

Time-dependent analysis of an $M/M/c$ preemptive priority system with two priority classes

Jori Selen^{1,2}  · Brian Fralix³

Received: 29 July 2016 / Revised: 27 June 2017 / Published online: 18 July 2017
© The Author(s) 2017. This article is an open access publication

Abstract We analyze the time-dependent behavior of an $M/M/c$ priority queue having two customer classes, class-dependent service rates, and preemptive priority between classes. More particularly, we develop a method that determines the Laplace transforms of the transition functions when the system is initially empty. The Laplace transforms corresponding to states with at least c high-priority customers are expressed explicitly in terms of the Laplace transforms corresponding to states with at most $c - 1$ high-priority customers. We then show how to compute the remaining Laplace transforms recursively, by making use of a variant of Ramaswami's formula from the theory of $M/G/1$ -type Markov processes. While the primary focus of our work is on deriving Laplace transforms of transition functions, analogous results can be derived for the stationary distribution; these results seem to yield the most explicit expressions known to date.

Keywords Static priority · Time-dependent analysis · Laplace transforms · Multi-dimensional Markov process

Mathematics Subject Classification 60J27 · 60J35 · 60K25

✉ Jori Selen
j.selen@tue.nl

¹ Department of Mathematics and Computer Science, Eindhoven University of Technology, Eindhoven, Netherlands

² Department of Mechanical Engineering, Eindhoven University of Technology, Eindhoven, Netherlands

³ Department of Mathematical Sciences, Clemson University, Clemson, SC, USA

1 Introduction

Priority models with multiple servers constitute an important class of queueing systems, having applications in areas as diverse as manufacturing, wireless communication and the service industry. Studies of these models date back to at least the 1950s (see, for example, Cobham [7], Davis [8], and Jaiswal [15, 16]), yet many properties of these systems still do not appear to be well understood. Recent work addressing priority models includes Sleptchenko et al. [27] and Wang et al. [29]. We refer the reader to [29] for more specific examples of applications of priority queueing models.

Our contribution to this stream of the literature is an analysis of the time-dependent behavior of a Markovian multi-server queue with two customer classes, class-dependent service rates and preemptive priority between classes. To the best of our knowledge, the joint stationary distribution of the $M/M/1$ 2-class preemptive priority system was first studied in Miller [24], who makes use of matrix-geometric methods to study the joint stationary distribution of the number of high- and low-priority customers in the system. More particularly, in [24] this queueing system is modeled as a quasi-birth-and-death (QBD) process having infinitely many levels, with each level containing infinitely many phases. Miller then shows how to recursively compute the elements of the rate matrix of this QBD process: Once enough elements of this rate matrix have been found, the joint stationary distribution can be approximated by appropriately truncating this matrix.

This single-server model is featured in many works that have recently appeared in the literature. In Sleptchenko et al. [27] an exact, recursive procedure is given for computing the joint stationary distribution of an $M/M/1$ preemptive priority queue that serves an arbitrary finite number of customer classes. The $M/M/1$ 2-class priority model is briefly discussed in Katehakis et al. [20], where they explain how the successive lumping technique can be used to study $M/M/1$ 2-class priority models when both customer classes experience the same service rate. Interesting asymptotic properties of the stationary distribution of the $M/M/1$ 2-class preemptive priority model can be found in the work of Li and Zhao [23].

Multi-server preemptive priority systems with two customer classes have also received some attention in the literature. One of the earlier references allowing for different service requirements between customer classes is Gail et al. [13]; see also the references therein. In [13], the authors derive the generating function of the joint stationary probabilities by expressing it in terms of the stationary probabilities associated with states where there is no queue. A combination of a generating function approach and the matrix-geometric approach is used in Sleptchenko et al. [26] to compute the joint stationary distribution of an $M/M/c$ 2-class preemptive priority queue. The $M/PH/c$ queue with an arbitrary number of preemptive priority classes is studied in Harchol-Balter et al. [14] using a recursive dimensionality reduction technique that leads to an accurate approximation of the mean sojourn time per customer class. Furthermore, in Wang et al. [29] the authors present a procedure for finding, for an $M/M/c$ 2-class priority model, the generating function of the distribution of the number of low-priority customers present in the system in stationarity.

Our work deviates from all of the above approaches, in that we construct a procedure for computing the Laplace transforms of the transition functions of the $M/M/c$ 2-class

preemptive priority model. Our method first makes use of a slight tweak of the *clearing analysis on phases* (CAP) method featured in Doroudi et al. [10], in that we show how CAP can be modified to study Laplace transforms of transition functions. The specific dynamics of our priority model allow us to take the analysis a few steps further, by showing each Laplace transform can be expressed *explicitly* in terms of transforms corresponding to states contained within a strip of states that is infinite in only one direction. Finally, we show how to compute these remaining transforms recursively, by making use of a slight modification of Ramaswami's formula [25]. While the focus of our work is on Laplace transforms of transition functions, analogous results can be derived for the stationary distribution of the $M/M/c$ 2-class preemptive priority model as well. We are not aware of any studies that obtain explicit expressions for the Laplace transforms of the transition functions, or even the stationary distribution, as we do here; these results seem to yield the most explicit expressions known to date.

The Laplace transforms we derive can easily be numerically inverted to retrieve the transition functions with the help of the algorithms of Abate and Whitt [3,4] or den Iseger [9]. These transition functions can be used to study—as a function of time—key performance measures such as the mean number of customers of each priority class in the system; the mean total number of customers in the system; or the probability that an arriving customer has to wait in the queue. These time-dependent performance measures can, for example, be used to analyze and dimension priority systems when one is interested in the behavior of such systems over a finite time horizon. Using the equilibrium distribution as an approximation of the time-dependent behavior to dimension the system can result in either over- or underdimensioning since it neglects transient effects such as the initial number of customers in the system. Prior to our work, the time-dependent behavior of multi-server queues could not be analyzed using any methods from previous work; until now one would have to resort to simulation in order to study the system's time-dependent behavior. Having explicit expressions for the Laplace transforms of the transition functions greatly simplifies the computation of some performance measures. For instance, these transforms yield explicit expressions for the Laplace transforms of the distribution of the number of low-priority customers in the system at time t .

We now present some numerical examples of the time-dependent performance measures, where we will make use of the notation introduced in Sect. 2. In Fig. 1, we plot the mean number of low-priority customers in the system as a function of time and we also depict the equilibrium values. Similarly, in Fig. 2 we plot the time-dependent and equilibrium delay probabilities for each priority class. The Laplace transforms used to obtain Figs. 1 and 2 can be computed numerically using the approach discussed in Sect. 5.5: here we used an error tolerance of $\epsilon = 10^{-8}$. Once these transforms have been found, numerical inversion can be done via the Euler summation algorithm of [3] where we again used an error tolerance of 10^{-8} . From Fig. 1 we can also informally derive the mixing times of each scenario. It seems that the mixing time vastly increases with an increase in the load. As expected, in Fig. 2 we see that the delay probability of a high-priority customer is much lower than the delay probability of a low-priority customer. Furthermore, as time passes, the delay probability of the high-priority customer tends to the delay probability in an $M/M/c$ queue with only high-priority customers, which was verified to be correct. Finally, in Table 1 we show

Fig. 1 Mean number of low-priority customers in the system as a function of time for increasing load of the high-priority customers. Equilibrium values are shown at $t = \infty$. Parameter settings are $c = 10$, $\rho_1 = 1/3$, and ρ_2 varies

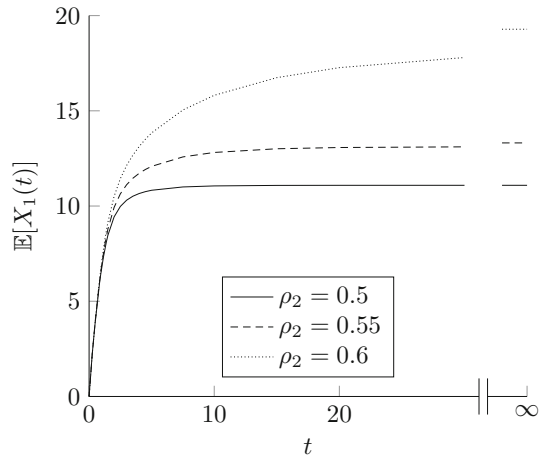


Fig. 2 Probability that an arriving customer has to wait in the queue as a function of time for the two priority classes. Equilibrium values are shown at $t = \infty$. Parameter settings are $c = 10$, $\rho_1 = 1/3$, and $\rho_2 = 1/2$

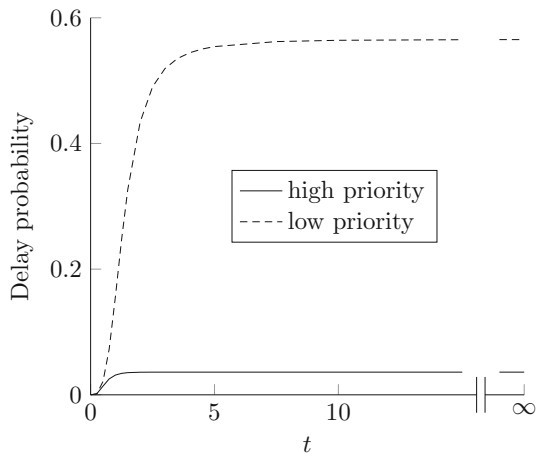


Table 1 Computation time in seconds required to calculate, for all states in \mathbb{S}_k , the stationary probabilities or the Laplace transforms for a specific $\alpha = 1/2 + 1/2i$

Number of servers	10	20	30	50	70	100
Stationary probabilities	0.35	0.66	1.1	2.9	7.1	21
Laplace transforms	0.52	1.3	2.7	8.7	23	62

We use the numerical implementation outlined in Sect. 5.5 with accuracies $\epsilon = 10^{-8}$. Parameter settings are $\rho_1 = 1/3$, $\rho_2 = 1/2$ and c varies

the computation times of the algorithm, which was implemented in Matlab and run on a 64-bit desktop with an Intel Core i7-3770 processor. The computation time scales reasonably well with the number of servers, and therefore the algorithm can be used to evaluate any practical instance.

This paper is organized as follows. Section 2 describes both the $M/M/c$ 2-class preemptive priority queueing system, as well as the two-dimensional Markov process used to model the dynamics of this system. In the same section, we introduce relevant notation and terminology and detail the outline of the approach. In Sects. 3–5, we describe this approach for calculating the Laplace transforms of the transition functions. We discuss the simplifications in the single-server case in Sect. 6. In Sect. 7, we summarize our contributions and comment on the derivation of the stationary distribution. The appendices provide supporting results on combinatorial identities and single-server queues used in deriving the expressions for the Laplace transforms.

2 Model description and outline of approach

We consider a queueing system consisting of c servers, where each server processes work at unit rate. This system serves customers from two different customer classes, referred to here as class-1 and class-2 customers. The class index indicates the priority rank, meaning that among the servers, class-2 customers have preemptive priority over class-1 customers in service. Recall that the term ‘preemptive priority’ means that whenever a class-2 customer arrives at the system, one of the servers currently serving a class-1 customer immediately drops that customer and begins serving the new class-2 arrival, and the dropped class-1 customer waits in the system until a server is again available to receive further processing, i.e., the priority rule is preemptive resume. Therefore, if there are currently i class-1 customers and j class-2 customers in the system, the number of class-2 customers in service is $\min(c, j)$, while the number of class-1 customers in service is $\max(\min(i, c - j), 0)$.

Class- n customers arrive in a Poisson manner with rate $\lambda_n, n = 1, 2$, and the Poisson arrival processes of the two populations are assumed to be independent. Each class- n arrival brings an exponentially distributed amount of work with rate μ_n , independently of everything else. We denote the total arrival rate by $\lambda := \lambda_1 + \lambda_2$, the load induced by class- n customers as $\rho_n := \lambda_n / (c\mu_n)$, and the load induced by both customer classes as $\rho := \rho_1 + \rho_2$.

The dynamics of this queueing system can be described with a continuous-time Markov chain (CTMC). For each $t \geq 0$, let $X_n(t)$ represent the number of class- n customers in the system at time t , and define $X(t) := (X_1(t), X_2(t))$. Then, $X := \{X(t)\}_{t \geq 0}$ is a CTMC on the state space $\mathbb{S} = \mathbb{N}_0^2$. Given any two distinct elements $x, y \in \mathbb{S}$, the element $q(x, y)$ of the transition rate matrix \mathbf{Q} associated with X denotes the transition rate from state x to state y . The row sums of \mathbf{Q} are 0, meaning for each $x \in \mathbb{S}, q(x, x) = -\sum_{y \neq x} q(x, y) =: -q(x)$, where $q(x)$ represents the rate of each exponential sojourn time in state x . For our queueing system, the nonzero transition rates of \mathbf{Q} are given by

$$\begin{aligned} q((i, j), (i + 1, j)) &= \lambda_1, \quad i, j \geq 0, \\ q((i, j), (i, j + 1)) &= \lambda_2, \quad i, j \geq 0, \\ q((i, j), (i - 1, j)) &= \max(\min(i, c - j), 0)\mu_1, \quad i \geq 1, j \geq 0, \\ q((i, j), (i, j - 1)) &= \min(c, j)\mu_2, \quad i \geq 0, j \geq 1. \end{aligned}$$

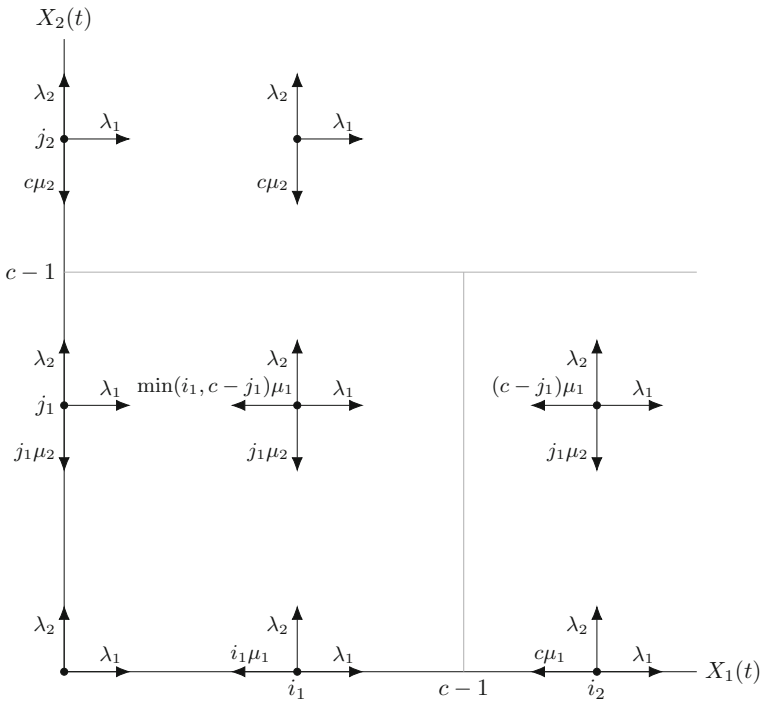


Fig. 3 Transition rate diagram of the Markov process X

Figure 3 displays the transition rate diagram.

We further associate with the Markov process X the collection of transition functions $\{p_{x,y}(\cdot)\}_{x,y \in \mathbb{S}}$, where for each $x, y \in \mathbb{S}$ (with possibly $x = y$) the function $p_{x,y} : [0, \infty) \rightarrow [0, 1]$ is defined as

$$p_{x,y}(t) := \mathbb{P}(X(t) = y \mid X(0) = x), \quad t \geq 0. \tag{2.1}$$

Each transition function $p_{x,y}(\cdot)$ has a Laplace transform $\pi_{x,y}(\cdot)$ that is well-defined on the subset of complex numbers $\mathbb{C}_+ := \{\alpha \in \mathbb{C} : \text{Re}(\alpha) > 0\}$ as

$$\pi_{x,y}(\alpha) := \int_0^\infty e^{-\alpha t} p_{x,y}(t) dt, \quad \alpha \in \mathbb{C}_+. \tag{2.2}$$

We restrict our interest to transition functions of X when $X(0) = (0, 0)$ with probability one (w.p.1), and so we drop the first subscript on both transition functions and Laplace transforms, i.e., $p_x(t) := p_{(0,0),x}(t)$ for each $t \geq 0$ and $\pi_x(\alpha) := \pi_{(0,0),x}(\alpha)$ for each $\alpha \in \mathbb{C}_+$. Our goal is to derive efficient numerical methods for calculating each Laplace transform $\pi_x(\alpha)$, $x \in \mathbb{S}$. We often refer to the Laplace transform $\pi_x(\alpha)$ associated with the state x as the Laplace transform for state x .

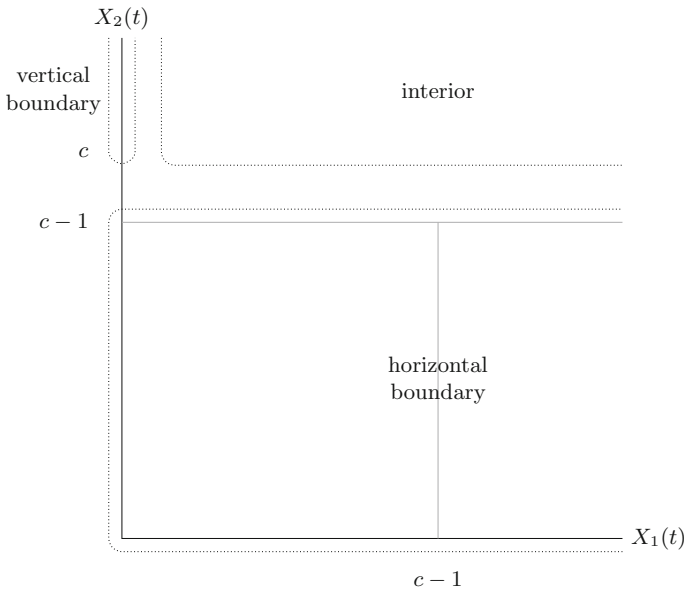


Fig. 4 Terminology of the various sets of states

2.1 Notation and terminology

It helps to decompose the state space \mathbb{S} into a countable number of levels, where, for each integer $i \geq 0$, the i -th level is the set $\{(i, 0), (i, 1), \dots\}$. We further decompose the i -th level into an upper level and a lower level: Upper level i is defined as $U_i := \{(i, c), (i, c + 1), \dots\}$, while lower level i is simply $L_i := \{(i, 0), (i, 1), \dots, (i, c - 1)\}$ and the union of lower levels L_0, L_1, \dots, L_i is denoted by $L_{\leq i} = \bigcup_{k=0}^i L_k$. The set of all states in phase j is denoted by $P_j := \{(0, j), (1, j), \dots\}$.

We sometimes refer to upper level U_0 as the *vertical boundary*. The union of upper levels

$$U_1 \cup U_2 \cup \dots \tag{2.3}$$

is called the *interior* of the state space. Finally, the union

$$L_0 \cup L_1 \cup \dots \tag{2.4}$$

is called the *horizontal boundary* or the *horizontal strip of boundary states*. Figure 4 depicts these sets.

The indicator function $\mathbb{1}\{A\}$ equals 1 if A is true and 0 otherwise. Given an arbitrary CTMC Z , we let $\mathbb{E}_z[f(Z)]$ represent the expectation of a functional of Z , conditional on $Z(0) = z$, and $\mathbb{P}_z(\cdot)$ denotes the conditional probability associated with $\mathbb{E}_z[\cdot]$. In our analysis, it should be clear from the context what is being conditioned on when we write $\mathbb{P}_z(\cdot)$ or $\mathbb{E}_z[\cdot]$.

We will also need to make use of hitting-time random variables. We define for each set $A \subset \mathbb{S}$,

$$\tau_A := \inf\{t > 0 : \lim_{s \uparrow t} X(s) \neq X(t) \in A\} \tag{2.5}$$

as the first time X makes a transition into the set A (so note $X(0) \in A$ does not imply $\tau_A = 0$) and τ_x should be understood to mean $\tau_{\{x\}}$.

2.2 Notation for $M/M/1$ queues

Most of the formulas we derive contain quantities associated with an ordinary $M/M/1$ queue. Given an $M/M/1$ queueing system with arrival rate λ and service rate μ , let $Q_{\lambda,\mu}(t)$ denote the total number of customers in the system at time t . Under the measure $\mathbb{P}_n(\cdot)$, which, in this case, represents conditioning on $Q_{\lambda,\mu}(0) = n$, let $B_{\lambda,\mu}$ denote the busy period duration induced by these customers. Under $\mathbb{P}_1(\cdot)$, the Laplace–Stieltjes transform of $B_{\lambda,\mu}$ is given by

$$\phi_{\lambda,\mu}(\alpha) := \mathbb{E}_1[e^{-\alpha B_{\lambda,\mu}}] = \frac{\lambda + \mu + \alpha - \sqrt{(\lambda + \mu + \alpha)^2 - 4\lambda\mu}}{2\lambda}. \tag{2.6}$$

Recall that under $\mathbb{P}_n(\cdot)$, $B_{\lambda,\mu}$ is equal in distribution to the sum of n i.i.d. copies of $B_{\lambda,\mu}$ under the measure $\mathbb{P}_1(\cdot)$; see, for example, [28, p. 32]. Thus, for each integer $n \geq 1$ we have

$$\mathbb{E}_n[e^{-\alpha B_{\lambda,\mu}}] = \phi_{\lambda,\mu}(\alpha)^n. \tag{2.7}$$

We will also need to make use of the following quantities in Sects. 5 and 6. Suppose $\{\Lambda_\theta(t)\}_{t \geq 0}$ is a homogeneous Poisson process with rate θ that is independent of $\{Q_{\lambda,\mu}(t)\}_{t \geq 0}$. For each integer $i \geq 0$, define

$$w_i^{(\lambda,\mu,\theta)}(\alpha) := \mathbb{E}_1[e^{-\alpha B_{\lambda,\mu}} \mathbb{1}\{\Lambda_\theta(B_{\lambda,\mu}) = i\}]. \tag{2.8}$$

Lemma 7 of Appendix C shows that the $w_i^{(\lambda,\mu,\theta)}(\alpha)$ terms satisfy a recursion, and, in Lemma 8 of Appendix C, we give explicit expressions for these terms by solving the recursion.

The following quantities associated with $M/M/1$ queues will appear at many places of the analysis. To increase readability, we adopt the notation used in [10] and define the quantities

$$\phi_2 := \phi_{\lambda_2,c\mu_2}(\lambda_1 + \alpha), \quad r_2 := \rho_2 \phi_{\lambda_2,c\mu_2}(\lambda_1 + \alpha), \tag{2.9}$$

and

$$\Omega_2 := \frac{\rho_2 \phi_{\lambda_2,c\mu_2}(\lambda_1 + \alpha)}{\lambda_2(1 - \rho_2 \phi_{\lambda_2,c\mu_2}(\lambda_1 + \alpha)^2)}. \tag{2.10}$$

Further results for $M/M/1$ queues are presented in Appendix C.

2.3 Outline of our approach

Our approach for computing the Laplace transforms of the transition functions of X when $X(0) = (0, 0)$ w.p.1 is divided into three parts:

1. For each integer $i \geq 0$, we use a slight modification of the CAP method [10] to write each Laplace transform for each state in U_i , i.e., $\pi_{(i,c-1+j)}(\alpha)$, $j \geq 1$, in terms of the Laplace transforms $\pi_{(k,c-1)}(\alpha)$, $0 \leq k \leq i$, as well as additional coefficients $\{v_{k,l}\}_{i \geq k \geq l \geq 0}$ that satisfy a recursion.
2. In Sect. 4, we obtain an explicit expression for the coefficients $\{v_{k,l}\}_{k \geq l \geq 0}$. This in turn shows that for each $i \geq 0$, each Laplace transform for each state in U_i , i.e., $\pi_{(i,c-1+j)}(\alpha)$, $j \geq 1$, can be explicitly expressed in terms of $\pi_{(k,c-1)}(\alpha)$, $0 \leq k \leq i$.
3. In Sect. 5, we derive a recursion with which we can determine the Laplace transforms for the states in the horizontal boundary. Specifically, we derive a modification of Ramaswami’s formula [25] to recursively compute the remaining Laplace transforms $\pi_{(i,j)}(\alpha)$, $i \geq 0$, $0 \leq j \leq c - 1$. The techniques we use to derive this recursion are exactly the same as the techniques recently used in [17] to study block-structured Markov processes. Only the Ramaswami-like recursion is needed to compute all Laplace transforms: Once the values for the Laplace transforms of the states in the horizontal boundary are known, all other transforms can be stated explicitly without using additional recursions.

3 A slight modification of the CAP method

The following theorem is used in multiple ways throughout our analysis. It appears in [17, Theorem 2.1] and can be derived by taking the Laplace transform of both sides of the equation at the top of page 124 of [21]. Equation (3.2) is the Laplace transform version of [10, Theorem 1].

Theorem 1 *Suppose A and B are disjoint subsets of \mathbb{S} with $x \in A$. Then for each $y \in B$,*

$$\pi_{x,y}(\alpha) = \sum_{z \in A} \pi_{x,z}(\alpha)(q(z) + \alpha) \mathbb{E}_z \left[\int_0^{\tau_A} e^{-\alpha t} \mathbb{1}\{X(t) = y\} dt \right], \quad (3.1)$$

or, equivalently,

$$\pi_{x,y}(\alpha) = \sum_{z \in A} \pi_{x,z}(\alpha) \sum_{z' \in A^c} q(z, z') \mathbb{E}_z \left[\int_0^{\tau_A} e^{-\alpha t} \mathbb{1}\{X(t) = y\} dt \right]. \quad (3.2)$$

3.1 Laplace transforms for states along the vertical boundary

In this subsection, we employ Theorem 1 to express each Laplace transform $\pi_{(0,c-1+j)}(\alpha)$, $j \geq 1$, in terms of $\pi_{(0,c-1)}(\alpha)$.

Using Theorem 1 with $A = U_0^c$ we obtain, for $j \geq 1$,

$$\begin{aligned} &\pi_{(0,c-1+j)}(\alpha) \\ &= \sum_{z \in U_0^c} \pi_z(\alpha) \sum_{z' \in U_0} q(z, z') \mathbb{E}_{z'} \left[\int_0^{\tau_{U_0^c}} e^{-\alpha t} \mathbb{1}\{X(t) = (0, c - 1 + j)\} dt \right] \\ &= \pi_{(0,c-1)}(\alpha) \lambda_2 \mathbb{E}_{(0,c)} \left[\int_0^{\tau_{U_0^c}} e^{-\alpha t} \mathbb{1}\{X(t) = (0, c - 1 + j)\} dt \right]. \end{aligned} \tag{3.3}$$

From the transition rate diagram in Fig. 3, we find that the expectation in (3.3) can be interpreted as an expectation associated with an $M/M/1$ queue having arrival rate λ_2 and service rate $c\mu_2$. Indeed, $\tau_{U_0^c}$ is equal in distribution to the minimum of the busy period—initialized by one customer—of this $M/M/1$ queue and an exponential random variable with rate λ_1 that is independent of the queue. Alternatively, $\tau_{U_0^c}$ can be thought of as being equal in distribution to the busy period duration of an $M/M/1$ clearing model, with arrival rate λ_2 , service rate $c\mu_2$, and clearings that occur in a Poisson manner with rate λ_1 . Applying Lemma 6 of Appendix C shows that

$$\lambda_2 \mathbb{E}_{(0,c)} \left[\int_0^{\tau_{U_0^c}} e^{-\alpha t} \mathbb{1}\{X(t) = (0, c - 1 + j)\} dt \right] = r_2^j. \tag{3.4}$$

Substituting (3.4) into (3.3) then yields

$$\pi_{(0,c-1+j)}(\alpha) = \pi_{(0,c-1)}(\alpha) r_2^j, \quad j \geq 1. \tag{3.5}$$

3.2 Laplace transforms for states within the interior

We next develop a recursion for the Laplace transforms for the states within the interior. First, we express the transforms in upper level U_i in terms of the transforms in upper level U_{i-1} and in state $(i, c - 1)$. Second, we use this result to express the transforms in upper level U_i in terms of the transforms for the states $(0, c - 1), (1, c - 1), \dots, (i, c - 1)$ and some additional coefficients.

Employing again Theorem 1, now with $A = U_i^c$, yields, for $i, j \geq 1$,

$$\begin{aligned} &\pi_{(i,c-1+j)}(\alpha) \\ &= \sum_{z \in U_i^c} \pi_z(\alpha) \sum_{z' \in U_i} q(z, z') \mathbb{E}_{z'} \left[\int_0^{\tau_{U_i^c}} e^{-\alpha t} \mathbb{1}\{X(t) = (i, c - 1 + j)\} dt \right] \\ &= \sum_{k=1}^{\infty} \pi_{(i-1,c-1+k)}(\alpha) \lambda_1 \mathbb{E}_{(i,c-1+k)} \left[\int_0^{\tau_{U_i^c}} e^{-\alpha t} \mathbb{1}\{X(t) = (i, c - 1 + j)\} dt \right] \\ &\quad + \pi_{(i,c-1)}(\alpha) \lambda_2 \mathbb{E}_{(i,c)} \left[\int_0^{\tau_{U_i^c}} e^{-\alpha t} \mathbb{1}\{X(t) = (i, c - 1 + j)\} dt \right]. \end{aligned} \tag{3.6}$$

The expectation

$$\mathbb{E}_{(i,c-1+k)} \left[\int_0^{\tau_{U_i^c}} e^{-\alpha t} \mathbb{1}\{X(t) = (i, c - 1 + j)\} dt \right], \quad j, k \geq 1, \quad (3.7)$$

has the same interpretation as the expectation in (3.3), except now the $M/M/1$ queue starts with k customers at time 0. Using Lemma 6 of Appendix C, we obtain

$$\lambda_1 \mathbb{E}_{(i,c-1+k)} \left[\int_0^{\tau_{U_i^c}} e^{-\alpha t} \mathbb{1}\{X(t) = (i, c - 1 + j)\} dt \right] = \Upsilon(j, k), \quad j, k \geq 1, \quad (3.8)$$

where, for $j, k \geq 1$,

$$\Upsilon(j, k) := \begin{cases} \lambda_1 \Omega_2 r_2^{j-k} (1 - (r_2 \phi_2)^k), & 1 \leq k \leq j - 1, \\ \lambda_1 \Omega_2 \phi_2^{k-j} (1 - (r_2 \phi_2)^j), & k \geq j. \end{cases} \quad (3.9)$$

Substituting (3.8) into (3.6) and simplifying yields a recursion. Specifically, for $i \geq 0$ and $j \geq 1$,

$$\pi_{(i+1,c-1+j)}(\alpha) = r_2^j \pi_{(i+1,c-1)}(\alpha) + \sum_{k=1}^{\infty} \Upsilon(j, k) \pi_{(i,c-1+k)}(\alpha), \quad (3.10)$$

with initial conditions $\pi_{(0,c-1+j)}(\alpha) = \pi_{(0,c-1)}(\alpha) r_2^j, j \geq 1$.

The recursion (3.10) can be solved, i.e., $\pi_{(i,c-1+j)}(\alpha)$ can be expressed in terms of $\pi_{(0,c-1)}(\alpha), \pi_{(1,c-1)}(\alpha), \dots, \pi_{(i,c-1)}(\alpha)$.

Theorem 2 (Interior) For $i \geq 0, j \geq 1$,

$$\pi_{(i,c-1+j)}(\alpha) = \sum_{k=0}^i v_{i,k} (1 - V_2)^k \binom{j-1+k}{k} r_2^j, \quad (3.11)$$

where the quantities $\{v_{i,j}\}_{i \geq j \geq 0}$ satisfy the following recursive scheme: for $i \geq 0$,

$$v_{i+1,0} = \pi_{(i+1,c-1)}(\alpha), \quad (3.12a)$$

$$v_{i+1,j} = V_1 \left(v_{i,j-1} + V_2 \sum_{k=j}^i v_{i,k} \right), \quad 1 \leq j \leq i + 1, \quad (3.12b)$$

with initial condition $v_{0,0} = \pi_{(0,c-1)}(\alpha)$. Here $V_1 = \frac{\lambda_1 \Omega_2}{1 - r_2 \phi_2}$, and $V_2 = r_2 \phi_2$.

Throughout we follow the convention that all empty sums, such as $\sum_{k=1}^0 (\cdot)$, represent the number zero.

Proof Clearly, when $i = 0$, (3.11) agrees with (3.5). Proceeding by induction, assume (3.11) holds among upper levels U_0, U_1, \dots, U_i for some $i \geq 0$. Substituting (3.11) into (3.10) yields

$$\pi_{(i+1,c-1+j)}(\alpha) = r_2^j \pi_{(i+1,c-1)}(\alpha) + \sum_{k=1}^{\infty} \mathcal{R}(j, k) \sum_{l=0}^i v_{i,l} (1 - V_2)^l \binom{k-1+l}{l} r_2^k. \tag{3.13}$$

Next, interchange the order of the two summations and apply Lemma 1 of Appendix A to get

$$\begin{aligned} (3.13) &= r_2^j \pi_{(i+1,c-1)}(\alpha) + \sum_{l=0}^i v_{i,l} (1 - V_2)^l \lambda_1 \Omega_2 \binom{j-1+l+1}{l+1} r_2^j \\ &\quad + \sum_{l=0}^i v_{i,l} (1 - V_2)^l V_1 V_2 \sum_{m=1}^l \frac{1}{(1 - V_2)^{l-m}} \binom{j-1+m}{m} r_2^j. \end{aligned} \tag{3.14}$$

To further simplify the right-hand side of (3.14), increase the summation index of the first summation by one by setting $k = l + 1$, multiply its summands by $\frac{1-r_2\phi_2}{1-r_2\phi_2}$, and change the order of the double summation. This yields

$$\begin{aligned} (3.14) &= r_2^j \pi_{(i+1,c-1)}(\alpha) + \sum_{k=1}^{i+1} V_1 v_{i,k-1} (1 - V_2)^k \binom{j-1+k}{k} r_2^j \\ &\quad + \sum_{k=1}^i V_1 V_2 \sum_{l=k}^i v_{i,l} (1 - V_2)^k \binom{j-1+k}{k} r_2^j \\ &= r_2^j \pi_{(i+1,c-1)}(\alpha) \\ &\quad + \sum_{k=1}^{i+1} V_1 \left(v_{i,k-1} + V_2 \sum_{l=k}^i v_{i,l} \right) (1 - V_2)^k \binom{j-1+k}{k} r_2^j, \end{aligned} \tag{3.15}$$

which shows $\pi_{(i+1,c-1+j)}(\alpha)$ satisfies (3.11), completing the induction step. □

4 Deriving an explicit expression for $v_{i,j}$

Theorem 2 suggests that the Laplace transforms for the states within the interior can be computed recursively:

1. *Initialization step* Determine $\pi_{(0,c-1)}(\alpha)$, which yields each Laplace transform $\pi_{(0,c-1+j)}(\alpha)$, $j \geq 1$.
2. *Recursive step on i* Given $\pi_{(k,c-1)}(\alpha)$ for $0 \leq k \leq i$ and the coefficients $\{v_{k,l}\}_{i \geq k \geq l \geq 0}$
 - a. Compute $\pi_{(i+1,c-1)}(\alpha)$.
 - b. Compute $\{v_{i+1,l}\}_{i+1 \geq l \geq 0}$.

c. Once steps 2a. and 2b. are completed, all Laplace transform values $\pi_{(i+1,c-1+j)}(\alpha)$, $j \geq 1$, are known.

Our next result, i.e., Theorem 3, shows that for each $i \geq 0$, the $\{v_{i,l}\}_{i \geq l \geq 0}$ terms can be expressed explicitly in terms of $\pi_{(k,c-1)}(\alpha)$, $0 \leq k \leq i$. If our goal is to only compute $\pi_{(i,c-1+j)}(\alpha)$ for some large i , then Theorem 3 allows us to avoid computing all intermediate $\{v_{k,l}\}_{i-1 \geq k \geq l \geq 0}$ terms, which means we can avoid computing an additional $O(i^2)$ terms. Not only that, knowing exactly how these $v_{i,j}$ coefficients look could aid in future questions asked by researchers interested in the $M/M/c$ 2-class priority queue.

Readers should keep in mind that the expressions we have derived for the $v_{i,j}$ coefficients do contain binomial coefficients, and one should be careful to avoid roundoff errors while computing these expressions.

Theorem 3 (Coefficients) *The coefficients $\{v_{i,j}\}_{i \geq j \geq 0}$ from Theorem 2 are as follows: for $i \geq j \geq 0$,*

$$v_{i,j} = V_1^j \pi_{(i-j,c-1)}(\alpha) + \sum_{k=j+1}^i V_1^k \pi_{(i-k,c-1)}(\alpha) \sum_{l=1}^{k-j} \frac{j}{k-j} \binom{k-j}{l} \binom{k-1}{l-1} V_2^l. \tag{4.1}$$

Proof From (3.12) we find, for each $i \geq 0$, that $v_{i,0} = \pi_{(i,c-1)}(\alpha)$ and $v_{i,i} = V_1^i \pi_{(0,c-1)}(\alpha)$; these expressions agree with (4.1).

Next, assume for some integer $i \geq 0$ that $v_{i,j}$ satisfies (4.1) for $0 \leq j \leq i$. Our aim is to show $v_{i+1,j}$ also satisfies (4.1) for $0 \leq j \leq i+1$: We do this by substituting (4.1) into (3.12b) and simplifying. There are three cases to consider: (i) $j = 1$; (ii) $2 \leq j \leq i-1$; and (iii) $j = i$. We focus on case (ii), with cases (i) and (iii) following similarly.

We first examine the $V_1 v_{i,j-1}$ term in (3.12b) by substituting (4.1). Here,

$$\begin{aligned} V_1 v_{i,j-1} &= V_1^j \pi_{(i+1-j,c-1)}(\alpha) \\ &+ \sum_{k=j}^i V_1^{k+1} \pi_{(i-k,c-1)}(\alpha) \sum_{l=1}^{k+1-j} \frac{j-1}{k+1-j} \binom{k+1-j}{l} \binom{k-1}{l-1} V_2^l \\ &= V_1^j \pi_{(i+1-j,c-1)}(\alpha) \\ &+ \sum_{k=j+1}^{i+1} V_1^k \pi_{(i+1-k,c-1)}(\alpha) \sum_{l=1}^{k-j} \frac{j-1}{k-j} \binom{k-j}{l} \binom{k-2}{l-1} V_2^l. \end{aligned} \tag{4.2}$$

Next, write (4.2) in a form where the binomial coefficients match the ones in (4.1):

$$(4.2) = V_1^j \pi_{(i+1-j,c-1)}(\alpha) + \sum_{k=j+1}^{i+1} V_1^k \pi_{(i+1-k,c-1)}(\alpha) \sum_{l=1}^{k-j} \frac{(j-1)(k-l)}{(k-j)(k-1)} \binom{k-j}{l} \binom{k-1}{l-1} V_2^l. \tag{4.3}$$

The remaining terms on the right-hand side of (3.12b) can be further simplified by substituting (4.1). Doing so reveals that

$$\begin{aligned}
 V_1 V_2 \sum_{k=j}^i v_{i,k} &= \sum_{k=j}^i V_1^{k+1} \pi_{(i-k,c-1)}(\alpha) V_2 \\
 &+ \sum_{k=j}^{i-1} \sum_{l=k+1}^i V_1^{l+1} \pi_{(i-l,c-1)}(\alpha) \sum_{m=1}^{l-k} \frac{k}{l-k} \binom{l-k}{m} \binom{l-1}{m-1} V_2^{m+1}.
 \end{aligned}
 \tag{4.4}$$

Swapping the order of the triple summation in (4.4) gives

$$\begin{aligned}
 (4.4) &= \sum_{k=j}^i V_1^{k+1} \pi_{(i-k,c-1)}(\alpha) V_2 \\
 &+ \sum_{l=j+1}^i V_1^{l+1} \pi_{(i-l,c-1)}(\alpha) \sum_{m=1}^{l-j} \sum_{k=j}^{l-m} \frac{k}{l-k} \binom{l-k}{m} \binom{l-1}{m-1} V_2^{m+1}.
 \end{aligned}
 \tag{4.5}$$

The inner-most summation over k of (4.5) can be evaluated using Lemma 3 of Appendix A:

$$\sum_{k=j}^{l-m} \frac{k}{l-k} \binom{l-k}{m} = \frac{l-m+jm}{m(l+1-j)} \binom{l+1-j}{m+1}.
 \tag{4.6}$$

Next, substitute (4.6) back into (4.5) and focus on the inner-most double summation of (4.5). This gives

$$\begin{aligned}
 &\sum_{m=1}^{l-j} \sum_{k=j}^{l-m} \frac{k}{l-k} \binom{l-k}{m} \binom{l-1}{m-1} V_2^{m+1} \\
 &= \sum_{m=1}^{l-j} \frac{l-m+jm}{m(l+1-j)} \binom{l+1-j}{m+1} \binom{l-1}{m-1} V_2^{m+1} \\
 &= \sum_{m=1}^{l-j} \frac{l-m+jm}{(l+1-j)l} \binom{l+1-j}{m+1} \binom{l}{m} V_2^{m+1} \\
 &= \sum_{m=2}^{l+1-j} \frac{l+1+jm-j-m}{(l+1-j)l} \binom{l+1-j}{m} \binom{l}{m-1} V_2^m.
 \end{aligned}
 \tag{4.7}$$

Substituting (4.7) into (4.5) and changing the two outer summation indices shows

$$\begin{aligned}
 (4.5) &= \sum_{l=j+1}^{i+1} V_1^l \pi_{(i+1-l, c-1)}(\alpha) V_2 \\
 &+ \sum_{l=j+2}^{i+1} V_1^l \pi_{(i+1-l, c-1)}(\alpha) \sum_{m=2}^{l-j} \frac{l+jm-j-m}{(l-j)(l-1)} \binom{l-j}{m} \binom{l-1}{m-1} V_2^m.
 \end{aligned}
 \tag{4.8}$$

Furthermore, since

$$V_2 = \frac{l+j \cdot 1 - j - 1}{(l-j)(l-1)} \binom{l-j}{1} \binom{l-1}{1-1} V_2^1,
 \tag{4.9}$$

we can merge the single summation with the double summation in (4.8). In other words,

$$(4.8) = \sum_{l=j+1}^{i+1} V_1^l \pi_{(i+1-l, c-1)}(\alpha) \sum_{m=1}^{l-j} \frac{l+jm-j-m}{(l-j)(l-1)} \binom{l-j}{m} \binom{l-1}{m-1} V_2^m.
 \tag{4.10}$$

Finally, summing (4.3) and (4.10) produces (4.1), as

$$\begin{aligned}
 &V_1 v_{i, j-1} + V_1 V_2 \sum_{k=j}^i v_{i, k} \\
 &= V_1^j \pi_{(i+1-j, c-1)}(\alpha) \\
 &+ \sum_{k=j+1}^{i+1} V_1^k \pi_{(i+1-k, c-1)}(\alpha) \sum_{l=1}^{k-j} \frac{(j-1)(k-l)}{(k-j)(k-1)} \binom{k-j}{l} \binom{k-1}{l-1} V_2^l \\
 &+ \sum_{k=j+1}^{i+1} V_1^k \pi_{(i+1-k, c-1)}(\alpha) \sum_{l=1}^{k-j} \frac{k+jl-j-l}{(k-j)(k-1)} \binom{k-j}{l} \binom{k-1}{l-1} V_2^l.
 \end{aligned}
 \tag{4.11}$$

Summing the coefficients in front of the binomial coefficient terms proves case (ii). Cases (i) and (iii) follow similarly. □

We now have an explicit expression for the coefficients. Substitute the expressions for $\{v_{i, j}\}_{i \geq j \geq 0}$ into (3.11) to obtain, for $j \geq 1$,

$$\begin{aligned}
 \pi_{(i,c-1+j)}(\alpha) &= \sum_{k=0}^i \pi_{(i-k,c-1)}(\alpha) (V_1(1-V_2))^k \binom{j-1+k}{k} r_2^j \\
 &+ \sum_{k=0}^{i-1} \sum_{l=k+1}^i \pi_{(i-l,c-1)}(\alpha) V_1^l \\
 &\cdot \sum_{m=1}^{l-k} \frac{k}{l-k} \binom{l-k}{m} \binom{l-1}{m-1} V_2^m (1-V_2)^k \binom{j-1+k}{k} r_2^j. \tag{4.12}
 \end{aligned}$$

Swapping the order of the double summation and grouping coefficients in front of each Laplace transform reveals the dependence of $\pi_{(i,c-1+j)}(\alpha)$ on the Laplace transforms $\pi_{(0,c-1)}(\alpha), \pi_{(1,c-1)}(\alpha), \dots, \pi_{(i,c-1)}(\alpha)$:

$$\begin{aligned}
 \pi_{(i,c-1+j)}(\alpha) &= r_2^j \pi_{(i,c-1)}(\alpha) + \sum_{l=1}^i V_1^l \pi_{(i-l,c-1)}(\alpha) \left[(1-V_2)^l \binom{j-1+l}{l} r_2^j \right. \\
 &+ \left. \sum_{k=0}^{l-1} (1-V_2)^k \binom{j-1+k}{k} r_2^j \sum_{m=1}^{l-k} \frac{k}{l-k} \binom{l-k}{m} \binom{l-1}{m-1} V_2^m \right]. \tag{4.13}
 \end{aligned}$$

From this expression, we see that for each fixed $i \geq 0$, as $j \rightarrow \infty, \pi_{(i,c-1+j)}(\alpha)$ behaves in a manner analogous to that found in Theorem 3.1 of [23], which addresses, when $c = 1$, the asymptotic behavior of the stationary distribution as the number of high-priority customers approaches infinity, while the number of low-priority customers is fixed.

The explicit expression (4.13) can be used to obtain an expression for the Laplace transforms of the number of class-1 customers in the system. That is,

$$\begin{aligned}
 &\int_0^\infty e^{-\alpha t} \mathbb{P}(X_1(t) = i \mid X(0) = (0, 0)) dt \\
 &= \int_0^\infty e^{-\alpha t} \sum_{j=0}^\infty \mathbb{P}(X(t) = (i, j) \mid X(0) = (0, 0)) dt \\
 &= \sum_{j=0}^\infty \pi_{(i,j)}(\alpha) = \sum_{j=0}^{c-1} \pi_{(i,j)}(\alpha) + \sum_{j=1}^\infty \pi_{(i,c-1+j)}(\alpha), \tag{4.14}
 \end{aligned}$$

where we can simplify the final infinite sum as

$$\begin{aligned}
 \sum_{j=1}^\infty \pi_{(i,c-1+j)}(\alpha) &= \frac{r_2}{1-r_2} \pi_{(i,c-1)}(\alpha) + \sum_{l=1}^i V_1^l \pi_{(i-l,c-1)}(\alpha) \left[\frac{(1-V_2)^l r_2}{(1-r_2)^{l+1}} \right. \\
 &+ \left. \sum_{k=0}^{l-1} \frac{(1-V_2)^k r_2}{(1-r_2)^{k+1}} \sum_{m=1}^{l-k} \frac{k}{l-k} \binom{l-k}{m} \binom{l-1}{m-1} V_2^m \right], \tag{4.15}
 \end{aligned}$$

via the identity

$$\sum_{j=1}^{\infty} \binom{j-1+k}{k} r_2^j = \frac{r_2}{(1-r_2)^{k+1}}. \tag{4.16}$$

5 Laplace transforms for states in the horizontal boundary

In the previous section, we showed how to express each Laplace transform for the states on the vertical boundary and within the interior explicitly in terms of transforms for the states in the horizontal boundary. So, it remains to determine the transforms for the states in the horizontal boundary. In this section, we show that the latter Laplace transforms satisfy a variant of Ramaswami’s formula, which will allow us to numerically compute these transforms recursively.

The approach we use to compute the above-mentioned variant of Ramaswami’s formula makes, like the CAP method, repeated use of Theorem 1. This approach is highly analogous to the approach used in [17] to study block-structured Markov processes, yet slightly modified since we are interested in recursively computing Laplace transforms only associated with states within the horizontal boundary. This idea of restricting ourselves to a subset of the state space seems similar in spirit to the censoring approach featured in the work of Li and Zhao [22], but it is not currently obvious to the authors if their approach is applicable to our setting.

We first introduce some relevant notation. Define the $1 \times c$ row vectors $\boldsymbol{\pi}_i(\alpha)$ as

$$\boldsymbol{\pi}_i(\alpha) := [\pi_{(i,0)}(\alpha) \ \pi_{(i,1)}(\alpha) \ \cdots \ \pi_{(i,c-1)}(\alpha)], \quad i \geq 0. \tag{5.1}$$

To properly state the Ramaswami-like formula satisfied by these row vectors, we need to define additional matrices. First, we define the $c \times c$ transition rate submatrices corresponding to lower levels L_i , $i \geq c$, as $\mathbf{A}_1 := \lambda_1 \mathbf{I}$, $\mathbf{A}_{-1} := \text{diag}(c\mu_1, (c-1)\mu_1, \dots, \mu_1)$ and

$$\mathbf{A}_0 := \begin{bmatrix} -\lambda_2 & \lambda_2 & & & & \\ \mu_2 & -(\lambda_2 + \mu_2) & \lambda_2 & & & \\ & 2\mu_2 & -(\lambda_2 + 2\mu_2) & \lambda_2 & & \\ & & & \ddots & & \\ & & & & (c-1)\mu_2 & -(\lambda_2 + (c-1)\mu_2) \end{bmatrix} = \mathbf{A}_1 - \mathbf{A}_{-1}, \tag{5.2}$$

where \mathbf{I} is the $c \times c$ identity matrix and $\text{diag}(\mathbf{x})$ is a square matrix with the vector \mathbf{x} along its main diagonal. We further define the $c \times c$ level-dependent transition rate submatrices associated with L_i , $1 \leq i \leq c-1$, as $\mathbf{A}_{-1}^{(i)} := \text{diag}(\mathbf{x}^{(i)})$ with $(\mathbf{x}^{(i)})_j := \min(i, c-j)\mu_1$, $0 \leq j \leq c-1$, $\mathbf{A}_0^{(i)} := \mathbf{A}_0 + \mathbf{A}_{-1} - \mathbf{A}_{-1}^{(i)}$, and for L_0 we have $\mathbf{A}_0^{(0)} := \mathbf{A}_0 + \mathbf{A}_{-1}$.

Next, we define the collection of $c \times c$ matrices $\{\mathbf{W}_m(\alpha)\}_{m \geq 0}$. Each element of $\mathbf{W}_m(\alpha)$ is equal to 0 except for element $(\mathbf{W}_m(\alpha))_{c-1, c-1}$, which is defined as $(\mathbf{W}_m(\alpha))_{c-1, c-1} := \lambda_2 w_m^{(\lambda_2, c\mu_2, \lambda_1)}(\alpha)$.

We also need the collection of $c \times c$ matrices $\{\mathbf{G}_{i,j}(\alpha)\}_{i > j \geq 0}$, where the (k, l) -th element of $\mathbf{G}_{i,j}(\alpha)$ is defined as

$$(\mathbf{G}_{i,j}(\alpha))_{k,l} := \mathbb{E}_{(i,k)}[e^{-\alpha \tau_{L_j}} \mathbb{1}\{X(\tau_{L_j}) = (j, l)\}], \quad 0 \leq k, l \leq c - 1. \tag{5.3}$$

Finally, we will need the collection of $c \times c$ matrices $\{\mathbf{N}_i(\alpha)\}_{i \geq 1}$, whose elements are defined as follows:

$$(\mathbf{N}_i(\alpha))_{k,l} := \mathbb{E}_{(i,k)}\left[\int_0^{\tau_{L_{i-1}}} e^{-\alpha t} \mathbb{1}\{X(t) = (i, l)\} dt\right], \quad 0 \leq k, l \leq c - 1. \tag{5.4}$$

Note that $\mathbf{N}_i(\alpha) = \mathbf{N}_c(\alpha)$ for $i \geq c$, and we therefore denote $\mathbf{N}(\alpha) := \mathbf{N}_c(\alpha)$.

5.1 A Ramaswami-like recursion

The following theorem shows that the vectors of transforms $\{\pi_i(\alpha)\}_{i \geq 0}$ satisfy a recursion analogous to Ramaswami’s formula [25].

Theorem 4 (Horizontal boundary) *For each integer $i \geq 0$, we have*

$$\begin{aligned} \pi_{i+1}(\alpha) &= \pi_i(\alpha) \mathbf{A}_1 \mathbf{N}_{i+1}(\alpha) \\ &\quad + \sum_{k=0}^i \pi_k(\alpha) \sum_{l=i+1}^{\infty} \mathbf{W}_{l-k}(\alpha) \mathbf{G}_{l,i+1}(\alpha) \mathbf{N}_{i+1}(\alpha), \end{aligned} \tag{5.5}$$

where we use the convention $\mathbf{G}_{i+1, i+1}(\alpha) = \mathbf{I}$.

Proof This result can be proven by making use of the approach found in [17]. Using Theorem 1 with $A = L_{\leq i}$, we see that for $i \geq 0$ and $0 \leq j \leq c - 1$,

$$\pi_{(i+1, j)}(\alpha) = \sum_{z \in L_{\leq i}} \pi_z(\alpha) \sum_{z' \in L_{\leq i}^c} q(z, z') \mathbb{E}_{z'}\left[\int_0^{\tau_{L_{\leq i}}} e^{-\alpha t} \mathbb{1}\{X(t) = (i + 1, j)\} dt\right]. \tag{5.6}$$

Due to the structure of the transition rates, many terms in the summation of (5.6) are zero. In particular, (5.6) can be stated more explicitly as

$$\begin{aligned} \pi_{(i+1, j)}(\alpha) &= \sum_{k=0}^i \pi_{(k, c-1)}(\alpha) \lambda_2 \mathbb{E}_{(k, c)}\left[\int_0^{\tau_{L_{\leq i}}} e^{-\alpha t} \mathbb{1}\{X(t) = (i + 1, j)\} dt\right] \\ &\quad + \sum_{m=0}^{c-1} \pi_{(i, m)}(\alpha) \lambda_1 \mathbb{E}_{(i+1, m)}\left[\int_0^{\tau_{L_{\leq i}}} e^{-\alpha t} \mathbb{1}\{X(t) = (i + 1, j)\} dt\right]. \end{aligned} \tag{5.7}$$

We now simplify each expectation appearing within the first sum on the right-hand side of (5.7). Summing over all ways in which the process reaches phase $c - 1$ again yields

$$\begin{aligned} & \mathbb{E}_{(k,c)} \left[\int_0^{\tau_{L \leq i}} e^{-\alpha t} \mathbb{1}\{X(t) = (i + 1, j)\} dt \right] \\ &= \mathbb{E}_{(k,c)} \left[\int_{\tau_{P_{c-1}}}^{\tau_{L \leq i}} e^{-\alpha t} \mathbb{1}\{X(t) = (i + 1, j)\} dt \right] \\ &= \sum_{l=i+1}^{\infty} \mathbb{E}_{(k,c)} \left[\mathbb{1}\{X(\tau_{P_{c-1}}) = (l, c - 1)\} e^{-\alpha \tau_{P_{c-1}}} \right. \\ & \quad \left. \cdot \int_{\tau_{P_{c-1}}}^{\tau_{L \leq i}} e^{-\alpha(t - \tau_{P_{c-1}})} \mathbb{1}\{X(t) = (i + 1, j)\} dt \right]. \end{aligned} \tag{5.8}$$

Applying the strong Markov property to each expectation appearing in (5.8) shows that

$$\begin{aligned} (5.8) &= \sum_{l=i+1}^{\infty} \mathbb{E}_{(k,c)} \left[\mathbb{1}\{X(\tau_{P_{c-1}}) = (l, c - 1)\} e^{-\alpha \tau_{P_{c-1}}} \right] \\ & \quad \cdot \mathbb{E}_{(l,c-1)} \left[\int_0^{\tau_{L \leq i}} e^{-\alpha t} \mathbb{1}\{X(t) = (i + 1, j)\} dt \right] \\ &= \sum_{l=i+1}^{\infty} w_{l-k}^{(\lambda_2, c\mu_2, \lambda_1)}(\alpha) (\mathbf{G}_{l,i+1}(\alpha) \mathbf{N}_{i+1}(\alpha))_{c-1,j}, \end{aligned} \tag{5.9}$$

where the last equality follows from the definitions of $\mathbf{G}_{l,i+1}(\alpha)$ and $\mathbf{N}_{i+1}(\alpha)$, and Lemma 8 of Appendix C. The expectations appearing within the second sum of (5.7) can easily be simplified by recognizing that they are elements of $\mathbf{N}_{i+1}(\alpha)$. Hence, we ultimately obtain

$$\begin{aligned} \pi_{(i+1,j)}(\alpha) &= \sum_{k=0}^i \pi_{(k,c-1)}(\alpha) \sum_{l=i+1}^{\infty} \lambda_2 w_{l-k}^{(\lambda_2, c\mu_2, \lambda_1)}(\alpha) (\mathbf{G}_{l,i+1}(\alpha) \mathbf{N}_{i+1}(\alpha))_{c-1,j} \\ & \quad + \sum_{m=0}^{c-1} \pi_{(i,m)}(\alpha) (\mathbf{A}_1)_{m,m} (\mathbf{N}_{i+1}(\alpha))_{m,j}, \end{aligned} \tag{5.10}$$

which, in matrix form, is (5.5). □

It remains to derive computable representations of $\{\mathbf{G}_{i,j}(\alpha)\}_{i>j \geq 0}$, as well as the matrices $\{\mathbf{N}_i(\alpha)\}_{1 \leq i \leq c}$.

5.2 Computing the $\mathbf{G}_{i,j}(\alpha)$ matrices

The next proposition shows that each $\mathbf{G}_{i,j}(\alpha)$ matrix can be expressed entirely in terms of the subset $\{\mathbf{G}_{i+1,i}(\alpha)\}_{0 \leq i \leq c-1}$.

Proposition 1 For each pair of integers i, j satisfying $i > j \geq 0$, we have

$$\mathbf{G}_{i,j}(\alpha) = \mathbf{G}_{i,i-1}(\alpha)\mathbf{G}_{i-1,i-2}(\alpha) \cdots \mathbf{G}_{j+1,j}(\alpha). \tag{5.11}$$

Furthermore, for each integer $k \geq 0$ we also have

$$\mathbf{G}_{c+k,c-1+k}(\alpha) = \mathbf{G}_{c,c-1}(\alpha). \tag{5.12}$$

Proof Equation (5.11) can be derived by applying the strong Markov property in an iterative manner, while (5.12) follows from the homogeneous structure of X along all lower levels $L_i, i \geq c$. \square

In light of Proposition 1, our goal now is to determine the subset of matrices $\{\mathbf{G}_{i+1,i}(\alpha)\}_{0 \leq i \leq c-1}$. We first focus on showing that $\mathbf{G}(\alpha) := \mathbf{G}_{c,c-1}(\alpha)$ is the solution to a fixed-point equation.

Proposition 2 The matrix $\mathbf{G}(\alpha)$ satisfies

$$\mathbf{G}(\alpha) = (\alpha\mathbf{I} - \mathbf{A}_0 - \mathbf{W}_0(\alpha))^{-1} \left(\mathbf{A}_{-1} + \mathbf{A}_1\mathbf{G}(\alpha)^2 + \sum_{l=1}^{\infty} \mathbf{W}_l(\alpha)\mathbf{G}(\alpha)^{l+1} \right). \tag{5.13}$$

Proof Assuming the matrix $\alpha\mathbf{I} - \mathbf{A}_0 - \mathbf{W}_0(\alpha)$ is invertible, Eq. (5.13) follows from a one-step analysis argument. To show that $\alpha\mathbf{I} - \mathbf{A}_0 - \mathbf{W}_0(\alpha)$ is indeed invertible, define the $c \times c$ matrix $\mathbf{H}(\alpha)$ with elements

$$(\mathbf{H}(\alpha))_{i,j} := \mathbb{E}_{(c,i)} \left[\int_0^{\tau_{(L_c \cup U_c)^c}} e^{-\alpha t} \mathbb{1}\{X(t) = (c, j)\} dt \right], \quad 0 \leq i, j \leq c-1, \tag{5.14}$$

and use again one-step analysis to establish

$$(\alpha\mathbf{I} - \mathbf{A}_0 - \mathbf{W}_0(\alpha))\mathbf{H}(\alpha) = \mathbf{I}, \tag{5.15}$$

which proves the claim. \square

The next proposition shows that through successive substitutions one can obtain $\mathbf{G}(\alpha)$ from (5.13).

Proposition 3 Suppose the sequence of matrices $\{\mathbf{Z}(n, \alpha)\}_{n \geq 0}$ satisfies the recursion

$$\begin{aligned} \mathbf{Z}(n+1, \alpha) &= (\alpha\mathbf{I} - \mathbf{A}_0 - \mathbf{W}_0(\alpha))^{-1} \\ &\cdot \left(\mathbf{A}_{-1} + \mathbf{A}_1\mathbf{Z}(n, \alpha)^2 + \sum_{l=1}^{\infty} \mathbf{W}_l(\alpha)\mathbf{Z}(n, \alpha)^{l+1} \right) \end{aligned} \tag{5.16}$$

with initial condition $\mathbf{Z}(0, \alpha) = \mathbf{0}$. Then,

$$\lim_{n \rightarrow \infty} \mathbf{Z}(n, \alpha) = \mathbf{G}(\alpha). \tag{5.17}$$

Proof This proof makes use of Proposition 2, and is completely analogous to the proofs of [18, Theorems 3.1 and 4.1] and [17, Theorem 3.4]. It is therefore omitted. □

Now that we have a method for approximating $\mathbf{G}(\alpha)$, it remains to find a method for computing $\mathbf{G}_{i+1,i}(\alpha)$, $0 \leq i \leq c - 2$. The next proposition shows that these matrices can be computed recursively.

Proposition 4 *For each integer i satisfying $0 \leq i \leq c - 2$, we have*

$$\begin{aligned} \mathbf{G}_{i+1,i}(\alpha) = & \left(\alpha \mathbf{I} - \mathbf{A}_0^{(i+1)} - \mathbf{A}_1 \mathbf{G}_{i+2,i+1}(\alpha) \right. \\ & \left. - \sum_{l=i+1}^{\infty} \mathbf{W}_{l-(i+1)}(\alpha) \mathbf{G}_{l,i+1}(\alpha) \right)^{-1} \mathbf{A}_{-1}^{(i+1)}. \end{aligned} \tag{5.18}$$

Proof Similar to Proposition 2, (5.18) follows by a one-step analysis. The inverse in (5.18) exists because the matrix $\mathbf{A}_{-1}^{(i+1)}$ is a diagonal matrix whose diagonal elements are all positive. □

We now have an iterative procedure for computing all $\{\mathbf{G}_{i+1,i}(\alpha)\}_{0 \leq i \leq c-1}$ matrices: first compute $\mathbf{G}(\alpha)$ from Proposition 3, then use Proposition 4 to compute $\mathbf{G}_{c-1,c-2}(\alpha)$, then $\mathbf{G}_{c-2,c-3}(\alpha)$, and so on, stopping at $\mathbf{G}_{1,0}(\alpha)$.

5.3 Computing the $\mathbf{N}_i(\alpha)$ matrices

The matrices $\{\mathbf{N}_i(\alpha)\}_{1 \leq i \leq c}$ can be expressed in terms of $\{\mathbf{G}_{i,j}(\alpha)\}_{i \geq j \geq 0}$.

Proposition 5 *For each integer i satisfying $1 \leq i \leq c$, we have*

$$\mathbf{N}_i(\alpha) = \left(\alpha \mathbf{I} - \mathbf{A}_0^{(i)} - \mathbf{A}_1 \mathbf{G}_{i+1,i}(\alpha) - \sum_{l=i}^{\infty} \mathbf{W}_{l-i}(\alpha) \mathbf{G}_{l,i}(\alpha) \right)^{-1}, \tag{5.19}$$

where we use the convention $\mathbf{A}_0^{(c)} = \mathbf{A}_0$.

Proof The proof uses one-step analysis and is analogous to the proofs of Propositions 2 and 4. It is therefore omitted. □

5.4 Computing $\pi_0(\alpha)$

It remains to devise a method for computing the vector $\pi_0(\alpha)$ so that the Ramaswami-like recursion from Theorem 4 can be properly initialized. The following is an adaptation of [17, Section 3.3]. We define the $c \times c$ matrix $\mathbf{N}_0(\alpha)$ whose elements are given by

$$(\mathbf{N}_0(\alpha))_{i,j} := \mathbb{E}_{(0,i)} \left[\int_0^{\tau(0,0)} e^{-\alpha t} \mathbb{1}\{X(t) = (0, j)\} dt \right], \quad 0 \leq i, j \leq c - 1. \tag{5.20}$$

In the derivation to follow, we require the notation $(\mathbf{A})^{[i,j]}$ which represents the matrix \mathbf{A} with row i and column j removed (meaning it is a $(c - 1) \times (c - 1)$ matrix), while keeping the indexing of entries exactly as in \mathbf{A} . Similarly, $(\mathbf{A})^{[i,\cdot]}$ has row i removed from \mathbf{A} (meaning it is a $(c - 1) \times c$ matrix) and $(\mathbf{A})^{[\cdot,j]}$ has column j removed from \mathbf{A} (meaning it is a $c \times (c - 1)$ matrix).

Proposition 6 *We have*

$$\begin{aligned}
 (\mathbf{N}_0(\alpha))^{[0,0]} &= \left(\alpha \mathbf{I}^{[0,0]} - (\mathbf{A}_0^{(0)})^{[0,0]} - (\mathbf{A}_1)^{[0,\cdot]} (\mathbf{G}_{1,0}(\alpha))^{[\cdot,0]} \right. \\
 &\quad \left. - \sum_{l=0}^{\infty} (\mathbf{W}_l(\alpha))^{[0,\cdot]} (\mathbf{G}_{l,0}(\alpha))^{[\cdot,0]} \right)^{-1}. \tag{5.21}
 \end{aligned}$$

Proof Similar to the proofs of Propositions 2, 4 and 5, we use a one-step analysis and the strong Markov property to prove the result. □

We employ (3.2) and Proposition 5 to determine the elements of the row vector $\pi_0(\alpha)$. Set $A = \{(0, 0)\}$ in Theorem 1: then, for $1 \leq j \leq c - 1$ and $c \geq 2$,

$$\begin{aligned}
 \pi_{(0,j)}(\alpha) &= \sum_{z \in A} \pi_z(\alpha) \sum_{z' \in A} q(z, z') \mathbb{E}_{z'} \left[\int_0^{\tau_A} e^{-\alpha t} \mathbb{1}\{X(t) = (0, j)\} dt \right] \\
 &= \pi_{(0,0)}(\alpha) \left(\lambda_1 \mathbb{E}_{(1,0)} \left[\int_0^{\tau(0,0)} e^{-\alpha t} \mathbb{1}\{X(t) = (0, j)\} dt \right] \right. \\
 &\quad \left. + \lambda_2 \mathbb{E}_{(0,1)} \left[\int_0^{\tau(0,0)} e^{-\alpha t} \mathbb{1}\{X(t) = (0, j)\} dt \right] \right) \\
 &= \pi_{(0,0)}(\alpha) \left(\lambda_1 \sum_{l=1}^{c-1} (\mathbf{G}_{1,0}(\alpha))_{0,l} (\mathbf{N}_0(\alpha))_{l,j} + \lambda_2 (\mathbf{N}_0(\alpha))_{1,j} \right). \tag{5.22}
 \end{aligned}$$

Hence, the transforms $\pi_{(0,j)}(\alpha)$, $1 \leq j \leq c - 1$, can be expressed in terms of $\pi_{(0,0)}(\alpha)$. In Sect. 5.5, we describe a numerical procedure to determine $\pi_{(0,0)}(\alpha)$.

Remark 1 (An alternative method for determining $\pi_0(\alpha)$) The Kolmogorov forward equations can also be used to derive $\pi_0(\alpha)$. The transition functions are known to satisfy these equations, since $\sup_{x \in \mathbb{S}} q(x) < \infty$. Taking the Laplace transform of the Kolmogorov forward equations for the states in L_0 yields, after using (3.5),

$$\pi_0(\alpha) (\alpha \mathbf{I} - \mathbf{A}_0^{(0)}) - \pi_0(\alpha) \mathbf{W}_0(\alpha) - \pi_1(\alpha) \mathbf{A}_{-1}^{(1)} = \mathbf{e}_0, \tag{5.23}$$

where \mathbf{e}_i is a row vector with all elements equal to zero except for the i -th element which is unity. Finally, using Theorem 4 to express $\pi_1(\alpha)$ in terms of $\pi_0(\alpha)$ yields

$$\begin{aligned}
 &-\pi_0(\alpha) \left(-\alpha \mathbf{I} + \mathbf{A}_0^{(0)} + \mathbf{W}_0(\alpha) \right. \\
 &\quad \left. + (\mathbf{A}_1 + \sum_{l=1}^{\infty} \mathbf{W}_l(\alpha) \mathbf{G}_{l,1}(\alpha)) \mathbf{N}_1(\alpha) \mathbf{A}_{-1}^{(1)} \right) = \mathbf{e}_0. \tag{5.24}
 \end{aligned}$$

If one is instead interested in the stationary distribution and in particular the stationary probabilities of L_0 , i.e., $\lim_{\alpha \downarrow 0} \alpha \boldsymbol{\pi}_0(\alpha)$, (5.24) results in a homogeneous system of equations, but it is not clear that this system still has a unique solution. On the other hand, the approach outlined earlier for determining $\boldsymbol{\pi}_0(\alpha)$ can straightforwardly be employed to obtain $\lim_{\alpha \downarrow 0} \alpha \boldsymbol{\pi}_0(\alpha)$.

5.5 Numerical implementation

In order to compute the vectors $\{\boldsymbol{\pi}_i(\alpha)\}_{i \geq 0}$, we first need to compute the matrices $\{\mathbf{G}_{i+1,i}(\alpha)\}_{0 \leq i \leq c-1}$, $\{\mathbf{N}_i(\alpha)\}_{1 \leq i \leq c}$, and $(\mathbf{N}_0(\alpha))^{[0,0]}$.

The first step is to compute $\mathbf{G}(\alpha) = \mathbf{G}_{c,c-1}(\alpha)$. Proposition 3 shows that this matrix can be approximated by using the recursion (5.16). Using this recursion requires us to truncate the infinite sum appearing within the recursion. One way of applying this truncation is as follows: given a fixed tolerance ϵ , pick an integer κ_ϵ large enough so that

$$\sum_{l=\kappa_\epsilon+1}^{\infty} |w_l^{(\lambda_2, c\mu_2, \lambda_1)}(\alpha)| \leq \epsilon/\lambda_2. \tag{5.25}$$

Once κ_ϵ has been found, we can use the approximation

$$\sum_{l=1}^{\kappa_\epsilon} \mathbf{W}_l(\alpha) \mathbf{Z}(n, \alpha)^{l+1} \approx \sum_{l=1}^{\infty} \mathbf{W}_l(\alpha) \mathbf{Z}(n, \alpha)^{l+1}, \tag{5.26}$$

since the modulus of each element of the matrix on the left-hand side of (5.26) can be shown to be within ϵ of what is being approximated. Here we used that the matrices $\mathbf{W}_l(\alpha)$ only have one element and the absolute value of each element of $\mathbf{Z}(n, \alpha)$ (and $\mathbf{G}(\alpha)$) is less than or equal to 1. Hence, we propose using the recursion

$$\begin{aligned} \mathbf{Z}(n+1, \alpha) &= (\alpha \mathbf{I} - \mathbf{A}_0 - \mathbf{W}_0(\alpha))^{-1} \\ &\cdot \left(\mathbf{A}_{-1} + \mathbf{A}_1 \mathbf{Z}(n, \alpha)^2 + \sum_{l=1}^{\kappa_\epsilon} \mathbf{W}_l(\alpha) \mathbf{Z}(n, \alpha)^{l+1} \right) \end{aligned} \tag{5.27}$$

to approximate $\mathbf{G}(\alpha)$. Notice that we can determine κ_ϵ satisfying (5.25) by writing the left-hand side of (5.25) as

$$\sum_{l=\kappa_\epsilon+1}^{\infty} |w_l^{(\lambda_2, c\mu_2, \lambda_1)}(\alpha)| = \sum_{l=1}^{\infty} |w_l^{(\lambda_2, c\mu_2, \lambda_1)}(\alpha)| - \sum_{l=1}^{\kappa_\epsilon} |w_l^{(\lambda_2, c\mu_2, \lambda_1)}(\alpha)|. \tag{5.28}$$

An explicit expression for the infinite sum on the right-hand side of (5.28) can be derived with the help of Lemma 8 of Appendix C and the generating function of the Catalan numbers; the finite sum can be computed numerically.

Once $\mathbf{G}(\alpha)$ has been found, we can use Proposition 4 to compute each $\mathbf{G}_{i+1,i}(\alpha)$ matrix, for $0 \leq i \leq c - 2$. For this computation, we use the same truncation procedure as outlined above.

The next step is to compute the matrices $\{\mathbf{N}_i(\alpha)\}_{1 \leq i \leq c}$ and $(\mathbf{N}_0(\alpha))^{[0,0]}$ using Propositions 5 and 6, respectively. For both computations, we again use the above truncation procedure.

It remains to recursively determine $\{\pi_i(\alpha)\}_{i \geq 0}$ using Theorem 4, where we again use the truncation procedure for the infinite sum. As we have seen in Sect. 5.4, this recursion should be properly initialized by the value of $\pi_{(0,0)}(\alpha)$. The random-product representation in Theorem 6 of Appendix B shows that all Laplace transforms $\pi_x(\alpha)$ satisfy, for each $x \in \mathbb{S}$,

$$\pi_x(\alpha) = \pi_{(0,0)}(\alpha)\psi_x(\alpha), \tag{5.29}$$

where $\pi_{(0,0)}(\alpha)$ is an unknown transform and $\psi_{(0,0)}(\alpha) = 1$. It is clear that the $\psi_x(\alpha)$ can be computed using the same procedure as for $\pi_x(\alpha)$, and (5.22) shows that $\psi_{(0,0)}(\alpha), \psi_{(0,1)}(\alpha), \dots, \psi_{(0,c-1)}(\alpha)$ are computable expressions.

We can calculate $\pi_{(0,0)}(\alpha)$ from the normalization condition

$$\sum_{x \in \mathbb{S}} \pi_x(\alpha) = \pi_{(0,0)}(\alpha) \sum_{x \in \mathbb{S}} \psi_x(\alpha) = \frac{1}{\alpha}, \tag{5.30}$$

which yields

$$\pi_{(0,0)}(\alpha) = \frac{1}{\alpha \sum_{x \in \mathbb{S}} \psi_x(\alpha)}. \tag{5.31}$$

Since we cannot compute this infinite sum, we determine $\psi_{(i,j)}(\alpha)$ for all (i, j) in a sufficiently large bounding box $\mathbb{S}_k := \{(i, j) \in \mathbb{S} : 0 \leq i \leq k\}$ for some $k \geq 0$. Notice that (4.15) allows \mathbb{S}_k to be an infinitely large rectangle. The choice of k in \mathbb{S}_k clearly influences the quality of the approximation. A simple procedure to choose k is the following. Define

$$\Psi_k := \sum_{x \in \mathbb{S}_k} \psi_x(\alpha). \tag{5.32}$$

Pick ϵ small and positive and continue increasing k until $\frac{|\Psi_{k+1} - \Psi_k|}{|\Psi_k|} < \epsilon$. Then, set $\pi_{(0,0)}(\alpha) = 1/(\alpha \Psi_{k+1})$ to normalize the Laplace transforms $\pi_x(\alpha) = \pi_{(0,0)}(\alpha)\psi_x(\alpha)$.

For $c = 1$ we can normalize the solution as outlined above, or we can explicitly determine the value of $\pi_{(0,0)}(\alpha)$; see the next section.

6 The single-server case

We now turn our attention to the case where $c = 1$, i.e., the case where the system consists of a single server. In this case, the analysis of the Laplace transforms $\pi_{(i,0)}(\alpha), i \geq 0$, simplifies considerably. The expressions for the Laplace transforms for the states in the interior and on the vertical boundary are identical to the multi-server case.

From [12, Corollary 2.1]—see also Theorem 6 in Appendix B—we have

$$\pi_{(0,0)}(\alpha) = \frac{1}{(q((0, 0)) + \alpha)(1 - \mathbb{E}_{(0,0)}[e^{-\alpha\tau_{(0,0)}}])}. \tag{6.1}$$

In light of (6.1), to evaluate $\pi_{(0,0)}(\alpha)$ the only thing that needs to be determined is the expectation $\mathbb{E}_{(0,0)}[e^{-\alpha\tau_{(0,0)}}]$. This quantity is the Laplace–Stieltjes transform of the sum of two independent exponential random variables: One is the exponential random variable E_λ having rate λ , the other is the busy period B of an $M/G/1$ queue having arrival rate λ and hyperexponential service times having cumulative distribution function $F(\cdot)$. More specifically,

$$F(t) = \frac{\lambda_1}{\lambda}(1 - e^{-\mu_1 t}) + \frac{\lambda_2}{\lambda}(1 - e^{-\mu_2 t}), \quad t \geq 0. \tag{6.2}$$

The Laplace–Stieltjes transform $\varphi(\alpha)$ of B is known to satisfy the Kendall functional equation

$$\varphi(\alpha) = \frac{\lambda_1}{\lambda} \frac{\mu_1}{\mu_1 + \alpha + \lambda(1 - \varphi(\alpha))} + \frac{\lambda_2}{\lambda} \frac{\mu_2}{\mu_2 + \alpha + \lambda(1 - \varphi(\alpha))}. \tag{6.3}$$

Furthermore, $\varphi(\alpha)$ can be determined numerically through successive substitutions of (6.3), starting with $\varphi(\alpha) = 0$; see [1, Section 1] for details. Using independence of E_λ and B ,

$$\mathbb{E}_{(0,0)}[e^{-\alpha\tau_{(0,0)}}] = \mathbb{E}[e^{-\alpha(E_\lambda+B)}] = \frac{\lambda}{\lambda + \alpha} \varphi(\alpha), \tag{6.4}$$

meaning that (see also [2, Eq. (36)]) for $c = 1$,

$$\pi_{(0,0)}(\alpha) = \frac{1}{\lambda(1 - \varphi(\alpha)) + \alpha}. \tag{6.5}$$

We now turn our attention to the horizontal boundary. When $c = 1$, the matrices $\mathbf{G}(\alpha)$ and $\mathbf{N}(\alpha)$ become scalars, which we denote as $G(\alpha)$ and $N(\alpha)$, respectively. More precisely,

$$G(\alpha) := \mathbb{E}_{(i+1,0)}[e^{-\alpha\tau_{Li}}], \quad N(\alpha) := \mathbb{E}_{(i,0)}\left[\int_0^{\tau_{Li-1}} e^{-\alpha t} \mathbb{1}\{X(t) = (i, 0)\} dt\right], \tag{6.6}$$

which are independent of $i \geq 1$.

Proposition 7 *The scalar $G(\alpha)$ is a solution to*

$$(\lambda + \mu_1 + \alpha)G(\alpha) = \mu_1 + \lambda_1 G(\alpha)^2 + \lambda_2 G(\alpha)\phi_{\lambda_2, \mu_2}(\lambda_1(1 - G(\alpha)) + \alpha). \tag{6.7}$$

Proof From Proposition 2, we easily find that when $c = 1$,

$$(\lambda + \mu_1 + \alpha)G(\alpha) = \mu_1 + \lambda_1 G(\alpha)^2 + \lambda_2 \sum_{l=0}^{\infty} w_l^{(\lambda_2, \mu_2, \lambda_1)}(\alpha) G(\alpha)^{l+1}. \tag{6.8}$$

The infinite series appearing in (6.8) can be simplified using Lemma 9 of Appendix C; doing so yields (6.7). \square

Even though we cannot use (6.7) to write down an explicit expression for $G(\alpha)$, we can still use it to devise an iterative scheme for computing $G(\alpha)$. The next result shows that $N(\alpha)$ can be expressed in terms of $G(\alpha)$.

Proposition 8 *We have*

$$N(\alpha) = \frac{1}{\alpha + \lambda + \mu_1 - \lambda_1 G(\alpha) - \lambda_2 \phi_{\lambda_2, \mu_2}(\lambda_1(1 - G(\alpha)) + \alpha)}. \quad (6.9)$$

Proof Using Proposition 5, we observe that when $c = 1$,

$$N(\alpha) = \left(\alpha + \lambda + \mu_1 - \lambda_1 G(\alpha) - \lambda_2 \sum_{l=0}^{\infty} w_l^{(\lambda_2, \mu_2, \lambda_1)}(\alpha) G(\alpha)^l \right)^{-1}. \quad (6.10)$$

The proof is completed by applying Lemma 9 of Appendix C to (6.10). \square

We now focus on the recursion for the horizontal boundary. When $c = 1$, Theorem 4 reduces to

$$\begin{aligned} \pi_{(i+1,0)}(\alpha) &= \lambda_1 \pi_{(i,0)}(\alpha) N(\alpha) \\ &+ \lambda_2 \sum_{k=0}^i \pi_{(k,0)}(\alpha) \sum_{l=i+1}^{\infty} w_{l-k}^{(\lambda_2, \mu_2, \lambda_1)}(\alpha) G(\alpha)^{l-(i+1)} N(\alpha). \end{aligned} \quad (6.11)$$

For the inner-most sum over l we have

$$\sum_{l=i+1}^{\infty} w_{l-k}^{(\lambda_2, \mu_2, \lambda_1)}(\alpha) G(\alpha)^{l-(i+1)} = G(\alpha)^{k-(i+1)} \sum_{m=i+1-k}^{\infty} w_m^{(\lambda_2, \mu_2, \lambda_1)}(\alpha) G(\alpha)^m. \quad (6.12)$$

Clearly, $i + 1 - k \geq 1$. So let us try to evaluate the tail of the generating function of $\{w_m^{(\lambda_2, \mu_2, \lambda_1)}(\alpha)\}_{m \geq 0}$ evaluated at the point $G(\alpha)$, i.e.,

$$\begin{aligned} &\sum_{m=i+1-k}^{\infty} w_m^{(\lambda_2, \mu_2, \lambda_1)}(\alpha) G(\alpha)^m \\ &= \phi_{\lambda_2, \mu_2}(\lambda_1(1 - G(\alpha)) + \alpha) - \sum_{m=0}^{i-k} w_m^{(\lambda_2, \mu_2, \lambda_1)}(\alpha) G(\alpha)^m, \end{aligned} \quad (6.13)$$

which follows from an application of Lemma 9 of Appendix C. The remaining finite summation is easy to compute since each $w_m^{(\lambda_2, \mu_2, \lambda_1)}(\alpha)$ term, by Lemma 8 of Appendix C, can be stated in terms of $b_K(\cdot)$ functions, and these satisfy the recursion found in Lemma 4 of Appendix A. These observations allow us to state the following

theorem, which yields a practical method for recursively computing Laplace transforms of the form $\pi_{(i,0)}(\alpha)$.

Theorem 5 (Horizontal boundary, single server) *When $c = 1$, the Laplace transforms of the transition functions on the horizontal boundary satisfy the following recursion: for $i \geq 0$,*

$$\begin{aligned} \pi_{(i+1,0)}(\alpha) &= \lambda_1 \pi_{(i,0)}(\alpha) N(\alpha) \\ &+ \lambda_2 \sum_{k=0}^i \pi_{(k,0)}(\alpha) G(\alpha)^{k-(i+1)} \left(\phi_{\lambda_2, \mu_2}(\lambda_1(1 - G(\alpha)) + \alpha) \right. \\ &\left. - \sum_{l=0}^{i-k} w_l^{(\lambda_2, \mu_2, \lambda_1)}(\alpha) G(\alpha)^l \right) N(\alpha). \end{aligned} \tag{6.14}$$

7 Conclusion

In this paper, we analyzed an $M/M/c$ priority system with two customer classes, class-dependent service rates and a preemptive resume priority rule. This queueing system can be modeled as a two-dimensional Markov process for which we analyzed the time-dependent behavior. More precisely, we obtained expressions for the Laplace transforms of the transition functions under the condition that the system is initially empty.

Using a slight modification of the CAP method, we showed that the Laplace transforms for the states with at least c high-priority customers can be expressed in terms of a finite sum of the Laplace transforms for the states with exactly $c - 1$ high-priority customers. This expression contained coefficients that satisfy a recursion. We solved this recursion to obtain an explicit expression for each coefficient. In doing so, each Laplace transform for the states on the vertical boundary and in the interior can easily be calculated from the values of the Laplace transforms for the states in the horizontal boundary.

Next, we developed a Ramaswami-like recursion for the Laplace transforms for the states in the horizontal boundary. The recursion required the collections of matrices $\{\mathbf{G}_{i+1,i}\}_{0 \leq i \leq c-1}$ and $\{\mathbf{N}_i(\alpha)\}_{1 \leq i \leq c}$ for which we showed that they can be determined iteratively. We demonstrated two ways in which the initial value of the recursion, i.e., the vector $\pi_0(\alpha)$, can be calculated. Finally, we discussed the numerical implementation of our approach for the horizontal boundary.

In the single-server case, the expressions for the Laplace transforms for the states on the vertical boundary and in the interior were identical to the multi-server case. The expressions for the horizontal boundary, however, simplified considerably. Specifically, the initial value $\pi_{(0,0)}(\alpha)$ of the recursion could be determined by comparing the queueing system to an $M/G/1$ queue with hyperexponentially distributed service times. Moreover, the calculation of $G(\alpha)$ and $N(\alpha)$, which are now scalars, simplified greatly.

We now comment on how our expressions for the Laplace transforms of the transitions functions can be used to determine the stationary distribution. It is clear from the

transition rate diagram in Fig. 3 that the Markov process X is irreducible. Moreover, it is well-known that X is positive recurrent if and only if $\rho < 1$. In that case, X has a unique stationary distribution $\mathbf{p} := [p_x]_{x \in \mathbb{S}}$. To compute each p_x term from $\pi_x(\alpha)$, simply note that

$$p_x = \lim_{\alpha \downarrow 0} \alpha \pi_x(\alpha). \tag{7.1}$$

Using this observation, we see that the procedure for finding \mathbf{p} is highly analogous to the one we presented for finding the Laplace transforms of the transition functions.

Acknowledgements The authors thank both the editor and the referees for their constructive comments on a previous version of our manuscript, as these helped us improve both the content and the readability of our paper. The first author is supported by a free competition Grant from the Netherlands Organisation for Scientific Research (NWO), while the second author gratefully acknowledges the support of the National Science Foundation (NSF), via Grant NSF-CMMI-1435261.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

Appendix A: Combinatorial identities

In this appendix, we collect some combinatorial identities that are used throughout the paper. These lemmas are likely known, but we prove them here to make the paper self-contained.

Lemma 1 For $j \geq 1, l \geq 0$ and $\Upsilon(j, k)$ defined in (3.9), we have

$$\begin{aligned} & \sum_{k=1}^{\infty} \Upsilon(j, k) \binom{k-1+l}{l} r_2^k \\ &= \lambda_1 \Omega_2 \binom{j-1+l+1}{l+1} r_2^j + \frac{\lambda_1 \Omega_2}{1-r_2 \phi_2} r_2 \phi_2 \sum_{m=1}^l \frac{1}{(1-r_2 \phi_2)^{l-m}} \binom{j-1+m}{m} r_2^j. \end{aligned} \tag{A.1}$$

Proof The result is nearly identical to [10, Lemma 1], so we omit its proof. □

Lemma 2 We have the identity

$$\sum_{k=0}^K \frac{(K-k+2l)!(k+2m)!}{(K-k)!k!} = (2l)!(2m)! \binom{K+2l+2m+1}{2l+2m+1}. \tag{A.2}$$

Proof Rewriting the fractions as two binomial coefficients,

$$(A.2) = (2l)!(2m)! \sum_{k=0}^K \binom{K-k+2l}{K-k} \binom{k+2m}{k}. \tag{A.3}$$

The rest of the proof follows from a direct application of [11, Chapter 2, Eq. (12.16)]. □

Lemma 3 *We have the identity*

$$\sum_{k=j}^{l-m} \frac{k}{l-k} \binom{l-k}{m} = \frac{l-m+jm}{m(l+1-j)} \binom{l+1-j}{m+1}. \tag{A.4}$$

Proof Change the summation variable to $n = l - k$ to get

$$\sum_{k=j}^{l-m} \frac{k}{l-k} \binom{l-k}{m} = \sum_{n=m}^{l-j} \frac{l-n}{n} \binom{n}{m}. \tag{A.5}$$

Splitting the summation and using the identity

$$\sum_{n=L}^U \binom{n}{L} = \binom{U+1}{L+1} \tag{A.6}$$

produces

$$\begin{aligned} \text{(A.5)} &= \sum_{n=m}^{l-j} \frac{l}{n} \binom{n}{m} - \sum_{n=m}^{l-j} \binom{n}{m} = \frac{l}{m} \sum_{n=m}^{l-j} \binom{n-1}{m-1} - \sum_{n=m}^{l-j} \binom{n}{m} \\ &= \frac{l}{m} \sum_{n=m-1}^{l-j-1} \binom{n}{m-1} - \sum_{n=m}^{l-j} \binom{n}{m} = \frac{l}{m} \binom{l-j}{m} - \binom{l+1-j}{m+1} \\ &= \frac{l(m+1)}{m(l+1-j)} \binom{l+1-j}{m+1} - \binom{l+1-j}{m+1}, \end{aligned} \tag{A.7}$$

proving the claim. □

Lemma 4 *Define*

$$b_K(z) := \sum_{k=0}^K C_k \binom{K+k}{K-k} z^k, \tag{A.8}$$

where $C_k := \frac{1}{k+1} \binom{2k}{k}$ are the Catalan numbers. The sequence $\{b_K(z)\}_{K \geq 0}$ satisfies two recursions. For $K \geq 2$ it satisfies

$$(K+1)b_K(z) = (2K-1)(1+2z)b_{K-1}(z) - (K-2)b_{K-2}(z), \tag{A.9}$$

alternatively, for $K \geq 0$, it satisfies

$$b_{K+1}(z) = b_K(z) + z \sum_{l=0}^K b_l(z) b_{K-l}(z) \tag{A.10}$$

with $b_0(z) = 1$ and $b_1(z) = 1 + z$.

Proof The terms $b_K(z)$ appear in [5, Section 3.3], where the authors derive (A.9). We believe that a small typographical error appears in their recursion that we have fixed here. To do so, in [5], substitute (23) into (24) to obtain the correct form of (16).

We now derive the recursion (A.10). Since we already have the explicit expression (A.8) for $b_K(z)$, we will substitute this into (A.10) and show that $b_{K+1}(z)$ again is given by (A.8).

Rewrite the first term on the right-hand side of (A.10) as

$$b_K(z) = \sum_{k=0}^K C_k \binom{K+k}{K-k} z^k = \sum_{k=0}^K \frac{K+1-k}{K+1+k} C_k \binom{K+1+k}{K+1-k} z^k. \tag{A.11}$$

Substituting (A.8) into the finite sum of (A.10) gives

$$\begin{aligned} z \sum_{l=0}^K b_l(z) b_{K-l}(z) &= z \sum_{l=0}^K \left(\sum_{k=0}^l C_k \binom{l+k}{l-k} z^k \right) \left(\sum_{m=0}^{K-l} C_m \binom{K-l+m}{K-l-m} z^m \right) \\ &= \sum_{l=0}^K \sum_{k=0}^l \sum_{m=0}^{K-l} C_k C_m \binom{l+k}{l-k} \binom{K-l+m}{K-l-m} z^{k+m+1}. \end{aligned} \tag{A.12}$$

Switch the order of the triple summation to obtain

$$\begin{aligned} \text{(A.12)} &= \sum_{k=0}^K \sum_{m=0}^{K-k} C_k C_m z^{k+m+1} \sum_{l=k}^{K-m} \binom{l+k}{l-k} \binom{K-l+m}{K-l-m} \\ &= \sum_{k=0}^K \sum_{m=0}^{K-k} C_k C_m z^{k+m+1} \sum_{n=0}^{K-k-m} \binom{n+2k}{n} \binom{K-k-m-n+2m}{K-k-m-n}, \end{aligned} \tag{A.13}$$

where we introduced $n = l - k$. Employing Lemma 2 for the inner-most summation results in

$$\text{(A.13)} = \sum_{k=0}^K \sum_{m=0}^{K-k} C_k C_m z^{k+m+1} \binom{K+1+k+m}{K-(k+m)}. \tag{A.14}$$

The double summation sums over all $k, m \geq 0$ such that $0 \leq k + m \leq K$. An equivalent summation is over the diagonals $k + m = d$ with $0 \leq d \leq K$ and $k, m \geq 0$:

$$\begin{aligned} \text{(A.14)} &= \sum_{d=0}^K \sum_{k,m \geq 0: k+m=d} C_k C_m z^{k+m+1} \binom{K+1+k+m}{K-(k+m)} \\ &= \sum_{d=0}^K \sum_{k=0}^d C_k C_{d-k} z^{d+1} \binom{K+1+d}{K-d}. \end{aligned} \tag{A.15}$$

Using the identity $C_{d+1} = \sum_{k=0}^d C_k C_{d-k}$, setting $k = d + 1$ and rewriting the binomial coefficient produces

$$\begin{aligned}
 \text{(A.15)} &= \sum_{d=0}^K C_{d+1} z^{d+1} \binom{K+1+d}{K-d} = \sum_{k=1}^{K+1} C_k \binom{K+k}{K+1-k} z^k \\
 &= \sum_{k=1}^{K+1} \frac{2k}{K+1+k} C_k \binom{K+1+k}{K+1-k} z^k.
 \end{aligned}
 \tag{A.16}$$

Finally, summing (A.11) and (A.16) yields

$$\begin{aligned}
 &b_K(z) + z \sum_{l=0}^K b_l(z) b_{K-l}(z) \\
 &= \sum_{k=0}^K \frac{K+1-k}{K+1+k} C_k \binom{K+1+k}{K+1-k} z^k + \sum_{k=1}^{K+1} \frac{2k}{K+1+k} C_k \binom{K+1+k}{K+1-k} z^k \\
 &= \sum_{k=0}^{K+1} C_k \binom{K+1+k}{K+1-k} z^k = b_{K+1}(z),
 \end{aligned}
 \tag{A.17}$$

proving the recursion (A.10) is correct. □

Appendix B: Random-product representation

The results we present in Appendix C make use of the *random-product representation* theory recently developed and discussed in [6, 12]. Using this theory to study a given Markov process X requires selecting an additional Markov process $\tilde{X} := \{\tilde{X}(t)\}_{t \geq 0}$ that shares the same state space \mathbb{S} as X . The elements of the transition rate matrix \mathbf{Q} of \tilde{X} must satisfy the following two properties; see [6, Section 1] and [12, Section 2]:

- (i) For each $x \in \mathbb{S}$, $\tilde{q}(x) := -\tilde{q}(x, x) = \sum_{y \neq x} \tilde{q}(x, y) = q(x)$;
- (ii) For each $x, y \in \mathbb{S}$, $x \neq y$, $\tilde{q}(x, y) > 0$ if and only if $q(y, x) > 0$.

Associate with the Markov process X its transition times $\{T_n\}_{n \geq 0}$, where $T_0 := 0$ and T_n represents the n -th transition time of X . From the transition times, we create the embedded discrete-time Markov chain $\{X_n\}_{n \geq 0}$ as $X_n := X(T_n)$, $n \geq 0$. The sequences $\{\tilde{T}_n\}_{n \geq 0}$ and $\{\tilde{X}_n\}_{n \geq 0}$ are constructed and defined similarly.

We will also need to make use of discrete-time hitting-time random variables. We define, for each set $A \subset \mathbb{S}$,

$$\eta_A := \inf\{n \geq 1 : X_n \in A\}
 \tag{B.1}$$

as the first time the embedded chain make a transition into A . η_x should be understood to mean $\eta_{\{x\}}$. The continuous-time hitting-time random variable $\tilde{\tau}_A$ is defined anal-

ogously to the definition in Sect. 2.1, and $\tilde{\eta}_A$ represents the first time the embedded DTMC of \tilde{X} reaches the set A .

The random-product representation can be used to determine the Laplace transforms of the transition functions, when the process is assumed to start in a fixed state $x \in \mathbb{S}$. We will employ the following theorem, which originally appeared in [12, Corollary 2.1 and Theorem 2.1].

Theorem 6 *Suppose $y \in \mathbb{S}$, where $y \neq x$. Then the Laplace transform $\pi_{x,y}(\alpha)$ of $p_{x,y}(\cdot)$ satisfies*

$$\pi_{x,y}(\alpha) = \pi_{x,x}(\alpha) \mathbb{E}_y \left[e^{-\alpha \tilde{\tau}_x} \prod_{l=1}^{\tilde{\eta}_x} \frac{q(\tilde{X}_l, \tilde{X}_{l-1})}{\tilde{q}(\tilde{X}_{l-1}, \tilde{X}_l)} \right], \tag{B.2}$$

where

$$\pi_{x,x}(\alpha) = \frac{1}{(q(x) + \alpha)(1 - \mathbb{E}_x[e^{-\alpha \tau_x}])}. \tag{B.3}$$

Appendix C: Results for $M/M/1$ queues

Here we derive key quantities associated with $M/M/1$ queues. The first lemma is a restatement of [10, Lemma 4], which was inspired by Problems 22 and 23 of [19, Chapter 7].

Lemma 5 *Suppose $\{X(t)\}_{t \geq 0}$ is an $M/M/1$ queueing model with exponential clearings. Arrivals occur according to a Poisson process with rate λ , each service is exponentially distributed with rate μ and there is an external Poisson process having rate θ of clearing instants, where, whenever a clearing occurs, all customers in the system at the clearing time are removed. Then, for $j \geq 1$,*

$$\mathbb{E}_1 \left[\mathbb{1}\{\tau_j < \tau_0\} e^{-\alpha \tau_j} \right] = \frac{\left(\frac{\lambda}{\mu} \phi_{\lambda,\mu}(\theta + \alpha)\right)^{j-1} \left(1 - \frac{\lambda}{\mu} \phi_{\lambda,\mu}(\theta + \alpha)^2\right)}{1 - \left(\frac{\lambda}{\mu} \phi_{\lambda,\mu}(\theta + \alpha)^2\right)^j}. \tag{C.1}$$

Proof The time τ_0 is the minimum of the busy period of a regular $M/M/1$ queue, i.e., $B_{\lambda,\mu}$, and an exponential random variable with parameter θ . Note that the two random variables are independent. So,

$$\mathbb{E}_1 \left[\mathbb{1}\{\tau_j < \tau_0\} e^{-\alpha \tau_j} \right] = \mathbb{E}_1 \left[\mathbb{1}\{\tau_j < B_{\lambda,\mu}\} \mathbb{1}\{\tau_j < E_\theta\} e^{-\alpha \tau_j} \right], \tag{C.2}$$

where E_θ is an exponential random variable with parameter θ . Under this description, the time τ_j is equal to the time τ_j^* to reach state j in a regular $M/M/1$ queue. Furthermore, conditioning on both τ_j^* and $B_{\lambda,\mu}$ yields

$$(C.2) = \mathbb{E}_1 \left[\mathbb{1}\{\tau_j^* < B_{\lambda,\mu}\} \mathbb{1}\{\tau_j^* < E_\theta\} e^{-\alpha \tau_j^*} \right]$$

$$\begin{aligned}
 &= \mathbb{E}_1 \left[\mathbb{E} \left[\mathbb{1} \{ \tau_j^* < B_{\lambda, \mu} \} \mathbb{1} \{ \tau_j^* < E_\theta \} e^{-\alpha \tau_j^*} \mid B_{\lambda, \mu}, \tau_j^* \right] \right] \\
 &= \mathbb{E}_1 \left[\mathbb{1} \{ \tau_j^* < B_{\lambda, \mu} \} e^{-\alpha \tau_j^*} \mathbb{E} \left[\mathbb{1} \{ \tau_j^* < E_\theta \} \mid B_{\lambda, \mu}, \tau_j^* \right] \right] \\
 &= \mathbb{E}_1 \left[\mathbb{1} \{ \tau_j^* < B_{\lambda, \mu} \} e^{-(\theta + \alpha) \tau_j^*} \right]. \tag{C.3}
 \end{aligned}$$

The remainder of the proof follows from the proof of [10, Lemma 4]. □

The following lemma is a minor generalization of [10, Theorem 2], in that we verify it is still valid for $\alpha \in \mathbb{C}_+$.

Lemma 6 *Suppose $\{X(t)\}_{t \geq 0}$ is the clearing model of Lemma 5. Then for each $j, k \geq 1$,*

$$\begin{aligned}
 &\mathbb{E}_k \left[\int_0^{\tau_0} e^{-\alpha t} \mathbb{1} \{ X(t) = j \} dt \right] \\
 &= \begin{cases} \frac{\frac{\lambda}{\mu} \phi_{\lambda, \mu}(\theta + \alpha)}{\lambda(1 - \frac{\lambda}{\mu} \phi_{\lambda, \mu}(\theta + \alpha)^2)} \left(\frac{\lambda}{\mu} \phi_{\lambda, \mu}(\theta + \alpha) \right)^{j-k} \left(1 - \left(\frac{\lambda}{\mu} \phi_{\lambda, \mu}(\theta + \alpha)^2 \right)^k \right), & 1 \leq k \leq j - 1, \\ \frac{\frac{\lambda}{\mu} \phi_{\lambda, \mu}(\theta + \alpha)}{\lambda(1 - \frac{\lambda}{\mu} \phi_{\lambda, \mu}(\theta + \alpha)^2)} \phi_{\lambda, \mu}(\theta + \alpha)^{k-j} \left(1 - \left(\frac{\lambda}{\mu} \phi_{\lambda, \mu}(\theta + \alpha)^2 \right)^j \right), & k \geq j. \end{cases} \tag{C.4}
 \end{aligned}$$

Proof Define the Laplace transform of the transition functions

$$\pi_{0,j}(\alpha) := \int_0^\infty e^{-\alpha t} p_{0,j}(t) dt. \tag{C.5}$$

The state space for this single-server queue is $\mathbb{S} = \mathbb{N}_0$. Select $A = \{0\}$ and $B = \mathbb{S} \setminus A$ and apply Theorem 1 to obtain, for $j \geq 1$,

$$\begin{aligned}
 \pi_{0,j}(\alpha) &= \pi_{0,0}(\alpha)(\lambda + \alpha) \mathbb{E}_0 \left[\int_0^{\tau_0} e^{-\alpha t} \mathbb{1} \{ X(t) = j \} dt \right] \\
 &= \pi_{0,0}(\alpha) \lambda \mathbb{E}_1 \left[\int_0^{\tau_0} e^{-\alpha t} \mathbb{1} \{ X(t) = j \} dt \right]. \tag{C.6}
 \end{aligned}$$

We can use the random-product representation of Appendix B to derive another expression for $\pi_{0,j}(\alpha)$. Construct a Markov process $\tilde{X} := \{\tilde{X}(t)\}_{t \geq 0}$ with transition rates $\tilde{q}(i, i - 1) = \lambda + \theta$ and $\tilde{q}(i, i + 1) = \mu$ for $i \geq 1$. \tilde{X} also has transitions from state 0 to every other state, but these do not factor into the calculations so there is no need to formally define them here. From Theorem 6,

$$\begin{aligned}
 \pi_{0,j}(\alpha) &= \pi_{0,0}(\alpha) \mathbb{E}_j \left[e^{-\alpha \tilde{\tau}_0} \prod_{l=1}^{\tilde{\eta}_0} \frac{q(\tilde{X}_l, \tilde{X}_{l-1})}{\tilde{q}(\tilde{X}_{l-1}, \tilde{X}_l)} \right] \\
 &= \pi_{0,0}(\alpha) \mathbb{E}_1 \left[e^{-\alpha B_{\mu, \lambda + \theta}} \left(\frac{\lambda}{\lambda + \theta} \right)^{D_{\mu, \lambda + \theta}} \right]^j
 \end{aligned}$$

$$= \pi_{0,0}(\alpha) \left(\frac{\lambda}{\mu} \phi_{\lambda,\mu}(\theta + \alpha) \right)^j. \tag{C.7}$$

Combining (C.6) with (C.7) gives

$$\mathbb{E}_1 \left[\int_0^{\tau_0} e^{-\alpha t} \mathbb{1}\{X(t) = j\} dt \right] = \frac{1}{\lambda} \left(\frac{\lambda}{\mu} \phi_{\lambda,\mu}(\theta + \alpha) \right)^j. \tag{C.8}$$

For now, abbreviate $\phi := \phi_{\lambda,\mu}(\theta + \alpha)$ and $r := \frac{\lambda}{\mu} \phi_{\lambda,\mu}(\theta + \alpha)$. The remaining expected values can be computed. First, for $2 \leq k \leq j$,

$$\begin{aligned} \mathbb{E}_1 \left[\int_0^{\tau_0} e^{-\alpha t} \mathbb{1}\{X(t) = j\} dt \right] &= \mathbb{E}_1 \left[\mathbb{1}\{\tau_k < \tau_0\} \int_0^{\tau_0} e^{-\alpha t} \mathbb{1}\{X(t) = j\} dt \right] \\ &= \mathbb{E}_1 \left[\mathbb{1}\{\tau_k < \tau_0\} e^{-\alpha \tau_k} \int_{\tau_k}^{\tau_0} e^{-\alpha(t-\tau_k)} \mathbb{1}\{X(t) = j\} dt \right] \\ &= \mathbb{E}_1 \left[\mathbb{1}\{\tau_k < \tau_0\} e^{-\alpha \tau_k} \right] \mathbb{E}_k \left[\int_0^{\tau_0} e^{-\alpha t} \mathbb{1}\{X(t) = j\} dt \right]. \end{aligned} \tag{C.9}$$

Furthermore, from Lemma 5,

$$\mathbb{E}_1 \left[\mathbb{1}\{\tau_k < \tau_0\} e^{-\alpha \tau_k} \right] = \frac{r^{k-1}(1 - r\phi)}{1 - (r\phi)^k}, \tag{C.10}$$

meaning

$$\mathbb{E}_k \left[\int_0^{\tau_0} e^{-\alpha t} \mathbb{1}\{X(t) = j\} dt \right] = \frac{r}{\lambda(1 - r\phi)} r^{j-k} (1 - (r\phi)^k). \tag{C.11}$$

Deriving the expected values when $k > j$ is a little more straightforward. Here,

$$\begin{aligned} \mathbb{E}_k \left[\int_0^{\tau_0} e^{-\alpha t} \mathbb{1}\{X(t) = j\} dt \right] &= \mathbb{E}_k \left[\mathbb{1}\{\tau_j < \tau_0\} e^{-\alpha \tau_j} \int_{\tau_j}^{\tau_0} e^{-\alpha(t-\tau_j)} \mathbb{1}\{X(t) = j\} dt \right] \\ &= \mathbb{E}_k \left[\mathbb{1}\{\tau_j < \tau_0\} e^{-\alpha \tau_j} \right] \mathbb{E}_j \left[\int_0^{\tau_0} e^{-\alpha t} \mathbb{1}\{X(t) = j\} dt \right] \\ &= \frac{r}{\lambda(1 - r\phi)} \phi^{k-j} (1 - (r\phi)^j), \end{aligned} \tag{C.12}$$

which proves the claim. □

Lemma 7 *The expectation $w_i^{(\lambda,\mu,\theta)}(\alpha)$, defined in (2.8), satisfies for $i \geq 1$ the recursion*

$$w_i^{(\lambda,\mu,\theta)}(\alpha) = \frac{\theta}{\lambda + \mu + \theta + \alpha} w_{i-1}^{(\lambda,\mu,\theta)}(\alpha) + \frac{\lambda}{\lambda + \mu + \theta + \alpha} \sum_{k=0}^i w_{i-k}^{(\lambda,\mu,\theta)}(\alpha) w_k^{(\lambda,\mu,\theta)}(\alpha) \tag{C.13}$$

and $w_0^{(\lambda, \mu, \theta)}(\alpha) = \phi_{\lambda, \mu}(\theta + \alpha)$.

Proof Conditioning on the length of the busy period, we have, for $i = 0$,

$$\begin{aligned} w_0^{(\lambda, \mu, \theta)}(\alpha) &= \mathbb{E}_1[e^{-\alpha B_{\lambda, \mu}} \mathbb{1}\{A_\theta(B_{\lambda, \mu}) = 0\}] = \mathbb{E}_1[e^{-\alpha B_{\lambda, \mu}} \mathbb{1}\{B_{\lambda, \mu} < E_\theta\}] \\ &= \int_0^\infty \mathbb{P}(t < E_\theta) e^{-\alpha t} f_{B_{\lambda, \mu}}(t) dt = \phi_{\lambda, \mu}(\theta + \alpha). \end{aligned} \tag{C.14}$$

The recursion for $i \geq 1$ follows from a one-step analysis and the strong Markov property. Since the birth–and–death process starts with 1 customer, the first event occurs after $E_{\lambda+\mu+\theta}$ time and is either an arrival according to the Poisson process with probability $\theta/(\lambda + \mu + \theta)$ or an arrival of an additional customer with probability $\lambda/(\lambda + \mu + \theta)$. If the former occurs, due to the strong Markov property, one less Poisson point needs to arrive. If the latter occurs, exactly i Poisson points need to arrive in two busy periods (due to the homogeneous structure of the birth–and–death process). This reasoning establishes the recursion (C.13). \square

Lemma 8 *The $\{w_i^{(\lambda, \mu, \theta)}(\alpha)\}_{i \geq 0}$ of Lemma 7 are given by $w_0^{(\lambda, \mu, \theta)}(\alpha) = \phi_{\lambda, \mu}(\theta + \alpha)$ and for $i \geq 1$,*

$$w_i^{(\lambda, \mu, \theta)}(\alpha) = W_1^i \phi_{\lambda, \mu}(\theta + \alpha) \sum_{k=0}^{i-1} C_k \binom{i-1+k}{i-1-k} W_2^k = W_1^i \phi_{\lambda, \mu}(\theta + \alpha) b_{i-1}(W_2), \tag{C.15}$$

where $C_k := \frac{1}{k+1} \binom{2k}{k}$ are the Catalan numbers, $b_K(z)$ is defined in Lemma 4, and

$$\begin{aligned} W_1 &= \frac{\theta}{\lambda(1 - 2\phi_{\lambda, \mu}(\theta + \alpha)) + \mu + \theta + \alpha}, \\ W_2 &= \frac{\lambda\phi_{\lambda, \mu}(\theta + \alpha)}{\lambda(1 - 2\phi_{\lambda, \mu}(\theta + \alpha)) + \mu + \theta + \alpha}. \end{aligned} \tag{C.16}$$

Proof For brevity, define $w_i := w_i^{(\lambda, \mu, \theta)}(\alpha)$. Rewrite (C.13) as

$$(\lambda + \mu + \theta + \alpha)w_{i+1} = \theta w_i + \lambda \left(\sum_{k=1}^i w_{i+1-k} w_k + 2w_0 w_{i+1} \right). \tag{C.17}$$

Using $w_0 = \phi_{\lambda, \mu}(\theta + \alpha)$, this reduces to

$$w_{i+1} = W_1 w_i + \frac{W_2}{\phi_{\lambda, \mu}(\theta + \alpha)} \sum_{k=1}^i w_{i+1-k} w_k. \tag{C.18}$$

Straightforwardly substituting (C.15) into (C.18) and dividing by $W_1^{i+1} \phi_{\lambda, \mu}(\theta + \alpha)$ results in

$$b_i(W_2) = b_{i-1}(W_2) + W_2 \sum_{k=1}^i b_{i-k}(W_2) b_{k-1}(W_2). \tag{C.19}$$

Now, change the summation index by setting $l = k - 1$ to retrieve

$$b_i(W_2) = b_{i-1}(W_2) + W_2 \sum_{l=0}^{i-1} b_{i-1-l}(W_2) b_l(W_2), \tag{C.20}$$

so that Lemma 4 proves the claim (C.15). □

Lemma 9 *The generating function of the $\{w_i^{(\lambda, \mu, \theta)}(\alpha)\}_{i \geq 0}$ is, for $|z| < 1$,*

$$\sum_{i=0}^{\infty} w_i^{(\lambda, \mu, \theta)}(\alpha) z^i = \phi_{\lambda, \mu}(\theta(1 - z) + \alpha). \tag{C.21}$$

Proof We use the definition of $w_i^{(\lambda, \mu, \theta)}(\alpha)$ in Lemma 7 and condition on the length of the busy period:

$$\begin{aligned} \sum_{i=0}^{\infty} w_i^{(\lambda, \mu, \theta)}(\alpha) z^i &= \sum_{i=0}^{\infty} \mathbb{E}_1 \left[e^{-\alpha B_{\lambda, \mu}} \mathbb{1}\{\Lambda_{\theta}(B_{\lambda, \mu}) = i\} \right] z^i \\ &= \sum_{i=0}^{\infty} \mathbb{E}_1 \left[e^{-\alpha B_{\lambda, \mu}} \mathbb{1}\{\Lambda_{\theta}(B_{\lambda, \mu}) = i\} z^i \right] \\ &= \mathbb{E}_1 \left[e^{-\alpha B_{\lambda, \mu}} z^{\Lambda_{\theta}(B_{\lambda, \mu})} \right] \\ &= \int_0^{\infty} \mathbb{E}_1 \left[e^{-\alpha B_{\lambda, \mu}} z^{\Lambda_{\theta}(B_{\lambda, \mu})} \mid B_{\lambda, \mu} = t \right] f_{B_{\lambda, \mu}}(t) dt \\ &= \int_0^{\infty} e^{-\alpha t} \mathbb{E} \left[z^{\Lambda_{\theta}(t)} \right] f_{B_{\lambda, \mu}}(t) dt \\ &= \int_0^{\infty} e^{-(\theta(1-z)+\alpha)t} f_{B_{\lambda, \mu}}(t) dt \\ &= \phi_{\lambda, \mu}(\theta(1 - z) + \alpha), \end{aligned} \tag{C.22}$$

where we used the probability generating function of a Poisson distribution with parameter θt . □

References

1. Abate, J., Whitt, W.: Solving probability transform functional equations for numerical inversion. *Oper. Res. Lett.* **12**(5), 275–281 (1992)
2. Abate, J., Whitt, W.: Transient behavior of the $M/G/1$ workload process. *Oper. Res.* **42**(4), 750–764 (1994)
3. Abate, J., Whitt, W.: Numerical inversion of Laplace transforms of probability distributions. *ORSA J. Comput.* **7**(1), 36–43 (1995)
4. Abate, J., Whitt, W.: A unified framework for numerically inverting Laplace transforms. *INFORMS J. Comput.* **18**(4), 408–421 (2006)

5. Abate, J., Whitt, W.: Integer sequences from queueing theory. *J. Integer Seq.* **13**(5), 1–21 (2010)
6. Buckingham, P., Fralix, B.: Some new insights into Kolmogorov's criterion, with applications to hysteretic queues. *Markov Process. Relat. Fields* **21**(2), 339–368 (2015)
7. Cobham, A.: Priority assignment in waiting line problems. *Oper. Res.* **2**(1), 70–76 (1954)
8. Davis, R.: Waiting-time distribution of a multi-server, priority queueing system. *Oper. Res.* **14**(1), 133–136 (1966)
9. den Iseger, P.: Numerical transform inversion using Gaussian quadrature. *Prob. Eng. Inf. Sci.* **20**, 1–44 (2006)
10. Doroudi, S., Fralix, B., Harchol-Balter, M.: Clearing analysis on phases: exact limiting probabilities for skip-free, unidirectional, quasi-birth-death processes (2015). ArXiv preprint [arXiv:1503.05899v3](https://arxiv.org/abs/1503.05899v3)
11. Feller, W.: *An Introduction to Probability Theory and Its Applications*, vol. I, 3 revised edn. Wiley, New York (1968)
12. Fralix, B.: When are two Markov chains similar? *Stat. Prob. Lett.* **107**, 199–203 (2015)
13. Gail, H., Hantler, S., Taylor, B.: On a preemptive Markovian queue with multiple servers and two priority classes. *Math. Oper. Res.* **17**(2), 365–391 (1992)
14. Harchol-Balter, M., Osogami, T., Scheller-Wolf, A., Wierman, A.: Multi-server queueing systems with multiple priority classes. *Queueing Syst.* **51**(3–4), 331–360 (2005)
15. Jaiswal, N.: Preemptive resume priority queue. *Oper. Res.* **9**(5), 732–742 (1961)
16. Jaiswal, N.: *Priority Queues*. Academic Press Inc, New York (1968)
17. Joyner, J., Fralix, B.: A new look at block-structured Markov processes. In: Working Paper (2016). <http://bfralix.people.clemson.edu/preprints/BlockStructuredPaper8June.pdf>
18. Joyner, J., Fralix, B.: A new look at Markov processes of $G/M/1$ -type. *Stoch. Models* **32**(2), 253–274 (2016)
19. Karlin, S., Taylor, H.: *A First Course in Stochastic Processes*, 2nd edn. Academic Press, San Diego (1975)
20. Katehakis, M., Smit, L., Spieksma, F.: A comparative analysis of the successive lumping and the lattice path counting algorithms. *J. Appl. Prob.* **53**(1), 106–120 (2016)
21. Latouche, G., Ramaswami, V.: *Introduction to Matrix Analytic Methods in Stochastic Modeling*. Society for Industrial and Applied Mathematics, Philadelphia (1999)
22. Li, Q.L., Zhao, Y.Q.: The RG -factorization in block-structured Markov renewal processes. In: Zhu, X. (ed.) *Observation, Theory and Modeling of Atmospheric Variability*, pp. 545–568. World Scientific, Singapore (2004)
23. Li, H., Zhao, Y.: Exact tail asymptotics in a priority queue—characterizations of the preemptive model. *Queueing Syst.* **63**(1–4), 355–381 (2009)
24. Miller, D.: Computation of steady-state probabilities for $M/M/1$ priority queues. *Oper. Res.* **29**(5), 945–958 (1981)
25. Ramaswami, V.: A stable recursion for the steady state vector in Markov chains of $M/G/1$ type. *Stoch. Models* **4**(1), 183–188 (1988)
26. Sleptchenko, A., van Harten, A., van der Heijden, M.: An exact solution for the state probabilities of the multi-class, multi-server queue with preemptive priorities. *Queueing Syst.* **50**(1), 81–107 (2005)
27. Sleptchenko, A., Selen, J., Adan, I., van Houtum, G.: Joint queue length distribution of multi-class, single-server queues with preemptive priorities. *Queueing Syst.* **81**(4), 379–395 (2015)
28. Takács, L.: *Introduction to the Theory of Queues*. Oxford University Press Inc, New York (1962)
29. Wang, J., Baron, O., Scheller-Wolf, A.: $M/M/c$ queue with two priority classes. *Oper. Res.* **63**(3), 733–749 (2015)