scheme discussed in Abate and Whitt [5] for the busy period LST of the work-conserving M/G/1 queue, but also iterative schemes for both the preemptive-repeat-different queue, and the preemptive-repeat-identical queue.

2 The Queue-Length Process and the Workload Process

The time-dependent behavior of both the queue-length process and the workload process is fairly tractable for not only the preemptive-resume queue, but also the preemptive-repeat-different queue and the preemptive-repeat-identical queue. We will observe that while the same argument can be used to study the time-dependent behavior of the queue-length process in isolation of each of these three queues, each queue needs to be considered separately when we seek to study the time-dependent behavior of both the queue-length process and the workload process, as well as the distribution of the busy period of each of these three queueing systems.

2.1 The Queue-Length Process

Our first goal in this section is to study the marginal distributions of $\{Q(t); t \geq 0\}$ in isolation. In particular, we will focus on deriving, for each $\alpha \in \mathbb{C}$,

$$\phi_{n_0;n}(\alpha, \mathbf{0}) = \int_0^\infty e^{-\alpha t} \mathbb{P}_{n_0}(Q(t) = n) dt.$$

These functions can be expressed in terms of the Laplace-Stieltjes transform $\pi : \mathbb{C}_+ \to \mathbb{C}$ of the busy period τ_0 with respect to $\mathbb{P}_1$:

$$\pi(\alpha) := \mathbb{E}_1[e^{-\alpha \tau_0}], \quad \alpha \in \mathbb{C}_+$$

where for each integer $k \geq 0$, and each $s \geq 0$,

$$\tau_k(s) := \inf\{t \geq s : Q(t-) \neq k, Q(t) = k\}$$

and we follow the convention that $\tau_k := \tau_k(0)$. Clearly, for each integer $n_0 \geq 1$,

$$\mathbb{E}_{n_0}[e^{-\alpha \tau_0}] = \pi(\alpha)^{n_0}, \quad \alpha \in \mathbb{C}_+$$

for each of the three preemptive queueing systems we will consider.

Our first result shows that $\phi_{n_0,0}$ can be expressed in terms of π.

Proposition 2.1 *For each integer $n_0 \geq 0$,*

$$\phi_{n_0;0}(\alpha) = \frac{\pi(\alpha)^{n_0}}{\alpha + \lambda(1 - \pi(\alpha))}. \tag{1}$$

Proof We begin by verifying (1) for the case where $n_0 = 0$. Using Fubini's theorem, we get

$$\begin{aligned}
\phi_{0;0}(\alpha) &= \int_0^\infty e^{-\alpha t} \mathbb{P}_0(Q(t) = 0) dt \\
&= \mathbb{E}_0 \left[\int_0^\infty e^{-\alpha t} \mathbf{1}_{\{Q(t)=0\}} dt \right] \\
&= \mathbb{E}_0 \left[\int_0^{T_1} e^{-\alpha t} \mathbf{1}_{\{Q(t)=0\}} dt \right] + \mathbb{E}_0 \left[\int_{\tau_0}^\infty e^{-\alpha t} \mathbf{1}_{\{Q(t)=0\}} dt \right] \\
&= \frac{1}{\lambda + \alpha} + \frac{\lambda}{\lambda + \alpha} \pi(\alpha) \phi_{0;0}(\alpha)
\end{aligned}$$

and solving for $\phi_{0;0}(\alpha)$ yields

$$\phi_{0;0}(\alpha) = \frac{1}{\alpha + \lambda(1 - \pi(\alpha))}.$$

Furthermore, when $Q(0) = n_0 \geq 1$, another application of Fubini's theorem yields

$$\begin{aligned}
\phi_{n_0;0}(\alpha) = \int_0^\infty e^{-\alpha t} \mathbb{P}_{n_0}(Q(t) = 0) dt &= \mathbb{E}_{n_0} \left[\int_0^\infty e^{-\alpha t} \mathbf{1}_{\{Q(t)=0\}} dt \right] \\
&= \mathbb{E}_{n_0} \left[\int_{\tau_0}^\infty e^{-\alpha t} \mathbf{1}_{\{Q(t)=0\}} dt \right] \\
&= \phi(\alpha)^{n_0} \phi_{0;0}(\alpha)
\end{aligned}$$

proving the claim. $\diamondsuit$

Our next result shows that $\phi_{n_0;n}$ can be expressed in terms of π. Our proof strategy will involve showing the $\phi_{n_0;n}(\alpha)$ terms satisfy a particular recursive scheme that can be solved easily. The same recursive scheme shows up repeatedly throughout this study, and so we now state an elementary proposition that addresses this recursion.

Proposition 2.2 *Suppose $\{a_n\}_{n\geq 1}$ and $\{b_n\}_{n\geq 1}$ are two sequences of complex numbers, and the sequence of complex numbers $\{x_n\}_{n\geq 0}$ satisfies, for each integer $n \geq 1$,*

$$x_n = b_n + a_n x_{n-1}.$$

Then for each integer $n \geq 1$,

$$x_n = \left[\prod_{\ell=1}^{n} a_\ell\right] x_0 + \sum_{k=1}^{n} b_k \prod_{\ell=k+1}^{n} a_\ell.$$

Proof This can be proven quickly using induction: we omit the details. $\diamondsuit$

Theorem 2.1 *For each integer $n_0 \geq 0$ and each integer $n \geq 0$, we have that for each $\alpha \in \mathbb{C}_+$,*

$$\phi_{n_0,n}(\alpha, \mathbf{0}) = \pi(\alpha)^{n_0}\left[\frac{1}{\alpha+\lambda(1-\pi(\alpha))}\right]\left[\frac{\lambda(1-\pi(\alpha))}{\alpha+\lambda(1-\pi(\alpha))}\right]^n + \sum_{k=1}^{n}\mathbf{1}_{\{n_0\geq k\}}\pi(\alpha)^{n_0-k}\left[\frac{1-\pi(\alpha)}{\alpha+\lambda(1-\pi(\alpha))}\right]\left[\frac{\lambda(1-\pi(\alpha))}{\alpha+\lambda(1-\pi(\alpha))}\right]^{n-k}. \tag{2}$$

This result, when $n_0 = 0$, was established for the preemptive-resume queue in the work of Kella et al. [16] using different techniques, and it can be inferred from [14] for the preemptive-resume queue as well, under any initial condition. Indeed, the idea behind the proof we present of Theorem 2.1 is virtually identical to the idea used to establish the main results found in [14], but the particular mechanics of this proof will be easier to modify later when we study both the queue-length process and the workload process.

Proof We begin by deriving $\phi_{0;n}(\alpha, \mathbf{0})$ for each integer $n \geq 1$. For each real $t \geq 0$,

$$\mathbf{1}_{\{Q(t)=n\}} = \int_{(0,t]} \mathbf{1}_{\{Q(s-)=n-1\}}\mathbf{1}_{\{\tau_{n-1}(s)>t,Q(t)=n\}}A(ds). \tag{3}$$

Taking the expected value of both sides of (3), while applying the Campbell-Mecke formula to the right-hand-side gives

$$\mathbb{P}_0(Q(t)=n) = \int_0^t \mathbb{P}_0(Q(s)=n-1)\mathbb{P}_1(\tau_0 > t-s, Q(t-s)=n)\lambda ds \tag{4}$$

and after multiplying both sides of (4) by $e^{-\alpha t}$, for $\alpha \in \mathbb{C}_+$, then integrating with respect to t over $[0,\infty)$, we get

$$\phi_{0;n}(\alpha, \mathbf{0}) = \lambda\mathbb{E}_1\left[\int_0^{\tau_0} e^{-\alpha t}\mathbf{1}_{\{Q(s)=1\}}ds\right]\phi_{0;n-1}(\alpha, \mathbf{0}) \tag{5}$$

where $\phi_{0;0}(\alpha, \mathbf{0}) := \phi_{0;0}(\alpha)$. Hence, for each integer $n \geq 1$,

$$\phi_{0;n}(\alpha, \mathbf{0}) = \left[\lambda\mathbb{E}_1\left[\int_0^{\tau_0} e^{-\alpha s}\mathbf{1}_{\{Q(s)=1\}}ds\right]\right]^n \phi_{0;0}(\alpha) = \left[\lambda\mathbb{E}_1\left[\int_0^{\tau_0} \mathbf{1}_{\{Q(s)=1\}}ds\right]\right]^n \frac{1}{\alpha+\lambda(1-\pi(\alpha))}. \tag{6}$$

We can also solve for the unknown expected value found in both (5) and (6). Summing both sides of (5) over the integers $n \geq 1$ gives

$$\frac{1}{\alpha} - \phi_{0;0}(\alpha) = \frac{\lambda}{\alpha}\mathbb{E}_1\left[\int_0^{\tau_0} e^{-\alpha s}\mathbf{1}_{\{Q(s)=1\}}ds\right] \tag{7}$$

and since $\phi_{0;0}(\alpha)$ was found in Proposition 2.1, (7) contains only one unknown: solving for this unknown yields

$$\mathbb{E}_1\left[\int_0^{\tau_0} e^{-\alpha s}\mathbf{1}_{\{Q(s)=1\}}ds\right] = \frac{1}{\lambda}\left[1 - \frac{\alpha}{\alpha+\lambda(1-\pi(\alpha))}\right] = \frac{1-\pi(\alpha)}{\alpha+\lambda(1-\pi(\alpha))}. \tag{8}$$

Plugging (8) into (6) then gives, for each integer $n \geq 1$,

$$\phi_{0;n}(\alpha, \mathbf{0}) = \frac{1}{\alpha+\lambda(1-\pi(\alpha))}\left[\frac{\lambda(1-\pi(\alpha))}{\alpha+\lambda(1-\pi(\alpha))}\right]^n$$

proving (2) for the case where $n_0 = 0$.

It remains to establish (2) for the case where $n_0 \geq 1$ and $n \geq 1$. For each integer $n \geq 1$,

$$\mathbf{1}_{\{Q(t)=n\}} = \mathbf{1}_{\{n_0 \geq n\}}\mathbf{1}_{\{Q(t)=n,\tau_{n-1}>t\}} + \int_{(0,t]} \mathbf{1}_{\{Q(s-)=n-1\}}\mathbf{1}_{\{\tau_{n-1}(s)>t,Q(t)=n\}}A(ds) \tag{9}$$

and after taking the expected value of both sides and simplifying, we get

$$\mathbb{P}_{n_0}(Q(t)=n) = \mathbf{1}_{\{n_0\geq n\}}\mathbb{P}_{n_0}(Q(t)=n,\tau_{n-1}>t) + \lambda\int_0^t \mathbb{P}_{n_0}(Q(s)=n-1)\mathbb{P}_1(\tau_0>t-s, Q(t-s)=n)ds. \tag{10}$$

Multiplying both sides of (10) by $e^{-\alpha t}$, then integrating with respect to t over $[0,\infty)$ gives

$$\phi_{n_0;n}(\alpha,\mathbf{0}) = \mathbf{1}_{\{n_0\geq n\}}\pi(\alpha)^{n_0-n}\frac{(1-\pi(\alpha))}{\alpha+\lambda(1-\pi(\alpha))} + \frac{\lambda(1-\pi(\alpha))}{\alpha+\lambda(1-\pi(\alpha))}\phi_{n_0;n-1}(\alpha,\mathbf{0}). \tag{11}$$

Finally, we can solve this recursion by applying Proposition 2.2 to (11) and conclude that for each integer $n \geq 1$,

$$\begin{aligned}\phi_{n_0,n}(\alpha,\mathbf{0}) &= \pi(\alpha)^{n_0}\left[\frac{1}{\alpha+\lambda(1-\pi(\alpha))}\right]\left[\frac{\lambda(1-\pi(\alpha))}{\alpha+\lambda(1-\pi(\alpha))}\right]^n \\ &\quad + \sum_{k=1}^{\min(n_0,n)} \pi(\alpha)^{n_0-k}\left[\frac{1-\pi(\alpha)}{\alpha+\lambda(1-\pi(\alpha))}\right]\left[\frac{\lambda(1-\pi(\alpha))}{\alpha+\lambda(1-\pi(\alpha))}\right]^{n-k}\end{aligned}$$

proving the claim. ♢

2.2 The Workload Process

Our next objective is to study the joint distribution of $Q(t)$ and $\{W_k(t)\}_{k\geq1}$, but unlike our study of $Q(t)$, we will need to use a separate derivation for each of the three preemptive queueing systems under consideration.

An important quantity associated with each preemptive queue is the function $N : \mathbb{C}_+ \times \mathbb{C}_+ \to \mathbb{C}$, defined as

$$N(\alpha,\gamma) := \mathbb{E}_1\left[\int_0^{\tau_0} e^{-\alpha s}\mathbf{1}_{\{Q(s)=1\}}e^{-\gamma W_1(s)}dt\right].$$

This function will repeatedly appear in our study of the $\phi_{n_0,n}$ functions. Each time we derive a computable expression for $N(\alpha, \gamma)$, we will seek to calculate the quantity $N(\alpha, \gamma, x)$, defined for each $x > 0$ as

$$N(\alpha, \gamma, x) := \mathbb{E}\left[\int_0^{\tau_0} e^{-\alpha t}\mathbf{1}(Q(t) = 1)e^{-\gamma W_1(t)}dt \mid Q(0) = 1, W_1(0) = x\right].$$

Note that $W_1(0)$ is simply the amount of work possessed by the single customer present at time zero, as we are also conditioning on $Q(0) = 1$. Once we have determined $N(\alpha, \gamma, x)$ for each $x \geq 0$, we can use a simple conditioning argument to conclude that

$$N(\alpha, \gamma) = \int_{(0,\infty)} N(\alpha, \gamma, x)dF(x).$$

2.2.1 Preemptive-Resume Queues

While the time-dependent behavior of the preemptive-resume queue is much better understood than the time-dependent behavior of both the preemptive-repeat-different and the preemptive-repeat-identical queues, our objective here is to first present a procedure for studying the time-dependent behavior of the preemptive-resume queue that can be adapted to study both preemptive-repeat queues.

Our first result in this subsection focuses on the derivation of $N(\alpha, \gamma)$ for the preemptive-resume queue.

Proposition 2.3 *For each $\alpha, \gamma \in \mathbb{C}_+$,*

$$N(\alpha, \gamma) = \frac{\beta(\gamma) - \pi(\alpha)}{\alpha + \lambda(1 - \pi(\alpha)) - \gamma} \tag{12}$$

Proof The key to calculating $N(\alpha, \gamma)$ is to first focus on deriving $N(\alpha, \gamma, x)$. Observe first that when $Q(0) = 1$ and $W_1(0) = x$,

$$\int_0^{\tau_0} e^{-\alpha t}\mathbf{1}(Q(t) = 1)e^{-\gamma W_1(t)}dt = \int_0^{\min(T_1,x)} e^{-\alpha y}e^{-\gamma(x-y)}dy + \mathbf{1}(T_1 \leq x)\int_{T_1}^{\tau_0} e^{-\alpha t}\mathbf{1}(Q(t) = 1)e^{-\gamma W_1(t)}dt \tag{13}$$

Taking the expected value of both sides of (13) gives

$$\begin{aligned} N(\alpha, \gamma, x) &= \int_0^x e^{-\alpha y}e^{-\gamma(x-y)}e^{-\lambda y}dy + \int_0^x e^{-\alpha y}\pi(\alpha)N(\alpha, \gamma, x - y)\lambda e^{-\lambda y}dy \\ &= e^{-(\alpha+\lambda)x}\int_0^x e^{(\alpha+\lambda-\gamma)y}dy + e^{-(\alpha+\lambda)x}\pi(\alpha)\int_0^x N(\alpha, \gamma, y)e^{(\alpha+\lambda)y}dy. \end{aligned} \tag{14}$$

Multiplying both sides of (14) by $e^{(\alpha+\lambda)x}$, then taking the partial derivative of both sides with respect to x yields

$$\frac{\partial}{\partial x}N(\alpha, \gamma, x) + (\alpha + \lambda(1 - \pi(\alpha)))N(\alpha, \gamma, x) = e^{-\gamma x}. \tag{15}$$

Equation (15) is essentially a first-order linear ODE, with initial condition $N(\alpha, \gamma, 0) = 0$. Solving this ODE yields

$$N(\alpha, \gamma, x) = \frac{e^{-\gamma x} - e^{-(\alpha+\lambda(1-\pi(\alpha)))x}}{\alpha + \lambda(1 - \pi(\alpha)) - \gamma}. \tag{16}$$

Integrating both sides of (16) over $[0, \infty)$ with respect to $dF(x)$ gives

$$N(\alpha, \gamma) = \frac{\beta(\gamma) - \beta(\alpha + \lambda(1 - \pi(\alpha)))}{\alpha + \lambda(1 - \pi(\alpha)) - \gamma} = \frac{\beta(\gamma) - \pi(\alpha)}{\alpha + \lambda(1 - \pi(\alpha)) - \gamma} \tag{17}$$

where the last equality follows from the well-known fixed point equation $\pi(\alpha) = \beta(\alpha + \lambda(1 - \pi(\alpha))$ satisfied by π when the service discipline of the M/G/1 queue is a work-conserving service discipline. This completes the proof of the claim. $\diamondsuit$

The next result shows how we can modify the proof of Theorem 2.1 to derive $\phi_{n_0;n}(\alpha, \boldsymbol{\gamma})$.

Theorem 2.2 *For each integer $n_0 \geq 0$, each integer $n \geq 0$, each $\alpha \in \mathbb{C}_+$, and each $\boldsymbol{\gamma} \in \mathbb{C}_+^\infty$,*

$$\begin{aligned} &\phi_{n_0;n}(\alpha, \boldsymbol{\gamma}) \\ &= \phi_{n_0;0}(\alpha) \left[\prod_{\ell=1}^{n} \lambda N(\alpha, \gamma_\ell) \right] + \sum_{k=1}^{n} \mathbf{1}_{\{n_0 \geq k\}} \pi(\alpha)^{n_0 - k} \left[\prod_{\ell=1}^{k-1} \beta(\gamma_\ell) \right] N(\alpha, \gamma_k) \left[\prod_{\ell=k+1}^{n} \lambda N(\alpha, \gamma_\ell) \right] \end{aligned} \tag{18}$$

Proof The key to establishing this result is to derive the following recursive scheme: for each integer $n \geq 1$,

$$\phi_{n_0;n}(\alpha, \boldsymbol{\gamma}) = \mathbf{1}_{\{n_0 \geq n\}} \pi(\alpha)^{n_0 - n} \left[\prod_{\ell=1}^{n-1} \beta(\gamma_\ell) \right] N(\alpha, \gamma_n) + \lambda N(\alpha, \gamma_n) \phi_{n_0;n-1}(\alpha, \boldsymbol{\gamma}). \tag{19}$$

Once this recursion has been established, it can be solved using Proposition 2.2 to get (18).

It remains, then, to derive the recursion. First we observe that for each $t \geq 0$, each integer $n \geq 1$, and each $\boldsymbol{\gamma} \in \mathbb{C}_+$,

$$\begin{aligned} &\mathbf{1}(Q(t) = n) e^{-\sum_{k=1}^{n} \gamma_k W_k(t)} \\ &= \mathbf{1}(n_0 \geq n) \mathbf{1}(\tau_n \leq t, \tau_{n-1} > t) e^{-\sum_{k=1}^{n} \gamma_k W_k(t)} \\ &\quad + \int_{(0,t]} \mathbf{1}(Q(s-) = n-1) e^{-\sum_{k=1}^{n-1} \gamma_k W_k(s-)} \mathbf{1}(Q(t) = n, \tau_{n-1}(s) > t) e^{-\gamma_n W_n(t)} A(ds). \end{aligned} \tag{20}$$

Taking the expected value of both sides of (20), while further applying the Campbell-Mecke formula to the right-hand-side gives

$$\begin{aligned} &\mathbb{E}_{n_0}[\mathbf{1}(Q(t) = n) e^{-\sum_{k=1}^{n} \gamma_k W_k(t)}] \\ &= \mathbf{1}(n_0 \geq n) \left[\prod_{k=1}^{n-1} \beta(\gamma_k) \right] \mathbb{E}_{n_0}[\mathbf{1}(\tau_1 \leq t, \tau_0 > t, Q(t) = 1) e^{-\gamma_k W_1(t)}] \\ &\quad + \lambda \int_0^t \mathbb{E}_{n_0}[\mathbf{1}(Q(s) = n-1) e^{-\sum_{k=1}^{n-1} \gamma_k W_k(s)}] \mathbb{E}_1 \left[\mathbf{1}(Q(t-s) = 1, \tau_0 > t-s) e^{-\gamma_n W_1(t-s)} \right] ds. \end{aligned} \tag{21}$$

Multiplying both sides of (21) by $e^{-\alpha t}$ for $\alpha \in \mathbb{C}_+$, then integrating both sides over $[0, \infty)$ gives

$$\begin{aligned} \overline{\phi}_{n_0,n}(\alpha, \boldsymbol{\gamma}) = \mathbf{1}(n \leq n_0) &\left[\prod_{k=1}^{n-1} \beta(\gamma_k) \right] \pi(\alpha)^{n_0 - n} \mathbb{E}_1 \left[\int_0^{\tau_0} e^{-\alpha t} \mathbf{1}(Q(t) = 1) e^{-\gamma_n W_1(t)} dt \right] \\ &+ \lambda \overline{\phi}_{n_0,n-1}(\alpha, \boldsymbol{\gamma}) \mathbb{E}_1 \left[\int_0^{\tau_0} e^{-\alpha t} \mathbf{1}(Q(t) = 1) e^{-\alpha \gamma_n W_1(t)} dt \right] \end{aligned}$$

i.e.

$$\overline{\phi}_{n_0,n}(\alpha, \boldsymbol{\gamma}) = \mathbf{1}(n \leq n_0) \left[\prod_{k=1}^{n-1} \beta(\gamma_k) \right] \pi(\alpha)^{n_0 - n} N(\alpha, \gamma_n) + \lambda N(\alpha, \gamma_n) \overline{\phi}_{n_0,n-1}(\alpha, \boldsymbol{\gamma})$$

which is (19). $\diamondsuit$

We conclude this subsection by stating the following corollary, which gives us the Laplace transform of the joint Laplace-Stieltjes transform of $Q(t)$ and $W(t)$.

Corollary 2.1 *For each $z \in D(0,1)$, each $\alpha \in \mathbb{C}_+$, and each $\gamma \in \mathbb{C}_+$,*

$$\int_0^\infty \mathbb{E}_{n_0}[z^{Q(t)} e^{-\gamma W(t)}] e^{-\alpha t} dt = \frac{\phi_{n_0;0}(\alpha)}{1 - \lambda N(\alpha,\gamma) z} + \frac{N(\alpha,\gamma) z}{1 - \lambda N(\alpha,\gamma) z} \sum_{k=1}^{n_0} \pi(\alpha)^{n_0 - k} (z\beta(\gamma))^{k-1}.$$

Proof This result follows quickly from Theorem 2.2, since

$$\int_0^\infty \mathbb{E}_{n_0}[z^{Q(t)} e^{-\gamma W(t)}] = \sum_{n=0}^\infty z^n \phi_{n_0;n}(\alpha, \gamma \mathbf{e}).$$

$\diamondsuit$

Readers should observe that the Laplace transform of the joint LST of $Q(t)$ and $W(t)$ was also recently derived, for the case where $Q(0) = 0$, in Theorem 2 of [13], and our formula (when $n_0 = 0$) agrees with the corresponding transform from [13]. In fact, it was further shown in [13] that when $Q(0) = 0$, the joint LST of $Q(t)$ and $W(t)$ is the same for all symmetric M/G/1 service disciplines: the preemptive-resume queue is an example of a symmetric M/G/1 queue.

2.2.2 Preemptive-Repeat-Different Queues

The joint Laplace-Stieltjes transform of $Q(t)$ and $\{W_k(t)\}_{k\geq 1}$ can be analyzed for the preemptive-repeat-different queue as well using a similar proof technique, in fact the analysis is in some ways simpler than the analysis used above to study the workload process of the preemptive-resume queue.

Our first result in this section is an expression for $N(\alpha, \gamma)$.

Proposition 2.4 *For each $\alpha, \gamma \in \mathbb{C}_+$,*

$$N(\alpha,\gamma) = \frac{\beta(\gamma) - \beta(\alpha+\lambda)}{(\alpha + \lambda - \gamma)(1 - \frac{\lambda}{\lambda+\alpha}(1 - \beta(\lambda+\alpha))\pi(\alpha))} \tag{22}$$

Proof We begin by expressing $N(\alpha,\gamma,x)$ in terms of $N(\alpha,\gamma)$. Conditioning on the fact that $Q(0) = 1$ and $W_1(0) = x$, for $x > 0$, we get

$$\begin{aligned}\int_0^{\tau_0} e^{-\alpha t} \mathbf{1}(Q(t) = 1) e^{-\gamma W_1(t)} dt &= \int_0^x e^{-\alpha y} \mathbf{1}(A(y) = 0) e^{-\gamma(x-y)} dy \\ &\quad + \mathbf{1}(T_1 \leq x) \int_{T_1}^{\tau_0} e^{-\alpha t} \mathbf{1}(Q(t) = 1) W_1(t) dt.\end{aligned} \tag{23}$$

Taking the expected value of both sides of (23), while further simplifying the right-hand-side gives

$$N(\alpha,\gamma,x) = \int_0^x e^{-\alpha y} e^{-\lambda y} e^{-\gamma(x-y)} dy + \pi(\alpha) \left[\int_0^x e^{-\alpha y} \lambda e^{-\lambda y} dy \right] N(\alpha,\gamma)$$

or

$$N(\alpha,\gamma,x) = \frac{e^{-\gamma x} - e^{-(\alpha+\lambda)x}}{\alpha + \lambda - \gamma} + \frac{\lambda}{\lambda+\alpha}(1 - e^{-(\lambda+\alpha)x})\pi(\alpha) N(\alpha,\gamma). \tag{24}$$

Integrating both sides of (24) over $[0,\infty)$ with respect to $dF(x)$ yields

$$N(\alpha,\gamma) = \frac{\beta(\gamma) - \beta(\alpha+\lambda)}{\alpha + \lambda - \gamma} + \frac{\lambda}{\lambda+\alpha}(1 - \beta(\lambda+\alpha))\pi(\alpha) N(\alpha,\gamma)$$

and solving for $N(\alpha,\gamma)$ gives (22). $\diamondsuit$

Having Proposition 2.4 in hand, we are now ready to derive the $\phi_{n_0;n}$ functions associated with the preemptive-repeat-different queue.

Theorem 2.3 *For each integer $n_0 \geq 0$, and each integer $n \geq 1$,*

$$\phi_{n_0;n}(\alpha, \boldsymbol{\gamma}) = \left[\prod_{\ell=1}^{n-1} \beta(\gamma_\ell)\right] N(\alpha, \gamma_n) \sum_{k=1}^{n} \mathbf{1}_{\{n_0 \geq k\}} \pi(\alpha)^{n_0-k} \left[\frac{\lambda(1-\pi(\alpha))}{\alpha + \lambda(1-\pi(\alpha))}\right]^{n-k} + \left[\prod_{\ell=1}^{n-1} \beta(\gamma_\ell)\right] \frac{N(\alpha, \gamma_n)}{N(\alpha, 0)} \pi(\alpha)^{n_0} \left[\frac{1}{\alpha + \lambda(1-\pi(\alpha))}\right] \left[\frac{\lambda(1-\pi(\alpha))}{\alpha + \lambda(1-\pi(\alpha))}\right]^{n}. \quad (25)$$

Before we give the proof, it is instructive to use this result to determine the stationary joint distribution of $Q(t)$ and $\{\{W_k(t)\}_{k \geq 1}$. Multiplying both sides by α, then letting $\alpha \downarrow 0$ yields

$$\mathbb{E}[\mathbf{1}_{\{Q(\infty)=n\}} e^{-\sum_{k=1}^{Q(\infty)} \gamma_k W_k(\infty)}] = \left[\prod_{\ell=1}^{n-1} \beta(\gamma_\ell)\right] \frac{N(0, \gamma_n)}{N(0,0)} \mathbb{P}(Q(\infty) = n)$$

which implies that conditional on $Q(\infty) = n$, the random variables $W_1(\infty), W_2(\infty), \ldots, W_n(\infty)$ are independent, where $W_k(\infty)$ has LST β, and $W_n(\infty)$ has LST

$$\frac{N(0, \gamma_n)}{N(0,0)} = \frac{\lambda(\beta(\gamma) - \beta(\lambda))}{(\lambda - \gamma)(1 - \beta(\lambda))}.$$

Proof The proof of this result is actually much simpler in many respects than the derivation of the $\phi_{n_0;n}$ functions associated with the preemptive-resume queue. For each integer $n \geq 1$,

$$\mathbf{1}_{\{Q(t)=n\}} e^{-\sum_{k=1}^{n} \gamma_k W_k(t)} = \mathbf{1}_{\{n_0 \geq n\}} e^{-\sum_{k=1}^{n-1} \gamma_k W_k(0)} \mathbf{1}_{\{\tau_n \leq t, \tau_{n-1} > t, Q(t)=n\}} e^{-\gamma_n W_n(t)} + \int_{(0,t]} \mathbf{1}_{\{Q(s-)=n-1\}} e^{-\sum_{k=1}^{n-1} \gamma_k W_k(t)} \mathbf{1}_{\{\tau_{n-1}(s) > t, Q(t)=n\}} e^{-\gamma_n W_n(t)} A(ds). \quad (26)$$

Taking the expected value of both sides of (26), while applying the Campbell-Mecke formula to the right-hand-side gives

$$\begin{aligned} &\mathbb{E}_{n_0}\left[\mathbf{1}_{\{Q(t)=n\}} e^{-\sum_{k=1}^{n} \gamma_k W_k(t)}\right] \\ &= \mathbf{1}_{\{n_0 \geq n\}} \left[\prod_{k=1}^{n-1} \beta(\gamma_k)\right] \mathbb{E}_{n_0}[\mathbf{1}_{\{\tau_n \leq t, \tau_{n-1} > t, Q(t)=n\}} e^{-\gamma_n W_n(t)}] \\ &\quad + \lambda \left[\prod_{k=1}^{n-1} \beta(\gamma_k)\right] \int_0^t \mathbb{P}_{n_0}(Q(s) = n-1) \mathbb{E}_1\left[\mathbf{1}_{\{Q(t-s)=1\}} e^{-\gamma_n W_1(t-s)}\right] ds \end{aligned} \quad (27)$$

and after multiplying both sides of the equality by $e^{-\alpha t}$, then integrating with respect to t over $[0, \infty)$, we get

$$\phi_{n_0;n}(\alpha, \boldsymbol{\gamma}) = \mathbf{1}_{\{n_0 \geq n\}} \pi(\alpha)^{n_0-n} \left[\prod_{\ell=1}^{n-1} \beta(\gamma_\ell)\right] N(\alpha, \gamma_n) + \lambda \left[\prod_{\ell=1}^{n-1} \beta(\gamma_\ell)\right] N(\alpha, \gamma_n) \phi_{n_0;n-1}(\alpha, \mathbf{0}) \quad (28)$$

implying

$$\begin{aligned} \phi_{n_0;n}(\alpha, \boldsymbol{\gamma}) &= \mathbf{1}_{\{n_0 \geq n\}} \pi(\alpha)^{n_0-n} \left[\prod_{\ell=1}^{n-1} \beta(\gamma_\ell)\right] N(\alpha, \gamma_n) \\ &\quad + \lambda \left[\prod_{\ell=1}^{n-1} \beta(\gamma_\ell)\right] N(\alpha, \gamma_n) \sum_{k=1}^{\min(n_0, n-1)} (1-\pi(\alpha)) \pi(\alpha)^{n_0-k} \left[\frac{1}{\alpha + \lambda(1-\pi(\alpha))}\right] \left[\frac{\lambda(1-\pi(\alpha))}{\alpha + \lambda(1-\pi(\alpha))}\right]^{n-1-k} \\ &\quad + \lambda \left[\prod_{\ell=1}^{n-1} \beta(\gamma_\ell)\right] N(\alpha, \gamma_n) \pi(\alpha)^{n_0} \left[\frac{1}{\alpha + \lambda(1-\pi(\alpha))}\right] \left[\frac{\lambda(1-\pi(\alpha))}{\alpha + \lambda(1-\pi(\alpha))}\right]^{n-1} \end{aligned} \quad (29)$$

and since

$$\lambda N(\alpha, 0) = \frac{\lambda(1-\pi(\alpha))}{\alpha + \lambda(1-\pi(\alpha))} \tag{30}$$

we arrive, after some algebra, at the claim. $\diamondsuit$

Remark It is interesting to notice that the joint distribution of $Q(t)$ and $\{W_k(t)\}_{k\geq 1}$ for the 'nonpreemptive-LIFO with restart' model of Horvath et al. [15] is exactly the same as the joint distribution of $Q(t)$ and $\{W_k(t)\}_{k\geq 1}$ for the preemptive-repeat-different model. In fact, we can recover Theorem 1 of [15] from Theorem 2.3 above by multiplying both sides by α, letting $\alpha \downarrow 0$, and setting $\gamma_k = 0$ for $1 \leq k \leq n-1$.

2.2.3 Preemptive-Repeat-Identical Queues

The workload process of the preemptive-repeat-identical queue behaves in a manner that makes its analysis more difficult than that for the preemptive-resume queue, and the preemptive-repeat-different queue, in that we will have to first study the joint distribution of $Q(t)$ and $\{W_k(t)\}_{k\geq 1}$ for a related preemptive-repeat-identical queueing system where each customer brings an amount of work that has a discrete distribution. Once we derive results for this system, analogous results will be established for the original preemptive-repeat-identical system through a natural limiting procedure.

Given our original preemptive-repeat-identical queueing system, we define, for each integer $m \geq 1$, the sequence of random variables $\{B_n^{(m)}\}_{n\geq 1}$, where for each integer $n \geq 1$,

$$B_n^{(m)} := \sum_{k=1}^{\infty} \frac{k}{2^m} \mathbf{1}_{\{(k-1)/2^m < B_n \leq k/2^m\}}.$$

For each integer $m \geq 1$, the mth queueing system is a single-server queue operating under the preemptive-repeat-identical discipline, where customers arrive in accordance to the arrival process $\{A(t); t \geq 0\}$ (meaning each of these systems have the same arrival process) and the nth arrival to the system arrives at time T_n, and brings an amount of work $B_n^{(m)}$ for processing. Finally, we let $Q^{(m)}(t)$ denote the number of customers present in the mth queueing system at time t, and for each integer $k \geq 1$, we let $W_k^{(m)}(t)$ denote the remaining amount of work possessed by the customer present in slot k at time t.

Given each customer in the mth queueing system brings a random amount of work for proccessing that has a discrete distribution, it is also useful to associate with each such customer a label that tells us precisely how much work was originally brought to the system by the arriving customer. We say that an arriving customer to the mth queueing system is a type-j customer with probability

$$p_{m,j} := F(j/2^m) - F((j-1)/2^m).$$

By introducing this labeling scheme, we may say that each type-j customer bring to the system a deterministic amount of work $x_{m,j} := j/2^m$ for processing. For each integer $k \geq 1$, and each $t \geq 0$, let $L_k(t)$ denote the customer type of the customer present in slot k at time t.

Next, we introduce the functions $\Delta_{n_0;n}^{(m)} : \mathbb{C}_+ \times \mathbb{C}_+^\infty \times \mathbb{N}^\infty \to \mathbb{C}_+$, defined as follows: for each $\alpha \in \mathbb{C}_+$, each $\boldsymbol{\gamma} \in \mathbb{C}_+^\infty$, and each $\mathbf{i} = (i_1, i_2, \ldots) \in \mathbb{N}^\infty$ (where $\mathbb{N}$ denotes the set of positive integers)

$$\Delta_{n_0;n}^{(m)}(\alpha, \boldsymbol{\gamma}, \mathbf{i}) := \mathbb{E}_{n_0}[\mathbf{1}(Q^{(m)}(t) = n, L_1^{(m)}(t) = i_1, \ldots, L_n^{(m)}(t) = i_n) e^{-\sum_{k=1}^n \gamma_k W_k^{(m)}(t)}].$$

Our first objective is to derive computable expression for the $\Delta_{n_0:n}^{(m)}$ functions, and we will see shortly that these functions be expressed in terms of the $N(\alpha, \gamma, x)$ quantities defined above. In the process of studying this joint distribution, we will first need to examine the joint distribution of

$Q(t)$ and $\{L_k(t)\}_{k\geq 1}$. Throughout, we let π_m denote the LST of the busy period of the mth queueing system, and we let

$$N_m(\alpha,\gamma,x) := \mathbb{E}_{(1,x)}\left[\int_0^{\tau_0^{(m)}} e^{-\alpha t}\mathbf{1}_{\{Q^{(m)}(t)=1\}}e^{-\gamma W_1^{(m)}(t)}dt\right].$$

Proposition 2.5 *For each $\alpha \in \mathbb{C}_+$, and each $\boldsymbol{i} \in E^\infty$, we have for each integer $n \geq 1$,*

$$\begin{aligned}\Delta^{(m)}_{n_0;n}(\alpha,\mathbf{0},\mathbf{i}) &= \phi^{(m)}_{n_0;0}(\alpha)\left[\prod_{\ell=1}^{n}\lambda p_{m,i_\ell}N_m(\alpha,0,x_{m,i_\ell})\right] \\ &+\sum_{k=1}^{n}\mathbf{1}_{\{n_0\geq k\}}\pi_m(\alpha)^{n_0-k}\left[\prod_{\ell=1}^{k}p_{m,i_\ell}\right]N_m(\alpha,0,x_{m,i_k})\left[\prod_{\ell=k+1}^{n}\lambda p_{m,i_\ell}N_m(\alpha,0,x_{m,i_\ell})\right]\end{aligned} \tag{31}$$

Proof This result can be proven using a method analogous to that used to establish Theorem 2.1, in that the key to proving the claim is to show that the $\Delta^{(m)}_{n_0;n}(\alpha,\mathbf{0},\mathbf{i})$ terms satisfy the following recursion: for each integer $n \geq 1$,

$$\Delta^{(m)}_{n_0;n}(\alpha,\mathbf{0},\mathbf{i}) = \mathbf{1}_{\{n_0\geq n\}}\pi_m(\alpha)^{n_0-n}\left[\prod_{\ell=1}^{n}p_{m,i_\ell}\right]N_m(\alpha,0,x_{m,i_n}) + \lambda p_{m,i_n}N_m(\alpha,0,x_{m,i_n})\Delta^{(m)}_{n_0;n-1}(\alpha,\mathbf{0},\mathbf{i}). \tag{32}$$

Once this recursion has been established, we can use Proposition 2.2 to arrive at (31).

It remains to establish (32). Fix $\mathbf{i} \in \mathbb{N}^\infty$, and observe that for each integer $n \geq 1$,

$$\begin{aligned}&\mathbf{1}_{\{Q^{(m)}(t)=n,L_1^{(m)}(t)=i_1,\ldots,L_n(m)(t)=i_n\}} \\ &= \mathbf{1}_{\{n_0\geq n\}}\mathbf{1}_{\{L_1^{(m)}(0)=i_1,\ldots,L_n^{(m)}(0)=i_n,\tau_n^{(m)}\leq t,\tau_{n-1}^{(m)}>t,Q^{(m)}(t)=n\}} \\ &\quad+\int_0^t \mathbf{1}_{\{Q^{(m)}(s-)=n-1,L_1^{(m)}(s-)=i_1,\ldots,L_{n-1}^{(m)}(s-)=i_{n-1}\}}\mathbf{1}_{\{Q^{(m)}(t)=n,\tau_{n-1}^{(m)}(s)>t\}}A_{m,i_n}(ds)\end{aligned} \tag{33}$$

where $\{A_{m,i}(t); t \geq 0\}$ denotes the thinned Poisson process of type-i arrivals. Taking the expected value of both sides of (33), while further applying the Campbell-Mecke formula to the stochastic integral on the right-hand-side yields

$$\begin{aligned}&\mathbb{P}_{n_0}(Q^{(m)}(t)=n,L_1^{(m)}(t)=i_1,\ldots,L_n^{(m)}(t)=i_n) \\ &= \mathbf{1}_{\{n_0\geq n\}}\mathbb{P}(L_1^{(m)}(0)=i_1,\ldots,L_n^{(m)}(0)=i_n,\tau_n^{(m)}\leq t,\tau_{n-1}^{(m)}>t,Q^{(m)}(t)=n) \\ &\quad+\lambda p_{m,i_n}\int_0^t \mathbb{P}(Q^{(m)}(s)=n-1,L_1^{(m)}(s)=i_1,\ldots,L_{n-1}^{(m)}(s)=i_{n-1})\mathbb{P}_{(1,x_{m,i_m})}(Q^{(m)}(t-s)=1,\tau_0^{(m)}>t-s)ds.\end{aligned} \tag{34}$$

Multiplying both sides of (34) by $e^{-\alpha t}$, then integrating with respect to t over $[0,\infty)$ gives

$$\Delta^{(m)}_{n_0;n}(\alpha,\mathbf{0},\mathbf{i}) = \mathbf{1}_{\{n_0\geq n\}}\left[\prod_{\ell=1}^{n}p_{m,i_\ell}\right]\pi_m(\alpha)^{n_0-n}N_m(\alpha,0,x_{m,i_n}) + \lambda p_{m,i_n}N_m(\alpha,0,x_{m,i_n})\Delta^{(m)}_{n_0;n-1}(\alpha,\mathbf{0},\mathbf{i})$$

which establishes the recursion. $\diamondsuit$

The next result completes the derivation of the $\Delta^{(m)}_{n_0;n}$ functions associated with the mth preemptive-repeat-identical queue.

Proposition 2.6 *For each $\alpha \in \mathbb{C}_+$, and each $\boldsymbol{\gamma} \in \mathbb{C}_+$,*

$$\begin{aligned}\Delta^{(m)}_{n_0;n}(\alpha,\boldsymbol{\gamma},\mathbf{i}) &= \mathbf{1}_{\{n_0\geq n\}}\pi_m(\alpha)^{n_0-n}\left[\prod_{\ell=1}^{n} p_{m,i_\ell}\right]e^{-\sum_{k=1}^{n-1}\gamma_k x_{m,i_k}}N_m(\alpha,\gamma_n,x_{m,i_n})\\ &+\phi^{(m)}_{n_0;0}(\alpha)\left[\prod_{\ell=1}^{n-1}\lambda p_{m,i_\ell}N_m(\alpha,0,x_{m,i_\ell})\right]\lambda p_{m,i_n}e^{-\sum_{k=1}^{n-1}\gamma_k x_{m,i_k}}N_m(\alpha,\gamma_n,x_{m,i_n})\\ &+\sum_{k=1}^{\min(n_0,n-1)}\pi_m(\alpha)^{n_0-k}\left[\prod_{\ell=1}^{k}p_{m,i_\ell}\right]N_m(\alpha,0,x_{m,i_k})\left[\prod_{\ell=k+1}^{n-1}\lambda p_{m,i_\ell}N_m(\alpha,0,x_{m,i_\ell})\right]\\ &\times\lambda p_{m,i_n}e^{-\sum_{\ell=1}^{n-1}\gamma_\ell x_{m,i_\ell}}N_m(\alpha,\gamma_n,x_{m,i_n}) \qquad (35)\end{aligned}$$

Proof Fix $t>0$, and observe that for each $\boldsymbol{\gamma}\in\mathbb{C}^\infty_+$, and each $\mathbf{i}\in\mathbb{N}^\infty$, we have for each integer $n\geq 1$,

$$\begin{aligned}&\mathbf{1}_{\{Q^{(m)}(t)=n,L_1^{(m)}(t)=i_1,\ldots,L_n^{(m)}(t)=i_n\}}e^{-\sum_{k=1}^{n}\gamma_k W_k^{(m)}(t)}\\ &=\mathbf{1}_{\{L_1^{(m)}(0)=i_1,\ldots,L_n^{(m)}(0)=i_n,\tau_n^{(m)}\leq t,\tau_{n-1}^{(m)}>t,Q^{(m)}(t)=n\}}e^{-\sum_{k=1}^{n-1}\gamma_k x_{m,i_k}}e^{-\gamma_n W_n^{(m)}(t)}\\ &+\int_{(0,t]}\mathbf{1}_{\{Q^{(m)}(s-)=n-1,L_1^{(m)}(s-)=i_1,\ldots,L_{n-1}^{(m)}(s-)=i_{n-1}\}}e^{-\sum_{k=1}^{n-1}\gamma_k x_{m,i_k}}e^{-\gamma_n W_n^{(m)}(t)}\mathbf{1}_{\{Q^{(m)}(t)=n,\tau_{n-1}^{(m)}(s)>t\}}A_{m,i_n}(ds). \qquad (36)\end{aligned}$$

After taking the expected value of both sides of (36), while applying the Campbell-Mecke formula to the integral on the right-hand-side, we get

$$\begin{aligned}&\mathbb{E}_{n_0}\left[\mathbf{1}_{\{Q^{(m)}(t)=n,L_1^{(m)}(t)=i_1,\ldots,L_n^{(m)}(t)=i_n\}}e^{-\sum_{k=1}^{n}\gamma_k W_k^{(m)}(t)}\right]\\ &=\left[\prod_{\ell=1}^{n}p_{m,i_\ell}\right]e^{-\sum_{k=1}^{n-1}\gamma_k x_{m,i_k}}\mathbb{E}_{n_0}\left[e^{-\gamma_n W_n(t)}\mathbf{1}_{\{\tau_n^{(m)}\leq t,\tau_{n-1}^{(m)}>t,Q^{(m)}(t)=n\}}\mid L_n^{(m)}(0)=i_n\right]\\ &+\lambda p_{m,i_n}\int_0^t\mathbb{P}_{n_0}(Q^{(m)}(s)=n-1)e^{-\sum_{k=1}^{n-1}\gamma_k x_{m,i_k}}\mathbb{E}_{(1,x_{m,i_n})}\left[e^{-\gamma_n W_1(t-s)}\mathbf{1}_{\{\tau_0^{(m)}>t-s,Q^{(m)}(t-s)=1\}}\right]ds. \qquad (37)\end{aligned}$$

Multiplying both sides of (37) by $e^{-\alpha t}$, the integrating with respect to t over $[0,\infty)$ gives

$$\begin{aligned}\Delta^{(m)}_{n_0;n}(\alpha,\boldsymbol{\gamma},\mathbf{i}) &= \mathbf{1}_{\{n_0\geq n\}}\pi_m(\alpha)^{n_0-n}\left[\prod_{\ell=1}^{n}p_{m,i_\ell}\right]e^{-\sum_{k=1}^{n-1}\gamma_k x_{m,i_k}}N_m(\alpha,\gamma_n,x_{m,i_n})\\ &+\lambda p_{m,i_n}e^{-\sum_{k=1}^{n-1}\gamma_k x_{m,i_k}}N_m(\alpha,\gamma_n,x_{m,i_n})\Delta^{(m)}_{n_0;n-1}(\alpha,\mathbf{0},\mathbf{i}). \qquad (38)\end{aligned}$$

Applying Proposition 2.5 to (38) proves the claim. $\diamondsuit$

We are now ready to study the joint distribution of $Q(t)$ and $\{W_k(t); t\geq 0\}$ for the original preemptive-repeat-identical queueing system. A proper limiting argument is somewhat cumbersome to state properly, as it requires a significant amount of new notation, so we will provide an overall sketch of the idea in order to help readers better understand how it works.

Letting $\{Q^{(m)}(t); t\geq 0\}$ and $\{\{W_k^{(m)}(t); t\geq 0\}\}$ denote the queue-length process and the remaining amounts of work of each customer present in each slot at time t, one can see from the construction of these queueing systems that on the set where no arrivals and service completions ever occur simultaneously, and the number of arrivals in the interval $(0,t]$ is finite (this event occurs with probability one), there exists an integer $m_0\geq 1$ (which may depend on the outcome in the set) where the order at which customers depart from the system, among those arriving in the interval $(0,t]$, is the same for each integer $m\geq m_0$. Once this fact has been observed, a bit of thought reveals that for each $m \geq m_0$, the customer found in slot k at time t is the same for all queueing systems $m \geq m_0$ (here m_0 may depend on t as well as the outcome), and so $Q^{(m)}(t) = Q(t)$ (again these statements are true for all outcomes on a set having probability one), and by the dominated convergence theorem, for each $\alpha \in \mathbb{C}_+$, and each $\boldsymbol{\gamma} \in \mathbb{C}_+^\infty$,

$$\lim_{n\to\infty} \mathbb{E}_{n_0}[\mathbf{1}_{\{Q^{(m)}(t)=n\}} e^{-\sum_{k=1}^n \gamma_k W_k^{(m)}(t)}] = \mathbb{E}_{n_0}[\mathbf{1}_{\{Q(t)=n\}} e^{-\sum_{k=1}^n \gamma_k W_k(t)}].$$

The next step of the argument is to show that for each $\alpha \in \mathbb{C}_+$, and each $\gamma \in \mathbb{C}_+$, when $Q^{(m)}(0) = 1$ for each $m \geq 1$, with that initial customer possessing an amount of work

$$B_0^{(m)} := \sum_{k=1}^{\infty} (k/2^m) \mathbf{1}_{\{(k-1)/2^m < B_0 \leq k/2^m\}}$$

where B_0 has CDF F and is independent of everything else,

$$\int_0^{\tau_0^{(m)}} e^{-\alpha t} \mathbf{1}_{\{Q^{(m)}(t)=1\}} e^{-\gamma W_1^{(m)}(t)} dt \overset{a.s.}{\to} \int_0^{\tau_0} e^{-\alpha t} \mathbf{1}_{\{Q(t)=1\}} e^{-\gamma W_1(t)} dt \tag{39}$$

as $m \to \infty$, and

$$e^{-\gamma B_0^{(m)}} \int_0^{\tau_0^{(m)}} e^{-\alpha t} \mathbf{1}_{\{Q^{(m)}(t)=1\}} dt \overset{a.s}{\to} e^{-\gamma B_0} \int_0^{\tau_0} e^{-\alpha t} \mathbf{1}_{\{Q(t)=1\}} dt \tag{40}$$

as $m \to \infty$. The argument required to show this is somewhat tedious, but the basic idea is to consider a decreasing sequence $\{\epsilon_r\}_{r\geq 1} \subset (0, 1/Re(\alpha))$ that converges to zero as $r \to \infty$, and for each integer $r \geq 1$, set t_r as that unique positive integer satisfying

$$\int_{t_r}^{\infty} e^{-Re(\alpha)t} dt = \epsilon_r$$

so clearly t_r depends on ϵ_r, and $t_r \to \infty$ as $r \to \infty$. Next, note that for each $r \geq 1$,

$$\begin{aligned}
&-\epsilon_r + \int_0^{\min(\tau_0^{(m)}, t_r)} e^{-Re(\alpha)t} \cos(Im(\alpha)t) \mathbf{1}_{\{Q^{(m)}(t)=1\}} e^{-Re(\gamma) W_1^{(m)}(t)} \cos(Im(\gamma) W_1^{(m)}(t)) dt \\
&\leq \int_0^{\tau_0^{(m)}} e^{-Re(\alpha)t} \cos(Im(\alpha)t) \mathbf{1}_{\{Q^{(m)}(t)=1\}} e^{-Re(\gamma) W_1^{(m)}(t)} \cos(Im(\gamma) W_1^{(m)}(t)) dt \\
&\leq \epsilon_r + \int_0^{\min(\tau_0^{(m)}, t_r)} e^{-Re(\alpha)t} \cos(Im(\alpha)t) \mathbf{1}_{\{Q^{(m)}(t)=1\}} e^{-Re(\gamma) W_1^{(m)}(t)} \cos(Im(\gamma) W_1^{(m)}(t)) dt
\end{aligned}$$

and on the event where there are a finite number of arrivals in $(0, t_r]$ for each $r \geq 1$, and no arrivals and service completions ever occur simultaneously,

$$\begin{aligned}
&-\epsilon_r + \int_0^{\min(\tau_0, t_r)} e^{-Re(\alpha)t} \cos(Im(\alpha)t) \mathbf{1}_{\{Q^{(m)}(t)=1\}} e^{-Re(\gamma) W_1(t)} \cos(Im(\gamma) W_1(t)) dt \\
&\leq \liminf_{m\to\infty} \int_0^{\tau_0^{(m)}} e^{-Re(\alpha)t} \cos(Im(\alpha)t) \mathbf{1}_{\{Q^{(m)}(t)=1\}} e^{-Re(\gamma) W_1^{(m)}(t)} \cos(Im(\gamma) W_1^{(m)}(t)) dt \\
&\leq \limsup_{m\to\infty} \int_0^{\tau_0^{(m)}} e^{-Re(\alpha)t} \cos(Im(\alpha)t) \mathbf{1}_{\{Q^{(m)}(t)=1\}} e^{-Re(\gamma) W_1^{(m)}(t)} \cos(Im(\gamma) W_1^{(m)}(t)) dt \\
&\leq \epsilon_r + \int_0^{\min(\tau_0, t_r)} e^{-Re(\alpha)t} \cos(Im(\alpha)t) \mathbf{1}_{\{Q^{(m)}(t)=1\}} e^{-Re(\gamma) W_1(t)} \cos(Im(\gamma) W_1(t)) dt
\end{aligned}$$

for each $r \geq 1$. Letting $r \to \infty$ then yields

$$\begin{aligned}
&\int_0^{\tau_0} e^{-Re(\alpha)t} \cos(Im(\alpha)t) \mathbf{1}_{\{Q^{(m)}(t)=1\}} e^{-Re(\gamma)W_1(t)} \cos(Im(\gamma)W_1(t))dt \\
&\leq \liminf_{m\to\infty} \int_0^{\tau_0^{(m)}} e^{-Re(\alpha)t} \cos(Im(\alpha)t) \mathbf{1}_{\{Q^{(m)}(t)=1\}} e^{-Re(\gamma)W_1^{(m)}(t)} \cos(Im(\gamma)W_1^{(m)}(t))dt \\
&\leq \limsup_{m\to\infty} \int_0^{\tau_0^{(m)}} e^{-Re(\alpha)t} \cos(Im(\alpha)t) \mathbf{1}_{\{Q^{(m)}(t)=1\}} e^{-Re(\gamma)W_1^{(m)}(t)} \cos(Im(\gamma)W_1^{(m)}(t))dt \\
&\leq \int_0^{\tau_0} e^{-Re(\alpha)t} \cos(Im(\alpha)t) \mathbf{1}_{\{Q^{(m)}(t)=1\}} e^{-Re(\gamma)W_1(t)} \cos(Im(\gamma)W_1(t))dt
\end{aligned}$$

which proves

$$\begin{aligned}
&\int_0^{\tau_0^{(m)}} e^{-Re(\alpha)t} \cos(Im(\alpha)t) \mathbf{1}_{\{Q^{(m)}(t)=1\}} e^{-Re(\gamma)W_1^{(m)}(t)} \cos(Im(\gamma)W_1^{(m)}(t))dt \\
&\overset{a.s.}{\to} \int_0^{\tau_0} e^{-Re(\alpha)t} \cos(Im(\alpha)t) \mathbf{1}_{\{Q^{(m)}(t)=1\}} e^{-Re(\gamma)W_1(t)} \cos(Im(\gamma)W_1(t))dt
\end{aligned}$$

as $m \to \infty$. A similar argument can be used three additional times, in order to account for the other cosine and sine terms associated with the real and complex parts of each complex number occurring in the integrand, which will eventually lead to (39). Once (39) has been proven, (40) follows as a simple consequence of (39), coupled with $B_1^{(m)} \to B_1$ almost-surely as $m \to \infty$. Finally, by the dominated convergence theorem, we get

$$\lim_{m\to\infty} \mathbb{E}\left[\int_0^{\tau_0^{(m)}} e^{-\alpha t} \mathbf{1}_{\{Q^{(m)}(t)=1\}} e^{-\gamma W_1^{(m)}(t)} dt\right] = \mathbb{E}\left[\int_0^{\tau_0} e^{-\alpha t} \mathbf{1}_{\{Q(t)=1\}} e^{-\gamma W_1(t)} dt\right] \quad (41)$$

and

$$\lim_{m\to\infty} \mathbb{E}\left[e^{-\gamma B_0^{(m)}} \int_0^{\tau_0^{(m)}} e^{-\alpha t} \mathbf{1}_{\{Q^{(m)}(t)=1\}} dt\right] = \mathbb{E}\left[e^{-\gamma B_0} \int_0^{\tau_0} e^{-\alpha t} \mathbf{1}_{\{Q(t)=1\}} dt\right]. \quad (42)$$

In other words,

$$\lim_{m\to\infty} N_m(\alpha,\gamma) = N(\alpha,\gamma), \qquad \lim_{m\to\infty} M_m(\alpha,\gamma) = M(\alpha,\gamma).$$

We now have all of the necessary ingredients needed to state our main result for the joint distribution of $Q(t)$ and $\{W_k(t)\}_{k\geq 1}$ for the preemptive-repeat-identical queueing system.

Theorem 2.4 *For each $\alpha \in \mathbb{C}_+$, and each $\boldsymbol{\gamma} \in \mathbb{C}_+^\infty$, we have that for each integer $n_0 \geq 0$, and each integer $n \geq 1$,*

$$\begin{aligned}
\phi_{n_0;n}(\alpha,\boldsymbol{\gamma}) &= \phi_{n_0;0}(\alpha)\lambda^n \left[\prod_{\ell=1}^{n-1} M(\alpha,\gamma_\ell)\right] N(\alpha,\gamma_n) \\
&\quad + \sum_{k=1}^{n} \mathbf{1}_{\{n_0\geq k\}} \pi(\alpha)^{n_0-k} \left[\prod_{\ell=1}^{k-1} \beta(\gamma_\ell)\right] \left[\prod_{\ell=k}^{n-1} M(\alpha,\gamma_\ell)\right] \lambda^{n-k} N(\alpha,\gamma_n)
\end{aligned}$$

where

Proof From Proposition 2.6, it is easy to show that

$$\phi^{(m)}_{n_0;n}(\alpha,\boldsymbol{\gamma}) = \sum_{(i_1,\ldots,i_n)\in E^n}\left[\prod_{\ell=1}^{n} p^{(m)}_{i_\ell}\right]\int_0^\infty e^{-\alpha t}\mathbb{E}_{n_0}\left[\mathbf{1}_{\{Q^{(m)}(t)=n,L_1(t)=i_1,\ldots,L_n(t)=i_n\}}e^{-\sum_{k=1}^{n}\gamma_k W_k^{(m)}(t)}\right]dt$$

$$= \phi^{(m)}_{n_0;0}(\alpha)\lambda^n\left[\prod_{\ell=1}^{n-1} M_m(\alpha,\gamma_\ell)\right]N_m(\alpha,\gamma_n)$$

$$+\sum_{k=1}^{n}\mathbf{1}_{\{n_0\geq k\}}\pi_m(\alpha)^{n_0-k}\left[\prod_{\ell=1}^{k-1}\beta_m(\gamma_\ell)\right]\left[\prod_{\ell=k}^{n-1}M_m(\alpha,\gamma_\ell)\right]\lambda^{n-k}N_m(\alpha,\gamma_n)$$

Letting $m\to\infty$ proves the claim. $\diamondsuit$

Our next objective is to find simpler expressions for both $N(\alpha,\gamma)$ and $M(\alpha,\gamma)$. The following proposition provides us with a computable expression for both $N(\alpha,\gamma,x)$ and $M(\alpha,\gamma,x)$.

Proposition 2.7 *For each $\alpha,\gamma\in\mathbb{C}_+$, and each real $x>0$, we have*

$$N(\alpha,\gamma,x) = \frac{e^{-\gamma x}-e^{-(\lambda+\alpha)x}}{(\lambda+\alpha-\gamma)(1-\frac{\lambda}{\lambda+\alpha}\pi(\alpha)(1-e^{-(\lambda+\alpha)x}))}. \tag{43}$$

Likewise,

$$M(\alpha,\gamma,x) = \frac{e^{-\gamma x}-e^{-(\gamma+\lambda+\alpha)x}}{(\lambda+\alpha)(1-\frac{\lambda}{\lambda+\alpha}\pi(\alpha)(1-e^{-(\lambda+\alpha)x}))}. \tag{44}$$

Proof Assuming $Q(0)=1$ and $W_1(0)=x$, we see that

$$\int_0^{\tau_0} e^{-\alpha s}\mathbf{1}_{\{Q(s)=1\}}e^{-\gamma W_1(s)}ds = \int_0^{\min(T_1,x)} e^{-\alpha s}e^{-\gamma(x-s)}ds + \mathbf{1}_{\{T_1\leq x\}}\int_{T_1+\tau_1(T_1)}^{\tau_0} e^{-\alpha s}\mathbf{1}_{\{Q(s)=1\}}e^{-\gamma W_1(s)}ds \tag{45}$$

Taking the expected value of both sides of (45) and simplifying further gives

$$N(\alpha,\gamma,x) = \int_0^x e^{-\alpha s}e^{-\lambda s}e^{-\gamma(x-s)}ds + \int_0^x e^{-\alpha y}\pi(\alpha)N(\alpha,\gamma,x)\lambda e^{-\lambda y}dy$$

$$= \frac{e^{-\gamma x}(1-e^{-(\lambda+\alpha-\gamma)x})}{\lambda+\alpha-\gamma} + \pi(\alpha)\frac{\lambda}{\lambda+\alpha}(1-e^{-(\lambda+\alpha)x})N(\alpha,\gamma,x)$$

and after solving for $N(\alpha,\gamma,x)$, we arrive at (43). Likewise,

$$M(\alpha,\gamma,x) = e^{-\gamma x}N(\alpha,0,x)$$

which proves (44). $\diamondsuit$

Corollary 2.2 *For each $\alpha,\gamma\in\mathbb{C}_+$,*

$$N(\alpha,\gamma) = \mathbb{E}\left[\frac{e^{-\gamma B}-e^{-(\lambda+\alpha)B}}{(\lambda+\alpha-\gamma)(1-\frac{\lambda}{\lambda+\alpha}\pi(\alpha)(1-e^{-(\lambda+\alpha)B}))}\right]$$

and

$$M(\alpha,\gamma) = \mathbb{E}\left[\frac{e^{-\gamma B}-e^{-(\gamma+\lambda+\alpha)B}}{(\lambda+\alpha)(1-\frac{\lambda}{\lambda+\alpha}\pi(\alpha)(1-e^{-(\lambda+\alpha)B}))}\right]$$

We close this section by verifying that our results can be used to rederive the steady-state results recently found in [9]. First,

$$N(0,0) = \mathbb{E}_1\left[\int_0^{\tau_0} \mathbf{1}_{\{Q(t)=1\}}dt\right] = \frac{\beta(-\lambda)-1}{\lambda}$$

which in turn implies that for each integer $n \geq 0$,

$$\mathbb{P}(Q(\infty)=n) = (\lambda N(0,0))^n(1-\lambda N(0,0)) = (\beta(-\lambda)-1)^n(2-\beta(-\lambda)).$$

This agrees with the results found in [9], and note in particular $0 < \mathbb{P}(Q(\infty)=0) < 1$ if and only if $\beta(-\lambda) = \mathbb{E}[e^{\lambda B}] < 2$, meaning the amount of work brought by each customer must be light-tailed, and have a finite moment generating function that is also well-defined at λ. Clearly $\mathbb{E}[e^{\lambda B}] > 1$ since $F(0) = 0$.

Next, observe that for each integer $n \geq 1$,

$$\begin{aligned}\mathbb{E}[\mathbf{1}_{\{Q(\infty)=n\}}e^{-\sum_{k=1}^n \gamma_k W_k(\infty)}] &= (2-\beta(-\lambda))\left[\prod_{\ell=1}^{n-1}\lambda M(0,\gamma_\ell)\right]N(0,\gamma_n)\\ &= (2-\beta(-\lambda))(\beta(-\lambda)-1)^n\left[\prod_{\ell=1}^{n-1}\frac{M(0,\gamma_\ell)}{N(0,0)}\right]\frac{N(0,\gamma_n)}{N(0,0)}.\end{aligned}$$

This formula reveals that, conditional on $Q(\infty) = n$,

$$\mathbb{E}[e^{-\gamma_k W_k(\infty)} \mid Q(\infty)=n] = \frac{M(0,\gamma_k)}{N(0,0)} = \frac{\beta(\gamma_k-\lambda)-\beta(\gamma_k)}{\beta(-\lambda)-1}$$

for $1 \leq k \leq n-1$, and

$$\mathbb{E}[e^{-\gamma_n W_n(\infty)} \mid Q(\infty)=n] = \frac{N(0,\gamma_n)}{N(0,0)} = \frac{\lambda(\beta(\gamma_n-\lambda)-1)}{(\lambda-\gamma_n)(\beta(-\lambda)-1)}.$$

This is in agreement with the results found in [9].

3 Calculating $\pi(\alpha)$

We conclude by discussing the problem of calculating $\pi(\alpha)$ for each $\alpha \in \mathbb{C}_+$. For preemptive-resume queues, it is very well-known that π satisfies the following fixed-point equation: for each $\alpha \in \mathbb{C}_+$,

$$\pi(\alpha) = \beta(\alpha + \lambda(1-\pi(\alpha))). \tag{46}$$

Unfortunately, (46) cannot in general be used to derive a useful expression for $\pi(\alpha)$, yet it is still possible to use (46) to approximate $\pi(\alpha)$, assuming $\beta(\alpha)$ can be evaluated easily.

Suppose first that $\alpha > 0$. In this case (as explained at the beginning of [5]) we can approximate $\pi(\alpha)$ by working with the function $T_\alpha : [0,1] \to [0,1]$, defined as

$$T_\alpha(x) := \beta(\alpha + \lambda(1-x)), \quad x \in [0,1]$$

with $T_\alpha^{(n)} : [0,1] \to [0,1]$ defined, for each integer $n \geq 2$, as the n-fold convolution of T_α with itself. T_α is clearly a nondecreasing, continuous function on $[0,1]$, and when $\alpha > 0$, $\pi(\alpha)$ satisfies $T_\alpha^{(n)}(x) \to \pi(\alpha)$ as $n \to \infty$. In particular, when $0 \leq x < \pi(\alpha)$,

$$T_\alpha^{(n)}(x) \leq T_\alpha^{(n+1)}(x)$$

for each integer $n \geq 1$. Likewise, when $\pi(\alpha) < x \leq 1$,

$$T_\alpha^{(n)}(x) \geq T_\alpha^{(n+1)}(x)$$

for each integer $n \geq 1$, which means $\pi(\alpha)$ can be calculated using the stopping criterion $T_\alpha^{(n)}(1) - T_\alpha^{(n)}(0) < \epsilon$ for some chosen error tolerance $\epsilon > 0$.

This justification no longer works when $\alpha \in \mathbb{C}$ has a complex component, and so the main point of [5] is to justify, using probabilistic methods, that $T_\alpha^{(n)}(x) \to \pi(\alpha)$ as $n \to \infty$ for each $x \in [0,1]$, and each $\alpha \in \mathbb{C}_+$. Our objective in this section is to show that a very simple coupling argument can be used to derive a convergence scheme for calculating $\pi(\alpha)$ for not only the preemptive-resume case, but both preemptive-repeat cases. The same coupling argument works for all three preemptive queues we analyze.

3.1 Constructing the Coupling

For each integer $N \geq 1$, let $\{Q_N(t); t \geq 0\}$ be the queue-lengh process of a preemptive queue fed by the same Poisson arrival process $\{A(t); t \geq 0\}$, but assume this queueing system cannot handle more than N customers at a time. Next, label each arrival as either altruistic with probability p, or selfish with probability $1-p$, independently of everything else. Both altruistic and selfish customers enter the system without hesitation whenever they observe $N-1$ customers or less in the system upon arrival: however, if an altruistic customer observes N customers in the system upon arrival, he/she departs without affecting the system in any way, and if a selfish customer observes N customers in the system upon arrival, he/she enters the system, and this action causes the server to shut down permanently, and everyone in the system is stuck there forever (meaning $Q_N(t) = N+1$ for all t past the arrival time of the selfish customer). Assume that $Q^{(N)}(0) = 1$ for each $N \geq 1$, and let

$$\tau_{N;k} := \inf\{t \geq 0 : Q^{(N)}(t) = k\}$$

for each integer $k \in \{0, 1, \ldots, N+1\}$. Finally, for each integer $N \geq 1$, we define the LST $\pi_N : \mathbb{C}_+ \to \mathbb{C}$ as

$$\pi_N(\alpha) := \mathbb{E}_1[e^{-\alpha \tau_{N;0}}].$$

Theorem 3.1 *The sequence of hitting times $\{\tau_{N;0}\}_{N \geq 1}$ converges almost-surely to τ_0 as $N \to \infty$. As a consequence, for each $\alpha \in \mathbb{C}_+$,*

$$\lim_{N \to \infty} \mathbb{E}[e^{-\alpha \tau_{N;0}}] = \mathbb{E}[e^{-\alpha \tau_0}].$$

Proof Fix an outcome $\omega \in B$, where

$$B := \left\{ \lim_{m \to \infty} T_m = \infty \right\}.$$

Clearly $\mathbb{P}(B) = 1$, so in order to prove this claim, it suffices to show $\tau_{N;0}(\omega) \to \tau_0(\omega)$ as $N \to \infty$, for each $\omega \in B$.

Case 1: Suppose first that $\omega \in B$ satisfies $\tau_0(\omega) < \infty$. Given such an ω, there must exists an integer n_0 (possibly depending on ω) such that for each integer $n \geq n_0$,

$$T_n(\omega) \geq \tau_0(\omega).$$

Define the random variable M as

$$M(\omega) := \sup_{t \in [0, \tau_0(\omega)]} Q(t)$$

and note that for such ω,

$$M(\omega) \leq n_0(\omega) < \infty$$

which, from the construction of $\{Q^{(N)}(t); t \geq 0\}$, implies $\tau_{N;0}(\omega) = \tau_0(\omega)$ for each $N \geq n_0(\omega)$, and so

$$\lim_{N\to\infty} \tau_{N;0}(\omega) = \tau_0(\omega).$$

Case 2: Next, suppose $\omega \in B$ satisfies $\tau_0(\omega) = \infty$, but $M(\omega) < \infty$. In this case, for each integer $N \geq M(\omega)$, $\tau_{N;0}(\omega) = \tau_0(\omega) = \infty$ and so again, we trivially get

$$\lim_{N\to\infty} \tau_{N;0}(\omega) = \tau_0(\omega).$$

Case 3: Finally, suppose $\omega \in B$ satisfies $\tau_0(\omega) = \infty$, and $M(\omega) = \infty$. In this case, we can see that for each integer $N \geq 1$, $\{Q(t,\omega); 0 \leq t \leq \tau_{N;N+1}(\omega)\} = \{Q^{(N)}(t,\omega); 0 \leq t \leq \tau_{N;N+1}(\omega)\}$, which implies

$$\tau_{N+1}(\omega) = \tau_{N;N+1}(\omega) \leq \tau_{N;0}(\omega)$$

and since $\{\tau_{N;N+1}(\omega)\}_{N\geq 1}$ must be a subsequence of $\{T_k(\omega)\}$, it follows that

$$\liminf_{N\to\infty} \tau_{N;0}(\omega) \geq \liminf_{N\to\infty} \tau_{N;N+1}(\omega) = \liminf_{N\to\infty} \tau_{N+1}(\omega) = \infty$$

so again, $\tau_{N;0}(\omega) \to \tau_0(\omega)$ as $N \to \infty$. This proves $\tau_{N;0}$ converges almost surely to τ_0 as $N \to \infty$, and the remaining part of the theorem follows quickly from the dominated convergence theorem. $\diamondsuit$

We are now ready to consider the problem of calculating $\pi(\alpha)$ for each $\alpha \in \mathbb{C}_+$. For each $\alpha \in \mathbb{C}_+$, and each $x > 0$, we define the function $\varphi : \mathbb{C}_+ \times (0,\infty) \to \mathbb{C}$ as

$$\varphi(\alpha, x) := \mathbb{E}_1[e^{-\alpha\tau_0} \mid B_0 = x]$$

where B_0 denotes the amount of work possessed by the single customer present at time zero. Likewise, for each integer $N \geq 1$ we define the function $\varphi_N : \mathbb{C}_+ \times (0,\infty) \to \mathbb{C}$ as

$$\varphi_N(\alpha, x) := \mathbb{E}_1[e^{-\alpha\tau_0^{(N)}} \mid B_0 = x].$$

Clearly, for each $\alpha \in \mathbb{C}_+$, it follows that for each integer $N \geq 1$,

$$\pi(\alpha) = \int_{(0,\infty)} \varphi(\alpha, x)dF(x), \qquad \pi_N(\alpha) = \int_{(0,\infty)} \varphi_N(\alpha, x)dF(x).$$

3.2 The Preemptive Resume Case

First, observe that when $N = 1$ and $B_0 = x$, the preemptive-resume queue can only empty if no selfish customers arrive during his/her service time. In other words,

$$\varphi_1(\alpha, x) = e^{-\alpha x}e^{-\lambda(1-p)x} = e^{-(\alpha+\lambda(1-p))x}$$

from which we get

$$\pi_1(\alpha) = \int_{(0,\infty)} e^{-(\alpha+\lambda(1-p))x}dF(x) = \beta(\alpha + \lambda(1-p)).$$

Next, suppose $N \geq 1$: conditioning on the first arrival time after time zero yields

$$\begin{aligned}
\varphi_{N+1}(\alpha, x) &= e^{-\alpha x}e^{-\lambda x} + \int_0^x e^{-\alpha y}\pi_N(\alpha)\varphi_{N+1}(\alpha, x-y)\lambda e^{-\lambda y}dy \\
&= e^{-(\lambda+\alpha)x} + \pi_N(\alpha)\int_0^x e^{-\alpha(x-y)}\varphi_{N+1}(\alpha, y)\lambda e^{-\lambda(x-y)}dy \\
&= e^{-(\lambda+\alpha)x} + \lambda\pi_N(\alpha)e^{-(\lambda+\alpha)x}\int_0^x e^{(\lambda+\alpha)y}\varphi_{N+1}(\alpha, y)dy.
\end{aligned}$$

Multiplying both sides of the equation by $e^{(\lambda+\alpha)x}$, then taking the partial derivative of both sides with respect to x yields, after simplifying,

$$\frac{\partial}{\partial x}\varphi_{N+1}(\alpha,x)+(\lambda+\alpha)\varphi_{N+1}(\alpha,x)=\lambda\pi_N(\alpha)\varphi_{N+1}(\alpha,x).$$

Solving this linear, first-order ODE yields

$$\varphi_{N+1}(\alpha,x)=e^{-(\alpha+\lambda(1-\pi_N(\alpha)))x}$$

and so

$$\pi_{N+1}(\alpha)=\beta(\alpha+\lambda(1-\pi_N(\alpha))). \tag{47}$$

This coincides with the recursion found in [5]. Finally, note that our coupling argument yields $\lim_{N\to\infty}\pi_N(\alpha)=\pi(\alpha)$ as $N\to\infty$ for each $\alpha\in\mathbb{C}_+$, and the argument used to establish (47) can be used to show $\pi(\alpha)$ satisfies, for $\alpha\in\mathbb{C}_+$,

$$\pi(\alpha)=\beta(\alpha+\lambda(1-\pi(\alpha))).$$

3.3 The Preemptive Repeat Different Case

We next use our coupling construction to devise a recursive procedure for calculating $\pi(\alpha)$ for the preemptive-repeat different queue. In [15], the authors show how to express $\pi(\alpha)$ explicitly in terms of λ and the service time LST β for each positive real number $\alpha\geq 0$, and we will build on their work slightly by addressing the case where $\alpha\in\mathbb{C}_+$.

It is not difficult to show that $\pi(\alpha)$ is a root of a quadratic polynomial, even when $\alpha\in\mathbb{C}_+$. Proceeding as in [15], we find that for each $x>0$,

$$\begin{aligned}\varphi(\alpha,x)&=e^{-(\alpha+\lambda)x}+\int_0^x e^{-\alpha y}\pi(\alpha)^2\lambda e^{-\lambda y}dy\\&=e^{-(\alpha+\lambda)x}+\left[\lambda\int_0^x e^{-(\alpha+\lambda)y}dy\right]\pi(\alpha)^2\\&=e^{-(\alpha+\lambda)x}+\frac{\lambda}{\lambda+\alpha}(1-e^{-(\alpha+\lambda)x})\pi(\alpha)^2.\end{aligned}$$

Integrating both sides over $[0,\infty)$ with respect to $dF(x)$ gives

$$\pi(\alpha)=\beta(\alpha+\lambda)+\frac{\lambda}{\lambda+\alpha}(1-\beta(\alpha+\lambda))\pi(\alpha)^2$$

which means $\pi(\alpha)$ satisfies

$$\lambda(1-\beta(\alpha+\lambda))\pi(\alpha)^2-(\alpha+\lambda)\pi(\alpha)+(\alpha+\lambda)\beta(\alpha+\lambda)=0.$$

The challenge now is to rigorously determine which root corresponds to $\pi(\alpha)$. If we restrict π so that it is defined on $(0,\infty)$ instead of $\mathbb{C}_+$, we get (as shown in [15])

$$\pi(\alpha)=\frac{\lambda+\alpha-\sqrt{(\lambda+\alpha)^2-4\lambda(1-\beta(\alpha+\lambda))(\lambda+\alpha)\beta(\alpha+\lambda)}}{2\lambda(1-\beta(\alpha+\lambda))} \tag{48}$$

because this restricted version of π is both continuous and nonincreasing on $(0,\infty)$.

We suspect (48) is still true for each $\alpha\in\mathbb{C}_+$, but this needs to be verified rigorously. Fortunately, our coupling construction can be used to get around this issue. Observe first that for each $\alpha\in\mathbb{C}_+$, and each $x>0$, the same argument we used for the preemptive-resume queue reveals

$$\varphi_1(\alpha,x)=e^{-(\alpha+\lambda(1-p))x}$$

and so $\pi_1(\alpha) = \beta(\alpha + \lambda(1-p))$. Next, for each integer $N \geq 1$, conditioning on the first arrival time after time zero gives

$$\begin{aligned}\varphi_{N+1}(\alpha, x) &= e^{-\alpha x} e^{-\lambda x} + \int_0^x e^{-\alpha y} \pi_N(\alpha) \pi_{N+1}(\alpha) \lambda e^{-\lambda y} dy \\ &= e^{-(\lambda+\alpha)x} + \lambda \left[\int_0^x e^{-(\lambda+\alpha)y} dy \right] \pi_N(\alpha) \pi_{N+1}(\alpha) \\ &= e^{-(\lambda+\alpha)x} + \frac{\lambda(1 - e^{-(\lambda+\alpha)x})}{\lambda + \alpha} \pi_N(\alpha) \pi_{N+1}(\alpha).\end{aligned}$$

Integrating both sides over $(0, \infty)$ with respect to $dF(x)$ yields

$$\pi_{N+1}(\alpha) = \beta(\lambda + \alpha) + \frac{\lambda(1 - \beta(\lambda + \alpha))}{\lambda + \alpha} \pi_N(\alpha) \pi_{N+1}(\alpha)$$

i.e.

$$\pi_{N+1}(\alpha) = \frac{\beta(\lambda + \alpha)}{1 - \frac{\lambda(1-\beta(\lambda+\alpha))}{\lambda+\alpha} \pi_N(\alpha)}.$$

We have already established from our coupling that $\pi_N(\alpha) \to \pi(\alpha)$ as $N \to \infty$ for each $\alpha \in \mathbb{C}_+$, so one way to calculate $\pi(\alpha)$ is to use this recursion to determine which root the $\pi_N(\alpha)$ terms appears to converge toward for large N, then use that root as the value for $\pi(\alpha)$.

3.4 The Preemptive Repeat Identical Case

It remains to derive a recursive procedure for calculating the LST $\pi(\alpha)$ associated with the preemptive repeat identical queue. Again, we begin by observing that for each $\alpha \in \mathbb{C}_+$, and each $x > 0$,

$$\varphi_1(\alpha, x) = e^{-\alpha x} e^{-\lambda(1-p)x} = e^{-(\alpha + \lambda(1-p))x}$$

which implies

$$\pi_1(\alpha) = \beta(\alpha + \lambda(1-p)).$$

Next, for each integer $N \geq 1$, conditioning on B_0 gives

$$\begin{aligned}\varphi_{N+1}(\alpha, x) &= e^{-\alpha x} e^{-\lambda x} + \int_0^x e^{-\alpha y} \pi_N(\alpha) \varphi_{N+1}(\alpha, x) \lambda e^{-\lambda y} dy \\ &= e^{-(\lambda+\alpha)x} + \lambda \left[\int_0^x e^{-(\lambda+\alpha)y} dy \right] \pi_N(\alpha) \varphi_{N+1}(\alpha, x) \\ &= e^{-(\lambda+\alpha)x} + \frac{\lambda(1 - e^{-(\lambda+\alpha)x})}{\lambda + \alpha} \pi_N(\alpha) \varphi_{N+1}(\alpha, x)\end{aligned}$$

which yields

$$\varphi_{N+1}(\alpha, x) = \frac{e^{-(\lambda+\alpha)x}}{1 - \frac{\lambda}{\lambda+\alpha}(1 - e^{-(\lambda+\alpha)x}) \pi_N(\alpha)}.$$

Finally, integrating over $[0, \infty)$ with respect to $dF(x)$ yields

$$\pi_{N+1}(\alpha) = \mathbb{E} \left[\frac{e^{-(\lambda+\alpha)B_0}}{1 - \frac{\lambda}{\lambda+\alpha}(1 - e^{-(\lambda+\alpha)B_0}) \pi_N(\alpha)} \right]. \tag{49}$$

Unfortunately, the expected value found on the right-hand-side of (49) may not be tractable in general, so numerical integration methods will be needed in order to approximate the expectation at each step of the recursion. On the other hand, if B_0 is a finite, discrete random variable, the recursion becomes somewhat easier to use. Obviously, the limit $\pi(\alpha)$ satisfies

$$\pi(\alpha) = \mathbb{E} \left[\frac{e^{-(\lambda+\alpha)B_0}}{1 - \frac{\lambda}{\lambda+\alpha}(1 - e^{-(\lambda+\alpha)B_0}) \pi(\alpha)} \right].$$

References

[1] Abate, J., and Whitt, W. (1987a). Transient behavior of the M/M/1 queue: staring at the origin. *Queueing Systems* **2**, 41-65.

[2] Abate, J., and Whitt, W. (1987b). Transient behavior of regulated Brownian motion I: starting at the origin. *Advances in Applied Probability* **19**, 560-598.

[3] Abate, J., and Whitt, W. (1987c). Transient behavior of regulated Brownian motion II: non-zero initial conditions. *Advances in Applied Probability* **19**, 599-631.

[4] Abate, J., and Whitt, W. (1988). The transient behavior of the M/M/1 queue via Laplace transforms. *Advances in Applied Probability* **20**, 145-178.

[5] Abate, J., and Whitt, W. (1992). Solving probability transform functional equations for numerical inversion. *Operations Research Letters* **12**, 275-281.

[6] Abate, J., and Whitt, W. (1994). Transient behavior of the M/G/1 workload process. *Operations Research* **42**, 750-764.

[7] Abate, J., and Whitt, W. (1996). Numerical inversion of Laplace transforms of probability distributions. *ORSA Journal of Computing* **7**, 36-43.

[8] Asmussen, S., and Glynn, P. (2017). On preemptive-repeat LIFO queues. *Queueing Systems* **87**, 1-22.

[9] Bergquist, J., and Sigman, K. (2022). Stationary workload and service times for some nonwork-conserving M/G/1 preemptive LIFO queues. *Stochastic Models* **38**, 515-544.

[10] Fralix, B. H. (2009). Another look at the transient behavior of the M/G/1 workload process. Unpublished EURANDOM report, accessible at https://www.eurandom.tue.nl/reports/2009/017-report.pdf

[11] Fralix, B. (2013). On the time-dependent moments of Markovian queues with reneging. *Queueing Systems* **75**, 149-168.

[12] Fralix, B., and Riaño, G. (2010). A new look at transient version of Little's law, with applications to M/G/1 preemptive Last-Come-First-Served queues. *Journal of Applied Probability* **47**, 459-473.

[13] Fralix, B., and Zwart, B. (2011). Time-dependent properties of symmetric queues. *Queueing Systems* **67**, 33-45.

[14] Fralix, B., Van Leeuwaarden, J., and Boxma, O. (2013). Factorization identities for a general class of reflected processes. *Journal of Applied Probability* **50**, 632-653.

[15] Horváth, I., Razumchik, R., and Telek, M. (2019). The resampling M/G/1 non-preemptive LIFO queue and its application for systems with uncertain service time. *Performance Evaluation* **134**, 102000.

[16] Kella, O., Zwart, B., and Boxma, O. (2005). Some time-dependent properties of symmetric M/G/1 queues. *Journal of Applied Probability* **42**, 223-234.

[17] Latouche, G., and Ramaswami, V. (1999) *Introduction to Matrix-Analytic Methods in Stochastic Modeling* ASA-SIAM Publiations, Philadelphia, PA, USA.