Performance Evaluation

# The Number of Overlapping Customers in a Queueing System with Non-Preemptive Priority

--Manuscript Draft--

| | |
|---|---|
| **Manuscript Number:** | |
| **Article Type:** | Full Length Article |
| **Keywords:** | Queues; Priority; Non-Preemptive; Overlaps; Heavy Traffic |
| **Corresponding Author:** | Jamol Pender<br>Cornell University<br>Ithaca, New York United States |
| **First Author:** | Jamol Pender |
| **Order of Authors:** | Jamol Pender |
| | Onno Boxma |
| | Uri Yechiali |
| **Abstract:** | We consider a 2-class non-preemptive priority M/M/1-type queueing system and derive the PGF of the joint distribution of the number of overlapping customers that a class-i customer (i=1,2) encounters with customers of her class and with customers of the other class. We further calculate the mean and second factorial moments of these quantities and obtain heavy-traffic limits of the overlap numbers. We finally derive the mean overlap numbers for an M-class system with general service times and provide a comparison between mean overlap numbers in the non-preemptive and preemptive cases, and the case without priorities. |

*Powered by Editorial Manager® and ProduXion Manager® from Aries Systems Corporation*

July 20, 2026

Editor-in-Chief of Performance Evaluation

Dear Editor,

We are pleased to submit our manuscript entitled "The Number of Overlapping Customers in a Queueing System with Non-Preemptive Priority" for consideration for publication in *Performance Evaluation*.

In this paper, we study a two-class non-preemptive priority M/M/1-type queueing system and develop a detailed characterization of customer overlap behavior. Specifically, we derive the probability generating function (PGF) of the joint distribution of the number of overlapping customers encountered by a tagged class-$i$ customer ($i = 1, 2$), distinguishing between overlaps with customers of the same class and those of the other class.

Using this framework, we compute the mean and second factorial moments of the overlap quantities, providing insight into both average behavior and variability. We further establish heavy-traffic limits for these overlap numbers, revealing their asymptotic structure under congestion. Beyond the two-class setting, we extend our analysis to an $M$-class system with general service time distributions, deriving expressions for mean overlap numbers in this more general setting.

Finally, we present a comparative analysis of overlap behavior across different service disciplines, including non-preemptive priority, preemptive priority, and systems without priorities. This comparison highlights the impact of prioritization mechanisms on interaction patterns among customers.

We believe that our results contribute to the growing literature on interaction-based performance metrics in queueing systems and will be of interest to researchers in applied probability, operations research, and service systems.

This manuscript is original, has not been published previously, and is not under consideration for publication elsewhere. All authors have approved the submission.

Thank you for your consideration. We look forward to your response.

Sincerely,

Jamol Pender, Onno Boxma, Uri Yechiali

Manuscript File Click here to view linked References

# The Number of Overlapping Customers in a Queueing System with Non-Preemptive Priority

Onno Boxma
Department of Mathematics and Computer Science
Eindhoven University of Technology
o.j.boxma@tue.nl

Jamol Pender
School of Operations Research and Information Engineering
Cornell University
jjp274@cornell.edu

Uri Yechiali
Department of Statistics and Operations Research
Tel Aviv University
uriy@tauex.tau.ac.il

July 20, 2026

**Abstract**

We consider a 2-class non-preemptive priority M/M/1-type queueing system and derive the PGF of the joint distribution of the number of overlapping customers that a class-i customer (i=1,2) encounters with customers of her class and with customers of the other class. We further calculate the mean and second factorial moments of these

quantities and obtain heavy-traffic limits of the overlap numbers. We finally derive the mean overlap numbers for an M-class system with general service times and provide a comparison between mean overlap numbers in the non-preemptive and preemptive cases, and the case without priorities.

# 1 Introduction

In a companion paper (Boxma et al. [6]) we have recently studied overlap numbers in the $M/M/1$ two class priority queue with preemptive priority for Class 1; in the present paper, we focus on the non-preemptive case.

Queueing systems with multiple customer classes arise naturally in a wide range of application areas, including healthcare and computer-communications. There is an extensive literature on queues with priority and on scheduling jobs in queueing systems; see for example Jaiswal [12], Kleinrock [15], Takagi [26], Ridley and Massey [25], Harchol-Balter [10], Weiss [27].

In recent years, a compelling new frontier in queueing theory has emerged, shifting the focus from classical performance measures such as waiting times and queue lengths to the subtle and complex interactions between customers within the system. Central to this innovative research is the study of **overlaps**: both the **number of overlaps**, which counts how many customers concurrently share time in the system with a particular tagged customer, and the **total overlapping time**, quantifying the cumulative duration of these shared intervals. These refined metrics open a window into the microscopic dynamics of congestion and customer experience, offering insights far richer than traditional aggregate statistics.

Recent work by Boxma [2], Boxma and Pender [3], Boxma et al. [5], Gao and Pender [9], Kang [13], Kang et al. [14], Palomo and Pender [21], Oz and Perlman [19], Palomo and Pender [20, 22, 23], Pender et al. [24], Xu et al. [28], Newman and Pender [17, 18], Boxma et al. [4], Xu et al. [29] has laid the foundational theoretical framework for analyzing these overlap phenomena. Notably, most of these studies consider systems without priority disciplines, with the exception of Oz and Perlman [19] (where overlap number is called co-exposure count) and Boxma et al. [6]. This emerging line of research holds great promise for unveiling new perspectives on the microscopic customer behaviors underpinning a wide range of queueing models, enriching our understanding of system dynamics and facilitating the design of more

responsive service systems.

In this paper, we analyze the two class M/M/1 priority queue with non-preemptive service. Unlike other studies, our goal is to analyze the distribution of the number of overlaps that each class experiences while in the system. This type of information is critical for understanding how differentiating the customers in a healthcare facility will potentially affect their exposition to infectious diseases. In the two class case, it is important to understand the number of overlaps that the prioritized class experiences versus the non-prioritized class. This differentiation will allow one to better understand the additional exposure that the non-prioritized class is subject to. One should also take the following into account, as also pointed out in Boxma et al. [6] and Oz and Perlman [19]: while a Class-1 customer benefits from its high priority by spending a relatively short time in the system and meeting relatively few Class-1 customers, it may encounter relatively many customers of Class 2.

## 1.1 Contributions of our work

Our main contributions are: **(i)** To the best of our knowledge, we derive new PGF (probability generating function) expressions for the joint distribution of numbers of customers and the type of the customer in service, for the non-preemptive priority queue with unequal service rates. **(ii)** We derive explicit expressions for the PGF's of the joint distribution of the numbers of overlaps a customer of Class $i = 1, 2$ experiences with customers of both classes. This is different from the preemptive case since we need to condition on the number of customers present and which class is in service. In the preemptive case, we only need to condition on the number of customers present in each class. **(iii)** We determine the means and second factorial moments of these quantities. **(iv)** We obtain heavy-traffic limits of the overlap numbers. **(v)** For the case of $M \geq 2$ priority classes and general service time distributions, we present a comparison of the mean overlap numbers for the non-preemptive case, the preemptive priority case and the case without priorities.

## 1.2 Organization of the paper

We first determine the PGF of the steady-state joint queue length distribution of high- and low-priority customers (Section 2). In Section 3 the PGF's of the joint distributions of overlap numbers of Class $i = 1, 2$ with

customers of both classes are expressed in terms of that joint queue length PGF; moments of overlap numbers are also derived in that section. Section 4 is devoted to an analysis of the heavy-traffic limits of the overlap numbers. In Section 5, where we allow $M \geq 2$ priority classes with general service time distributions, we present a comparison of mean overlap numbers; first between classes, and next between the non-preemptive and preemptive priority cases, and with the case of a multiclass system without priorities.

## 1.3 Notation

In this section, we provide the notation that will be used throughout the remainder of the paper. Indices $i$ and $j$ run over all of the classes (hence in most of the paper, $i = 1, 2$ and $j = 1, 2$).

- $\lambda_i$ : arrival rate of Class-$i$ customers.
- $\mu_i$ : service rate of Class-$i$ customers.
- $\mathcal{S}_i$ : service time of a Class-$i$ customer.
- $\rho_i := \lambda_i \mathbb{E}[\mathcal{S}_i]$ : traffic load of Class-$i$ customers; $\rho = \rho_1 + \rho_2$ is the total traffic load.
- $\mathcal{Q}_i$ : number of customers of Class $i$ in the system.
- $C_i$ : tagged customer of Class $i$.
- $\mathcal{O}_{i,j}$ : number of customers of Class $j$ that have some overlap with $C_i$.
- $\mathcal{O}_i^T$ : total number of overlaps $C_i$ encounters.
- $p_k(\cdot, \cdot)$ denotes the steady-state joint queue length distribution of Class-1 and Class-2 customers, with the server serving a Class-$k$ customer, $k = 1, 2$; hence $p_k(m, n) = \mathbb{P}(\mathcal{Q}_1 = m, \mathcal{Q}_2 = n, \text{Class k in service})$.
- $P_k(z_1, z_2)$ denotes the probability generating function of the steady-state joint queue length distribution $p_k(\cdot, \cdot)$, $k = 1, 2$.
- $P_1$ : busy period of Class-1 customers in isolation.
- $CT_2$ : completion time of a Class-2 customer.
- $I(E)$ denotes an indicator function: If event $E$ is true, then $I(E) = 1$ and else $I(E) = 0$.

# 2 The $M/M/1$ queue with two priority classes and non-preemptive priority

Consider a single server queue with two customer classes, and non-preemptive priority of Class 1 above Class 2. Class-$i$ customers arrive according to a Poisson process with rate $\lambda_i$ and with $\exp(\mu_i)$ distributed service requirements, $i = 1, 2$. We consider the system in steady state; the stability condition is that $\rho := \frac{\lambda_1}{\mu_1} + \frac{\lambda_2}{\mu_2} < 1$. Let $p(0,0)$ denote the probability of an empty system, and $p_k(i,j)$ the steady-state probability of having $i$ Class-1 customers and $j$ Class-2 customers, with the server serving a Class-$k$ customer, $k = 1, 2$; observe that, unlike in the preemptive case, the server can serve a Class-2 customer while there are Class-1 customers present (who apparently have arrived during the ongoing service). The PGF of the $p_k(i,j)$ are defined as $P_1(z_1, z_2) := \sum_{i=1}^{\infty} \sum_{j=0}^{\infty} p_1(i,j) z_1^i z_2^j$ and $P_2(z_1, z_2) := \sum_{i=0}^{\infty} \sum_{j=1}^{\infty} p_2(i,j) z_1^i z_2^j$. In Subsection 2.1, we determine these two PGF's, and in Subsection 2.2 the mean queue lengths and second factorial moments, also distinguishing between having a Class-1 or a Class-2 customer in service. In Section 3, we express the steady-state joint overlap PGF for Class-1 customers, and for Class-2 customers, in terms of those PGFs.

*Remark.* There is a huge literature on priority queues; see, e.g., Jaiswal [12] and Takagi [26]. An important first study on the two-class $M/M/1$ non-preemptive priority queue that we consider is Chapter 9 of Morse [16]. Morse [16] derives the joint queue length PGF, for the two options that a Class-1 or a Class-2 customer is in service, for the special case that both priority classes have the same service rate; for our case of unequal service rates, he provides the balance equations and partial results. Jaiswal [11] gives an explicit expression for $P_1(z_1, z_2) + P_2(z_1, z_2) + p(0,0)$ in his Equation 42. We have not been able to find in the literature explicit results for the PGF's $P_1(z_1, z_2)$ and $P_2(z_1, z_2)$ and second (factorial) moments of queue lengths, taking also into account which class is in service. Hence, and because these results seem of independent interest, we have decided to present their analysis here in some detail. However, our main goal is to use the results of the present section for our overlap studies in subsequent sections.

## 2.1 Queue length distributions

In this section, we analyze the steady-state balance equations of the two-class priority queue without preemption and the resulting generating functions. The balance equations for the two-class priority queue without preemption are as follows:

$$\begin{aligned}
p_1(i,j)(\lambda_1+\lambda_2+\mu_1) &= p_1(i-1,j)\lambda_1 I(i\geq 2) + p_1(i,j-1)\lambda_2 \\
&+ p_1(i+1,j)\mu_1 + p_2(i,j+1)\mu_2, \quad i,j\geq 1, \\
p_1(i,0)(\lambda_1+\lambda_2+\mu_1) &= p_1(i-1,0)\lambda_1 + p_1(i+1,0)\mu_1 + p_2(i,1)\mu_2, \quad i\geq 1, \\
p_2(i,j)(\lambda_1+\lambda_2+\mu_2) &= p_2(i-1,j)\lambda_1 + p_2(i,j-1)\lambda_2 I(j\geq 2), \quad i,j\geq 1, \\
p_2(0,j)[\lambda_1+\lambda_2+\mu_2] &= p_2(0,j-1)\lambda_2 + p_2(0,j+1)\mu_2 + p_1(1,j)\mu_1, \quad j\geq 1, \\
p(0,0)(\lambda_1+\lambda_2) &= p_1(1,0)\mu_1 + p_2(0,1)\mu_2. \qquad (1)
\end{aligned}$$

Taking probability generating functions in the first two balance equations yields the following expression

$$\begin{aligned}
&P_1(z_1,z_2)[\lambda_1+\lambda_2+\mu_1] = \lambda_1 z_1 P_1(z_1,z_2) + \lambda_2 z_2 P_1(z_1,z_2) \qquad (2) \\
+ \quad &\frac{\mu_1}{z_1}\sum_{i=1}^{\infty}\sum_{j=0}^{\infty} z_1^{i+1} z_2^j p_1(i+1,j) + \frac{\mu_2}{z_2}\sum_{i=1}^{\infty}\sum_{j=0}^{\infty} z_1^i z_2^{j+1} p_2(i,j+1) + \lambda_1 z_1 p(0,0).
\end{aligned}$$

We find it helpful to introduce the following two functions

$$R_1(z) := \sum_{j=0}^{\infty} p_1(1,j) z^j, \qquad (3)$$

$$R_2(z) := \sum_{j=1}^{\infty} p_2(0,j) z^j, \qquad (4)$$

after which we can rewrite Equation (2) as the following expression

$$K(z_1,z_2)P_1(z_1,z_2) = \frac{\mu_2}{z_2}P_2(z_1,z_2) + \lambda_1 z_1 p(0,0) - \left[\mu_1 R_1(z_2) + \frac{\mu_2}{z_2}R_2(z_2)\right], \qquad (5)$$

where

$$K(z_1,z_2) := \lambda_1(1-z_1) + \lambda_2(1-z_2) + \mu_1\left(1-\frac{1}{z_1}\right). \qquad (6)$$

As argued in Boxma et al. [6], for each $z_2$ with $|z_2| \leq 1$, the quadratic (in $z_1$) function $K(z_1, z_2)$ has a unique zero $z_{1-}(z_2)$ *inside* the unit circle. It plays a key role in the analysis. We denote it by $\pi(z_2)$, and observe that it can be interpreted as PGF of the number of Class-2 arrivals during a busy period $P_1$ of Class-1 customers in isolation: $\pi(z_2) = \mathbb{E}[e^{-\lambda_2(1-z_2)P_1}]$. We conclude that, because $P_1(z_1, z_2)$ is analytic in the product of the two unit circles, the righthand side of Equation (5) must also be zero for $z_1 = \pi(z_2)$. Hence,

$$P_2(\pi(z_2), z_2) = R_2(z_2) + \frac{\mu_1}{\mu_2} z_2 R_1(z_2) - \frac{\lambda_1 \pi(z_2) z_2}{\mu_2} p(0,0). \tag{7}$$

Taking generating functions in the third and fourth balance equations yields:

$$\begin{aligned} & P_2(z_1, z_2)[\lambda_1 + \lambda_2 + \mu_2] = \lambda_1 z_1 P_2(z_1, z_2) + \lambda_2 z_2 P_2(z_1, z_2) \\ + \quad & \mu_1 \sum_{j=1}^{\infty} z_2^j p_1(1, j) + \frac{\mu_2}{z_2} \sum_{j=1}^{\infty} z_2^{j+1} p_2(0, j+1) + \lambda_2 z_2 p(0,0), \end{aligned} \tag{8}$$

which yields (using (7) for the last equality):

$$\begin{aligned} & P_2(z_1, z_2)\,[\lambda_1(1 - z_1) + \lambda_2(1 - z_2) + \mu_2] \\ = \quad & \mu_1[R_1(z_2) - p_1(1,0)] + \frac{\mu_2}{z_2}[R_2(z_2) - z_2 p_2(0,1)] + \lambda_2 z_2 p(0,0) \\ = \quad & \frac{\mu_2}{z_2} P_2(\pi(z_2), z_2) - \mu_1 p_1(1,0) - \mu_2 p_2(0,1) + \lambda_1 \pi(z_2) p(0,0) + \lambda_2 z_2 p(0,0). \end{aligned} \tag{9}$$

Finally, using the fifth balance equation, we get

$$\begin{aligned} & P_2(z_1, z_2) = [\lambda_1(1 - z_1) + \lambda_2(1 - z_2) + \mu_2]^{-1} \\ \times \quad & \left[ \frac{\mu_2}{z_2} P_2(\pi(z_2), z_2) + [\lambda_1(\pi(z_2) - 1) + \lambda_2(z_2 - 1)] p(0,0) \right]. \end{aligned} \tag{10}$$

Observe that $P_2(z_1, z_2)$, the steady state PGF of the numbers of customers of both classes during the service of a Class-2 customer, can be written as $\mu_2/[\lambda_1(1 - z_1) + \lambda_2(1 - z_2) + \mu_2]$ times a function of $z_2$ only. This may be explained as follows: The first factor is the PGF of the numbers of arrivals of both customer classes during the past part of the $\exp(\mu_2)$ service time, while at the beginning of that Class-2 service there were no Class-1 customers present.

Taking $z_1 = \pi(z_2)$ in (10) gives, when using that $\lambda_1(1-\pi(z_2)) + \lambda_2(1-z_2) = -\mu_1\left(1-\frac{1}{\pi(z_2)}\right)$:

$$P_2(\pi(z_2), z_2) = p(0,0)\frac{\lambda_1(\pi(z_2)-1) + \lambda_2(z_2-1)}{\mu_2\left(1-\frac{1}{z_2}\right) - \mu_1\left(1-\frac{1}{\pi(z_2)}\right)}. \tag{11}$$

It follows from the normalizing condition that $p(0,0) = 1-\rho$, but we could also conclude it from first principles. Having determined $P_2(\pi(z_2), z_2)$, we obtain $P_2(z_1, z_2)$ from (10):

$$\begin{aligned} P_2(z_1,z_2) &= p(0,0)[\lambda_1(1-z_1) + \lambda_2(1-z_2) + \mu_2]^{-1} \\ &\times \left[\lambda_1(\pi(z_2)-1) + \lambda_2(z_2-1) + \frac{\mu_2}{z_2}\frac{\mu_1\left(1-\frac{1}{\pi(z_2)}\right)}{\mu_2\left(1-\frac{1}{z_2}\right) - \mu_1\left(1-\frac{1}{\pi(z_2)}\right)}\right]. \end{aligned} \tag{12}$$

Finally, $P_1(z_1, z_2)$ follows using (5) and (7):

$$\begin{aligned} P_1(z_1,z_2) &= \left[\lambda_1(1-z_1) + \lambda_2(1-z_2) + \mu_1\left(1-\frac{1}{z_1}\right)\right]^{-1} \\ &\times \left[\frac{\mu_2}{z_2}(P_2(z_1,z_2) - P_2(\pi(z_2),z_2)) + \lambda_1 p(0,0)[z_1 - \pi(z_2)]\right] \\ &= \left[\lambda_1(1-z_1) + \lambda_2(1-z_2) + \mu_1\left(1-\frac{1}{z_1}\right)\right]^{-1} \\ &\quad [p(0,0)\lambda_1(z_1 - \pi(z_2)) \\ &+ p(0,0)\frac{\mu_2}{z_2}\frac{\lambda_1(1-\pi(z_2)) + \lambda_2(1-z_2)}{\lambda_1(1-z_1) + \lambda_2(1-z_2) + \mu_2}\frac{\lambda_1(1-z_1) + \lambda_2(1-z_2) + \mu_1\left(1-\frac{1}{\pi(z_2)}\right)}{\mu_2(1-\frac{1}{z_2}) - \mu_1\left(1-\frac{1}{\pi(z_2)}\right)}]. \end{aligned} \tag{13}$$

In the last step, we use Equations (11) and (12).

*Remark.* If $\mu_1 \uparrow \infty$, then $P_1(z_1, z_2)$ tends to zero (not surprising, as the probability of a Class-1 customer being in service tends to zero), while we obtain after some rewriting that

$$\lim_{\mu_1\uparrow\infty} P_2(z_1,z_2) = \rho_2\frac{\mu_2}{\lambda_1(1-z_1) + \lambda_2(1-z_2) + \mu_2}\left[\frac{1-\rho_2}{1-\rho_2 z_2} - (1-\rho_2)(1-z_2)\right], \tag{14}$$

and hence

$$\lim_{\mu_1\uparrow\infty} P_2(z_1,1) = \rho_2 \frac{\mu_2}{\lambda_1(1-z_1)+\mu_2}, \tag{15}$$

$$\lim_{\mu_1\uparrow\infty} P_2(1,z_2) = \frac{(1-\rho_2)\rho_2 z_2}{1-\rho_2 z_2}. \tag{16}$$

The various expressions can be interpreted. For example, the term outside the large brackets in (14) is the probability that a Class-2 customer is in service times the joint PGF of the numbers of arrivals of both classes during the past part (still $\exp(\mu_2)$!) of a Class-2 service (which also immediately explains (15)); and the PGF in (16) is recognized as the PGF of the number of customers in a non-empty $M_{\lambda_2}/M_{\mu_2}/1$ queue.

If $\mu_2 \uparrow \infty$, then $P_2(z_1,z_2)$ tends to zero, while

$$\lim_{\mu_2\uparrow\infty} P_1(z_1,z_2) = \frac{(1-\rho_1)\lambda_1(z_1-\pi(z_2))}{\lambda_1(1-z_1)+\lambda_2(1-z_2)+\mu_1(1-\frac{1}{z_1})}, \tag{17}$$

$$\lim_{\mu_2\uparrow\infty} P_1(z_1,1) = \frac{(1-\rho_1)\rho_1 z_1}{1-\rho_1 z_1}, \tag{18}$$

$$\lim_{\mu_2\uparrow\infty} P_1(1,z_2) = \rho_1 \frac{1-\pi(z_2)}{(1-z_2)\lambda_2 \mathbb{E}[P_1]}. \tag{19}$$

The limit in (18) is the PGF of the number of customers in a non-empty $M_{\lambda_1}/M_{\mu_1}/1$ queue, while the quotient in (19) is the PGF of the number of Class-2 arrivals during the past part of a busy period of that $M_{\lambda_1}/M_{\mu_1}/1$ queue.

*Remark.* In the special case $\mu_1 = \mu_2 = \mu$, we have

$$P_1(z,z) = \frac{(1-\rho)z}{1-\rho z}\frac{1}{1-z}\left[\rho_1(\pi(z)-z) + \frac{1-\pi(z)}{z-\pi(z)}\left(\frac{1}{1+\rho-\rho z}\frac{1}{\pi(z)} - 1\right)\right], \tag{20}$$

$$P_2(z,z) = \frac{(1-\rho)z}{1+\rho-\rho z}\frac{\pi(z)-1}{\pi(z)(z-\pi(z))}, \tag{21}$$

and after an elaborate calculation, in which we use (see below (6)) that $\lambda_1\pi(z)^2 - (\lambda_1+\lambda_2(1-z)+\mu)\pi(z)+\mu = 0$, we verify that

$$P_1(z,z) + P_2(z,z) + 1 - \rho = \frac{1-\rho}{1-\rho z}, \tag{22}$$

as it should be; indeed, the total number of customers is now distributed as the number of customers in an $M/M/1$ queue with traffic load $\rho$.

Below we give expressions for the marginal queue length PGF's, which we differentiate subsequently to obtain moments. We replace $z_1$ (and thereafter also $z_2$) by $x$ to make the expressions slightly easier to read. We give some details for the simplification of $P_1(x,1)$ and leave similar calculations for the other PGF's to the reader. It follows from (13) that

$$\begin{aligned} P_1(x,1) &= \frac{1-\rho}{\lambda_1(1-x)-\mu_1\frac{1-x}{x}} \\ &\times \left[\lambda_1(x-1)+\frac{\mu_2\lambda_1(1-x)}{\lambda_1(1-x)+\mu_2}\lim_{z_2\to 1}\frac{\lambda_1(1-\pi(z_2))+\lambda_2(1-z_2)}{\mu_2\left(1-\frac{1}{z_2}\right)-\mu_1\left(1-\frac{1}{\pi(z_2)}\right)}\right] \\ &= (1-\rho)\frac{\rho_1 x}{1-\rho_1 x}\left[1+\frac{\rho_2}{1-\rho}\frac{\mu_2}{\mu_2+\lambda_1(1-x)}\right]. \end{aligned} \tag{23}$$

In the second step, we divide the denominator and the term between square brackets by $\lambda_1(1-x)$, and subsequently we use l'Hopital's rule and $\pi'(1) = \frac{\lambda_2}{\mu_1(1-\rho_1)}$ to show that

$$\lim_{z_2\to 1}\frac{\lambda_1(1-\pi(z_2))+\lambda_2(1-z_2)}{\mu_2\left(1-\frac{1}{z_2}\right)-\mu_1\left(1-\frac{1}{\pi(z_2)}\right)} = \frac{\lambda_2}{\lambda_2-\mu_2(1-\rho_1)} = \frac{\rho_2}{\rho-1}. \tag{24}$$

Next comes $P_1(1,x)$. After some rewriting, we obtain

$$P_1(1,x) = (1-\rho)\frac{\lambda_1}{\lambda_2}\frac{1-\pi(x)}{1-x} + \frac{1-\rho}{\rho_2}\frac{\lambda_1\frac{1-\pi(x)}{1-x}+\lambda_2}{\mu_2+\lambda_2(1-x)}\frac{\lambda_2-\mu_1\frac{1-\pi(x)}{1-x}\frac{1}{\pi(x)}}{-\mu_2+\mu_1\frac{1-\pi(x)}{1-x}\frac{x}{\pi(x)}}. \tag{25}$$

However, for the determination of moments, the following representation seems better:

$$P_1(1,x) = (1-\rho)\frac{\lambda_1(1-\pi(x))}{\lambda_2(1-x)}\left[1+\frac{\mu_2}{\mu_2+\lambda_2(1-x)}\frac{1}{-x+\frac{\mu_2}{\mu_1}\frac{1-x}{1-\pi(x)}\pi(x)}\right]. \tag{26}$$

It follows from (12), using l'Hopital's rule, that

$$P_2(x,1) = \frac{\lambda_2}{\mu_2+\lambda_1(1-x)}, \tag{27}$$

$$P_2(1,x) = \frac{1-\rho}{\mu_2 + \lambda_2(1-x)} \left[ \lambda_1(\pi(x)-1) + \lambda_2(x-1) + \frac{\mu_2}{-x + \frac{\mu_2}{\mu_1}\pi(x)\frac{1-x}{1-\pi(x)}} \right]. \tag{28}$$

It readily follows from these equations that $P_1(1,1) = \rho_1$ (again use that $\lim_{x\uparrow 1} \frac{1-\pi(x)}{1-x} = \pi'(1) = \frac{\lambda_2}{\mu_1(1-\rho_1)}$) and $P_2(1,1) = \rho_2$, as expected; indeed, these are the fractions of time that the server is serving Class-1 respectively Class-2 customers.

## 2.2 Queue length moments

It follows from (23) that

$$\mathbb{E}[\mathcal{Q}_1 I(\text{Class 1 in service})] = \frac{\mathrm{d}}{\mathrm{dx}} P_1(x,1)|_{x=1} = \frac{\rho_1}{1-\rho_1} + \frac{\lambda_1}{\mu_2}\frac{\rho_1\rho_2}{1-\rho_1}, \tag{29}$$

and from (27) that

$$\mathbb{E}[\mathcal{Q}_1 I(\text{Class 2 in service})] = \frac{\mathrm{d}}{\mathrm{dx}} P_2(x,1)|_{x=1} = \frac{\lambda_1}{\mu_2}\rho_2; \tag{30}$$

from (26) (using that $\pi''(1) = \lambda_2^2 \mathbb{E}[P_1^2] = \frac{2\lambda_2^2}{\mu_1^2(1-\rho_1)^3}$ ) that

$$\begin{aligned} &\mathbb{E}[\mathcal{Q}_2 I(\text{Class 1 in service})] = \frac{\mathrm{d}}{\mathrm{dx}} P_1(1,x)|_{x=1} \qquad (31)\\ = \;& \frac{\rho_1\rho_2^2}{1-\rho_1}\frac{2-\rho}{1-\rho} + \frac{\lambda_2}{\mu_1}\frac{\rho_1}{(1-\rho_1)^2}\left[1 + \frac{\rho_1\rho_2}{1-\rho}\right], \end{aligned}$$

and from (28) that

$$\begin{aligned} &\mathbb{E}[\mathcal{Q}_2 I(\text{Class 2 in service})] = \frac{\mathrm{d}}{\mathrm{dx}} P_2(1,x)|_{x=1}\\ = \;& \rho_2^2 + \rho_2\frac{1-\rho}{1-\rho_1} + \frac{\rho_2^2}{1-\rho}\left[1 + \frac{\mu_2}{\mu_1}\frac{\rho_1}{1-\rho_1}\right]. \qquad (32) \end{aligned}$$

It follows from (29) and (30) that

$$\mathbb{E}[\mathcal{Q}_1] = \frac{\rho_1}{1-\rho_1} + \frac{\lambda_1}{\mu_2}\frac{\rho_2}{1-\rho_1}, \tag{33}$$

and from (31) and (32) that

$$\mathbb{E}[\mathcal{Q}_2] = \lambda_2 \frac{\frac{\rho_1}{\mu_1} + \frac{\rho_2}{\mu_2}}{(1-\rho)(1-\rho_1)} + \rho_2. \tag{34}$$

Equations (33) and (34) are in agreement with known results for the $M/G/1$ queue with non-preemptive priority for Class 1, cf. Chapter 9 of Adan and Resing [1].

We close this section by determining the second factorial moments of the queue lengths, viz., $\mathbb{E}[\mathcal{Q}_i(\mathcal{Q}_i - 1)I(\text{Class } j \text{ in service})]$, for $i, j = 1, 2$. This is done by twice differentiating the four generating functions given in (23), (26), (27) and (28) and then taking $x = 1$. In Section 3.3 we show how to express the second factorial moments of overlap numbers into these second factorial moments of queue lengths.
We have:

$$\begin{aligned}\mathbb{E}[\mathcal{Q}_1(\mathcal{Q}_1 - 1)I(\text{Class 1 in service})] &= \frac{\mathrm{d}^2}{\mathrm{d}x^2} P_1(x,1)|_{x=1} \\ &= \frac{2\rho_1^2}{(1-\rho_1)^2} + 2\frac{\lambda_1}{\mu_2}\frac{\rho_1\rho_2}{(1-\rho_1)^2} + 2\left(\frac{\lambda_1}{\mu_2}\right)^2 \frac{\rho_1\rho_2}{1-\rho_1},\end{aligned} \tag{35}$$

$$\mathbb{E}[\mathcal{Q}_1(\mathcal{Q}_1 - 1)I(\text{Class 2 in service})] = \frac{\mathrm{d}^2}{\mathrm{d}x^2} P_2(x,1)|_{x=1} = 2\left(\frac{\lambda_1}{\mu_2}\right)^2 \rho_2. \tag{36}$$

The other two second factorial moments require a quite lengthy calculation, in which one needs to determine the following first and second derivatives (cf. (26) and (28)):

$$\frac{\mathrm{d}}{\mathrm{d}x} \frac{1}{-x + \frac{\mu_2}{\mu_1}\frac{x-1}{\pi(x)-1}\pi(x)}|_{x=1} = \frac{\rho_2^2}{(1-\rho)^2}\left[1 + \frac{\mu_2}{\mu_1}\frac{\rho_1}{1-\rho_1}\right], \tag{37}$$

$$\frac{\mathrm{d}^2}{\mathrm{d}x^2} \frac{1}{-x + \frac{\mu_2}{\mu_1}\frac{x-1}{\pi(x)-1}\pi(x)}|_{x=1} = 2\left(\frac{\lambda_2}{\mu_1}\right)^2 \frac{\rho_2}{(1-\rho)^2}\frac{\rho_1}{(1-\rho_1)^3} + 2\frac{\rho_2}{(1-\rho)^3}\left(\rho_2 + \frac{\lambda_2}{\mu_1}\frac{\rho_1}{1-\rho_1}\right)^2. \tag{38}$$

Here we have not only used that $\pi(1) = 1$, $\pi'(1) = \frac{\lambda_2}{\mu_1(1-\rho_1)}$ and $\pi''(1) = \frac{2\lambda_2^2}{\mu_1^2(1-\rho_1)^3}$, but also that $\pi'''(1) = \lambda_2^3\mathbb{E}[P_1^3] = 6(\frac{\lambda_2}{\mu_1})^3\frac{1+\rho_1}{(1-\rho_1)^5}$ (cf. p. 251 of Cohen

[7]). We thus get

$$\begin{aligned}
&\mathbb{E}[\mathcal{Q}_2(\mathcal{Q}_2-1)I(\text{Class 1 in service})] = \frac{\mathrm{d}^2}{\mathrm{d}x^2}P_1(1,x)|_{x=1} \\
= \;& \frac{\lambda_1}{\lambda_2}[2\left(\frac{\lambda_2}{\mu_1}\right)^3\frac{1-\rho_1^2}{(1-\rho_1)^4} + 2\left(\frac{\lambda_2}{\mu_1}\right)^2\frac{\rho_2^2}{(1-\rho_1)^4} \\
+ \;& 2\left(\frac{\lambda_2}{\mu_1}\right)^2\frac{\rho_2^2}{(1-\rho)(1-\rho_1)^3}\left(1+\frac{\mu_2}{\mu_1}\frac{\rho_1}{1-\rho_1}\right) + \frac{\lambda_2}{\mu_1}\frac{1}{1-\rho_1}2\frac{(2-\rho)\rho_2^2}{1-\rho} \\
+ \;& 2\left(\frac{\lambda_2}{\mu_1}\right)^2\frac{\rho_1}{(1-\rho_1)^2}\frac{\rho_2^2}{1-\rho} + 2\left(\frac{\lambda_2}{\mu_1}\right)^2\frac{\rho_2}{1-\rho}\frac{\rho_1}{(1-\rho_1)^3} \\
+ \;& 2\frac{\lambda_2}{\mu_1}\frac{\rho_2}{(1-\rho_1)(1-\rho)^2}\left(\rho_2+\frac{\lambda_2}{\mu_1}\frac{\rho_1}{1-\rho_1}\right)^2], \qquad (39)
\end{aligned}$$

$$\begin{aligned}
&\mathbb{E}[\mathcal{Q}_2(\mathcal{Q}_2-1)I(\text{Class 2 in service})] = \frac{\mathrm{d}^2}{\mathrm{d}x^2}P_2(1,x)|_{x=1} \\
= \;& 2\rho_2^3 + 2\rho_2^2\frac{1-\rho}{1-\rho_1} + 2\frac{\rho_2^3}{1-\rho}\left[1+\frac{\mu_2}{\mu_1}\frac{\rho_1}{1-\rho_1}\right] + 2\frac{\lambda_2}{\mu_1}\frac{\rho_1\rho_2(1-\rho)}{(1-\rho_1)^3} \\
+ \;& 2\left(\frac{\lambda_2}{\mu_1}\right)^2\frac{\rho_2}{1-\rho}\frac{\rho_1}{(1-\rho_1)^3} + 2\frac{\rho_2}{(1-\rho)^2}\left(\rho_2+\frac{\lambda_2}{\mu_1}\frac{\rho_1}{1-\rho_1}\right)^2. \qquad (40)
\end{aligned}$$

# 3 The numbers of overlaps

Our goal in Subsection 3.1 is to determine the PGF of the numbers of overlaps of Class-1 customers with customers of both classes; in Subsection 3.2 we aim to do the same for Class-2 customers. We denote the number of overlapping Class-$j$ customers with a tagged Class-$i$ customer by $\mathcal{O}_{ij}$, $i,j=1,2$. Their first two moments are determined in Subsection 3.3.

## 3.1 The numbers of overlap for Class-1 customers

Suppose that the arriving Class-1 customer $C_1$ sees $i \geq 1$ Class-1 and $j$ Class-2 customers present, and that a Class-$k$ customer is being served; we say that $C_1$ sees $(i,j;k)$, $k=1,2$. Denote the service time of the $i$-th Class-$k$ customer by $\mathcal{S}_k^{(i)}$, $k=1,2$, and denote $C_1$'s service time by $\mathcal{S}_1^{(0)}$. If $k=1$, then $C_1$'s sojourn time consists of the sum $\mathcal{S}_1^{(1)}+\cdots+\mathcal{S}_1^{(i)}+\mathcal{S}_1^{(0)}$ of i.i.d.

service times which are $\exp(\mu_1)$ distributed (notice that the first service time is a residual service time, but still has the same distribution). Let $N_i(t)$ denote the number of Class-$i$ arrivals in an interval of length $t$, $i = 1, 2$; its PGF is $\mathrm{e}^{-\lambda_i(1-z)t}$. It is readily seen that

$$\mathbb{E}\left[z_1^{N_1\left(\mathcal{S}_1^{(1)}\right)} z_2^{N_2\left(\mathcal{S}_1^{(1)}\right)}\right] = \frac{\mu_1}{\mu_1 + \lambda_1(1-z_1) + \lambda_2(1-z_2)}, \tag{41}$$

and hence we have the following Proposition.

**Proposition 3.1.** *The conditional PGF of the steady-state joint distribution of the numbers of overlaps $(\mathcal{O}_{11}, \mathcal{O}_{12})$ of a Class-1 customer in the two-class $M/M/1$ queue with non-preemptive priority for Class 1 is given by*

$$\mathbb{E}\left[z_1^{\mathcal{O}_{11}} z_2^{\mathcal{O}_{12}} | C_1 \text{ sees } (i,j;1)\right] = z_1^i z_2^j \left(\frac{\mu_1}{\mu_1 + \lambda_1(1-z_1) + \lambda_2(1-z_2)}\right)^{i+1} \tag{42}$$

*and*

$$\begin{aligned} &\mathbb{E}\left[z_1^{\mathcal{O}_{11}} z_2^{\mathcal{O}_{12}} | C_1 \text{ sees } (i,j;2)\right] \\ &= z_1^i z_2^j \frac{\mu_2}{\mu_2 + \lambda_1(1-z_1) + \lambda_2(1-z_2)} \left(\frac{\mu_1}{\mu_1 + \lambda_1(1-z_1) + \lambda_2(1-z_2)}\right)^{i+1}. \end{aligned} \tag{43}$$

*Proof.* We start with the conditional PGF when there is a Class-1 customer in service. Thus, we have

$$\begin{aligned} \mathbb{E}\left[z_1^{\mathcal{O}_{11}} z_2^{\mathcal{O}_{12}} | C_1 \text{ sees } (i,j;1)\right] &= z_1^i z_2^j \mathbb{E}\left[z_1^{N_1\left(\mathcal{S}_1^{(1)}+\cdots+\mathcal{S}_1^{(i)}+\mathcal{S}_1^{(0)}\right)} z_2^{N_2\left(\mathcal{S}_1^{(1)}+\cdots+\mathcal{S}_1^{(i)}+\mathcal{S}_1^{(0)}\right)}\right] \\ &= z_1^i z_2^j \prod_{m=0}^{i} \mathbb{E}\left[z_1^{N_1\left(\mathcal{S}_1^{(m)}\right)} z_2^{N_2\left(\mathcal{S}_1^{(m)}\right)}\right] \\ &= z_1^i z_2^j \left(\frac{\mu_1}{\mu_1 + \lambda_1(1-z_1) + \lambda_2(1-z_2)}\right)^{i+1}. \end{aligned} \tag{44}$$

Now we define the conditional PGF when there is a Class-2 customer in

service. We have

$$\begin{aligned}
&\mathbb{E}\left[z_1^{\mathcal{O}_{11}} z_2^{\mathcal{O}_{12}} | C_1 \text{ sees } (i,j;2)\right] \\
&= z_1^i z_2^j \mathbb{E}\left[z_1^{N_1(\mathcal{S}_2)} z_2^{N_2(\mathcal{S}_2)}\right] \mathbb{E}\left[z_1^{N_1\left(\mathcal{S}_1^{(1)}+\cdots+\mathcal{S}_1^{(i)}+\mathcal{S}_1^{(0)}\right)} z_2^{N_2\left(\mathcal{S}_1^{(1)}+\cdots+\mathcal{S}_1^{(i)}+\mathcal{S}_1^{(0)}\right)}\right] \\
&= z_1^i z_2^j \mathbb{E}\left[z_1^{N_1(\mathcal{S}_2)} z_2^{N_2(\mathcal{S}_2)}\right] \prod_{m=0}^{i} \mathbb{E}\left[z_1^{N_1(\mathcal{S}_1^{(m)})} z_2^{N_2(\mathcal{S}_1^{(m)})}\right] \qquad (45) \\
&= z_1^i z_2^j \frac{\mu_2}{\mu_2 + \lambda_1(1-z_1) + \lambda_2(1-z_2)} \left(\frac{\mu_1}{\mu_1 + \lambda_1(1-z_1) + \lambda_2(1-z_2)}\right)^{i+1}.
\end{aligned}$$

□

Proposition 3.1 now enables us to prove the following theorem for the PGF of the number of overlaps experienced by a class-1 customer.

**Theorem 3.2.** *The PGF of the steady-state joint distribution of the numbers of overlaps $(\mathcal{O}_{11}, \mathcal{O}_{12})$ of a Class-1 customer in the two-class $M/M/1$ queue with non-preemptive priority for Class 1 is given by*

$$\begin{aligned}
\mathbb{E}[z_1^{\mathcal{O}_{11}} z_2^{\mathcal{O}_{12}}] &= \frac{\mu_1}{\mu_1 + \lambda_1(1-z_1) + \lambda_2(1-z_2)} P_1\left(\frac{z_1\mu_1}{\mu_1 + \lambda_1(1-z_1) + \lambda_2(1-z_2)}, z_2\right) \\
&+ \frac{\mu_1}{\mu_1 + \lambda_1(1-z_1) + \lambda_2(1-z_2)} \frac{\mu_2}{\mu_2 + \lambda_1(1-z_1) + \lambda_2(1-z_2)} \\
&\times P_2\left(\frac{z_1\mu_1}{\mu_1 + \lambda_1(1-z_1) + \lambda_2(1-z_2)}, z_2\right) \\
&+ (1-\rho)\frac{\mu_1}{\mu_1 + \lambda_1(1-z_1) + \lambda_2(1-z_2)}, \qquad (46)
\end{aligned}$$

*where $P_1(z_1, z_2)$ is given in (13) and $P_2(z_1, z_2)$ in (12).*

*Proof.* First note that we can break the joint PGF into the following infinite sum.

$$
\begin{aligned}
\mathbb{E}\left[z_1^{\mathcal{O}_{11}} z_2^{\mathcal{O}_{12}}\right] &= \sum_{i=0}^{\infty}\sum_{j=0}^{\infty} \mathbb{E}\left[z_1^{\mathcal{O}_{11}} z_2^{\mathcal{O}_{12}} | \mathcal{Q}_1 = i, \mathcal{Q}_2 = j; 1\right] \mathbb{P}(\mathcal{Q}_1 = i, \mathcal{Q}_2 = j; 1) \\
&+ \sum_{i=0}^{\infty}\sum_{j=0}^{\infty} \mathbb{E}\left[z_1^{\mathcal{O}_{11}} z_2^{\mathcal{O}_{12}} | \mathcal{Q}_1 = i, \mathcal{Q}_2 = j; 2\right] \mathbb{P}(\mathcal{Q}_1 = i, \mathcal{Q}_2 = j; 2) \\
&+ \sum_{i=0}^{\infty}\sum_{j=0}^{\infty} \mathbb{E}\left[z_1^{\mathcal{O}_{11}} z_2^{\mathcal{O}_{12}} | \mathcal{Q}_1 = i, \mathcal{Q}_2 = j; 0\right] \mathbb{P}(\mathcal{Q}_1 = i, \mathcal{Q}_2 = j; 0) \\
&= \sum_{i=0}^{\infty}\sum_{j=0}^{\infty} z_1^i z_2^j \left(\frac{\mu_1}{\mu_1 + \lambda_1(1-z_1) + \lambda_2(1-z_2)}\right)^{i+1} \mathbb{P}(\mathcal{Q}_1 = i, \mathcal{Q}_2 = j; 1) \\
&+ \sum_{i=0}^{\infty}\sum_{j=0}^{\infty} z_1^i z_2^j \frac{\mu_2}{\mu_2 + \lambda_1(1-z_1) + \lambda_2(1-z_2)} \\
&\times \left(\frac{\mu_1}{\mu_1 + \lambda_1(1-z_1) + \lambda_2(1-z_2)}\right)^{i+1} \mathbb{P}(\mathcal{Q}_1 = i, \mathcal{Q}_2 = j; 2) \\
&+ \sum_{i=0}^{\infty}\sum_{j=0}^{\infty} \left(\frac{\mu_1}{\mu_1 + \lambda_1(1-z_1) + \lambda_2(1-z_2)}\right)^{i+1} \mathbb{P}(\mathcal{Q}_1 = i, \mathcal{Q}_2 = j; 0) \\
&= \frac{\mu_1}{\mu_1 + \lambda_1(1-z_1) + \lambda_2(1-z_2)} P_1\left(\frac{z_1\mu_1}{\mu_1 + \lambda_1(1-z_1) + \lambda_2(1-z_2)}, z_2\right) \\
&+ \frac{\mu_1}{\mu_1 + \lambda_1(1-z_1) + \lambda_2(1-z_2)} \frac{\mu_2}{\mu_2 + \lambda_1(1-z_1) + \lambda_2(1-z_2)} \\
&\times P_2\left(\frac{z_1\mu_1}{\mu_1 + \lambda_1(1-z_1) + \lambda_2(1-z_2)}, z_2\right) \\
&+ (1-\rho)\frac{\mu_1}{\mu_1 + \lambda_1(1-z_1) + \lambda_2(1-z_2)}. \qquad (47)
\end{aligned}
$$

This completes the proof. □

## 3.2 The numbers of overlap for Class-2 customers

**Proposition 3.3.** *The conditional PGF of the steady-state joint distribution of the numbers of overlaps $(\mathcal{O}_{21}, \mathcal{O}_{22})$ of a Class-2 customer in the two-class*

*$M/M/1$ queue with non-preemptive priority for Class 1 is given by*

$$\begin{aligned}
&\mathbb{E}\left[z_1^{\mathcal{O}_{21}} z_2^{\mathcal{O}_{22}} | C_2 \text{ sees } (i,j;1)\right] \qquad (48)\\
&= z_2^j g_1^i(z_1, \lambda_2(1-z_2))\\
&\times \left(\frac{\mu_2}{\mu_2 + \lambda_1(1 - g_1(z_1, \lambda_2(1-z_2))) + \lambda_2(1-z_2)}\right)^j \frac{\mu_2}{\mu_2 + \lambda_1(1-z_1) + \lambda_2(1-z_2)},
\end{aligned}$$

*and*

$$\begin{aligned}
&\mathbb{E}\left[z_1^{\mathcal{O}_{21}} z_2^{\mathcal{O}_{22}} | C_2 \text{ sees } (i,j;2)\right] \qquad (49)\\
&= z_2^j g_1^i(z_1, \lambda_2(1-z_2))\\
&\times \left(\frac{\mu_2}{\mu_2 + \lambda_1(1 - g_1(z_1, \lambda_2(1-z_2))) + \lambda_2(1-z_2)}\right)^j \frac{\mu_2}{\mu_2 + \lambda_1(1-z_1) + \lambda_2(1-z_2)}.
\end{aligned}$$

*Proof.* Similar to Proposition 3.1, we find that if a Class-2 customer $C_2$ sees $(i,j;1)$, then its sojourn time consists of the sum of $i$ Class-1 busy periods and $j$ Class-2 *completion times* $CT_2$ plus its own service time. A completion time consists of a Class-2 service time plus the Class-1 busy periods generated by all Class-1 arrivals during that service time. With $g_1(z,s)$ denoting the joint transform of the number of customers $N_1$ served in a busy period of Class 1 and the length $P_1$ of that busy period, we have that the PGF of the numbers of arrivals of both classes during such a busy period $P_1$, including the Class-1 customer who started the busy period, is given by $g_1(z_1, \lambda_2(1-z_2))$. The PGF of the numbers of arrivals of both classes during $CT_2$ is then given by

$$\begin{aligned}
&\mathbb{E}\left[z_1^{N_1(CT_2)} z_2^{N_2(CT_2)}\right]\\
&= \int_{t=0}^{\infty} \mu_2 \mathrm{e}^{-\mu_2 t} \sum_{j=0}^{\infty} \mathrm{e}^{-\lambda_1 t} \frac{(\lambda_1 t)^j}{j!} g_1(z_1, \lambda_2(1-z_2))^j \sum_{k=0}^{\infty} \mathrm{e}^{-\lambda_2 t} \frac{(\lambda_2 t)^k}{k!} z_2^k \mathrm{d}t\\
&= \frac{\mu_2}{\mu_2 + \lambda_1(1 - g_1(z_1, \lambda_2(1-z_2))) + \lambda_2(1-z_2)}. \qquad (50)
\end{aligned}$$

Similarly, the PGF of the numbers of arrivals of both classes during $C_2$'s own service is $\frac{\mu_2}{\mu_2 + \lambda_1(1-z_1) + \lambda_2(1-z_2)}$. Hence (48) follows.

If $C_2$ sees $(i,j;2)$, then the remaining $\exp(\mu_2)$ ongoing service time is not preempted, but still this remaining service time gives rise to a completion time – because Class-1 arrivals during it will lead to Class-1 busy periods which are handled *before* $C_2$ is taken into service. Hence the above reasoning

leading to (48) can be repeated, and the expression in (49) is identical to that in (48). □

Using the conditional pgf of Proposition 3.3, we arrive at the following theorem.

**Theorem 3.4.** *The PGF of the steady-state joint distribution of the numbers of overlaps $(\mathcal{O}_{21}, \mathcal{O}_{22})$ of a Class-2 customer in the two-class $M/M/1$ queue with non-preemptive priority for Class 1 is given by*

$$\begin{aligned}
\mathbb{E}[z_1^{\mathcal{O}_{21}} z_2^{\mathcal{O}_{22}}] &= \frac{\mu_2}{\mu_2 + \lambda_1(1-z_1) + \lambda_2(1-z_2)} \\
&\times P_1\left(g_1(z_1, \lambda_2(1-z_2)), \frac{\mu_2 z_2}{\mu_2 + \lambda_1(1 - g_1(z_1, \lambda_2(1-z_2))) + \lambda_2(1-z_2)}\right) \\
&+ \frac{\mu_2}{\mu_2 + \lambda_1(1-z_1) + \lambda_2(1-z_2)} \\
&\times P_2\left(g_1(z_1, \lambda_2(1-z_2)), \frac{\mu_2 z_2}{\mu_2 + \lambda_1(1 - g_1(z_1, \lambda_2(1-z_2))) + \lambda_2(1-z_2)}\right) \\
&+ (1-\rho)\frac{\mu_2}{\mu_2 + \lambda_1(1-z_1) + \lambda_2(1-z_2)}, \qquad (51)
\end{aligned}$$

*where $P_1(z_1, z_2)$ is given in (13) and $P_2(z_1, z_2)$ in (12).*

*Proof.* Deconditioning in (48), respectively in (49), gives us the first, respectively second, term in the righthand side of (51). Finally, if an arriving Class-2 customer sees an empty system (which happens with probability $1-\rho$), we get an overlap contribution $\frac{\mu_2}{\mu_2+\lambda_1(1-z_1)+\lambda_2(1-z_2)}$. □

## 3.3 Moments

In this subsection we determine means and second moments of overlap numbers from the PGF's in Propositions 3.1 and 3.3. It should be noticed that *mean* values can also be obtained, even in greater generality, from an MVA (Mean Value Analysis) approach; for this we refer to Section 5. However, to obtain higher moments one really needs to take resort to the PGF's of Subsections 3.1 and 3.2.

We first express the mean overlap numbers $\mathbb{E}[\mathcal{O}_{ij}]$ (for the cases that $C_1$ sees a Class-1 customer in service upon arrival, or a Class-2, or an empty system) in terms of the mean queue lengths of both customer classes.

**Proposition 3.5.** *The first moments of $\mathcal{O}_{11}$, $\mathcal{O}_{12}$, have the following expressions in terms of the queue lengths $\mathcal{Q}_1$, $\mathcal{Q}_2$.*

$$\begin{aligned}\mathbb{E}[\mathcal{O}_{11}I(\mathrm{C}_1 \text{ sees } 1)] &= \rho_1\left[\rho_1 + (1+\rho_1)\mathbb{E}[\mathcal{Q}_1|\text{Class 1 in service}]\right] \\ &= \frac{\rho_1[1+2\rho_1-\rho_1^2]}{1-\rho_1} + \frac{\lambda_1}{\mu_2}\frac{\rho_1\rho_2(1+\rho_1)}{1-\rho_1}, \end{aligned} \tag{52}$$

$$\begin{aligned}\mathbb{E}[\mathcal{O}_{12}I(\mathrm{C}_1 \text{ sees } 1)] &= \rho_1\left[\mathbb{E}[\mathcal{Q}_2|\text{Class 1 in service}]\right] \\ &+ \rho_1[\frac{\lambda_2}{\mu_1}(\mathbb{E}[\mathcal{Q}_1|\text{Class 1 in service}]+1)] \\ &= \frac{\rho_1\rho_2^2}{1-\rho_1}\frac{2-\rho}{1-\rho} + \frac{\lambda_2}{\mu_1}\frac{\rho_1}{(1-\rho_1)^2}[1+\frac{\rho_1\rho_2}{1-\rho}] \\ &+ \frac{\lambda_2}{\mu_1}\left(\frac{2\rho_1-\rho_1^2}{1-\rho_1} + \frac{\lambda_1}{\mu_2}\frac{\rho_1\rho_2}{1-\rho_1}\right), \end{aligned} \tag{53}$$

$$\begin{aligned}\mathbb{E}[\mathcal{O}_{11}I(\mathrm{C}_1 \text{ sees } 2)] &= \rho_2\left[\frac{\lambda_1}{\mu_2} + \rho_1 + (1+\rho_1)\mathbb{E}[\mathcal{Q}_1|\text{Class 2 in service}]\right] \\ &= \rho_1\rho_2 + \frac{\lambda_1}{\mu_2}(\rho_2 + (1+\rho_1)\rho_2), \end{aligned} \tag{54}$$

$$\begin{aligned}\mathbb{E}[\mathcal{O}_{12}I(\mathrm{C}_1 \text{ sees } 2)] &= \rho_2\left[\rho_2 + \frac{\lambda_2}{\mu_1} + \mathbb{E}[\mathcal{Q}_2|\text{Class 2 in service}]\right] \\ &+ \rho_2\left[\frac{\lambda_2}{\mu_1}\mathbb{E}[\mathcal{Q}_1|\text{Class 2 in service}]\right] \\ &= \rho_2^2 + \rho_2\frac{\lambda_2}{\mu_1} \\ &+ \rho_2^2 + \rho_2\frac{1-\rho}{1-\rho_1} + \frac{\rho_2^2}{1-\rho}[1+\frac{\mu_2}{\mu_1}\frac{\rho_1}{1-\rho_1}] + \frac{\lambda_2}{\mu_1}\frac{\lambda_1}{\mu_2}\rho_2, \end{aligned} \tag{55}$$

$$\mathbb{E}[\mathcal{O}_{11}I(\mathrm{C}_1 \text{ sees } (0,0))] = (1-\rho)\frac{\lambda_1}{\mu_1}, \tag{56}$$

$$\mathbb{E}[\mathcal{O}_{12}I(\mathrm{C}_1 \text{ sees } (0,0))] = (1-\rho)\frac{\lambda_2}{\mu_1}. \tag{57}$$

*Proof.* Differentiating (42) and (43) once yields:

$$\mathbb{E}[\mathcal{O}_{11}|\mathrm{C}_1 \text{ sees } (\mathrm{i},\mathrm{j},1)] = i + \rho_1(i+1), \tag{58}$$

$$\mathbb{E}[\mathcal{O}_{12}|\mathrm{C}_1 \text{ sees } (\mathrm{i},\mathrm{j},1)] = j + \frac{\lambda_2}{\mu_1}(i+1). \tag{59}$$

$$\mathbb{E}[\mathcal{O}_{11}|\mathrm{C}_1 \text{ sees } (\mathrm{i},\mathrm{j},2)] = i + \frac{\lambda_1}{\mu_2} + \rho_1(i+1), \tag{60}$$

$$\mathbb{E}[\mathcal{O}_{12}|\mathrm{C}_1 \text{ sees } (\mathrm{i},\mathrm{j},2)] = j + \rho_2 + \frac{\lambda_2}{\mu_1}(i+1). \tag{61}$$

Now use the fact that

$$\mathbb{E}[\mathcal{O}_{11} I(\text{Class 1 in service})] = \sum_{i=1}^{\infty}\sum_{j=0}^{\infty} p_1(i,j)\mathbb{E}[\mathcal{O}_{11}|\mathrm{C}_1 \text{ sees } (\mathrm{i},\mathrm{j},1)], \tag{62}$$

to obtain the first equality in (52) from (58) (notice that $\rho_1$ equals the probability of having a Class-1 customer in service); the second equality follows from (29). In the same way, (53), (54) and (55) are obtained, successively using (59), (60) and (61) to obtain their first equality and (31), (30) and (32) to obtain their second equality. Notice that one could simplify some of the expressions a bit further, for example using that $\frac{\lambda_2}{\mu_1}\frac{\lambda_1}{\mu_2} = \rho_1\rho_2$. Finally, (56) and (57) follow by observing that $C_1$ sees an empty system with probability $1-\rho$, and during her service an average of $\frac{\lambda_i}{\mu_1}$ customers of Class $i$ arrive, $i = 1, 2$. □

**Proposition 3.6.** *The first moments of $\mathcal{O}_{21}$, $\mathcal{O}_{22}$ have the following expressions in terms of the mean (conditional) queue lengths of $\mathcal{Q}_1$, $\mathcal{Q}_2$.*

$$\begin{aligned}\mathbb{E}[\mathcal{O}_{21} I(\mathrm{C}_2 \text{ sees } 1)] &= \rho_1\left[\frac{\lambda_1}{\mu_2} + \frac{\mathbb{E}[\mathcal{Q}_1|\text{Class 1 in service}]}{1-\rho_1}\right. \\ &\left. + \frac{\lambda_1}{\mu_2}\frac{\mathbb{E}[\mathcal{Q}_2|\text{Class 1 in service}]}{1-\rho_1}\right] \qquad (63)\\ &= \rho_1\frac{\lambda_1}{\mu_2} + \frac{1}{1-\rho_1}\left[\frac{\rho_1}{1-\rho_1} + \frac{\lambda_1}{\mu_2}\frac{\rho_1\rho_2}{1-\rho_1}\right] \\ &+ \frac{\lambda_1}{\mu_2}\frac{1}{1-\rho_1}\left[\frac{\rho_1\rho_2^2}{1-\rho_1}\frac{2-\rho}{1-\rho} + \frac{\lambda_2}{\mu_1}\frac{\rho_1}{(1-\rho_1)^2}\left[1 + \frac{\rho_1\rho_2}{1-\rho}\right]\right], \qquad (64)\end{aligned}$$

$$\mathbb{E}[\mathcal{O}_{22}I(\mathrm{C}_2 \text{ sees } 1)] = \rho_1 \left[\frac{\lambda_2}{\mu_2} + (1 + \frac{\rho_2}{1-\rho_1})\mathbb{E}[\mathcal{Q}_2|\text{Class 1 in service}] + \frac{\lambda_2}{\mu_1(1-\rho_1)}\mathbb{E}[\mathcal{Q}_1|\text{Class 1 in service}]\right]$$

$$\begin{aligned} &= \rho_1\rho_2 + (1 + \frac{\rho_2}{1-\rho_1})[\frac{\rho_1\rho_2^2}{1-\rho_1}\frac{2-\rho}{1-\rho} + \frac{\lambda_2}{\mu_1}\frac{\rho_1}{(1-\rho_1)^2}[1 + \frac{\rho_1\rho_2}{1-\rho}]] \\ &+ \frac{\lambda_2}{\mu_1(1-\rho_1)}[\frac{\rho_1}{1-\rho_1} + \frac{\lambda_1}{\mu_2}\frac{\rho_1\rho_2}{1-\rho_1}], \end{aligned} \tag{65}$$

$$\mathbb{E}[\mathcal{O}_{21}I(\mathrm{C}_2 \text{ sees } 2)] = \rho_2 \left[\frac{\lambda_1}{\mu_2} + \frac{\mathbb{E}[\mathcal{Q}_1|\text{Class 2 in service}]}{1-\rho_1} + \frac{\lambda_1}{\mu_2}\frac{1}{1-\rho_1}\mathbb{E}[\mathcal{Q}_2|\text{Class 2 in service}]\right]$$

$$\begin{aligned} &= \rho_2\frac{\lambda_1}{\mu_2} + \frac{1}{1-\rho_1}\frac{\lambda_1}{\mu_2}\rho_2 \\ &+ \frac{\lambda_1}{\mu_2}\frac{1}{1-\rho_1}[\rho_2^2 + \rho_2\frac{1-\rho}{1-\rho_1} + \frac{\rho_2^2}{1-\rho}[1 + \frac{\mu_2}{\mu_1}\frac{\rho_1}{1-\rho_1}]], \end{aligned} \tag{66}$$

$$\mathbb{E}[\mathcal{O}_{22}I(\mathrm{C}_2 \text{ sees } 2)] = \rho_2 \left[\rho_2 + (1 + \frac{\rho_2}{1-\rho_1})\mathbb{E}[\mathcal{Q}_2|\text{Class 2 in service}] + \frac{\lambda_2}{\mu_1}\frac{1}{1-\rho_1}\mathbb{E}[\mathcal{Q}_1|\text{Class 2 in service}]\right]$$

$$\begin{aligned} &= \rho_2^2 + (1 + \frac{\rho_2}{1-\rho_1})[\rho_2^2 + \rho_2\frac{1-\rho}{1-\rho_1} + \frac{\rho_2^2}{1-\rho}[1 + \frac{\mu_2}{\mu_1}\frac{\rho_1}{1-\rho_1}]] \\ &+ \frac{\lambda_2}{\mu_1}\frac{1}{1-\rho_1}\frac{\lambda_1}{\mu_2}\rho_2, \end{aligned} \tag{67}$$

$$\mathbb{E}[\mathcal{O}_{21}I(\mathrm{C}_2 \text{ sees}(0,0))] = (1-\rho)\frac{\lambda_1}{\mu_2}, \tag{68}$$

$$\mathbb{E}[\mathcal{O}_{22}I(\mathrm{C}_2 \text{ sees}(0,0))] = (1-\rho)\frac{\lambda_2}{\mu_2}. \tag{69}$$

*Proof.* Differentiating (49) yields:

$$\begin{aligned}\mathbb{E}[\mathcal{O}_{21}|\mathrm{C}_2 \text{ sees } (\mathrm{i},\mathrm{j},1)]] &= \mathbb{E}[\mathcal{O}_{21}|\mathrm{C}_2 \text{ sees } (\mathrm{i},\mathrm{j},2)] \\ &= \frac{i}{1-\rho_1} + j\frac{\lambda_1}{\mu_2}\frac{1}{1-\rho_1} + \frac{\lambda_1}{\mu_2}, \qquad (70)\\ \mathbb{E}[\mathcal{O}_{22}|\mathrm{C}_2 \text{ sees } (\mathrm{i},\mathrm{j},1)]] &= \mathbb{E}[\mathcal{O}_{22}|\mathrm{C}_2 \text{ sees } (\mathrm{i},\mathrm{j},2)] \\ &= j + i\frac{\lambda_2}{\mu_1(1-\rho_1)} + j\frac{\rho_2}{1-\rho_1} + \frac{\lambda_2}{\mu_2}. \qquad (71)\end{aligned}$$

Now proceed exactly as in the proof of Proposition 3.5, combining (70) and (71) with (29)-(32). Finally, (68) and (69) follow by observing that $C_2$ sees an empty system with probability $1-\rho$, and during her service an average of $\frac{\lambda_i}{\mu_2}$ customers of Class $i$ arrive, $i = 1, 2$. □

By summing the expressions in (52), (54) and (56), respectively the expressions in (53), (55) and (57), we obtain the means of the overlap numbers $\mathcal{O}_{11}$ respectively $\mathcal{O}_{12}$:

$$\mathbb{E}[\mathcal{O}_{11}] = \frac{2\rho_1}{1-\rho_1} + 2\frac{\lambda_1}{\mu_2}\frac{\rho_2}{1-\rho_1}, \qquad (72)$$

$$\mathbb{E}[\mathcal{O}_{12}] = \rho_2[\frac{(2-\rho)\rho_2}{(1-\rho)(1-\rho_1)} + 1] + \frac{\lambda_2}{\mu_1}\frac{1-\rho_2}{(1-\rho)(1-\rho_1)}. \qquad (73)$$

And by summing the expressions in (64), (66) and (68), respectively the expressions in (65), (67) and (69), we obtain the means of the overlap numbers $\mathcal{O}_{21}$ respectively $\mathcal{O}_{22}$:

$$\mathbb{E}[\mathcal{O}_{21}] = \frac{\lambda_1}{\mu_2} + \frac{1}{1-\rho_1}[\frac{\rho_1}{1-\rho_1} + \frac{\lambda_1}{\mu_2}\frac{\rho_2}{1-\rho_1}] + \frac{\lambda_1}{\mu_2}\frac{1}{1-\rho_1}[\lambda_2\frac{\frac{\rho_1}{\mu_1}+\frac{\rho_2}{\mu_2}}{(1-\rho)(1-\rho_1)} + \rho_2], \qquad (74)$$

$$\mathbb{E}[\mathcal{O}_{22}] = \rho_2 + \frac{\lambda_2}{\mu_1(1-\rho_1)}[\frac{\rho_1}{1-\rho_1} + \frac{\lambda_1}{\mu_2}\frac{\rho_2}{1-\rho_1}] + (1+\frac{\rho_2}{1-\rho_1})[\lambda_2\frac{\frac{\rho_1}{\mu_1}+\frac{\rho_2}{\mu_2}}{(1-\rho)(1-\rho_1)} + \rho_2], \qquad (75)$$

where the terms between square brackets are the means of $\mathcal{Q}_1$ and $\mathcal{Q}_2$ as given in (33) and (34).

Of course one could regroup some terms and simplify, but the present representation makes the calculations easier to follow. It follows after some calculations that (72)-(75) agree with corresponding expressions in Section 5.2.

We now turn to the second moments of the overlap numbers. Differentiating (42) and (43) twice yields:

$$\mathbb{E}[\mathcal{O}_{11}(\mathcal{O}_{11}-1)|\mathrm{C}_1 \text{ sees } (\mathrm{i},\mathrm{j},1))] \;=\; i(i-1)+2i(i+1)\frac{\lambda_1}{\mu_1}+(i+1)(i+2)\frac{\lambda_1^2}{\mu_1^2}, \quad (76)$$

$$\mathbb{E}[\mathcal{O}_{12}(\mathcal{O}_{12}-1)|\mathrm{C}_1 \text{ sees } (\mathrm{i},\mathrm{j},1))] \;=\; j(j-1)+2j(i+1)\frac{\lambda_2}{\mu_1}+(i+1)(i+2)\frac{\lambda_2^2}{\mu_1^2}, \quad (77)$$

$$\begin{aligned}\mathbb{E}[\mathcal{O}_{11}(\mathcal{O}_{11}-1)|\mathrm{C}_1 \text{ sees } (\mathrm{i},\mathrm{j},2))] \;&=\; i(i-1)+2i\frac{\lambda_1}{\mu_2}+2i(i+1)\frac{\lambda_1}{\mu_1}\\ +\;& 2\frac{\lambda_1^2}{\mu_2^2}+2\frac{\lambda_1}{\mu_2}(i+1)\frac{\lambda_1}{\mu_1}+(i+1)(i+2)\frac{\lambda_1^2}{\mu_1^2}, \quad (78)\end{aligned}$$

$$\begin{aligned}\mathbb{E}[\mathcal{O}_{12}(\mathcal{O}_{12}-1)|\mathrm{C}_1 \text{ sees } (\mathrm{i},\mathrm{j},2))] \;&=\; j(j-1)+2j\frac{\lambda_2}{\mu_2}+2j(i+1)\frac{\lambda_2}{\mu_1}\\ +\;& 2\frac{\lambda_2^2}{\mu_2^2}+2\frac{\lambda_2}{\mu_2}(i+1)\frac{\lambda_2}{\mu_1}+(i+1)(i+2)\frac{\lambda_2^2}{\mu_1^2}, \quad (79)\end{aligned}$$

$$\begin{aligned}\mathbb{E}[\mathcal{O}_{11}\mathcal{O}_{12}|\mathrm{C}_1 \text{ sees } (\mathrm{i},\mathrm{j},1))] \;&=\; ij+i(i+1)\frac{\lambda_2}{\mu_1}+j(i+1)\frac{\lambda_1}{\mu_1}\\ +\;& (i+1)(i+2)\frac{\lambda_1}{\mu_1}\frac{\lambda_2}{\mu_1}, \quad (80)\end{aligned}$$

$$\begin{aligned}\mathbb{E}[\mathcal{O}_{11}\mathcal{O}_{12}|\mathrm{C}_1 \text{ sees } (\mathrm{i},\mathrm{j},2))] \;&=\; ij+i\frac{\lambda_2}{\mu_2}+i(i+1)\frac{\lambda_2}{\mu_1}+j\frac{\lambda_1}{\mu_2}+\frac{\lambda_1}{\mu_2}\frac{\lambda_2}{\mu_2}\\ +\;& \frac{\lambda_1}{\mu_2}(i+1)\frac{\lambda_2}{\mu_1}+j(i+1)\frac{\lambda_1}{\mu_1}+\frac{\lambda_2}{\mu_2}(i+1)\frac{\lambda_1}{\mu_1}+(i+1)(i+2)\frac{\lambda_1}{\mu_1}\frac{\lambda_2}{\mu_1}. \quad (81)\end{aligned}$$

Differentiating (48) and (49) twice yields:

$$\begin{aligned}&\mathbb{E}[\mathcal{O}_{21}(\mathcal{O}_{21}-1)|\mathrm{C}_2 \text{ sees } (\mathrm{i},\mathrm{j},1)] \;=\; \mathbb{E}[\mathcal{O}_{21}(\mathcal{O}_{21}-1)|\mathrm{C}_2 \text{ sees } (\mathrm{i},\mathrm{j},2)]\\ =\;& i(i-1)\frac{1}{(1-\rho_1)^2}+2i\frac{\rho_1}{(1-\rho_1)^3}\\ +\;& 2j\frac{\lambda_1}{\mu_2}\frac{\rho_1}{(1-\rho_1)^3}+j(j+1)(\frac{\lambda_1}{\mu_2})^2\frac{1}{(1-\rho_1)^2}+2\frac{\lambda_1^2}{\mu_2^2}\\ +\;& 2ij\frac{\lambda_1}{\mu_2}\frac{1}{(1-\rho_1)^2}+2i\frac{\lambda_1}{\mu_2}\frac{1}{1-\rho_1}+2j(\frac{\lambda_1}{\mu_2})^2\frac{1}{1-\rho_1}, \quad (82)\end{aligned}$$

$$
\begin{aligned}
&\mathbb{E}[\mathcal{O}_{22}(\mathcal{O}_{22}-1)|\mathrm{C}_2 \text{ sees } (\mathrm{i},\mathrm{j},1))] \;=\; \mathbb{E}[\mathcal{O}_{22}(\mathcal{O}_{22}-1)|\mathrm{C}_2 \text{ sees } (\mathrm{i},\mathrm{j},2)] \\
= \;& j(j-1)+i(i-1)(\frac{\lambda_2}{\mu_1(1-\rho_1)})^2+2i\frac{\lambda_2^2}{\mu_1^2(1-\rho_1)^3} \\
+ \;& j(j+1)(\frac{\rho_2}{1-\rho_1})^2+2j\frac{\lambda_1\lambda_2^2}{\mu_1^2\mu_2(1-\rho_1)^3}+2\rho_2^2 \\
+ \;& 2ij\frac{\lambda_2}{\mu_1(1-\rho_1)}+2j^2\frac{\rho_2}{\mu_1(1-\rho_1)}+2j\rho_2+2ij\frac{\lambda_2^2}{\mu_1\mu_2(1-\rho_1)^2} \\
+ \;& 2i\frac{\lambda_2^2}{\mu_1\mu_2(1-\rho_1)}+2j\frac{\lambda_2^2}{\mu_2^2(1-\rho_1)}, \qquad (83)
\end{aligned}
$$

$$
\begin{aligned}
&\mathbb{E}[\mathcal{O}_{21}\mathcal{O}_{22}|\mathrm{C}_1 \text{ sees } (\mathrm{i},\mathrm{j},1))] \;=\; \mathbb{E}[\mathcal{O}_{21}\mathcal{O}_{22}|\mathrm{C}_1 \text{ sees } (\mathrm{i},\mathrm{j},2))] \\
= \;& ij\frac{1}{1-\rho_1}+j^2\frac{\lambda_1}{\mu_2}\frac{1}{1-\rho_1}+j\frac{\lambda_1}{\mu_2}+i(i-1)\frac{\lambda_2}{\mu_1}(\frac{1}{1-\rho_1})^2+i\frac{\lambda_2}{\mu_1}\frac{1+\rho_1}{(1-\rho_1)^3} \\
+ \;& ij\frac{\rho_1\rho_2}{(1-\rho_1)^2}+j\frac{\rho_1\rho_2(1+\rho_1)}{(1-\rho_1)^3}+j(j+1)\frac{\lambda_1}{\mu_2}\frac{\rho_2}{(1-\rho_1)^2}+i\frac{\rho_1\rho_2}{1-\rho_1} \\
+ \;& j\frac{\rho_2}{1-\rho_1}+2\frac{\lambda_1}{\mu_2}\rho_2. \qquad (84)
\end{aligned}
$$

In (84) we have used that $\frac{\mathrm{d}^2}{\mathrm{d}z_1\mathrm{d}z_2}g_1(z_1,\lambda_2(1-z_2))=\lambda_2\mathbb{E}[N_1P_1]=\lambda_2\frac{1+\rho_1}{\mu_1(1-\rho_1)^3}$, where $N_1$ and $P_1$ denote the number of customers in a busy period and its length for an $M/M/1$ queue with arrival rate $\lambda_1$ and service rate $\mu_1$. In (82) we have used that $\mathbb{E}[N_1(N_1-1)]=\frac{2\rho_1}{(1-\rho_1)^3}$, and in (83) that $\mathbb{E}[P_1^2]=\frac{2}{\mu_1^2(1-\rho_1)^3}$.

We now have everything that is needed to determine the unconditional moments $\mathbb{E}[\mathcal{O}_{km}(\mathcal{O}_{km}-1)]$ and $\mathbb{E}[\mathcal{O}_{km}\mathcal{O}_{kl}]$, expressed in terms like $\mathbb{E}[\mathcal{Q}_{km}(\mathcal{Q}_{km}-1)I(\text{Class r in service})]$, $r=1,2$. As the calculations are fairly straightforward, but would occupy quite a bit of space, we just demonstrate how $\mathbb{E}[\mathcal{O}_{11}(\mathcal{O}_{11}-1)I(\text{Class 1 in service})]$ is determined, and leave the other terms to the reader. First rewrite the righthand side of (76) as $i(i-1)(1+\rho_1)^2+4i\rho_1(1+\rho_1)+2\rho_1^2$. Multiplying both sides of (76) by $p_1(i,j)$ and taking the sum $\sum_{i=1}^{\infty}\sum_{j=0}^{\infty}$ we obtain (with $\sum_{i=1}^{\infty}\sum_{j=0}^{\infty}p_1(i,j)=\rho_1$):

$$
\begin{aligned}
&\mathbb{E}[\mathcal{O}_{11}(\mathcal{O}_{11}-1)I(\text{Class 1 in service})] \\
= \;& (1+\rho_1)^2\mathbb{E}[\mathcal{Q}_1(\mathcal{Q}_1-1)I(\text{Class 1 in service})] \\
+ \;& 4\rho_1(1+\rho_1)\mathbb{E}[\mathcal{Q}_1 I(\text{Class 1 in service})]+2\rho_1^3. \qquad (85)
\end{aligned}
$$

# 4 Heavy-traffic limits

In our companion paper Boxma et al. [6], we have derived heavy-traffic limits for the overlap numbers $\mathcal{O}_{ij}$ and $\mathcal{O}_i^T$ in the case of a two-class $M/G/1$ queue with preemptive-resume priority for Class 1. In the present section we briefly sketch how one can also obtain such heavy-traffic limits in the related model with non-preemptive priority for Class 1. A comparison with Theorem 4.1 of Boxma et al. [6] will show that the heavy-traffic behavior of the overlap numbers is the same for both priority disciplines. This is not surprising, since the two priority disciplines only behave differently when a Class-1 customer arrives to find a Class-2 customer in service and no other Class-1 customer present.

The theorem below shows that all $\mathcal{O}_{ij}$ and $\mathcal{O}_i^T$, except for $\mathcal{O}_{11}$, after multiplication by $1-\rho$ tend to an exponentially distributed random variable when $\rho \uparrow 1$.

**Theorem 4.1.** *If the traffic load $\rho \uparrow 1$, then*

$$(1-\rho)\mathcal{O}_{11} \;\to\; 0, \tag{86}$$

$$(1-\rho)\mathcal{O}_{12} \;\to\; \exp\left(\frac{\mu_1}{\mu_2\rho_1+\mu_1\rho_2}\right), \tag{87}$$

$$(1-\rho)\mathcal{O}_1^T \;\to\; \exp\left(\frac{\mu_1}{\mu_2\rho_1+\mu_1\rho_2}\right), \tag{88}$$

$$(1-\rho)\mathcal{O}_{21} \;\to\; \exp\left(\frac{\lambda_2}{\mu_2\rho_1^2+\lambda_1\rho_2}\right), \tag{89}$$

$$(1-\rho)\mathcal{O}_{22} \;\to\; \exp\left(\frac{\mu_1}{2\mu_2\rho_1+2\mu_1\rho_2}\right), \tag{90}$$

$$(1-\rho)\mathcal{O}_2^T \;\to\; \exp\left(\frac{1}{2\rho_2+\frac{\lambda_1}{\mu_2}+2\frac{\mu_2}{\mu_1}\rho_1+\frac{\rho_1^2}{\rho_2}}\right). \tag{91}$$

*Proof.* For $\mathcal{O}_{11}$ we get a trivial limit, because Class-1 customers are only marginally affected by Class-2 customers, so that Class 1 does not experience heavy traffic when $\rho \uparrow 1$. Let us now focus on $\mathcal{O}_{12}$. Starting point is (46), which for $z_1 = 1$ gives, with $h(z_2) := \frac{\mu_1}{\mu_1+\lambda_2(1-z_2)}$:

$$\mathbb{E}[z_2^{\mathcal{O}_{12}}] = h(z_2)\left[P_1(h(z_2),z_2) + \frac{\mu_2}{\mu_2+\lambda_2(1-z_2)}P_2(h(z_2),z_2) + 1 - \rho\right]. \tag{92}$$

Now take $z_2 := \mathrm{e}^{-s(1-\rho)}$. If we can show that

$$\mathbb{E}[\mathrm{e}^{-s(1-\rho)\mathcal{O}_{12}}] \to \frac{1}{1+\frac{\mu_2\rho_1+\mu_1\rho_2}{\mu_1}s}, \tag{93}$$

then Feller's convergence theorem for Laplace transforms (Extended Continuity Theorem of Feller Feller [8]) implies (87). To study how $P_i(h(\mathrm{e}^{-s(1-\rho)}), \mathrm{e}^{-s(1-\rho)})$ behaves for $\rho \uparrow 1$, we consider (12) and (13). In (12) it is readily seen that, in both expressions between square brackets, only the third term makes a non-disappearing contribution. Write

$$\frac{\mu_2}{z_2}\frac{\mu_1(1-\frac{1}{\pi(z_2)})}{\mu_2(1-\frac{1}{z_2})-\mu_1(1-\frac{1}{\pi(z_2)})} = \frac{\mu_2}{z_2}\left[-1+\frac{\mu_2}{\mu_1}\frac{z_2-1}{\pi(z_2)-1}\frac{\pi(z_2)}{z_2}\right]^{-1}, \tag{94}$$

and use $\frac{z_2-1}{\pi(z_2)-1} \sim \frac{1}{\pi'(1)+(z_2-1)\frac{\pi"(1)}{2}}$ for $z_2 \to 1$, with $\pi'(1) = \lambda_2\mathbb{E}[P_1] = \frac{\lambda_2}{\mu_1(1-\rho_1)}$ and $\pi"(1) = \lambda_2^2\mathbb{E}[P_1^2] = \frac{2\lambda_2^2}{\mu_1^2(1-\rho_1)^3}$. Finally replacing $z_2$ by $\mathrm{e}^{-s(1-\rho)}$, we obtain $\lim_{\rho\uparrow 1}P_2(h(\mathrm{e}^{-s(1-\rho)}), \mathrm{e}^{-s(1-\rho)}) = \frac{\rho_2}{1+\frac{\lambda_2}{\mu_1(1-\rho_1)}s}$ (where one can replace $\frac{\rho_2}{1-\rho_1}$ by one, as $\rho_1+\rho_2 \uparrow 1$). Similarly, from (13) we obtain $\lim_{\rho\uparrow 1}P_1(h(\mathrm{e}^{-s(1-\rho)}), \mathrm{e}^{-s(1-\rho)}) = \frac{\rho_1}{1+\frac{\lambda_2}{\mu_1(1-\rho_1)}s}$, and (93) is proven – which in turn implies (87).

The procedure for $\mathcal{O}_{21}$ and $\mathcal{O}_{22}$ is very similar, and again (94) plays a key role. A slight complication is the appearance of the function $g_1(z_1, \lambda_2(1-z_2))$ in the PGF of $(\mathcal{O}_{21}, \mathcal{O}_{22})$ (see the beginning of Section 3.2; notice that $g_1(1, \lambda_2(1-z_2)) = \pi(z_2)$, but that we now also need the first and second derivative of $g_1(z_1, 1)$ at $z_1 = 1$, and hence the $M/G/1$ results $\mathbb{E}[N_1] = \frac{1}{1-\rho_1}$ and $\mathbb{E}[N_1(N_1-1)] = \frac{2\rho_1}{(1-\rho_1)^3}$. We refrain from giving the detailed calculations. □

# 5 Mean overlap numbers via Mean Value Analysis

While the derivation in Section 3 of the overlap number distributions is rather involved, it is quite straightforward to derive *mean* overlap numbers; one can even allow an arbitrary number of priority classes, say $M$, and general service time distributions for all classes. Below we first (Section 5.1) give a set of formulas from Boxma et al. [6] which hold for the case of a multi-class queue

with independent Poisson arrival streams of all classes, and then (Section 5.2) focus on the non-preemptive case. In Subsection 5.3 we compare mean overlap numbers for the cases of non-preemptive and preemptive-resume priority and for the case of no priority at all.

## 5.1 Multi-class queues

Consider a single server serving $M$ customer classes (with or without some form of priority). The customers of the various classes arrive according to independent Poisson arrival streams, with rate $\lambda_i$ for Class $i$, $i = 1, \ldots, M$. A generic service time $\mathcal{S}_i$ of a Class-$i$ customer has some general, not necessarily exponential, distribution. All service times and all interarrival times are independent. Let $\mathbb{E}[\mathcal{Q}_i]$ and $\mathbb{E}[T_i]$ denote the steady-state mean number of customers of Class $i$, and their mean sojourn time, $i = 1, 2, \ldots, M$; by Little's formula, $\mathbb{E}[\mathcal{Q}_i] = \lambda_i \mathbb{E}[T_i]$. The steady-state mean number of overlaps $\mathbb{E}[\mathcal{O}_{ij}]$ that a tagged customer of Class $i$ experiences with Class-$j$ customers is for $i, j = 1, \ldots, M$ given by:

$$\mathbb{E}[\mathcal{O}_{ij}] \;=\; \mathbb{E}[\mathcal{Q}_j] + \lambda_j \mathbb{E}[T_i] \;=\; \mathbb{E}[\mathcal{Q}_j] + \frac{\lambda_j}{\lambda_i}\mathbb{E}[\mathcal{Q}_i]. \tag{95}$$

Here PASTA is used in the first equality and Little's formula in the second. Introducing the steady-state total number of overlaps $\mathcal{O}_i^T := \sum_{j=1}^M \mathcal{O}_{ij}$ experienced by a Class-$i$ customer, the total arrival rate $\Lambda := \sum_{i=1}^M \lambda_i$, and the steady-state total number of customers in the system $\mathcal{Q} := \sum_{i=1}^M \mathcal{Q}_i$, it follows from (95) that

$$\mathbb{E}[\mathcal{O}_i^T] \;=\; \mathbb{E}[\mathcal{Q}] + \frac{\Lambda}{\lambda_i}\mathbb{E}[\mathcal{Q}_i], \quad i = 1, \ldots, M. \tag{96}$$

Finally, the mean total number of overlaps $\mathbb{E}[\mathcal{O}^T]$ experienced by an arbitrary arriving customer equals

$$\mathbb{E}[\mathcal{O}^T] = \sum_{i=1}^M \frac{\lambda_i}{\Lambda}\mathbb{E}[\mathcal{O}_i^T] \;=\; 2\mathbb{E}[\mathcal{Q}]. \tag{97}$$

This is a generally valid relation for the case of independent Poisson arrival processes, cf. Oz and Perlman [19]. In Boxma et al. [6] the above relations are used to study the ratios $\mathbb{E}[\mathcal{O}_i^T]/\mathbb{E}[\mathcal{O}_j^T]$ and $\mathbb{E}[\mathcal{O}_i^T]/\mathbb{E}[\mathcal{O}^T]$. In Subsection 5.3 we use the above formulas to compare mean overlap times for the non-preemptive, preemptive-resume and no-priority cases.

## 5.2 The non-preemptive case

We now focus on the case in which the customer classes are prioritized from 1 to $M$, and the service policy is non-preemptive. Below we give the mean sojourn time $\mathbb{E}[T_i]$ of a Class-$i$ customer, $i = 1, \ldots, M$, taken from Ch. 9 of Adan and Resing [1]. The mean number of Class-$i$ customers $\mathbb{E}[\mathcal{Q}_i]$ then follows via multiplication by $\lambda_i$. In the sequel, $\mathcal{S}_i$ denotes a generic service time of a Class-$i$ customer and $\mathbb{E}[R_i] := \frac{\mathbb{E}[\mathcal{S}_i^2]}{2\mathbb{E}[\mathcal{S}_i]}$ denotes the mean residual service time of a Class-$i$ customer, $i = 1, \ldots, M$. $\rho_i := \lambda_i \mathbb{E}[\mathcal{S}_i]$ is the load of Class $i$, $i = 1, \ldots, M$.
We have:

$$\mathbb{E}[T_i] \;=\; \frac{\sum_{j=1}^{M} \rho_j \mathbb{E}[R_j]}{(1 - \sum_{m=1}^{i} \rho_m)(1 - \sum_{m=1}^{i-1} \rho_m)} \;+\; \mathbb{E}[\mathcal{S}_i], \quad i = 1, \ldots, M. \tag{98}$$

For comparison purposes we also mention the corresponding expression for preemptive-resume:

$$\mathbb{E}[T_i] \;=\; \frac{\sum_{j=1}^{i} \rho_j \mathbb{E}[R_j]}{(1 - \sum_{m=1}^{i} \rho_m)(1 - \sum_{m=1}^{i-1} \rho_m)} \;+\; \frac{\mathbb{E}\mathcal{S}_i]}{1 - \sum_{m=1}^{i-1} \rho_m}, \quad i = 1, \ldots, M. \tag{99}$$

Let us now return to the case of *two* customer classes, with $\exp(\mu_i)$ distributed service times for Class $i = 1, 2$. Using either the results from Proposition 3.5 and 3.6 or the above relations, we have for the non-preemptive case:

$$\mathbb{E}[\mathcal{O}_{11}] = 2\frac{\lambda_1(\frac{\rho_1}{\mu_1} + \frac{\rho_2}{\mu_2})}{1 - \rho_1} + 2\rho_1, \tag{100}$$

$$\mathbb{E}[\mathcal{O}_{12}] = \lambda_2 \frac{(2 - \rho)(\frac{\rho_1}{\mu_1} + \frac{\rho_2}{\mu_2})}{(1 - \rho)(1 - \rho_1)} + \lambda_2(\mathbb{E}[\mathcal{S}_1] + \mathbb{E}[\mathcal{S}_2]), \tag{101}$$

$$\mathbb{E}[\mathcal{O}_{21}] = \lambda_1 \frac{(2 - \rho)(\frac{\rho_1}{\mu_1} + \frac{\rho_2}{\mu_2})}{(1 - \rho)(1 - \rho_1)} + \lambda_1(\mathbb{E}[\mathcal{S}_1] + \mathbb{E}[\mathcal{S}_2]), \tag{102}$$

$$\mathbb{E}[\mathcal{O}_{22}] = 2\frac{\lambda_2(\frac{\rho_1}{\mu_1} + \frac{\rho_2}{\mu_2})}{(1 - \rho)(1 - \rho_1)} + 2\rho_2. \tag{103}$$

Notice that the ratio between $\mathbb{E}[\mathcal{O}_{12}]$ and $\mathbb{E}[\mathcal{O}_{21}]$ equals $\lambda_2/\lambda_1$ or $\lambda_1 \mathbb{E}[\mathcal{O}_{12}] = \lambda_2 \mathbb{E}[\mathcal{O}_{21}]$.

In total, we have

$$\mathbb{E}[\mathcal{O}_1^T] = \frac{(2\lambda_1(1-\rho)+\lambda_2(2-\rho))(\frac{\rho_1}{\mu_1}+\frac{\rho_2}{\mu_2})}{(1-\rho)(1-\rho_1)} + \rho_1 + \rho_2 + \Lambda\mathbb{E}[\mathcal{S}_1], \tag{104}$$

$$\mathbb{E}[\mathcal{O}_2^T] = \frac{(\lambda_1(2-\rho)+2\lambda_2)(\frac{\rho_1}{\mu_1}+\frac{\rho_2}{\mu_2})}{(1-\rho)(1-\rho_1)} + \rho_1 + \rho_2 + \Lambda\mathbb{E}[\mathcal{S}_2]. \tag{105}$$

It is readily verified that

$$\mathbb{E}[\mathcal{O}^T] = \frac{\lambda_1}{\Lambda}\mathbb{E}[\mathcal{O}_1^T] + \frac{\lambda_2}{\Lambda}\mathbb{E}[\mathcal{O}_2^T] = 2\frac{(\lambda_1(1-\rho)+\lambda_2)(\frac{\rho_1}{\mu_1}+\frac{\rho_2}{\mu_2})}{(1-\rho)(1-\rho_1)} + 2\rho_1 + 2\rho_2, \tag{106}$$

and that hence, cf. (33) and (34),

$$\mathbb{E}[\mathcal{O}^T] = 2\mathbb{E}[\mathcal{Q}], \tag{107}$$

in agreement with (97).

## 5.3 Comparisons

In this subsection we compare overlap numbers for the cases of non-preemptive priority, preemptive-resume priority and no priority. When we refer to the case of no priority, we have an $M/G/1$ queue in mind with arrival rate $\Lambda$ and service time distribution a weighted sum, with weight factors $\lambda_m/\Lambda$, of the distribution of $\mathcal{S}_m$, $m = 1, 2, \ldots, M$. From Boxma [2] we have for the mean overlap number in this non-priority $M/G/1$ queue: $\mathbb{E}[\mathcal{O}_{\text{no-prio}}] = 2\mathbb{E}[X]$, with $\mathbb{E}[X]$ its mean number of customers. Hence

$$\mathbb{E}[\mathcal{O}_{\text{no-prio}}] = \frac{2\Lambda\sum_{m=1}^{M}\rho_m\mathbb{E}[R_m]}{1-\rho} + 2\rho. \tag{108}$$

We start with two observations for the situation with $M \geq 2$ priority classes. Firstly, it is obvious from (96) and (98) respectively (99) that $\mathbb{E}[\mathcal{O}_i^T]$ is strictly increasing in $i$ for both the non-preemptive and the preemptive-resume cases.

Secondly, a comparison of the expressions for $\mathbb{E}[\mathcal{O}_i^T]$ for non-preemptive priority and preemptive-resume priority based on (96), (98) and (99), or a comparison of Theorem 4.1 and Theorem 4.1 of Boxma et al. [6], makes it clear that $\lim_{\rho\uparrow 1}(1-\rho)\mathbb{E}[\mathcal{O}_i^T]$ is the same for both disciplines. So just as in

Boxma et al. [6] we have the following heavy-traffic limits: for $\rho \uparrow 1$,

$$(1-\rho)\mathbb{E}[\mathcal{O}_i^T] \to \frac{\lambda_M \sum_{m=1}^M \rho_m \mathbb{E}[R_m]}{\rho_M}, \quad i = 1, \ldots, M_1, \tag{109}$$

$$(1-\rho)\mathbb{E}[\mathcal{O}_M^T] \to \frac{(\lambda_M + \Lambda) \sum_{m=1}^M \rho_m \mathbb{E}[R_m]}{\rho_M}, \tag{110}$$

$$(1-\rho)\mathbb{E}[\mathcal{O}^T] \to \frac{2 \sum_{m=1}^M \rho_m \mathbb{E}[R_m]}{\mathbb{E}[\mathcal{S}_M]}. \tag{111}$$

It follows from (108) that, for $\rho \uparrow 1$,

$$(1-\rho)\mathbb{E}[\mathcal{O}_{\rm no-prio}] \to 2\Lambda \sum_{m=1}^M \rho_m \mathbb{E}[R_m]. \tag{112}$$

We thus obtain the following ratios for mean overlap numbers in the non-priority and non-preemptive priority cases: for $\rho \uparrow 1$,

$$\frac{\mathbb{E}[\mathcal{O}_{\rm no-prio}]}{\mathbb{E}[\mathcal{O}_i^T]} \to 2\Lambda \mathbb{E}[\mathcal{S}_M], \quad i = 1, \ldots, M-1, \tag{113}$$

$$\frac{\mathbb{E}[\mathcal{O}_{\rm no-prio}]}{\mathbb{E}[\mathcal{O}_M^T]} \to \frac{2\Lambda \rho_M}{\lambda_M + \Lambda}, \tag{114}$$

$$\frac{\mathbb{E}[\mathcal{O}_{\rm no-prio}]}{\mathbb{E}[\mathcal{O}^T]} \to \Lambda \mathbb{E}[\mathcal{S}_M]. \tag{115}$$

Regarding (115), remember that the mean overlap number for the whole queue is twice the mean queue length. As observed in Boxma et al. [6], (115) implies that if $\mathbb{E}[\mathcal{S}_M]$ is larger (smaller) than the average over all mean service times, then the mean overlap number in the priority queue is smaller (larger) than in the corresponding queue without priorities; and this makes sense when realizing that SRPT (Shortest Remaining Processing Time first) minimizes mean queue length.

In the remainder of the section we focus on the case of $M = 2$ classes. We use (96), (98) and (99) to compare mean overlap numbers. Below P indicates preemptive-resume and NP indicates non-preemptive priority. We give some detail for the difference between $\mathbb{E}[\mathcal{O}_1^T]|_P - \mathbb{E}[\mathcal{O}_1^T]|_{NP}$; we omit those details for the other differences, as the approach is the same each time. In particular, mean queue lengths are each time obtained by applying Little's formula to

the mean sojourn times of (98) and (99).

$$\mathbb{E}[\mathcal{O}_1^T]|_P - \mathbb{E}[\mathcal{O}_1^T]|_{NP} = \mathbb{E}[\mathcal{Q}]|_P + \frac{\Lambda}{\lambda_1}\mathbb{E}[\mathcal{Q}]|_P - \mathbb{E}[\mathcal{Q}]|_{NP} - \frac{\Lambda}{\lambda_1}\mathbb{E}[\mathcal{Q}]|_{NP}$$

$$= \frac{\rho_2}{1-\rho_1}\left(\rho_1 - \lambda_1\mathbb{E}[R_2] - \Lambda\mathbb{E}[R_2]\right). \tag{116}$$

Similarly,

$$\mathbb{E}[\mathcal{O}_2^T]|_P - \mathbb{E}[\mathcal{O}_2^T]|_{NP} = \frac{\rho_2}{1-\rho_1}\left(\rho_1 - \lambda_1\mathbb{E}[R_2] + \frac{\Lambda}{\lambda_2}\rho_1\right), \tag{117}$$

$$\mathbb{E}[\mathcal{O}^T]|_P - \mathbb{E}[\mathcal{O}^T]|_{NP} = \frac{2\lambda_1\rho_2}{1-\rho_1}\left(\mathbb{E}[\mathcal{S}_1] - \mathbb{E}[R_2]\right). \tag{118}$$

We conclude from (118) that preemptive-resume yields a smaller (larger) mean total overlap number than non-preemptive if the mean service time for Class 1 is smaller (larger) than the mean residual service time of Class 2. Again, this makes sense from the perspective of the SRPT principle: It pays for Class 1 to interrupt a Class-2 service time if the mean of the remainder of that service time is larger than the mean Class-1 service time.

Comparing the priority disciplines with the case without priorities, we find:

$$\mathbb{E}[\mathcal{O}^T]|_{NP} - \mathbb{E}[\mathcal{O}_{no-prio}] = \frac{2\lambda_1\lambda_2}{1-\rho_1}\frac{\rho_1\mathbb{E}[R_1] + \rho_2\mathbb{E}[R_2]}{1-\rho}\left(\mathbb{E}[\mathcal{S}_1] - \mathbb{E}[S_2]\right). \tag{119}$$

We conclude that non-preemptive priority yields a smaller (larger) mean total overlap number than no priority if the mean service time for Class 1 is smaller (larger) than that of Class 2.

Adding the expressions in (118) and (119) gives the difference in mean total overlap number between preemptive-resume and the case of no priority.

# Acknowledgments

The research of Onno Boxma was supported by the NWO Gravitation project NETWORKS, grant number 024.002.003. Jamol Pender would like to acknowledge the gracious support of the National Science Foundation DMS Award # 2206286. Uri Yechiali would like to acknowledge the support of the Israel Science Foundation for Grant # 1968-23.

# References


[1] Ivo Adan and Jacques Resing. Queueing theory. *Eindhoven University of Technology*, Lecture Notes, 2002.

[2] Onno Boxma. The number of overlapping customers. *Operations Research Letters*, 57:107203, 2024.

[3] Onno Boxma and Jamol Pender. Overlap times in the G/G/1 queue via Laplace transforms. *Queueing Systems*, 109:10, 2025.

[4] Onno Boxma, Lucas van Kreveld, and Jamol Pender. The overlapping time distribution in the $M/G/\infty$ queue. *Probability in the Engineering and Informational Sciences, to appear.*

[5] Onno Boxma, Jamol Pender, and Uri Yechiali. The Maximum Overlap Time in the $M_\lambda/G/1/k$ Queue. *Submitted to Queueing Systems*, 2025.

[6] Onno Boxma, Jamol Pender, and Uri Yechiali. The number of overlapping customers in priority queues. *Journal of Applied Probability to Appear*, 2026.

[7] Jacob Willem Cohen. *The single server queue*. Elsevier, 2012.

[8] William Feller. *An introduction to probability theory and its applications, Volume 2*. John Wiley & Sons, 1991.

[9] Ruici Gao and Jamol Pender. Overlap times in tandem queues: Identically distributed station case. *arXiv preprint arXiv:2311.01261*, 2023.

[10] Mor Harchol-Balter. *Performance modeling and design of computer systems: queueing theory in action*. Cambridge University Press, 2013.

[11] Narendra Kumar Jaiswal. Time-dependent solution of the head-of-the-line priority queue.

[12] Narendra Kumar Jaiswal. *Priority queues*, volume 50. Academic Press, 1968.

[13] Kang Kang. *Essays on Service System Design: Efficiency, Fairness, and Safety*. PhD thesis, University of Minnesota, 2022.

[14] Kang Kang, Sherwin Doroudi, Mohammad Delasay, and Alexander Wickeham. A queueing-theoretic framework for evaluating transmission risks in service facilities during a pandemic. *Production and Operations Management*, 32(5):1453–1470, 2023.

[15] Leonard Kleinrock. *Queueing systems: theory*, volume 2. Wiley, 1974.

[16] Philip McCord Morse. *Queues, inventories and maintenance: the analysis of operational systems with variable demand and supply.* Courier Corporation, 2004.

[17] Laurel Newman and Jamol Pender. Contagious queues: Analyzing the spread of infection in the M/G/$\infty$ queue. *Submitted to Queueing Systems*, 2025.

[18] Laurel Newman and Jamol Pender. When lines become contagious: Analyzing the spread of infection in the M/M/1 queue. *Submitted to Stochastic Systems*, 2025.

[19] Binyamin Oz and Yael Perlman. Co-exposure management in queueing systems. *SSRN*, 2024.

[20] Sergio Palomo and Jamol Pender. Measuring the overlap with other customers in the single server queue. *2021 Winter Simulation Conference (WSC)*, pages 1–12, 2021.

[21] Sergio Palomo and Jamol Pender. Overlap times in the infinite server queue. *Probability in the Engineering and Informational Sciences*, 38 (1):21–27, 2024.

[22] Sergio Palomo and Jamol Pender. The maximum overlap time in the M/M/1 queue. *Statistics & Probability Letters*, 219:110322, 2025.

[23] Sergio D Palomo and Jamol Pender. Overlap times in the $GI^B/GI/\infty$ queue. *arXiv preprint arXiv:2302.07410*, 2023.

[24] Jamol Pender, Young Myoung Ko, and Jin Xu. The number of overlapping customers in Erlang-A queues: an asymptotic approach. *Probability in the Engineering and Informational Sciences*, pages 1–26, 2025.

[25] Ahmad Ridley and William Massey. Fluid approximations for a priority call center with time-varying arrivals. In *Proceedings of the 2003 Winter Simulation Conference, 2003*, volume 2, pages 1817–1823. IEEE, 2003.

[26] Hideaki Takagi. Queueing analysis: a foundation of performance evaluation. *Vol. 1: Vacation and priority systems, Part 1*, 1991.

[27] Gideon Weiss. *Scheduling and control of queueing networks*, volume 14. Cambridge University Press, 2021.

[28] Jin Xu, Young Myoung Ko, Min Kong, and Jamol Pender. Queueing management for reducing the overlaps of customers in service systems. *OR Spectrum*, pages 1–30, 2025.

[29] Jin Xu, Dening Huang, Young Myoung Ko, and Jamol Pender. Managing queueing systems with gathering-averse customers: staffing, pooling, and opening hours. *Queueing Systems*, 110(1):22, 2026.

**Declaration of interests**

☒The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

☐The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: