Check for updates

J. Appl. Probab. 1–26 (2026)
*doi:*10.1017/jpr.2026.10107

THE OVERLAPPING TIME DISTRIBUTION IN THE M/M/1 QUEUE

ONNO J. BOXMA,* *Eindhoven University of Technology*
JAMOL PENDER,** *Cornell University*

Abstract

In this paper, we analyze the distribution of the total overlapping time spent with other customers in the $M_\lambda/M_\mu/1$ queue with First-Come First-Served service discipline. We show that the Laplace–Stieltjes transform of the overlapping time reduces to an incomplete gamma function representation. We also calculate the transform of the joint distribution of the overlapping time and the number of overlaps. In addition, we prove a heavy-traffic limit for the total overlapping time with scaling $1/(1-\rho)^2$ to a Weibull $(1/2, 1/\mu)$ random variable. Finally, for the $M_\lambda/G/1$ case, we derive the first two moments of the overlapping time.

Keywords: M/M/1 queue; overlapping time distribution; heavy traffic

2020 Mathematics Subject Classification: Primary 60K25

1. Introduction

Overlap times in queues refer to the duration of time when two or more individuals simultaneously share a space in a queue. This sharing of space creates opportunities for customer interaction and, unfortunately, can lead to disease transmission. Understanding the dynamics and distributional behavior of overlap times is critical to modeling infectious disease spread because they capture moments when individuals are in close proximity, allowing for the transmission of disease through airborne particles, droplets, or surface contact. A complete understanding of overlap times helps researchers and public health officials design interventions, such as staggered entry, optimized queue length designs, or ventilation improvements, to reduce the risk of outbreaks in public and shared spaces.

Recently, overlap times have been studied from a mathematical perspective to understand the time that customers spend with one another in a queue. In particular, Palomo and Pender [22] analyzed overlap times in the $M_\lambda/M_\mu/1$ queue and computed the distribution of the overlap time of customers who are k spaces apart. Moreover, they derived a new Lindley-type recursion to efficiently calculate overlap times in the G/G/1 queue. Kang et al. [17] analyzed the Laplace–Stieltjes transform (LST) of overlap times in a queueing model where there are infected and susceptible customers. They computed the mean number of infected customers in the multi-server queue, providing a queueing version of R_0 in the SIR modeling literature.

Received 27 August 2025 UTC; accepted 30 April 2026 UTC.
* Postal address: Department of Mathematics and Computer Science, Eindhoven University of Technology, The Netherlands. Email address: o.j.boxma@tue.nl
** Postal address: School of Operations Research and Information Engineering, Cornell University, USA. Email address: jjp274@cornell.edu

© The Author(s), 2026. Published by Cambridge University Press on behalf of Applied Probability Trust. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.

https://doi.org/10.1017/jpr.2026.10107 Published online by Cambridge University Press

Palomo and Pender [23, 24] studied the distribution of the overlap time in the $G/G/\infty$ queue and the $G^B/G/\infty$ queue respectively. In [24] they also studied the number of overlapping customers that a tagged customer interacts with during her sojourn time. In some special cases they derived the distribution of the number of overlaps that a tagged customer will have during their sojourn time, which is very useful for understanding disease spread as it represents all of the people that will need to be contact-traced in a serious epidemic. Pender et al. [27] used fluid and diffusion limit theorems to derive new approximations for the number of overlapping customers in the Erlang-A queue, which has customer abandonment. In fact they were able to provide novel mean and variance estimates for the number of overlapping customers in the Erlang-A queue, allowing one to understand the impact of abandonment on the number of overlapping customers.

Moreover, recent work by Palomo and Pender [25] analyzes the maximum overlap time in the $M_\lambda/M_\mu/1$ queue and computes its distribution. They make the observation that the maximum overlap time is with the customer in front of or behind the tagged customer. This observation is coupled with the fact that the overlap time of the two adjacent customers is equal to the waiting time. The maximum overlap time analysis was extended to special cases of the G/G/1 queue via Laplace–Stieltjes transforms in Boxma and Pender [7]. The Laplace transform is derived in the case where either the inter-arrival distribution or service time distribution follows a mixed-Erlang distribution, which is a special case of a phase-type distribution that is still dense in the family of distributions of non-negative random variables.

In addition to overlap times, the number of overlaps, and maximum overlap time, the *overlapping time* was introduced by Xu et al. [31] as a way of measuring the sum of all the amounts of time that a tagged customer spends with all other customers. The total overlapping time is an important measure of infectious disease spread as it measures the amount of time that you spend with other customers. From the perspective of an infected customer, it is the amount of exposure you give to others, and from the perspective of a susceptible customer it is the amount of time that you spend with infected customers. Xu et al. [31] computed the mean of the overlapping time in various $M_\lambda/M_\mu/c$ queues as a metric for how much time customers will spend with each other in a queue. While the mean is an important metric for determining how to staff the queueing process, it is far more important to understand the distribution of the overlapping time to get a complete picture of the metric. In this work, we derive the distribution of the overlapping time via computing the LST of the overlapping time.

1.1. Contributions of our work

In this paper, we make the following contributions to the literature, focusing on the $M_\lambda/M_\mu/1$ queue.

- We derive the distribution, and the LST, of the overlapping time T_a of a tagged customer with the customers she meets at arrival, and of the overlapping time T_d with those who arrive during her sojourn. We also obtain the LST of the sum $T = T_a + T_d$, i.e. the total overlapping time, and of the vector (T_a, T_d).
- We derive various moments of T_a, T_d, and T.
- The heavy-traffic limiting distributions of T_a, T_d, and T are proved to be Weibull; we show that one has to scale by $(1-\rho)^2$ when the traffic load $\rho \uparrow 1$.
- We generalize some of the above results in several ways. Firstly, we obtain the joint transform of overlap numbers and overlapping times. Secondly, under the assumption

https://doi.org/10.1017/jpr.2026.10107 Published online by Cambridge University Press

that each customer has a probability q of being infectious, we derive the LST, moments, and heavy-traffic limit of the overlapping times of a tagged customer with infectious customers. We also explicitly compute the first two moments of T_a and T_d for the $\mathrm{M}_\lambda/\mathrm{G}/1$ queue. Finally, we determine the LST of T_a for the $\mathrm{G}/\mathrm{M}_\mu/1$ queue and the LST of T_d for the $\mathrm{D}/\mathrm{M}_\mu/1$ queue.

1.2. Organization of the paper

The remainder of the paper is organized as follows. We close this section with a concise model description and notations, and a brief review of the (incomplete) gamma function. Section 2 contains results on the marginal overlapping time T_a of a tagged customer with those customers already present, and on the marginal overlapping time T_d of the tagged customer with those arriving during her sojourn time. Section 3 discusses the joint distribution of T_a and T_d, and their sum, the total overlapping time $T = T_a + T_d$. The heavy-traffic limit of this joint distribution is analyzed in Section 4. Section 5 is devoted to determining the transform of the joint distribution of T_a, T_d and the numbers of customers seen at arrival and left behind by the tagged customer (the overlap *numbers*). As a by-product, we thus obtain the joint distribution of total overlap number and total overlapping time. In Section 6 we consider the case in which any customer is infectious with a fixed probability q; we are then interested in the overlapping time with infectious customers. In Section 7 we discuss overlapping times in the $\mathrm{M}_\lambda/\mathrm{G}/1$ FCFS queue, pointing out technical difficulties which make it very challenging to derive exact distributional results for overlapping times for the case of a general service time distribution. Overlapping times in the $\mathrm{G}/\mathrm{M}_\mu/1$ FCFS queue are briefly discussed in Section 8.

1.3. Model description and notations

We consider the classical $\mathrm{M}_\lambda/\mathrm{M}_\mu/1$ queue, with a single server, a Poisson arrival process with intensity λ, and independent, exponentially distributed service times $S_1, S_2, \ldots$ with mean $1/\mu$. A generic service time is denoted by S. Henceforth we shall often denote this system by $\mathrm{M}_\lambda/\mathrm{M}_\mu/1$. The customers are served in order of arrival, i.e. First-Come First-Served (FCFS). The load of the server is denoted by $\rho := \lambda/\mu$; we assume that $\rho < 1$, so that the system reaches steady state. Sometimes we shall also denote an $\mathrm{Exp}(\zeta)$ random variable by $\mathcal{E}_\zeta$; hence $S \sim \mathcal{E}_\mu$.

Our focus is on the overlapping times of a tagged customer C. If C finds X_a customers present upon arrival, then the total overlapping time T_a of C with these customers is the sum of the X_a overlap times (time simultaneously spent in the system) of C with those customers. If X_d customers arrive during the sojourn time of C in the system (and hence X_d customers are left behind when C leaves the system), then the total overlapping time T_d of C with these customers is the sum of the X_d overlap times of C with those customers. The main performance measure under investigation in this paper is C's *total overlapping time* T, where

$$T = T_a + T_d.$$

1.4. Review of the incomplete gamma function

The gamma function and incomplete gamma functions play a significant role in the remainder of this paper. Hence we briefly recall their definition and some properties. First we recall the integral definition of the gamma function:

$$\Gamma(a) = \int_0^\infty t^{a-1}\, \mathrm{e}^{-t}\, \mathrm{d}t \quad \text{for } a > 0;$$

https://doi.org/10.1017/jpr.2026.10107 Published online by Cambridge University Press

for integer a one has $\Gamma(a) = (a-1)!$. One can also partition the gamma function integral at a point $x \geq 0$ to obtain the lower and upper incomplete gamma functions respectively as

$$\gamma(a, x) = \int_0^x t^{a-1}\, \mathrm{e}^{-t}\, \mathrm{d}t, \tag{1}$$

$$\Gamma(a, x) = \int_x^\infty t^{a-1}\, \mathrm{e}^{-t}\, \mathrm{d}t = \mathrm{e}^{-x} \int_0^\infty (t+x)^{a-1}\, \mathrm{e}^{-t}\, \mathrm{d}t.$$

Moreover, we know that

$$\frac{1}{c^a}\gamma(a, cx) = \int_0^x t^{a-1}\, \mathrm{e}^{-ct}\, \mathrm{d}t,$$

$$\frac{1}{c^a}\Gamma(a, cx) = \int_x^\infty t^{a-1}\, \mathrm{e}^{-ct}\, \mathrm{d}t.$$

2. Marginal overlapping times in the $\mathrm{M}_\lambda/\mathrm{M}_\mu/1$ queue

In this section we determine the marginal distribution and LST of T_a (Section 2.1) and T_d (Section 2.2).

2.1. The marginal distribution of T_a

Assume that C arrives to find X_a customers. It is well known that its steady-state distribution is geometric with parameter ρ:

$$\mathbb{P}(X_a = k) = (1-\rho)\rho^k, \quad k = 0, 1, \ldots. \tag{2}$$

If $X_a = k \geq 1$, then C overlaps the remaining service time R with the customer in service, $R + S_2$ with the second customer, and so on, and finally $R + \sum_{j=2}^k S_j$ with the kth customer. Because of the memoryless property of the exponential distribution, R has the same distribution as S_1. Hence

$$T_a\Big|_{X_a=k} \stackrel{d}{=} \sum_{j=1}^k (k-j+1)S_j \stackrel{d}{=} \sum_{j=1}^k j \cdot S_j. \tag{3}$$

The last equality (in distribution) follows since all service times are i.i.d. (independent and identically distributed).

Obviously,

$$\mathbb{P}(j \cdot S_j > t) = \mathbb{P}\left(S_j > \frac{t}{j}\right) = \mathrm{e}^{-(\mu/j)t}. \tag{4}$$

The above equations (3) and (4) allow us to conclude that T_a, given $X_a = k$, is hypo-exponentially distributed:

$$T_a\Big|_{X_a=k} \stackrel{d}{=} \sum_{j=1}^k \mathcal{E}_{\mu/j} \sim \text{hypo-exp}\left(\mu, \frac{\mu}{2}, \ldots, \frac{\mu}{k}\right).$$

We summarize these findings in the following proposition.

https://doi.org/10.1017/jpr.2026.10107 Published online by Cambridge University Press

Proposition 1. *If we condition on X_a to be k, then the overlapping time upon arrival, T_a, is* hypo-exp$(\mu, \mu/2, \ldots, \mu/k)$*:*

$$\begin{aligned}\mathbb{P}(T_a > t | X_a = k) &= \boldsymbol{\alpha}\, \mathrm{e}^{t\mathcal{M}}\mathbf{1} \\ &= \sum_{j=1}^{k} \mathrm{e}^{-(\mu/j)t}\left(\prod_{i=1, i\neq j}^{k} \frac{\mu/i}{\mu/i - \mu/j}\right) \\ &= \sum_{j=1}^{k} \mathrm{e}^{-(\mu/j)t}\left(\prod_{i=1, i\neq j}^{k} \frac{j}{j-i}\right),\end{aligned}$$

where

$$\alpha = \begin{bmatrix}1 & 0 & 0 & \ldots & 0 & 0\end{bmatrix},$$

$$\mathcal{M} = \begin{bmatrix} -\mu & \mu & 0 & \ldots & 0 & 0 \\ 0 & -\frac{\mu}{2} & \frac{\mu}{2} & \ddots & 0 & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & \ddots & -\frac{\mu}{k-2} & \frac{\mu}{k-2} & 0 \\ 0 & 0 & \ldots & 0 & -\frac{\mu}{k-1} & \frac{\mu}{k-1} \\ 0 & 0 & \ldots & 0 & 0 & -\frac{\mu}{k} \end{bmatrix}.$$

The unconditional distribution tail of T_a follows by combining (2) and Proposition 1.

Proposition 2. *The tail distribution of T_a is given by*

$$\mathbb{P}(T_a > t) = \sum_{k=1}^{\infty} (1-\rho)\rho^k \sum_{j=1}^{k} \mathrm{e}^{-(\mu/j)t}\left(\prod_{i=1, i\neq j}^{k} \frac{j}{j-i}\right).$$

For several purposes we shall find it useful to also determine the LST of T_a. We have the following result.

Lemma 1. *The LST of T_a is given by*

$$\mathbb{E}\left[\mathrm{e}^{-\theta T_a}\right] = \sum_{k=0}^{\infty} (1-\rho)\left(\frac{\lambda}{\theta}\right)^k \frac{\Gamma(1+\mu/\theta)}{\Gamma(k+1+\mu/\theta)}.$$

Proof. We have

$$\begin{aligned}\mathbb{E}\left[\mathrm{e}^{-\theta T_a}\right] &= \sum_{k=0}^{\infty} (1-\rho)\rho^k \mathbb{E}\left[\mathrm{e}^{-\theta \sum_{j=1}^{k} jS_j}\right] \\ &= \sum_{k=0}^{\infty} (1-\rho)\rho^k \prod_{j=1}^{k}\left(\frac{\mu}{\mu + j\theta}\right)\end{aligned}$$

https://doi.org/10.1017/jpr.2026.10107 Published online by Cambridge University Press

$$= \sum_{k=0}^{\infty} (1-\rho)\rho^k \prod_{j=1}^{k} \left(\frac{\mu/\theta}{\mu/\theta + j} \right)$$

$$= \sum_{k=0}^{\infty} (1-\rho) \left(\frac{\lambda}{\theta} \right)^k \frac{\Gamma(1+\mu/\theta)}{\Gamma(k+1+\mu/\theta)}. \qquad \square$$

In Lemma 2 we present a different derivation of the LST of T_a. This alternative approach is particularly convenient when, in Section 3, we derive the joint LST of (T_a, T_d). After the proof of the lemma we will show agreement between the results of Lemmas 1 and 2.

Lemma 2. *The LST of T_a is given by*

$$\mathbb{E}\left[\mathrm{e}^{-\theta T_a}\right] = (1-\rho) \int_{t=0}^{\infty} \mu \mathrm{e}^{-\mu t}\, \mathrm{e}^{(\lambda/\theta)(1-\mathrm{e}^{-\theta t})}\, \mathrm{d}t = (\mu - \lambda) \left(\frac{\theta}{\lambda} \right)^{\mu/\theta} \frac{1}{\theta}\, \mathrm{e}^{\lambda/\theta} \gamma\left(\frac{\mu}{\theta}, \frac{\lambda}{\theta} \right). \tag{5}$$

Proof. We exploit the following. (i) The number X_a seen by the tagged customer upon arrival is geom (ρ)-distributed. (ii) If $X_a = k$, then the sojourn time Σ of the tagged customer is Erlang $(k+1, \mu)$-distributed, hence with density

$$\mu \frac{(\mu t)^k}{k!}\, \mathrm{e}^{-\mu t}.$$

(iii) A Poisson arrival process has the following well-known property (see Section 2.3 of Ross [28]), which we shall henceforth refer to as property $\mathcal{P}$. If there are k Poisson events in an interval of length t, then the joint distribution of the event epochs equals the joint distribution of the order statistics of k independent, uniformly on $(0, t)$-distributed random variables. Hence

$$\mathbb{E}\left[\mathrm{e}^{-\theta T_a}\right] = \sum_{k=0}^{\infty} (1-\rho)\rho^k \int_{t=0}^{\infty} \mathbb{E}\left[\mathrm{e}^{-\theta T_a} \mid X_a = k,\, \Sigma = t\right] \mathrm{d}\mathbb{P}(\Sigma < t \mid X_a = k)$$

$$= \sum_{k=0}^{\infty} (1-\rho)\rho^k \int_{t=0}^{\infty} \mu \frac{(\mu t)^k}{k!}\, \mathrm{e}^{-\mu t} \mathbb{E}\left[\mathrm{e}^{-\theta T_a} \mid X_a = k,\, \Sigma = t\right] \mathrm{d}t. \tag{6}$$

Now observe that the k service completions of the $X_a = k$ customers seen by C at arrival occur in $(0, t)$ according to a Poisson process with rate μ. Hence, exploiting the above-mentioned property $\mathcal{P}$,

$$\mathbb{E}\left[\mathrm{e}^{-\theta T_a} \mid X_a = k, S = t\right] = \left(\int_{u=0}^{t} \mathrm{e}^{-\theta u} \frac{\mathrm{d}u}{t} \right)^k = \left(\frac{1 - \mathrm{e}^{-\theta t}}{\theta t} \right)^k. \tag{7}$$

Substitution of (7) in (6) and summation over k gives the first equality in (5). The second equality follows via the substitution $z = (\lambda/\theta)\, \mathrm{e}^{-\theta t}$; it transforms the first integral in (5) into

$$(\mu - \lambda)\, \mathrm{e}^{\lambda/\theta} \left(\frac{\theta}{\lambda} \right)^{\mu/\theta} \int_{z=0}^{\lambda/\theta} z^{\mu/\theta}\, \mathrm{e}^{-z} \frac{\mathrm{d}z}{\theta z} = (\mu - \lambda) \left(\frac{\theta}{\lambda} \right)^{\mu/\theta} \frac{1}{\theta}\, \mathrm{e}^{\lambda/\theta} \int_{z=0}^{\lambda/\theta} z^{\mu/\theta - 1}\, \mathrm{e}^{-z}\, \mathrm{d}z. \tag{8}$$

Now observe (see (1)) that the last integral in (8) equals $\gamma(\mu/\theta, \lambda/\theta)$. $\square$

https://doi.org/10.1017/jpr.2026.10107 Published online by Cambridge University Press

It should be observed that many identities and asymptotic results are known for incomplete gamma functions; see Abramowitz and Stegun [1].

We now show the equivalence between the results of Lemmas 1 and 2. To this end, rewrite the part of (8) after the equality sign as follows, using the transformation $v = 1 - (\theta/\lambda)z$ and the Taylor series development of $e^{(\lambda/\theta)v}$:

$$(\mu-\lambda)\frac{1}{\theta}\int_{v=0}^{1}(1-v)^{\mu/\theta-1}\,e^{(\lambda/\theta)v}\,dv = (\mu-\lambda)\frac{1}{\theta}\sum_{k=0}^{\infty}\frac{(\lambda/\theta)^k}{k!}\int_{v=0}^{1}(1-v)^{\mu/\theta-1}v^k\,dv.$$

Finally, observe that the last integral equals the beta function

$$B\left(\frac{\mu}{\theta},k+1\right)=\frac{\Gamma(\mu/\theta)\Gamma(k+1)}{\Gamma(k+1+\mu/\theta)},$$

and use the fact that

$$\Gamma\left(1+\frac{\mu}{\theta}\right)=\frac{\mu}{\theta}\Gamma\left(\frac{\mu}{\theta}\right).$$

We close this subsection by deriving two expressions for the moments of T_a.

Proposition 3. *With X_a denoting a random variable that is* geom (ρ)*-distributed, we have*

$$\mathbb{E}\left[T_a^m\right]=\frac{m!}{\mu^m}\sum_{k=0}^{\infty}(1-\rho)\rho^k\sum_{j=1}^{k}\left(\prod_{i=1,i\neq j}^{k}\frac{j}{j-i}\right)j^m$$

$$=\mathbb{E}\left[S^m\right]\cdot\mathbb{E}\left[\sum_{j=1}^{X_a}\left(\prod_{i=1,i\neq j}^{X_a}\frac{j}{j-i}\right)j^m\right].$$

Proof. Use Proposition 2 and the fact that the mth moment of an Exp(μ/j) distribution equals $(m!/\mu^m)j^m$. □

Using the multinomial theorem, we also have the following alternative representation for the moments of T_a.

Proposition 4. *We have*

$$\mathbb{E}\left[T_a^m\right]=\frac{m!}{\mu^m}\sum_{k=0}^{\infty}(1-\rho)\rho^k\sum_{\substack{n_1+\cdots+n_k=m\\ n_1,\ldots,n_k\geq 0}}\prod_{j=1}^{k}j^{n_j}.$$

Proof. This follows directly from first conditioning on the queue length and observing the fact that

$$\mathbb{E}\left[S_j^{n_j}\right]=\frac{n_j!}{\mu^{n_j}}$$

and the multinomial theorem. We provide an outline of the proof below:

$$\mathbb{E}\left[T_a^m\right]=\sum_{k=0}^{\infty}\mathbb{E}\left[T_a^m\mid X_a=k\right]\mathbb{P}(X_a=k)$$

$$=\sum_{k=0}^{\infty}\mathbb{E}\left[\left(\sum_{j=1}^{k}jS_j\right)^m\right]\mathbb{P}(X_a=k)$$

https://doi.org/10.1017/jpr.2026.10107 Published online by Cambridge University Press

$$= \sum_{k=0}^{\infty} (1-\rho)\rho^k \mathbb{E}\left[\sum_{\substack{n_1+\cdots+n_k=m \\ n_1,\ldots,n_k\geq 0}} \binom{m}{n_1, n_2, \ldots, n_k} \prod_{j=1}^{k} (j \cdot S_j)^{n_j} \right]$$

$$= \sum_{k=0}^{\infty} (1-\rho)\rho^k \sum_{\substack{n_1+\cdots+n_k=m \\ n_1,\ldots,n_k\geq 0}} \binom{m}{n_1, n_2, \ldots, n_k} \prod_{j=1}^{k} j^{n_j} \cdot \mathbb{E}\big[S_j^{n_j}\big]$$

$$= \frac{m!}{\mu^m} \sum_{k=0}^{\infty} (1-\rho)\rho^k \sum_{\substack{n_1+\cdots+n_k=m \\ n_1,\ldots,n_k\geq 0}} \prod_{j=1}^{k} j^{n_j}.$$

This completes the proof. □

2.2. The marginal distribution of T_d

As is well known, the sojourn time Σ of tagged customer C in the $M_\lambda/M_\mu/1$ queue is $\mathrm{Exp}(\mu-\lambda)$-distributed. To obtain the distribution of the total overlapping time T_d of C with the customers who arrive during her sojourn time, we reason as follows. If $\Sigma = x$, then $N_\lambda(x)$ customers arrive during C's sojourn time, where $N_\lambda(x)$ is Poisson (λx)-distributed. Their arrival epochs are distributed as the order statistics of $N_\lambda(x)$ i.i.d. uniformly distributed random variables on $(0, x)$ (see property $\mathcal{P}$); we denote them by $U_{(j)}(0, x)$, for $j = 1, \ldots, N_\lambda(x)$. Hence

$$T_d\Big|_{\Sigma=x, N_\lambda(x)=n} \stackrel{d}{=} \sum_{j=1}^{n} (x - U_{(j)}(0, x)).$$

We shall use a well-known fact about the sum of i.i.d. uniformly distributed random variables.

Lemma 3. *Let G_n be the sum of n i.i.d. uniform (0, x) random variables. Then the CDF of G_n is equal to*

$$F_n(t) = \frac{1}{n!} \sum_{k=0}^{\lfloor t/x \rfloor} (-1)^k \binom{n}{k} \left(\frac{t}{x} - k\right)^n, \quad 0 \leq t \leq nx.$$

Proof. This is a well-known fact; the CDF is called the Irwin–Hall distribution [14, 16]. □

Denoting $\bar{F}_n(x) = 1 - F_n(x)$, we have the following expression for $\mathbb{P}(T_d > t)$.

Proposition 5. *The tail distribution of T_d is given by*

$$\mathbb{P}(T_d > t) = \int_{x=0}^{\infty} (\mu-\lambda)\, \mathrm{e}^{-(\mu-\lambda)x}\, \mathbb{E}\left[\bar{F}_{N_\lambda(x)}\left(\frac{t}{x}\right)\right] \mathrm{d}x.$$

Proof. We have

$$\mathbb{P}(T_d > t) = \int_{x=0}^{\infty} (\mu-\lambda)\, \mathrm{e}^{-(\mu-\lambda)x}\, \mathbb{P}\big(T_d > t \mid \Sigma = x\big)\, \mathrm{d}x$$

$$= \int_{x=0}^{\infty} (\mu-\lambda)\, \mathrm{e}^{-(\mu-\lambda)x} \sum_{n=0}^{\infty} \mathrm{e}^{-\lambda x} \frac{(\lambda x)^n}{n!} \mathbb{P}\left(\sum_{j=1}^{n} (x - U_{(j)}(0, x)) > t\right) \mathrm{d}x$$

https://doi.org/10.1017/jpr.2026.10107 Published online by Cambridge University Press

$$= \int_{x=0}^{\infty} (\mu - \lambda)\, \mathrm{e}^{-(\mu-\lambda)x} \sum_{n=0}^{\infty} \mathrm{e}^{-\lambda x} \frac{(\lambda x)^n}{n!} \mathbb{P}\left(\sum_{j=1}^{n} U_{(j)}(0, x) > t \right) \mathrm{d}x$$

$$= \int_{x=0}^{\infty} (\mu - \lambda)\, \mathrm{e}^{-(\mu-\lambda)x} \sum_{n=0}^{\infty} \mathrm{e}^{-\lambda x} \frac{(\lambda x)^n}{n!} \bar{F}_n\left(\frac{t}{x} \right) \mathrm{d}x$$

$$= \int_{x=0}^{\infty} (\mu - \lambda)\, \mathrm{e}^{-(\mu-\lambda)x} \mathbb{E}\left[\bar{F}_{N_\lambda(x)}\left(\frac{t}{x} \right) \right] \mathrm{d}x. \tag{9}$$

□

Remark 1. A more compact expression for $\mathbb{P}(T_d > t)$ immediately follows from the above proposition:

$$\mathbb{P}(T_d > t) = \mathbb{E}\left[\bar{F}_{N_\lambda(\mathcal{E}_{\mu-\lambda})}\left(\frac{t}{\mathcal{E}_{\mu-\lambda}} \right) \right].$$

We now derive an expression for the LST of T_d.

Lemma 4. *The LST of T_d is given by*

$$\mathbb{E}\left[\mathrm{e}^{-\theta T_d}\right] = (1-\rho) \int_{x=0}^{\infty} \mu\, \mathrm{e}^{-\mu x}\, \mathrm{e}^{(\lambda/\theta)(1-\mathrm{e}^{-\theta x})}\, \mathrm{d}x = (\mu - \lambda) \left(\frac{\theta}{\lambda} \right)^{\mu/\theta} \frac{1}{\theta}\, \mathrm{e}^{\lambda/\theta}\, \gamma\left(\frac{\mu}{\theta}, \frac{\lambda}{\theta} \right). \tag{10}$$

Proof. Conditioning on the sojourn time Σ of C being equal to x and subsequently on the number of arrivals during x being equal to n we can write (very similar to (9) above)

$$\mathbb{E}\left[\mathrm{e}^{-\theta T_d}\right] = \int_{x=0}^{\infty} (\mu - \lambda)\, \mathrm{e}^{-(\mu-\lambda)x} \mathbb{E}\left[\mathrm{e}^{-\theta T_d} \mid \Sigma = x\right] \mathrm{d}x$$

$$= \int_{x=0}^{\infty} (\mu - \lambda)\, \mathrm{e}^{-\mu x} \sum_{n=0}^{\infty} \frac{(\lambda x)^n}{n!} \left(\frac{1 - \mathrm{e}^{-\theta x}}{\theta x} \right)^n \mathrm{d}x \quad \text{(property } \mathcal{P}\text{)}$$

$$= \int_{x=0}^{\infty} (\mu - \lambda)\, \mathrm{e}^{-\mu x}\, \mathrm{e}^{(\lambda/\theta)(1-\mathrm{e}^{-\theta x})}\, \mathrm{d}x.$$

□

Remark 2. The expressions in (5) and (10) reveal that T_a and T_d are identically distributed in the $\mathrm{M}_\lambda/\mathrm{M}_\mu/1$ FCFS queue. This is a much stronger result than that of Xu et al. [31] or Oz and Perlman [21], who only show this for expectations. Moreover, it highlights an important distinction between the single server queue and the infinite server queue. In the infinite server queue T_a and T_d have the same mean; however, their distributions are not identical (see Boxma et al. [8]).

By differentiating the first expression in either (5) or (10) (which perhaps works best by taking the Taylor series development of $((1 - \mathrm{e}^{-\theta x})/\theta)$, it is readily seen that the first moments of T_a and T_d are given by

$$\mathbb{E}[T_a] = \mathbb{E}[T_d] = \frac{\rho}{\mu(1-\rho)^2}, \tag{11}$$

$$\mathbb{E}[T_a^2] = \mathbb{E}[T_d^2] = \frac{2\rho + 4\rho^2}{\mu^2(1-\rho)^4}, \tag{12}$$

$$\mathrm{Var}[T_a] = \mathrm{Var}[T_d] = \frac{2\rho + 3\rho^2}{\mu^2(1-\rho)^4}. \tag{13}$$

https://doi.org/10.1017/jpr.2026.10107 Published online by Cambridge University Press

In view of the fact that the study of overlapping times is still in its infancy, it seems justified to close this section by discussing an alternative way to study overlapping times. Suppose that C is customer zero, and that there was an infinite stream of customers $-1, -2, \ldots$ before C, and an infinite stream of customers $1, 2, \ldots$ after C. Define O_n as the overlap time of C with customer n, $n = \ldots, -2, -1, 1, 2, \ldots$. Then we can write

$$T_a = \sum_{n=1}^{\infty} O_{-n}, \quad T_d = \sum_{n=1}^{\infty} O_n. \tag{14}$$

Successive overlap times depend on each other, and it is not *a priori* clear how to obtain the distribution of T_a and T_d from (14). However, the fact that O_n contributes in the same way to the overlap time of C with customer n as to the overlap time of customer n with C gives us some insight into the fact that T_a and T_d have the same distribution. Furthermore, $\mathbb{E}[T_a]$ and $\mathbb{E}[T_d]$ easily follow from (14) and the following proposition.

Proposition 6. *The mean overlap time of C and customer k in the* $\mathrm{M}_\lambda/\mathrm{M}_\mu/1$ *FCFS queue is given by*

$$\mathbb{E}[O_k] = \frac{\rho^k}{\mu - \lambda}, \quad k = 1, 2, \ldots.$$

Proof. We have

$$\mathbb{E}[O_k] = \int_0^\infty \mathbb{P}(O_k > t)\, \mathrm{d}t = \int_0^\infty \mathrm{e}^{-(\mu-\lambda)t} \rho^k \, \mathrm{d}t = \frac{\rho^k}{\mu - \lambda}.$$

In the second equality we have used the memoryless property of the $\mathrm{Exp}(\mu - \lambda)$-distributed sojourn time of C, which implies that $\mathbb{P}(O_k > t \mid O_k > 0) = \mathrm{e}^{-(\mu-\lambda)t}$, and we have also used the fact that

$$\mathbb{P}(O_k > 0) = \left(\frac{\lambda}{\lambda + (\mu - \lambda)}\right)^k = \rho^k$$

(the probability that there are at least k arrivals during the sojourn time of C). □

It follows from (14) and the above proposition that the mean overlapping time with customers that arrive after C is equal to

$$\mathbb{E}[T_d] = \mathbb{E}\left[\sum_{k=1}^{\infty} O_k\right] = \sum_{k=1}^{\infty} \mathbb{E}[O_k] = \sum_{k=1}^{\infty} \frac{\rho^k}{\mu - \lambda} = \frac{\rho}{\mu(1-\rho)^2},$$

in agreement with (11).

3. The joint LST of T_a and T_d

In order to calculate the joint LST of T_a and T_d, we again exploit the following observations about the $\mathrm{M}_\lambda/\mathrm{M}_\mu/1$ queue.

(i) The number of customers X_a seen by the tagged customer upon arrival is Geometric (ρ)-distributed.

https://doi.org/10.1017/jpr.2026.10107 Published online by Cambridge University Press

(ii) If $X_a = k$, then the sojourn time Σ of the tagged customer is Erlang $(k+1, \mu)$-distributed, hence with density $\mu((\mu t)^k/k!)\,\mathrm{e}^{-\mu t}$.

(iii) Property $\mathcal{P}$, as stated in the proof of Lemma 2.

Now we are ready to give the joint LST of T_a and T_d.

Theorem 1. *The joint LST of (T_a, T_d) is given by*

$$\mathbb{E}\big[\mathrm{e}^{-\theta_1 T_a - \theta_2 T_d}\big] = (1-\rho)\int_{t=0}^{\infty} \mu \mathrm{e}^{-\mu t} \exp\left(\lambda \frac{1-\mathrm{e}^{-\theta_1 t}}{\theta_1}\right) \mathrm{e}^{-\lambda t} \exp\left(\lambda \frac{1-\mathrm{e}^{-\theta_2 t}}{\theta_2}\right) \mathrm{d}t. \tag{15}$$

Proof. We have

$$\begin{aligned}
\mathbb{E}\big[\mathrm{e}^{-\theta_1 T_a - \theta_2 T_d}\big] &= \sum_{k=0}^{\infty} \mathbb{P}(X_a = k)\mathbb{E}\big[\mathrm{e}^{-\theta_1 T_a - \theta_2 T_d} \mid X_a = k\big] \\
&= \sum_{k=0}^{\infty} (1-\rho)\rho^k \mathbb{E}\big[\mathrm{e}^{-\theta_1 T_a - \theta_2 T_d} \mid X_a = k,\, \Sigma = t\big]\, \mathrm{d}\mathbb{P}(\Sigma < t \mid X_a = k) \\
&= \sum_{k=0}^{\infty} (1-\rho)\rho^k \int_{t=0}^{\infty} \mu \frac{(\mu t)^k}{k!}\, \mathrm{e}^{-\mu t} \mathbb{E}\big[\mathrm{e}^{-\theta_1 T_a - \theta_2 T_d} \mid X_a = k,\, \Sigma = t\big]\, \mathrm{d}t. \qquad (16)
\end{aligned}$$

To determine

$$\mathbb{E}\big[\mathrm{e}^{-\theta_1 T_a - \theta_2 T_d} \mid X_a = k,\, \Sigma = t\big],$$

we first observe that T_a and T_d are conditionally independent, given that $X_a = k$ and $\Sigma = t$. Let us now twice use (iii), i.e. the Poisson property $\mathcal{P}$. For T_d we have n arrivals during $\Sigma = t$ with probability $\mathrm{e}^{-\lambda t}(\lambda t)^n/n!$, and T_d is the sum of the times from their arrival until the end of Σ. Moreover, T_d does not depend on X_a, given $\Sigma = t$. Hence, as in (7), we have

$$\mathbb{E}\big[\mathrm{e}^{-\theta_2 T_d} \mid X_a = k,\, \Sigma = t\big] = \sum_{n=0}^{\infty} \mathrm{e}^{-\lambda t} \frac{(\lambda t)^n}{n!} \left(\frac{1-\mathrm{e}^{-\theta_2 t}}{\theta_2 t}\right)^n = \mathrm{e}^{-\lambda t} \exp\left(\lambda \frac{1-\mathrm{e}^{-\theta_2 t}}{\theta_2}\right). \tag{17}$$

For T_a, we observe that the k service completions of the $X_a = k$ customers seen at arrival occur in $(0, t)$ according to a Poisson process with rate μ. Hence

$$\mathbb{E}\big[\mathrm{e}^{-\theta_1 T_a} \mid X_a = k,\, \Sigma = t\big] = \left(\frac{1-\mathrm{e}^{-\theta_1 t}}{\theta_1 t}\right)^k. \tag{18}$$

The proof now follows by combining (16), (17), and (18), also using the fact that $\rho\mu = \lambda$. □

By successively taking $\theta_2 = 0$, $\theta_1 = 0$, and $\theta_1 = \theta_2 = \theta$ in (15), we get the marginal LST of T_a as already given in (5), the marginal LST of T_d as already given in (10), and the LST of the total overlapping time $T := T_a + T_d$.

Corollary 1. *We have*

$$\mathbb{E}\big[\mathrm{e}^{-\theta T}\big] = (1-\rho)\int_{t=0}^{\infty} \mu \mathrm{e}^{-\mu t}\, \mathrm{e}^{-\lambda t} \exp\left(2\lambda \frac{1-\mathrm{e}^{-\theta t}}{\theta}\right) \mathrm{d}t. \tag{19}$$

https://doi.org/10.1017/jpr.2026.10107 Published online by Cambridge University Press

Remark 3. Just as in (5) and (10), one can implement the substitution $e^{-\theta t} = y$, with $dy/dt = -\theta y$. This yields the following expression:

$$\mathbb{E}\left[e^{-\theta T}\right] = (1-\rho)\frac{\mu}{\theta}\int_{y=0}^{1} y^{(\mu+\lambda)/\theta-1}\, e^{2\lambda(1-y)/\theta}\, dy = (1-\rho)e^{2\lambda/\theta}\frac{\mu}{\theta}\int_{y=0}^{1} y^{(\mu+\lambda)/\theta-1} e^{-2\lambda y/\theta} dy. \tag{20}$$

The integral in (20) again is an incomplete gamma function. In particular, we can write

$$\mathbb{E}\left[e^{-\theta T}\right] = (1-\rho)\frac{\mu}{\theta}\, e^{2\lambda/\theta}\left(\frac{\theta}{2\lambda}\right)^{(\mu+\lambda)/\theta} \gamma\left(\frac{\mu+\lambda}{\theta}, \frac{2\lambda}{\theta}\right).$$

In order to get moments of T_a, T_d, and T, it seems best to differentiate the expressions in Corollary 1, using the Taylor series development of $(1 - e^{-\theta t})/\theta$. We thus obtain (11), (12), and (13), as well as

$$\mathbb{E}[T_a T_d] = \frac{7\rho^2}{\mu^2(1-\rho)^4}, \tag{21}$$

$$\mathrm{Cov}[T_a, T_d] = \frac{6\rho^2}{\mu^2(1-\rho)^4}, \tag{22}$$

$$\mathrm{Corr}[T_a, T_d] = \frac{5\rho}{2+3\rho}, \tag{23}$$

$$\mathbb{E}[T] = \frac{2\rho}{\mu(1-\rho)^2}, \tag{24}$$

$$\mathbb{E}[T^2] = \frac{4\rho + 20\rho^2}{\mu^2(1-\rho)^4}, \tag{25}$$

$$\mathrm{Var}[T] = \frac{4\rho + 16\rho^2}{\mu^2(1-\rho)^4}. \tag{26}$$

4. Heavy-traffic limits

In this section we study T and (T_a, T_d) in the heavy-traffic limit $\rho \uparrow 1$. In the heavy-traffic analysis of queue lengths and waiting times, which typically have means that behave as a constant over $1-\rho$ for $\rho \uparrow 1$, one usually has to scale by $1-\rho$ to obtain a non-trivial heavy-traffic limit; an exception is when service times are heavy-tailed (see Boxma and Cohen [6]). However, the fact that the means of T_a, T_d, and T behave as $(1-\rho)^{-2}$ for $\rho \uparrow 1$ suggests that we now have to scale by $(1-\rho)^2$. Below we shall see that this indeed gives a valid heavy-traffic limit. Remarkably, the resulting limit random variable is a Weibull random variable.

Theorem 2. *Let $T = T_a + T_d$ be the overlapping time in the $M_\lambda/M_\mu/1$ queue. Then, as we let $\rho \to 1$, we have the following heavy-traffic limit result:*

$$\lim_{\rho\uparrow 1}\mathbb{E}\left[e^{-(1-\rho)^2\theta T}\right] = \int_{z=0}^{\infty} \frac{1}{2}\sqrt{\frac{\mu}{z}}\, e^{-\sqrt{\mu z}}\, e^{-\theta z}\, dz = \mathbb{E}\left[e^{-\theta W}\right],$$

where W has the Weibull $(1/2, 1/\mu)$ *distribution with density* $\frac{1}{2}\sqrt{\mu/z}\, e^{-\sqrt{\mu z}}$.

https://doi.org/10.1017/jpr.2026.10107 Published online by Cambridge University Press

Proof. It follows from (19) that

$$\mathbb{E}\left[\mathrm{e}^{-(1-\rho)^2\theta T}\right] = (1-\rho)\mu \int_{t=0}^{\infty} \mathrm{e}^{-\mu(1-\rho)t} \exp\left(2\lambda\left[\frac{1-\mathrm{e}^{-(1-\rho)^2\theta t}}{(1-\rho)^2\theta} - t\right]\right) \mathrm{d}t.$$

Using the Taylor series development of e^a for $a \downarrow 0$, we can write

$$\lim_{\rho\uparrow 1}\mathbb{E}\left[\mathrm{e}^{-(1-\rho)^2\theta \mathrm{T}}\right] = \lim_{\rho\uparrow 1} (1-\rho)\mu \int_{\mathrm{t}=0}^{\infty} \mathrm{e}^{-\mu(1-\rho)\mathrm{t}}\, \mathrm{e}^{-\rho\mu(1-\rho)^2\theta \mathrm{t}^2}\, \mathrm{dt} = \int_{\mathrm{y}=0}^{\infty} \mu \mathrm{e}^{-\mu \mathrm{y} - \mu\theta \mathrm{y}^2}\, \mathrm{dy}.$$

In the last step we have used the transformation $y = (1-\rho)t$. One can invert this LST by using the transformation $z = \mu y^2$: for $\rho \uparrow 1$,

$$\mathbb{E}\left[\mathrm{e}^{-(1-\rho)^2\theta T}\right] \to \int_{z=0}^{\infty} \frac{1}{2}\sqrt{\frac{\mu}{z}}\, \mathrm{e}^{-\sqrt{\mu z}}\, \mathrm{e}^{-\theta z}\, \mathrm{d}z.$$

Hence Feller's convergence theorem implies that $(1-\rho)^2 T$ converges, for $\rho \uparrow 1$, to a random variable with density

$$t(z) := \frac{1}{2}\sqrt{\frac{\mu}{z}}\, \mathrm{e}^{-\sqrt{\mu z}}.$$

It is easily seen that this is the density of distribution $T(z) := 1 - \mathrm{e}^{-\sqrt{\mu z}}$. This, in turn, is the distribution of the *square* of an $\mathrm{Exp}(\sqrt{\mu})$-distributed random variable; put differently, it is a Weibull $(1/2,\, 1/\mu)$-distributed random variable. Interestingly, the heavy-traffic limit in this $\mathrm{M}_\lambda/\mathrm{M}_\mu/1$ model has a subexponential tail. □

Proceeding in exactly the same way, it follows from (5) and (10) that

$$\begin{aligned}
&\lim_{\rho\uparrow 1}\mathbb{E}\left[\mathrm{e}^{-(1-\rho)^2\theta \mathrm{T_a}}\right] \\
&= \lim_{\rho\uparrow 1}\mathbb{E}\left[\mathrm{e}^{-(1-\rho)^2\theta \mathrm{T_d}}\right] \\
&= \lim_{\rho\uparrow 1} (1-\rho)\mu \int_{t=0}^{\infty} \mathrm{e}^{-\mu(1-\rho)t}\, \mathrm{e}^{-\frac{1}{2}\rho\mu(1-\rho)^2\theta t^2}\, \mathrm{d}t \\
&= \int_{y=0}^{\infty} \mu \mathrm{e}^{-\mu y - \frac{1}{2}\mu\theta y^2}\, \mathrm{d}y,
\end{aligned}$$

which yields the following corollary.

Corollary 2. *We have*

$$\lim_{\rho\uparrow 1} \mathbb{E}\left[\mathrm{e}^{-(1-\rho)^2\theta T_a}\right] = \lim_{\rho\uparrow 1} \mathbb{E}\left[\mathrm{e}^{-(1-\rho)^2\theta T_d}\right] = \int_{z=0}^{\infty} \sqrt{\frac{\mu}{2z}}\, \mathrm{e}^{-\sqrt{2\mu z}}\, \mathrm{e}^{-\theta z}\, \mathrm{d}z,$$

that is, both T_a and T_d, scaled by $(1-\rho)^2$, in the heavy-traffic limit have a Weibull $(1/2,\, 1/(2\mu))$ *distribution.*

As we have seen, T_a and T_d are dependent, and their correlation approaches one in the heavy-traffic case (see (23)). In general, it is hard to obtain the distribution of the sum of two Weibull-distributed random variables; but in this heavy-traffic limit, apparently

https://doi.org/10.1017/jpr.2026.10107 Published online by Cambridge University Press

$(1-\rho)^2(T_a+T_d)$ also has a Weibull distribution. Incidentally, for overlap *numbers* in the $M_\lambda/M_\mu/1$ queue, a somewhat similar result is found in Boxma [5]: the overlap number O is geometrically distributed, just like its two components, namely the number of customers X_a seen at arrival of the tagged customer and the number of customers X_d arriving during the sojourn of the tagged customer.

Remark 4. Another example, for the $M_\lambda/G/1$ queue, where a performance measure has a mean value which behaves as $(1-\rho)^{-2}$ for $\rho \uparrow 1$ is the area underneath the workload and queue length process during a busy period [3, 10, 15]. This is not so surprising, because overlapping time and area seem to be somewhat related quantities. To determine the heavy-traffic limiting distribution of the workload area in the $M_\lambda/M_\mu/1$ queue, one could use an explicit expression for its LST in equation (1.3) of Borovkov et al. [3]. We leave the highly involved analysis to a future study.

5. Generalizations

In this section we consider a generalization of the analysis in Section 3: we determine the transform of the joint distribution of X_a, X_d, T_a, and T_d. As a by-product this gives the transform of the joint distribution of the total overlap number $X_O = X_a + X_d$ and total overlapping time $T = T_a + T_d$ in the $M_\lambda/M_\mu/1$ queue. Here the overlap number of a tagged customer is the number of customers with whom she shares some time in the system.

Theorem 3. *The joint transform of X_a, X_d, T_a, and T_d in the $M_\lambda/M_\mu/1$ queue is given by*

$$\mathbb{E}\Big[z_1^{X_a} z_2^{X_d}\, \mathrm{e}^{-\theta_1 T_a - \theta_2 T_d}\Big] = (1-\rho)\mu \int_{t=0}^{\infty} \mathrm{e}^{-(\mu+\lambda)t}\, \mathrm{e}^{\lambda(z_1/\theta_1)(1-\mathrm{e}^{-\theta_1 t})}\, \mathrm{e}^{\lambda(z_2/\theta_2)(1-\mathrm{e}^{-\theta_2 t})}\, \mathrm{d}t, \quad (27)$$

and the joint transform of total overlap number X_O and total overlapping time T in the $M_\lambda/M_\mu/1$ queue is given by

$$\mathbb{E}\big[z^{X_O}\, \mathrm{e}^{-\theta T}\big] = (1-\rho)\mu \int_{t=0}^{\infty} \mathrm{e}^{-(\mu+\lambda)t}\, \mathrm{e}^{2\lambda(z/\theta)(1-\mathrm{e}^{-\theta t})}\, \mathrm{d}t. \quad (28)$$

Proof. The proof is very similar to that of Theorem 1:

$$\begin{aligned}
&\mathbb{E}\Big[z_1^{X_a} z_2^{X_d}\, \mathrm{e}^{-\theta_1 T_a - \theta_2 T_d}\Big] \\
&= \sum_{k=0}^{\infty} z_1^k \mathbb{P}(X_a = k)\, \mathbb{E}\Big[z_2^{X_d}\, \mathrm{e}^{-\theta_1 T_a - \theta_2 T_d} \mid X_a = k\Big] \\
&= \sum_{k=0}^{\infty} (1-\rho)(\rho z_1)^k\, \mathbb{E}\Big[z_2^{X_d}\, \mathrm{e}^{-\theta_1 T_a - \theta_2 T_d} \mid X_a = k,\, \Sigma = t\Big]\, \mathrm{d}\mathbb{P}(\Sigma < t \mid X_a = k) \\
&= \sum_{k=0}^{\infty} (1-\rho)(\rho z_1)^k \int_{t=0}^{\infty} \mu \frac{(\mu t)^k}{k!}\, \mathrm{e}^{-\mu t}\, \mathbb{E}\Big[z_2^{X_d}\, \mathrm{e}^{-\theta_1 T_a - \theta_2 T_d} \mid X_a = k,\, \Sigma = t\Big]\, \mathrm{d}t
\end{aligned}$$

https://doi.org/10.1017/jpr.2026.10107 Published online by Cambridge University Press

$$
\begin{aligned}
&= \sum_{k=0}^{\infty} (1-\rho)(\rho z_1)^k \int_{t=0}^{\infty} \mu \frac{(\mu t)^k}{k!} \, \mathrm{e}^{-\mu t} \\
&\quad \times \mathbb{E}\big[\mathrm{e}^{-\theta_1 T_a} \mid X_a = k,\, \Sigma = t\big]\, \mathbb{E}\Big[z_2^{X_d}\, \mathrm{e}^{-\theta_2 T_d} \mid X_a = k,\, \Sigma = t\Big]\, \mathrm{d}t \\
&= \sum_{k=0}^{\infty} (1-\rho)(\rho z_1)^k \int_{t=0}^{\infty} \mu \frac{(\mu t)^k}{k!} \, \mathrm{e}^{-\mu t} \left(\frac{1-\mathrm{e}^{-\theta_1 t}}{\theta_1 t}\right)^k \\
&\quad \times \sum_{n=0}^{\infty} \mathrm{e}^{-\lambda t} \frac{(\lambda t)^n}{n!}\, z_2^n \left(\int_{u=0}^{t} \mathrm{e}^{-\theta_2 (t-u)} \frac{\mathrm{d}u}{t}\right)^n \mathrm{d}t.
\end{aligned}
$$

Equation (27) now follows by performing the two summations and realizing that $\rho\mu = \lambda$. Equation (28) subsequently follows from (27) by taking $z_1 = z_2 = z$ and $\theta_1 = \theta_2 = \theta$. □

Remark 5. The marginal transform (probability generating function) of the overlap number X_O for the $\mathrm{M}_\lambda/\mathrm{G}/1$ queue was already obtained in Boxma [5]. For the $\mathrm{M}_\lambda/\mathrm{M}_\mu/1$ queue, this gives

$$\mathbb{E}\big[z^{X_O}\big] = \frac{1-\rho}{1+\rho-2\rho z}, \tag{29}$$

corresponding to the geometric distribution

$$\mathbb{P}(X_O = k) = \left(1 - \frac{2\rho}{1+\rho}\right)\left(\frac{2\rho}{1+\rho}\right)^k, \quad k = 0, 1, \ldots,$$

with mean

$$\mathbb{E}\big[X_O\big] = \frac{2\rho}{1-\rho}$$

and variance

$$\mathrm{Var}\big[X_O\big] = \frac{2\rho + 2\rho^2}{(1-\rho)^2}.$$

Expression (29) also follows from (28) by realizing that

$$\lim_{\theta \downarrow 0} \frac{1-\mathrm{e}^{-\theta t}}{\theta} = t$$

and performing the ensuing integration in (28).

It readily follows from (28) by differentiation that

$$\mathrm{Cov}[X_O, T] = \frac{2\rho + 6\rho^2}{\mu(1-\rho)^3},$$

and finally the correlation between total overlap number X_O and total overlapping time T is found:

$$\mathrm{Corr}[X_O, T] = \frac{1+3\rho}{\sqrt{(2+2\rho)(1+4\rho)}}.$$

It is readily checked that this correlation is non-negative and increasing in ρ.

https://doi.org/10.1017/jpr.2026.10107 Published online by Cambridge University Press

In taking the heavy-traffic limit in (28) with respect to both X_O and T, we need to scale X_O by $1-\rho$ and T by $(1-\rho)^2$. Replacing z with e^{-s} in (28), we can write

$$\begin{aligned}
&\lim_{\rho\uparrow 1} \mathbb{E}\left[\mathrm{e}^{-(1-\rho)sX_O-(1-\rho)^2\theta T}\right] \\
&= \lim_{\rho\uparrow 1} \mu(1-\rho)\int_0^\infty \mathrm{e}^{-(\mu+\lambda)t}\exp\left(2\lambda\,\mathrm{e}^{-(1-\rho)s}\left(\frac{1-\mathrm{e}^{-(1-\rho)^2\theta t}}{(1-\rho)^2\theta}\right)\right)\mathrm{d}t \\
&= \lim_{\rho\uparrow 1} \mu(1-\rho)\int_0^\infty \mathrm{e}^{-\mu(1-\rho)t}\,\mathrm{e}^{-2\lambda(1-\rho)st}\,\mathrm{e}^{-\lambda(1-\rho)^2\theta t^2}\,\mathrm{d}t \\
&= \int_0^\infty \mu\mathrm{e}^{-\mu y}\,\mathrm{e}^{-2\mu sy}\,\mathrm{e}^{-\mu\theta y^2}\,\mathrm{d}y. \qquad (30)
\end{aligned}$$

In the third line of (30) we have omitted terms of smaller $(1-\rho)$ order in the Taylor series developments of exponents, and in the fourth line we have used the transformation $y=(1-\rho)t$ while also replacing λ with μ.

The expression in the last line of (30), to be called $\psi(s,\theta)=\mathbb{E}\left[\mathrm{e}^{-sG-\theta H}\right]$, is the double transform of $(G,H) := \lim_{\rho\uparrow 1}((1-\rho)X_O,(1-\rho)^2T)$. It is readily verified that

$$G\sim \mathrm{Exp}\left(\frac{1}{2}\right) \quad \text{and} \quad H\sim \mathrm{Weibull}\left(\frac{1}{2},\frac{1}{\mu}\right),$$

and that

$$\mathbb{E}[G]=2, \quad \mathrm{Var}[G]=4, \quad \mathbb{E}[H]=\frac{2}{\mu}, \quad \mathrm{Var}\,[H]=\frac{20}{\mu^2},$$
$$\mathbb{E}[GH]=\frac{12}{\mu}, \quad \mathrm{Cov}[G,H]=\frac{8}{\mu}, \quad \mathrm{Corr}[G,H]=\frac{1}{\sqrt{1.25}}$$

(which is consistent with the above moments of X_O and T).

6. Overlapping time in the $\mathbf{M_\lambda/M_\mu/1}$ queue infection model

In this section we analyze the infection queueing models of Newman and Pender [19, 20] where each customer has a probability q of being infectious. We are interested in the amount of time that the tagged customer C spends with infectious customers. We now let Z_a be the total time C spends with infectious customers who are already present at her arrival, and Z_d be the total time C spends with infectious customers who arrive during her sojourn. Below, we derive the LST of (Z_a, Z_d).

Theorem 4. *The joint LST of (Z_a, Z_d) is given by*

$$\begin{aligned}
&\mathbb{E}\left[\mathrm{e}^{-\theta_1 Z_a-\theta_2 Z_d}\right] \\
&\quad = (1-\rho)\int_{t=0}^\infty \mu\mathrm{e}^{-\mu t}\exp\left(\lambda\left(q\frac{1-\mathrm{e}^{-\theta_1 t}}{\theta_1}+(1-q)t\right)\right)\mathrm{e}^{-\lambda qt}\exp\left(\lambda q\frac{1-\mathrm{e}^{-\theta_2 t}}{\theta_2}\right)\mathrm{d}t.
\end{aligned}$$

https://doi.org/10.1017/jpr.2026.10107 Published online by Cambridge University Press

Proof. We have

$$\begin{aligned}\mathbb{E}\big[\mathrm{e}^{-\theta_1 Z_a-\theta_2 Z_d}\big] &= \sum_{k=0}^{\infty} \mathbb{P}(X_a = k)\mathbb{E}(\mathrm{e}^{-\theta_1 Z_a-\theta_2 Z_d} \mid X_a = k)\\ &= \sum_{k=0}^{\infty} (1-\rho)\rho^k \mathbb{E}\big[\mathrm{e}^{-\theta_1 Z_a-\theta_2 Z_d} \mid X_a = k,\, \Sigma = t\big]\, \mathrm{d}\mathbb{P}(\Sigma < t \mid X_a = k)\\ &= \sum_{k=0}^{\infty} (1-\rho)\rho^k \int_{t=0}^{\infty} \mu \frac{(\mu t)^k}{k!}\, \mathrm{e}^{-\mu t}\mathbb{E}\big[\mathrm{e}^{-\theta_1 Z_a-\theta_2 Z_d} \mid X_a = k,\, \Sigma = t\big]\, \mathrm{d}t. \quad (31)\end{aligned}$$

Just as in the proof of Theorem 1 we observe that Z_a and Z_d are conditionally independent, given that $X_a = k$ and $\Sigma = t$, and that Z_d does not depend on X_a, given Σ. The LST of Z_d given $\Sigma = t$ is equal to the LST of T_d given $\Sigma = t$, except that λ should be replaced by λq, which is the arrival rate of infectious customers (because for T_d it does not make a difference whether or not non-infected customers arrive after C joined the system). Hence, using (17),

$$\mathbb{E}\big[\mathrm{e}^{-\theta_2 Z_d} \mid \Sigma = t\big] = \mathrm{e}^{-\lambda q t} \exp\left(\lambda q\left(\frac{1-\mathrm{e}^{-\theta_2 t}}{\theta_2}\right)\right). \quad (32)$$

Finally, consider Z_a. We can write

$$\begin{aligned}\mathbb{E}\big[\mathrm{e}^{-\theta_1 Z_a} \mid X_a = k,\, \Sigma = t\big] &= \sum_{j=0}^{k} \binom{k}{j} q^j \left(\frac{1}{t}\int_0^t \mathrm{e}^{-\theta_1 u}\, \mathrm{d}u\right)^j (1-q)^{k-j}\\ &= \left[q\left(\frac{1-\mathrm{e}^{-\theta_1 t}}{\theta_1 t}\right) + 1 - q\right]^k. \quad (33)\end{aligned}$$

Combining (31), (32), and (33), we get

$$\begin{aligned}&\mathbb{E}\big[\mathrm{e}^{-\theta_1 Z_a-\theta_2 Z_d}\big]\\ &= \sum_{k=0}^{\infty} (1-\rho)\rho^k\\ &\quad \times \int_{t=0}^{\infty} \mu \frac{(\mu t)^k}{k!}\, \mathrm{e}^{-\mu t}\left[q\left(\frac{1-\mathrm{e}^{-\theta_1 t}}{\theta_1 t}\right) + 1 - q\right]^k \mathrm{e}^{-\lambda q t} \exp\left(\lambda q\left(\frac{1-\mathrm{e}^{-\theta_2 t}}{\theta_2}\right)\right) \mathrm{d}t,\end{aligned}$$

and summing over k gives the statement of the theorem. □

It immediately follows from the above theorem that the total overlapping time Z with infectious customers has LST

$$\begin{aligned}\mathbb{E}\big[\mathrm{e}^{-\theta Z}\big] &= \mathbb{E}\big[\mathrm{e}^{-\theta(Z_a+Z_d)}\big]\\ &= (1-\rho)\mu \int_0^{\infty} \mathrm{e}^{-\mu t+\lambda(1-q)t} \exp\left(\lambda q\left(\frac{1-\mathrm{e}^{-\theta t}}{\theta}\right)\right) \mathrm{e}^{-\lambda q t} \exp\left(\lambda q\left(\frac{1-\mathrm{e}^{-\theta t}}{\theta}\right)\right) \mathrm{d}t\\ &= (1-\rho)\mu \int_0^{\infty} \mathrm{e}^{-(\mu-\lambda)t}\, \mathrm{e}^{-2\lambda q t} \exp\left(2\lambda q\left(\frac{1-\mathrm{e}^{-\theta t}}{\theta}\right)\right) \mathrm{d}t. \quad (34)\end{aligned}$$

https://doi.org/10.1017/jpr.2026.10107 Published online by Cambridge University Press

The LSTs of Z_a and of Z_d immediately follow from Theorem 4 by taking $\theta_2 = 0$ (resp. $\theta_1 = 0$):

$$\mathbb{E}\left[e^{-\theta Z_a}\right] = \mathbb{E}\left[e^{-\theta Z_d}\right] = (1-\rho)\mu \int_0^\infty e^{-\mu t} \exp\left(\lambda q\left(\frac{1-e^{-\theta t}}{\theta}\right)\right) dt. \tag{35}$$

Moments of Z, Z_a, and Z_d can be derived in exactly the same way in which moments of T, T_a, and T_d were obtained in previous sections. In particular,

$$\mathbb{E}[Z_a] = \mathbb{E}[Z_d] = \frac{\rho q}{\mu(1-\rho)^2},$$

$$\mathbb{E}[Z] = \frac{2\rho q}{\mu(1-\rho)^2},$$

$$\mathbb{E}[Z^2] = \frac{4q\rho(1-\rho) + 24q^2\rho^2}{\mu^2(1-\rho)^4}.$$

For $q = 1$, we find agreement with results from Section 3.

The heavy-traffic behavior of Z, Z_a, and Z_d can be derived from (34) and (35) in exactly the same way in which the heavy-traffic limits in Section 4 were obtained. We find

$$\lim_{\rho\uparrow 1} \mathbb{E}\left[e^{-(1-\rho)^2\theta Z}\right] = \int_0^\infty \mu e^{-\mu y} e^{-\mu q\theta y^2} dy = \int_0^\infty \frac{1}{2}\sqrt{\frac{\mu}{qz}}\, e^{-\sqrt{\mu z/q}}\, e^{-\theta z} dz,$$

and

$$\begin{aligned}\lim_{\rho\uparrow 1} \mathbb{E}\left[e^{-(1-\rho)^2\theta Z_a}\right] &= \lim_{\rho\uparrow 1} \mathbb{E}\left[e^{-(1-\rho)^2\theta Z_d}\right] \\ &= \int_0^\infty \mu e^{-\mu y} e^{-\frac{1}{2}\mu q\theta y^2} dy \\ &= \int_0^\infty \sqrt{\frac{\mu}{2qz}}\, e^{-\sqrt{\mu z/q}}\, e^{-\theta z} dz.\end{aligned}$$

In other words, in the heavy-traffic limit, $(1-\rho)^2 Z$ has a Weibull $(1/2, q/\mu)$ distribution, and $(1-\rho)^2 Z_a$ and $(1-\rho)^2 Z_d$ have a Weibull $(1/2, q/(2\mu))$ distribution.

7. The $M_\lambda/G/1$ queue

In this section we briefly consider the overlapping time distribution in the $M_\lambda/G/1$ queue, with a general non-specified service time distribution. We obtain expressions for the LST of T_a and of T_d, and point out why it is difficult to obtain the LST of their joint distribution.

Consider the $M_\lambda/G/1$ queue with generic service time S. If a tagged customer meets a non-empty system, then we denote the residual part of the ongoing service by R; as before, X_a denotes the number of customers seen upon arrival by the tagged customer. We can write

$$\begin{aligned}&\mathbb{E}\left[e^{-\theta_1 T_a - \theta_2 T_d}\right] \\ &\quad = \mathbb{P}(X_a = 0)\mathbb{E}\left[e^{-\lambda S} \exp\left(\lambda \frac{1-e^{-\theta_2 S}}{\theta_2}\right)\right]\end{aligned}$$

https://doi.org/10.1017/jpr.2026.10107 Published online by Cambridge University Press

$$+\sum_{k=1}^{\infty}\int_{v=0}^{\infty}\mathbb{P}(X_a=k,\,R\in(v,\,v+\mathrm{d}v))$$

$$\times\,\mathbb{E}\Big[\mathrm{e}^{-\left[kv+\sum_{j=1}^{k-1}(k-j)S_j\right]\theta_1}\,\mathrm{e}^{-\lambda\left(v+\sum_{j=1}^{k-1}S_j+S\right)}\exp\left(\lambda\frac{1-\mathrm{e}^{-(v+\sum_{j=1}^{k-1}S_j+S)\theta_2}}{\theta_2}\right)\Big]. \tag{36}$$

Here the expectation term in the first line of the right-hand side follows by using property $\mathcal{P}$; that property is also used in the last line for the θ_2-term, while the θ_1-term is obtained by realizing that, given $X_a=k$ and $R=v$, we have $T_a=kv+\sum_{j=1}^{k-1}(k-j)S_j$.

For $\mathrm{M}_\lambda/\mathrm{M}_\mu/1$, X_a and R are independent, and we have an explicit expression for $\mathbb{P}(X_a=k,\,R\in(v,\,v+\mathrm{d}v))$. However, for any other service time distribution, X_a and R are dependent and we only have their transform $\mathbb{E}\big[z^{X_a}\,\mathrm{e}^{-\theta R}\big]$; see for example the work by Boxma [4] and Wishart [30]. This transform is given by

$$\mathbb{E}\big[z^{X_a}\,\mathrm{e}^{-\theta R}I(X_a>0)\big]=(1-\rho)z\frac{\lambda(1-z)}{\mathbb{E}\big[\mathrm{e}^{-\lambda(1-z)S}\big]-z}\frac{\mathbb{E}\big[\mathrm{e}^{-\theta S}\big]-\mathbb{E}\big[\mathrm{e}^{-\lambda(1-z)S}\big]}{\lambda(1-z)-\theta}, \tag{37}$$

with $I(\cdot)$ denoting an indicator function. It seems very hard to work out the inverse of this double transform, and subsequently (36).

Below we (i) determine the marginal LST of T_d (for which an alternative derivation is possible, resulting in a much simpler expression), (ii) consider the marginal LST of T_a (which is quite complicated), (iii) indicate how (36) simplifies for $\mathrm{M}_\lambda/\mathrm{E}_2/1$ and $\mathrm{M}_\lambda/\mathrm{D}/1$, and (iv) obtain expressions for $\mathbb{E}[T_a]$ and $\mathbb{E}[T_d]$ and indicate how higher moments can be derived.

(i) The marginal LST of T_d. To derive the marginal LST of T_d, we do not need to use (36). Instead we can exploit the following. Firstly, the LST of the sojourn time Σ in the $\mathrm{M}_\lambda/\mathrm{G}/1$ queue is known. Secondly, given that $\Sigma=t$ and that there are j arrivals in $(0,\,t)$, those arrival epochs are distributed as the order statistics of j random variables which are $U(0,\,t)$-distributed. Using property $\mathcal{P}$, we thus obtain

$$\mathbb{E}\big[\mathrm{e}^{-\theta T_d}\big]=\int_{t=0}^{\infty}\sum_{j=0}^{\infty}\mathrm{e}^{-\lambda t}\frac{(\lambda t)^j}{j!}\left(\frac{1-\mathrm{e}^{-\theta t}}{\theta t}\right)^j\mathrm{d}\mathbb{P}(\Sigma<t)=\int_{t=0}^{\infty}\mathrm{e}^{-\lambda t+(\lambda/\theta)(1-\mathrm{e}^{-\theta t})}\,\mathrm{d}\mathbb{P}(\Sigma<t). \tag{38}$$

For service time distributions for which the distribution of Σ is explicitly known, this might allow one to obtain a fairly tractable expression for the LST of T_d. Furthermore, moments of T_d can be derived in a quite straightforward manner (see (iv) below).

(ii) The marginal LST of T_a. From (36),

$$\mathbb{E}\big[\mathrm{e}^{-\theta T_a}\big]=\mathbb{P}(X_a=0)+\sum_{k=1}^{\infty}\prod_{j=1}^{k-1}\mathbb{E}\big[\mathrm{e}^{-(k-j)\theta S_j}\big]\int_0^{\infty}\mathrm{e}^{-k\theta v}\,\mathrm{d}_v\mathbb{P}(X_a=k,\,R<v)$$

$$=1-\rho+\sum_{k=1}^{\infty}\prod_{j=1}^{k-1}\mathbb{E}\big[\mathrm{e}^{-j\theta S}\big]\mathbb{E}\big[\mathrm{e}^{-k\theta R}I(X_a=k)\big]. \tag{39}$$

Inversion of (37) with respect to the generating function of X_a would give $\mathbb{E}\big[\mathrm{e}^{-\theta R}I(X_a=k)\big]$, which features in (39). However, even such an inversion in one of the two variables is not

https://doi.org/10.1017/jpr.2026.10107 Published online by Cambridge University Press

straightforward (it could be done formally via contour integration). For a general service time distribution, one could derive moments of T_a from (39); see (iv) below.

(iii) The $M_\lambda/E_2/1$ and $M_\lambda/D/1$ cases. In the case of an Erlang-2-distributed service time with mean $2/\mu$, one has

$$\mathbb{E}\big[e^{-\theta S}\big] = \left(\frac{\mu}{\mu+\theta}\right)^2.$$

Observing that the denominator of (37) contains a factor $1-z$ and the numerator a factor $\theta - \lambda(1-z)$, (37) simplifies to

$$\mathbb{E}\big[z^{X_a} e^{-\theta R} I(X_a > 0)\big] = \left(\frac{\mu}{\mu+\theta}\right)^2 \frac{\lambda(1-\rho)z}{\lambda^2 z^2 - (\lambda^2 + 2\lambda\mu)z + \mu^2}[2\mu + \theta + \lambda - \lambda z].$$

This can be rewritten in the form

$$\mathbb{E}\big[z^{X_a} e^{-\theta R} I(X_a > 0)\big] = \left(\frac{\mu}{\mu+\theta}\right)^2 z\left[\frac{A_1(\theta)}{z-z_1} + \frac{A_2(\theta)}{z-z_2}\right],$$

with z_1 and z_2 the two zeros of $\lambda^2 z^2 - (\lambda^2 + 2\lambda\mu)z + \mu^2 = 0$, while $A_1(\theta)$, $A_2(\theta)$ are also easily determined. It is now trivial to obtain $\mathbb{E}\big[e^{-\theta R} I(X_a = k)\big]$, and even $\mathbb{P}(X_a = k, R < v)$. A similar procedure can be followed for $M/E_k/1$ with $k \geq 3$, but now one has to find the zeros $z_1, \ldots, z_k$ of a kth-order polynomial in z.

If the service time equals a constant, say D, then equation (36) simplifies to

$$\begin{aligned}\mathbb{E}\big[e^{-\theta_1 T_a - \theta_2 T_d}\big] &= \mathbb{P}(X_a = 0)\, e^{-\lambda D} \exp\left(\lambda \frac{1 - e^{-\theta_2 D}}{\theta_2}\right) \\ &\quad + \sum_{k=1}^{\infty} \int_{v=0}^{\infty} \mathbb{P}(X_a = k, R \in (v, v + dv)) \\ &\quad \times e^{-[kv + \frac{1}{2}k(k-1)D]\theta_1}\, e^{-\lambda(v+kD)} \exp\left(\lambda \frac{1 - e^{-(v+kD)\theta_2}}{\theta_2}\right). \qquad (40)\end{aligned}$$

For $M_\lambda/D/1$, (37) becomes

$$\mathbb{E}\big[z^{X_a} e^{-\theta R} I(X_a > 0)\big] = (1-\rho)z \frac{\lambda(1-z)}{e^{-\lambda(1-z)} - z} \frac{e^{-\theta D} - e^{-\theta\lambda(1-z)}}{\lambda(1-z) - \theta}.$$

Even if one is able to invert this transform, for example by using contour integration expressions such as

$$\mathbb{P}(X_a = k) = \frac{1}{2\pi i} \int \frac{\mathbb{E}[z^{X_a}]}{z^{k+1}}\, dz,$$

then one still has to sum over k, with terms like $e^{-\frac{1}{2}k(k-1)D\theta_1}$, in order to evaluate (40), and this seems to be very difficult.

https://doi.org/10.1017/jpr.2026.10107 Published online by Cambridge University Press

(iv) The moments of T_a and T_d for $M_\lambda/G/1$. Via differentiation of the expression in (36), or (39) and (38), or by reasonably straightforward probabilistic reasoning, it follows that

$$\begin{aligned}\mathbb{E}[T_a] &= \sum_{k=1}^{\infty}\int_{v=0}^{\infty}\left(kv + \sum_{j=1}^{k-1}(k-j)\mathbb{E}[S_j]\right)\mathbb{P}(X_a = k, R \in (v, v+\mathrm{d}v)) \\ &= \sum_{k=1}^{\infty}\int_{v=0}^{\infty}\left(kv + \frac{1}{2}k(k-1)\mathbb{E}[S]\right)\mathbb{P}(X_a = k, R \in (v, v+\mathrm{d}v)) \\ &= \mathbb{E}[X_a R] + \frac{1}{2}\mathbb{E}[X_a(X_a-1)]\mathbb{E}[S], \end{aligned} \tag{41}$$

$$\begin{aligned}\mathbb{E}[T_d] &= \mathbb{P}(X_a = 0)\lambda\frac{\mathbb{E}[S^2]}{2} \\ &\quad + \sum_{k=1}^{\infty}\int_{v=0}^{\infty}\frac{\lambda}{2}\mathbb{E}\left[\left(v + \sum_{j=1}^{k-1}S_j + S\right)^2\right]\mathbb{P}(X_a = k, R \in (v, v+\mathrm{d}v)) \\ &= (1-\rho)\lambda\frac{\mathbb{E}[S^2]}{2} + \frac{\lambda}{2}\rho\mathbb{E}[R^2] + \rho\mathbb{E}[X_a R] \\ &\quad + \frac{\lambda}{2}\mathbb{E}[X_a(X_a-1)]\mathbb{E}[S]^2 + \frac{\lambda}{2}\mathbb{E}[X_a]\mathbb{E}[S^2]. \end{aligned} \tag{42}$$

From formula (2.7) of Boxma [4] we have

$$\mathbb{E}[X_a R] = \frac{\lambda^3(\mathbb{E}[S^2])^2}{4(1-\rho)} + \frac{\lambda^2\mathbb{E}[S^3]}{6} + \frac{\lambda\mathbb{E}[S^2]}{2}.$$

An expression for $\mathbb{E}[X_a(X_a-1)]$ rarely occurs in the literature. It readily follows from the well-known expression for the PGF $\mathbb{E}[z^{X_a}]$ (see equation (II.4.17) in Cohen [11]) that, for $M_\lambda/G/1$ FCFS,

$$\mathbb{E}[X_a(X_a-1)] = \frac{\lambda^4(\mathbb{E}[S^2])^2}{2(1-\rho)^2} + \frac{\lambda^3\mathbb{E}[S^3]}{3(1-\rho)} + \frac{\lambda^2\mathbb{E}[S^2]}{1-\rho}.$$

Now, working out the expressions in (41) and (42) reveals that

$$\mathbb{E}[T_a] = \mathbb{E}[T_d] = \frac{\lambda^3(\mathbb{E}[S^2])^2}{4(1-\rho)^2} + \frac{\lambda^2\mathbb{E}[S^3]}{6(1-\rho)} + \frac{\lambda\mathbb{E}[S^2]}{2(1-\rho)}, \tag{43}$$

and hence the mean total overlapping time in the $M_\lambda/G/1$ FCFS queue is given by

$$\mathbb{E}[T] = \frac{\lambda^3(\mathbb{E}[S^2])^2}{2(1-\rho)^2} + \frac{\lambda^2\mathbb{E}[S^3]}{3(1-\rho)} + \frac{\lambda\mathbb{E}[S^2]}{1-\rho}.$$

For $M_\lambda/M_\mu/1$ FCFS, this reduces to the expression obtained in (24).

To determine moments of T_d, it actually seems easier to differentiate the expression in (38) with respect to θ. In that way it readily follows that

$$\mathbb{E}[T_d] = \frac{\lambda}{2}\mathbb{E}[\Sigma^2], \tag{44}$$

$$\mathbb{E}[T_d^2] = \frac{\lambda^2}{4}\mathbb{E}[\Sigma^4] + \frac{\lambda}{3}\mathbb{E}[\Sigma^3]. \tag{45}$$

https://doi.org/10.1017/jpr.2026.10107 Published online by Cambridge University Press

Equation (44) is in agreement with Corollary 1 of Oz and Perlman [21], which states that $\mathbb{E}[T] = \lambda \mathbb{E}[\Sigma^2]$ for overtake-free queues with Poisson(λ) arrivals, where customers depart in the order of their arrival. In combination with an $M_\lambda/G/1$ result on page 256 of Cohen [11] we retrieve the second part of (43).

By differentiating (39) twice, and using the identities

$$\sum_{j=1}^{k-1} j^2 = \frac{1}{6}(k-1)k(2k-1) \quad \text{and} \quad \sum_{j=1}^{k-1} j^3 = \frac{1}{4}(k-1)^2k^2,$$

one can derive the second moment of T_a:

$$\begin{aligned}\mathbb{E}\big[T_a^2\big] &= \mathbb{E}\big[X_a^2R^2\big] + 2\mathbb{E}[S]\mathbb{E}\big[X_a^2(X_a-1)R\big] + \frac{1}{3}\mathbb{E}[S^2]\mathbb{E}\big[(X_a-1)X_a(2X_a-1)\big] \\ &\quad + \mathbb{E}[S^2]\mathbb{E}\left[\frac{1}{4}(X_a-1)^2X_a^2 - \frac{1}{6}(X_a-1)X_a(2X_a-1)\right]. \end{aligned} \tag{46}$$

Finally, by differentiating (36) once with respect to both z and θ, we get

$$\begin{aligned}\mathbb{E}\big[T_aT_d\big] &= \frac{1}{2}\mathbb{E}\big[X_aR^2\big] + \frac{1}{2}\mathbb{E}[S]\mathbb{E}\big[X_a^2R\big] + \frac{1}{4}\mathbb{E}[S]\mathbb{E}\big[(X_a^2 - X_a)R\big] \\ &\quad + \frac{1}{4}(\mathbb{E}[S])^2\mathbb{E}[X_a(X_a-1)] + \frac{1}{4}\mathbb{E}[S^2]\mathbb{E}[X_a(X_a-1)] \\ &\quad + \frac{1}{4}(\mathbb{E}[S])^2\mathbb{E}[X_a(X_a-1)(X_a-2)]. \end{aligned} \tag{47}$$

All these terms can be evaluated using (37). In particular, using (45)–(47), one can thus obtain the second moment of the total overlapping time for the $M_\lambda/G/1$ FCFS queue, using $\mathbb{E}[T^2] = \mathbb{E}[T_a^2] + 2\mathbb{E}[T_aT_d] + \mathbb{E}[T_d^2]$. We omit the details.

8. The $G/M_\mu/1$ queue

In this section we describe some extensions of our work to the $G/M_\mu/1$ queue, with generic interarrival time A and service rate μ where $\mu\mathbb{E}[A] > 1$. We show that understanding the overlapping time in this case is also quite non-trivial. We start with the LST of the overlapping time experienced by the tagged customer upon arrival, i.e. T_a. We let ρ^* be the unique solution to $x = \mathbb{E}[e^{-\mu(1-x)A}]$ in [0,1].

Lemma 5. *For the* $G/M_\mu/1$ *queue, let* $\mu\mathbb{E}[A] > 1$ *and* $\theta > 0$. *Then*

$$\mathbb{E}\big[e^{-\theta T_a}\big] = (1-\rho^*)\,{}_1F_1\left(1; 1+\frac{\mu}{\theta}; \frac{\rho^*\mu}{\theta}\right),$$

where ${}_1F_1(a;b;z)$ *denotes the confluent hypergeometric function of the first kind.*

https://doi.org/10.1017/jpr.2026.10107 Published online by Cambridge University Press

Proof. We have

$$\begin{aligned}
\mathbb{E}\big[\mathrm{e}^{-\theta T_a}\big] &= \sum_{k=0}^{\infty} \mathbb{P}(X_a = k)\mathbb{E}\big[\mathrm{e}^{-\theta T_a} \mid X_a = k\big] \\
&= \sum_{k=0}^{\infty} (1-\rho^*)(\rho^*)^k \prod_{j=1}^{k} \frac{\mu}{\mu + j\theta} \\
&= \sum_{k=0}^{\infty} (1-\rho^*)(\rho^*)^k \left(\frac{\mu}{\theta}\right)^k \prod_{j=1}^{k} \frac{1}{\mu/\theta + j} \\
&= (1-\rho^*) \sum_{k=0}^{\infty} \frac{(\rho^*)^k (\mu/\theta)^k}{(\mu/\theta+1)_k} \\
&= (1-\rho^*) \sum_{k=0}^{\infty} \frac{(\rho^*)^k (\mu/\theta)^k (1)_k}{(\mu/\theta+1)_k\, k!} \\
&= (1-\rho^*)\, {}_1F_1\left(1; 1+\frac{\mu}{\theta}; \frac{\rho^*\mu}{\theta}\right).
\end{aligned}$$

□

Remark 6. In the case where the arrival process is deterministic with $A \equiv \Delta$, yielding a $\mathrm{D}/\mathrm{M}_\mu/1$ queue, the fixed point for $x = \mathrm{e}^{-\mu(1-x)\Delta}$ equals

$$\rho^* = -\frac{W_{-1}(-\mathrm{e}^{-\Delta\mu}\Delta\mu)}{\Delta\mu},$$

where $W_{-1}(x)$ is defined as the principal branch of the Lambert W function when $-1/\mathrm{e} \le x \le 0$. Moreover, in the case where the arrival process is an Erlang $(2, \lambda)$, yielding an $\mathrm{E}_2/\mathrm{M}_\mu/1$ queue, we have the fixed point

$$\rho^* = \frac{\lambda}{\mu} + \frac{1}{2} - \sqrt{\frac{\lambda}{\mu} + \frac{1}{4}}.$$

It is challenging to calculate the LST for T_d for the $\mathrm{G}/\mathrm{M}_\mu/1$ queue. However, we show that it is possible to calculate it for the $\mathrm{D}/\mathrm{M}_\mu/1$ queue and show a connection to partial theta functions.

Lemma 6. *For the* $\mathrm{D}/\mathrm{M}_\mu/1$ *queue where* Δ *is the interarrival time and* $m = \lfloor x/\Delta \rfloor$*, then we have the following expression for the LST of* T_d*:*

$$\mathbb{E}\big[\mathrm{e}^{-\theta T_d}\big] = \sum_{m=0}^{\infty} \frac{\mu(1-\rho^*)\,\mathrm{e}^{-\mu(1-\rho^*)\Delta m}}{\mu(1-\rho^*)+\theta m}\big(1 - \mathrm{e}^{-\mu(1-\rho^*)\Delta}\,\mathrm{e}^{-\theta\Delta m}\big) \exp\left(-\theta\frac{m(m-1)}{2}\Delta\right).$$

Proof. When the sojourn time is equal to x, then it is only possible for $m = \lfloor x/\Delta \rfloor$ customers to enter the queue. This also implies that the overlapping time is equal to

$$T_d = \sum_{j=1}^{m} (x - j\Delta) = mx - \frac{m(m+1)}{2}\Delta = \frac{m(m-1)}{2}\Delta + m(x - m\Delta).$$

https://doi.org/10.1017/jpr.2026.10107 Published online by Cambridge University Press

Therefore, using the fact that the sojourn time in $G/M_\mu/1$ is $\mathrm{Exp}(\mu(1-\rho^*))$, we obtain the following LST expression for T_d:

$$\begin{aligned}\mathbb{E}\left[e^{-\theta T_d}\right] &= \int_0^\infty \mathbb{E}\left[e^{-\theta T_d} \mid \mathcal{S}=x\right]\mu(1-\rho^*)\, e^{-\mu(1-\rho^*)x}\, dx \\ &= \sum_{m=0}^{\infty} \int_{\Delta m}^{\Delta(m+1)} \mu(1-\rho^*)\, e^{-(\mu(1-\rho^*)+\theta m)x}\, e^{\theta m^2 \Delta} \exp\left(-\theta \frac{m(m-1)}{2}\Delta\right) dx \\ &= \sum_{m=0}^{\infty} \frac{\mu(1-\rho^*)\, e^{-\mu(1-\rho^*)\Delta m}}{\mu(1-\rho^*)+\theta m}\left(1 - e^{-\mu(1-\rho^*)\Delta}\, e^{-\theta \Delta m}\right) \exp\left(-\theta \frac{m(m-1)}{2}\Delta\right). \end{aligned}$$

□

It is important to note that this expression is connected to partial theta functions, which are a class of special functions that arise in number theory and combinatorics. These partial theta functions are given by

$$\psi(a, q) = \sum_{n=0}^{\infty} a^n q^{n(n-1)/2}, \quad |q| < 1$$

and are the partial sums of the standard Jacobi theta functions, which sum over the positive and negative integers [2, 29]. Recently, these partial theta functions have appeared in queueing models with exponential/geometric balking; see for example Pender [26]. Thus this correspondence provides a novel connection between stochastic service systems and classical objects of analytic number theory. Moreover, the quadratic decay of the summand ensures rapid convergence of the infinite series and makes the numerical evaluation via truncation both efficient and highly accurate.

9. Conclusion

There are several extensions worth pursuing here. Clearly, from Section 7, it is interesting to consider the $M_\lambda/G/1$ queue as an extension. However, as we have illustrated, this is very challenging since understanding the connection between the residual service time and the queue length is critical to understanding the overlapping time. Second, it would be interesting to extend our work to the $G/M_\mu/1$ queue beyond what we have shown in Section 8. Third, we would like to extend our analysis to the setting where the queueing process has a finite capacity, such as the $M_\lambda/M_\mu/1/k$ queue. Analyzing the overlapping time in the $M_\lambda/M_\mu/1/k$ queue would allow us to understand the impact of a finite capacity on the overlapping time. By understanding the relationship between queue capacity and overlapping time, public health officials can design queueing systems to achieve a targeted level of overlapping time.

Another extension would be to allow multiple servers, i.e. the $M_\lambda/M_\mu/c$ queue. The multiple server setting is interesting and difficult for two reasons. The first is that the tagged customer might leave before others who are present upon arrival. The second is that the customers after arrival of the tagged customer might leave before the tagged customer. This makes it quite difficult to analyze, except for the infinite server setting (see e.g. [8]), where customers do not interfere with each other.

Finally, it would be interesting to understand how a Hawkes arrival process would impact the distribution of the overlapping time or even overlap times in general; see for example Daw and Pender [12] and Koops et al. [18]. In Chen and Hong [9], there is a known difference between Hawkes arrival processes and renewal processes in their heavy-traffic behavior and

https://doi.org/10.1017/jpr.2026.10107 Published online by Cambridge University Press

beyond. Understanding this difference would be critical to understanding more complicated infectious disease behaviors and dynamics. This could also be expanded to bivariate Hawkes processes as in Daw et al. [13], where infected and susceptible customers follow a bivariate Hawkes process. We hope to consider these extensions in some follow-up work.

Acknowledgements

The authors gratefully acknowledge several stimulating discussions with Uri Yechiali. Google Gemini was used for editing the introduction.

Funding information

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 # 2510768.

Competing interests

There were no competing interests to declare which arose during the preparation or publication process of this article.

References

[1] Abramowitz, M. and Stegun, I. A. (1968). *Handbook of Mathematical Functions, with Formulas, Graphs, and Mathematical Tables* (Applied Mathematics Series **55**). US Government printing office.
[2] Andrews, G. E. and Warnaar, S. O. (2007). The product of partial theta functions. *Adv. Appl. Math.* **39**, 116–120.
[3] Borovkov, A. A., Boxma, O. J. and Palmowski, Z. (2003). On the integral of the workload process of the single server queue. *J. Appl. Prob.* **40**, 200–225.
[4] Boxma, O. (1984). Joint distribution of sojourn time and queue length in the M/G/1 queue with (in)finite capacity. *Europ. J. Operat. Res.* **16**, 246–256.
[5] Boxma, O. (2024). The number of overlapping customers. *Operat. Res. Lett.* **57**, 107203.
[6] Boxma, O. and Cohen, J. W. (1999). Heavy-traffic analysis for the GI/G/1 queue with heavy-tailed distributions. *Queueing Systems* **33**, 177–204.
[7] Boxma, O. and Pender, J. (2025). Overlap times in the G/G/1 queue via Laplace transforms. *Queueing Systems* **109**, 10.
[8] Boxma, O., van Kreveld, L. and Pender, J. The overlapping time distribution in the M/G/∞ queue. *Prob. Eng. Inf. Sci.*, to appear.
[9] Chen, X. and Hong, G. (2023). Steady-state analysis and online learning for queues with Hawkes arrivals. Available at https://arxiv.org/abs/2311.02577.
[10] Cohen, J. W. (1978). Properties of the process of level crossings during a busy cycle of the M/G/1 queueing system. *Math. Operat. Res.* **3**, 133–144.
[11] Cohen, J. W. (1982). *The Single Server Queue*. North-Holland, Amsterdam.
[12] Daw, A. and Pender, J. (2018). Queues driven by Hawkes processes. *Stoch. Systems* **8**, 192–229.
[13] Daw, A., Castellanos, A., Yom-Tov, G. B., Pender, J. and Gruendlinger, L. (2025). The co-production of service: Modeling services in contact centers using Hawkes processes. *Manag. Sci.* **71**, 2635–2656.
[14] Hall, P. (1927). The distribution of means for samples of size n drawn from a population in which the variate takes values between 0 and 1, all such values being equally probable. *Biometrika* **19**, 240–245.
[15] Iglehart, D. L. (1971). Functional limit theorems for the queue GI/G/1 in light traffic. *Adv. Appl. Prob.* **3**, 269–281.
[16] Irwin, J. O. (1927). On the frequency distribution of the means of samples from a population having any law of frequency with finite moments, with special reference to Pearson's Type II. *Biometrika* **19**, 225–239.
[17] Kang, K., Doroudi, S., Delasay, M. and Wickeham, A. (2023). A queueing-theoretic framework for evaluating transmission risks in service facilities during a pandemic. *Production Operat. Manag.* **32**, 1453–1470.
[18] Koops, D. T., Saxena, M., Boxma, O. J. and Mandjes, M. (2018). Infinite-server queues with Hawkes input. *J. Appl. Prob.* **55**, 920–943.

https://doi.org/10.1017/jpr.2026.10107 Published online by Cambridge University Press

[19] NEWMAN, L. AND PENDER, J. (2025). Contagious queues: Analyzing the spread of infection in the M/G/∞ queue. Working paper.
[20] NEWMAN, L. AND PENDER, J. (2025). When lines become contagious: Analyzing the spread of infection in the M/M/1 queue. Working paper.
[21] OZ, B. AND PERLMAN, Y. (2024). Co-exposure management in queueing systems. *SSRN*. Available at http://dx.doi.org/10.2139/ssrn.4842246.
[22] PALOMO, S. AND PENDER, J. (2021). Measuring the overlap with other customers in the single server queue. In *2021 Winter Simulation Conference (WSC '21)*, art. 204. IEEE Press.
[23] PALOMO, S. AND PENDER, J. (2023). Overlap times in the $GI^B/GI/\infty$ queue. Available at https://arxiv.org/abs/2302.07410.
[24] PALOMO, S. AND PENDER, J. (2024). Overlap times in the infinite server queue. *Prob. Eng. Inf. Sci.* **38**, 21–27.
[25] PALOMO, S. AND PENDER, J. (2025). The maximum overlap time in the M/M/1 queue. *Statist. Prob. Lett.* **219**, 110322.
[26] PENDER, J. (2025). Ramanujan's queue. Working paper.
[27] PENDER, J., KO, Y. M. AND XU, J. (2025). The number of overlapping customers in Erlang-A queues: An asymptotic approach. *Prob. Eng. Inf. Sci.* **39**, 344–369.
[28] ROSS, S. M. (1995). *Stochastic Processes*. John Wiley.
[29] WARNAAR, S. O. (2003). Partial theta functions I: Beyond the lost notebook. *Proc. London Math. Soc*. **87**, 363–395.
[30] WISHART, D. M. G. (1961). An application of ergodic theorems in the theory of queues. In *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability*, Vol. 2, *Contributions to Probability Theory*, pp. 581–593. University of California Press.
[31] XU, J., KO, Y. M., KONG, M. AND PENDER, J. (2026). Queueing management for reducing the overlaps of customers in service systems. *OR Spectrum* **48**, 325–354.

https://doi.org/10.1017/jpr.2026.10107 Published online by Cambridge University Press