Asymptotics of Waiting Time Distributions in the Accumulating Priority Queue

Joris Walraevens[1,*], Thomas Van Giel[2], Stijn De Vuyst[3] and Sabine Wittevrongel[1]

[1]SMACS Research Group, Department of Telecommunications and Information Processing, Ghent University
[2]Ghent University
[3]Department of Industrial Systems Engineering and Product Design, Ghent University

Abstract

We analyze the asymptotics of waiting time distributions in the two-class accumulating priority queue with general service times. The accumulating priority queue was suggested by Kleinrock in the 60s - he coined it time-dependent priority - to diversify waiting time objectives of different classes in a paramaterized way. It also avoids the typical starvation problem of regular priority queues. All customers build up priority linearly while waiting in the queue but at a class-dependent rate. At a service opportunity epoch, the customer with highest priority present is served. Stanford and colleagues recently calculated the Laplace-Stieltjes Transform (LST) of the waiting time distributions of the different classes, but only invert these LSTs numerically. In this paper, we analytically calculate the asymptotics of the corresponding distributions from these LSTs. We show that different singularities of the LST can play a role in the asymptotics, depending on the magnitude of service differentiation between both classes.

Keywords: Accumulating priority, dominant singularity analysis
Mathematics Subject Classification: 60K25, 90B22

1 Introduction

A static priority scheduling is a simple and efficient way to reduce waiting times of part of the customers in a queue (the high-priority customers) at the expense of the other customers (low-priority customers). Therefore, priority queues with diverse arrival and service time characteristics have been analyzed abundantly in the past (see e.g. [6, 16, 17, 20]). Such a priority scheduling is, in particular, effective if the fraction of high-priority customers is low, since the waiting times of these customer can be reduced vastly without significantly increasing the waiting times of the other customers. However, in case of numerous high-priority customers (even occasionally), queues with a priority scheduling suffer from starvation of the low-priority customers [11, 19].

[*]Corresponding author

Several solutions for the starvation problem have been proposed and implemented. Weighted Fair Queueing and Weighted Round Robin are popular scheduling disciplines in telecommunication networks, where the next customer (packet) to be served is chosen in some fair way [3]. Queues with these scheduling disciplines are usually not easy to analyze and therefore queueing theory and performance analysis researchers have turned to the analysis of theoretical ideal disciplines, like Generalized Processor Sharing and Discriminatory Processor Sharing [2, 13, 21]. Other scheduling disciplines that are less drastic than static priority include priority with priority jumps [10, 11], place reservation systems [4] and threshold-based scheduling [7].

In this paper, we study one particular scheduling that alleviates the starvation problem, namely the accumulating priority queue. In this queue, all customers build up priority linearly with the time they are waiting in the queue, but at class-dependent rates. This means that high-priority customers build up priority faster than low-priority customers, but also that low-priority customers that are already waiting a long time have built up a considerable amount of priority. At a service opportunity epoch, the customer in the queue with the highest priority level is served next. This scheduling was suggested by Kleinrock [6] as the 'time-dependent priority queue'. Stanford et al. [14] later coined it the 'accumulating priority queue'. In computer operating systems, the process is known as 'aging' [12, 22]. This scheduling can be generalized to non-linearly increasing priority levels, but Li et al. [9] have proved that many systems with common non-linear priority levels can be reduced to a system with linear levels in the sense that they lead to the same exact waiting times (and sample paths of buffer occupancies).

We study the asymptotics of the waiting time distributions in the two-class accumulating priority queue with start priority levels equal to 0 for all arriving customers. Stanford et al. [14] succeeded in calculating the Laplace-Stieltjes Transform (LST) of these distributions and these will be the start for our analysis. Using singularity analysis of these LSTs (cf., [1]), we calculate asymptotics of the waiting time density functions $w_i(t)$, $i = 1, 2$ of the two classes, i.e., we find expressions that are asymptotically equal to $w_i(t)$ for $t \to \infty$. The rates of increase of the two priority levels only impact the waiting time distributions through their ratio γ_b, which is therefore an important parameter that captures the priority differentiation between both classes. Therefore, we thoroughly study the influence of this parameter on the asymptotics and prove several interesting properties of the behavior of these asymptotics.

The paper is structured as follows. We first make the model we study explicit, summarize results of interest of [14], and lay out the methodology for analyzing asymptotic behavior in section 2. We also transform the LSTs from [14] somewhat in this section. Before calculating and analyzing the asymptotic behavior of the waiting times of both classes in sections 5 and 6, we do the same for some auxiliary variables, the accreditation periods of both classes in sections 3 and 4. In particular, we study the impact of the (relative) priority level rates on these asymptotics and illustrate our findings through some numerical examples. We end this paper with some conclusions and future work, and further discussion on the relevance of our results.

2 Preliminaries

In this section, we first explain the model. We then write down the expressions of the LSTs of the waiting time distributions as calculated in [14], and transform these expressions in a more explicit form. We also briefly look into two extreme cases. Finally, we elucidate our methodology for analysis of the asymptotic behavior of the variables of interest.

2.1 Model

We assume a continuous-time queue with two classes of customers, say, class 1 and class 2. Customers of class i arrive to the queueing system according to a Poisson process with rate λ_i, $i = 1, 2$. There is one server to serve the customers. Service times of class-i customers are distributed according to a general distribution with density function $b^{(i)}(t)$, mean $1/\mu_i$, and LST $\tilde{B}^{(i)}(s)$, $i = 1, 2$[1]. We further assume that the distributions of the service times are 'class-I' distributions[2], as coined in [1], i.e., the LST $\tilde{B}^{(i)}(s)$ has a dominant (right-most) singularity $-s_{B^{(i)}}$ in the left-half plane with $s_{B^{(i)}} > 0$ and the $\tilde{B}^{(i)}(s)$ going to ∞ for $s \to -s_{B^{(i)}}$.[3] The load of class-i is defined as $\rho_i = \lambda_i/\mu_i$, $i = 1, 2$.

The accumulating priority scheduling is defined as follows: a class-i customer accumulates priority at rate b_i starting at 0 upon arrival. So if a class-i customer arrived at time t' its accumulated priority at time $t \geq t'$ equals $b_i \cdot (t - t')$. Each time the server becomes available, the customer with the highest accumulated priority present starts service. Service is never interrupted. We assume $b_1 \geq b_2$, i.e., the priority level of class-2 customers does not grow faster than that of class-1 customers. The scheduling discipline is work conserving and as a result the stability condition is given by $\rho := \rho_1 + \rho_2 < 1$.

2.2 Reformulation of the Expressions for the LSTs from [14]

Define $\tilde{W}^{(i)}(s)$ as the LST of the stationary waiting time distribution. Stanford et al. [14] found the following results for these LSTs[4]:

$$\tilde{W}^{(i)}(s) = 1 - \rho + \rho \tilde{V}^{(i)}(s/b_i), \quad i = 1, 2, \tag{1}$$

with

$$\tilde{V}^{(1)}(s) = \frac{b_2}{b_1}\tilde{V}^{(2)}(s) + \frac{(1-\rho)(b_1-b_2)}{b_1(1-\sigma_1)}\tilde{V}^{(1,0)}(s) + \frac{(\rho-\sigma_1)(b_1-b_2)}{b_1(1-\sigma_1)}\tilde{V}^{(2)}(s)\tilde{V}^{(1,1)}(s),$$

$$\tilde{V}^{(2)}(s) = \frac{1}{-\tilde{\Gamma}_0^{(2)\prime}(0)} \frac{\left[1 + \tilde{\Gamma}_2^{(2)\prime}(0)\left(\lambda_2 + \lambda_1 \frac{b_2}{b_1}\right)\right]\left[1 - \tilde{\Gamma}_0^{(2)}(b_2 s)\right]}{b_2 s - \left(\lambda_2 + \lambda_1 \frac{b_2}{b_1}\right)\left[1 - \tilde{\Gamma}_2^{(2)}(b_2 s)\right]},$$

$$\tilde{V}^{(1,0)}(s) = \frac{1}{-\tilde{B}_0^{(2)\prime}(0)} \frac{\left[1 - \rho_1\left(1 - \frac{b_2}{b_1}\right)\right]\left[\tilde{\Gamma}_0^{(2)}(b_2 s) - \tilde{B}_0^{(2)}(b_1 s)\right]}{\left(1 - \frac{b_2}{b_1}\right)\left\{b_1 s - \lambda_1\left[1 - \tilde{B}^{(1)}(b_1 s)\right]\right\}},$$

[1] In this paper, we denote the LST corrsesponding to a density funtion $f(t)$ by $\tilde{F}(s)$.

[2] Not to be confused with class-1 customers.

[3] Other distribution types can be treated as well, but make the (writing down of the) analysis more cumbersome. We refer to [1] for more details.

[4] We follow notation of Stanford et al. [14] as closely as possible.

$$\tilde{V}^{(1,1)}(s) = \frac{1}{-\tilde{B}_2^{(2)\prime}(0)} \frac{\left[1-\rho_1\left(1-\frac{b_2}{b_1}\right)\right]\left[\tilde{\Gamma}_2^{(2)}(b_2 s)-\tilde{B}_2^{(2)}(b_1 s)\right]}{\left(1-\frac{b_2}{b_1}\right)\left\{b_1 s-\lambda_1\left[1-\tilde{B}^{(1)}(b_1 s)\right]\right\}},$$

$$\tilde{\Gamma}_i^{(2)}(s) = \tilde{B}_i^{(2)}\left(s+\lambda_1\left(1-\frac{b_2}{b_1}\right)\left[1-\tilde{\Gamma}^{(1)}(s)\right]\right), \quad i=0,2,$$

$$\tilde{\Gamma}^{(1)}(s) = \tilde{B}^{(1)}\left(s+\lambda_1\left(1-\frac{b_2}{b_1}\right)\left[1-\tilde{\Gamma}^{(1)}(s)\right]\right),$$

$$\tilde{B}_0^{(2)}(s) = \frac{\lambda_1\tilde{B}^{(1)}(s)+\lambda_2\tilde{B}^{(2)}(s)}{\lambda_1+\lambda_2},$$

$$\tilde{B}_2^{(2)}(s) = \frac{\lambda_1\frac{b_2}{b_1}\tilde{B}^{(1)}(s)+\lambda_2\tilde{B}^{(2)}(s)}{\lambda_1\frac{b_2}{b_1}+\lambda_2},$$

$$\sigma_1 = \frac{\rho_1(b_1-b_2)}{b_1}.$$

Explicit substitution of all these expressions in (1) shows that $\tilde{W}^{(i)}(s)$ only depends on b_1 and b_2 through their ratio $\gamma_b := b_2/b_1$. We find

$$\begin{aligned}\tilde{W}^{(1)}(s) = {} & \frac{1-\rho}{\left\{s-\lambda_1\left[1-\tilde{B}^{(1)}(s)\right]\right\}\left\{\gamma_b s-\lambda_1\gamma_b\left[1-\tilde{\Gamma}^{(1)}(\gamma_b s)\right]-\lambda_2\left[1-\tilde{\Gamma}^{(2)}(\gamma_b s)\right]\right\}} \\ & \cdot\left(\left\{s-\lambda_1\left[1-\tilde{\Gamma}^{(1)}(\gamma_b s)\right]\right\}\left\{\gamma_b s-\lambda_2(1-\gamma_b)\left[1-\tilde{B}^{(2)}(s)\right]\right\}\right. \\ & \left.-\lambda_2 s\left[\tilde{B}^{(2)}(s)-\tilde{\Gamma}^{(2)}(\gamma_b s)\right]\right), \end{aligned} \tag{2}$$

$$\tilde{W}^{(2)}(s) = (1-\rho)\frac{s+\lambda_1(1-\gamma_b)\left[1-\tilde{\Gamma}^{(1)}(s)\right]}{s-\lambda_1\gamma_b\left[1-\tilde{\Gamma}^{(1)}(s)\right]-\lambda_2\left[1-\tilde{\Gamma}^{(2)}(s)\right]}, \tag{3}$$

with

$$\tilde{\Gamma}^{(i)}(s) = \tilde{B}^{(i)}\left(s+\lambda_1(1-\gamma_b)\left[1-\tilde{\Gamma}^{(1)}(s)\right]\right), \quad i=1,2. \tag{4}$$

The random variables corresponding with $\tilde{\Gamma}^{(i)}(s)$ are called 'accreditation periods' in [14]. Note that $\tilde{\Gamma}^{(1)}(s)$ is implicitly defined, cf. (4) for $i=1$, and can, for general $\tilde{B}^{(1)}(s)$, not be calculated explicitly.[5] All other functions depend on this $\tilde{\Gamma}^{(1)}(s)$. Derivatives in $s=0$ can, however, be calculated explicitly. For instance, we have

$$\tilde{\Gamma}^{(1)\prime}(0) = -\frac{1}{\mu_1}\frac{1}{1-\rho_1(1-\gamma_b)}.$$

Numerical inversion of the LSTs is also feasible, cf. [14].

[5]For exponential service times of class 1, an explicit expression is possible.

2.3 Special Cases $\gamma_b = 0$ and $\gamma_b = 1$

We end this preliminary section with a discussion on the two limit cases $\gamma_b = 0$ ($b_2 = 0$ or $b_1 = \infty$) and $\gamma_b = 1$ ($b_1 = b_2$), since these reduce to well-known scheduling disciplines for which the asymptotics of the densities of the waiting times are well understood.

When $\gamma_b = 0$, class-1 customers have non-preemptive priority over class-2 customers. The expressions of the LSTs of the waiting times simplify to

$$\tilde{W}^{(1)}(s) = \frac{(1-\rho)s + \lambda_2 \left[1 - \tilde{B}^{(2)}(s)\right]}{s - \lambda_1 \left[1 - \tilde{B}^{(1)}(s)\right]}, \tag{5}$$

$$\tilde{W}^{(2)}(s) = (1-\rho)\frac{s + \lambda_1 \left[1 - \tilde{\Gamma}^{(1)}(s)\right]}{s - \lambda_2 \left[1 - \tilde{\Gamma}^{(2)}(s)\right]},$$

with

$$\tilde{\Gamma}^{(i)}(s) = \tilde{B}^{(i)}\left(s + \lambda_1 \left[1 - \tilde{\Gamma}^{(1)}(s)\right]\right), \quad i = 1, 2.$$

These are indeed consistent with the results for the M/G/1 non-preemptive priority queue, cf. for instance [16]. In the priority queue, $\tilde{\Gamma}^{(i)}(s)$ is the LST of the busy period of class-1 (high-priority) customers started by a class-i customer, $i = 1, 2$ (in case of class-2 this is to be regarded as an exceptional first service time). Comparison of the expression of this LST with (4) shows that the accreditation period in the accumulating priority queue is nothing more than a busy period in an equivalent system with (reduced) arrival rate $\lambda_1(1-\gamma_b)$ for the class-1 customers. This shows that the priority of class-1 customers, which is absolute for $\gamma_b = 0$, is more alleviated for larger $\gamma_b > 0$. Since asymptotics in the priority queue are well understood (cf. [1]) and this LST of the busy period plays a major role in them (see also further), this already demonstrates that asymptotics in the accumulating priority queue relate to that in the priority queue to some extent.

When $\gamma_b = 1$, all customers gain priority at the same rate and the system results in a FCFS system. The expressions of the LSTs of the waiting times simplify to

$$\tilde{W}^{(i)}(s) = \frac{(1-\rho)s}{s - \lambda_1 \left[1 - \tilde{B}^{(1)}(s)\right] - \lambda_2 \left[1 - \tilde{B}^{(2)}(s)\right]}, \quad i = 1, 2.$$

These are indeed consistent with the results for the M/G/1 queue where the service times are a probabilistic mixture of class-1 and class-2 service times.

2.4 Methodology

In this paper, we develop the asymptotics of the waiting time distributions in the accumulating priority queue, i.e., we establish approximate expressions for the density function of the waiting times, $i = 1, 2$, which are asymptotically exact for $t \to \infty$. This derivation is based on dominant singularity analysis of the corresponding LSTs. The asymptotic behavior of a distribution is determined by the *location* of the rightmost singularity of the LST (which is a non-positive real number) and the *type* of that singularity (pole with certain multiplicity, branchpoint, ...). More precisely if we find A and γ for an LST $\tilde{F}(s)$ with rightmost singularity $-s^*$ such that

$$\tilde{F}(s) \sim A \cdot (s+s^*)^{\gamma}, \quad s \to -s^*,$$

the asymptotics for the corresponding density function are given by

$$f(t) \sim \frac{Ae^{-s^*t}}{t^{\gamma+1}\Gamma(-\gamma)}, \quad t \to \infty$$

with $\Gamma(x)$ the Gamma function, according to the Heaviside Operational Principle ([1], p. 188).[6]

We focus, in particular, on the influence of the essential model parameter γ_b. This parameter summarizes the (inverse) priority level of class-1 customers ($\gamma_b = 0$ yields full priority while $\gamma_b = 1$ is the system without priority, cf. supra) and, in accordance with this, the rightmost singularities of the different LSTs demonstrate some monotonic behavior in γ_b. We will therefore add γ_b, where appropriate, as a variable to the different functions and constants in the remaining sections.

We start with the asymptotic inversion of the $\tilde{\Gamma}^{(i)}(s)$ functions which have a stochastic interpretation as the LSTs of busy periods in an M/G/1 queue and which are key to the asymptotics of the waiting times. Then we look at the class-2 waiting times first as the expression of the LST is simplest for this class, leading to easier derivations as well. Last, we investigate the class-1 waiting time distribution asymptotics. Throughout, we illustrate the properties by numerical examples.

3 Accreditation Periods Class 1

The function $\tilde{\Gamma}^{(1)}(s)$ as in (4) for $i = 1$ is the LST of an accreditation period started by a class-1 service time. The LST is in fact equal to the LST of a busy period in the M/G/1 queue with arrival rate $\lambda_1(1-\gamma_b)$ and service times with LST $\tilde{B}^{(1)}(s)$, cf. [16], p. 20, formula (2.4).

We first identify the rightmost singularity of $\tilde{\Gamma}^{(1)}(s)$ in the case $\gamma_b < 1$ and investigate how this singularity behaves as function of γ_b. Then we write down the asymptotics of the corresponding density. We end by proving some further properties for future reference and by briefly looking into the special case $\gamma_b = 1$.

3.1 Rightmost Singularity

For $\gamma_b < 1$, we can refer to Abate and Whitt [1][7]. The rightmost singularity $-s_{\gamma_1}$ of $\tilde{\Gamma}^{(1)}(s)$ is that s for which $\tilde{\Gamma}^{(1)\prime}(s)$ becomes infinite. This is in essence a consequence of the implicit function theorem, that says that an implicitly defined function ($\tilde{\Gamma}^{(1)}(s)$) is regular as long as its first derivative exists. We have

$$s_{\gamma_1} = \zeta_1 + \lambda_1(1-\gamma_b)(1-\tilde{B}^{(1)}(-\zeta_1)),$$

[6] $f(t) \sim g(t), t \to t_0 \Leftrightarrow \lim_{t\to t_0} f(t)/g(t) = 1$.

[7] Note that our formulas are not entirely identical as in Abate and Whitt [1], since they assumed $\mu_1 = 1$.

with $-\zeta_1$ the negative real root of the busy-period asymptotics equation

$$-\tilde{B}^{(1)\prime}(s) = \frac{1}{\lambda_1(1-\gamma_b)}, \tag{6}$$

cf., respectively, formulas (7.4) and (7.1) in Abate and Whitt [1]. Note that $\tilde{\Gamma}^{(1)}(-s_{\gamma_1}) = \tilde{B}^{(1)}(-\zeta_1)$.

3.2 Behavior of the Rightmost Singularity

The rightmost singularity $-s_{\gamma_1}(\gamma_b)$ of $\tilde{\Gamma}_1(s;\gamma_b)$ decreases monotonously with γ_b since

$$\begin{aligned} s'_{\gamma_1}(\gamma_b) &= \zeta'_1(\gamma_b)\left[1+\lambda_1(1-\gamma_b)\tilde{B}^{(1)\prime}(-\zeta_1(\gamma_b))\right] - \lambda_1\left[1-\tilde{B}^{(1)}(-\zeta_1(\gamma_b))\right] \\ &= -\lambda_1\left[1-\tilde{B}^{(1)}(-\zeta_1(\gamma_b))\right] > 0, \end{aligned}$$

where we used that $-\zeta_1$ is a solution of (6) in the second equality. This monotonous decrease is also intuitively clear as increasing γ_b means less accredited class-1 customers (less priority for class-1 customers).

Note that $\lim_{\gamma_b\to 1} s'_{\gamma_1}(\gamma_b) = +\infty$, since for $\gamma_b \to 1^-$, $-s_{\gamma_1}$ and $-\zeta_1$ converge to the rightmost singularity $-s_{B^{(1)}}$ of $\tilde{B}^{(1)}(s)$. The latter is consistent with the fact that the busy period in a system with zero arrival rate equals a single service time.

3.3 Asymptotics

We can write [1]

$$\tilde{\Gamma}^{(1)}(s) \sim \tilde{B}^{(1)}(-\zeta_1) - \sqrt{\frac{2}{\lambda_1^3(1-\gamma_b)^3\tilde{B}^{(1)\prime\prime}(-\zeta_1)}}(s+s_{\gamma_1})^{1/2}, \quad s \to -s_{\gamma_1}.$$

The Heaviside Operational Principle then leads to the following asymptotics for the density function $\gamma^{(1)}(t)$ of the accreditation periods of class 1:

$$\gamma^{(1)}(t) \sim \left[2\lambda_1^3(1-\gamma_b)^3\tilde{B}^{(1)\prime\prime}(-\zeta_1)\right]^{-1/2}\left(\pi t^3\right)^{-1/2} e^{-s_{\gamma_1}t},$$

cf. [1], formula (7.3).

3.4 Some Further Properties

Just like $s_{\gamma_1}(\gamma_b)$, $\zeta_1(\gamma_b)$ is an increasing function. From the busy periods asymptotics equation and the implicit function theorem we find

$$\zeta'_1(\gamma_b) = \frac{1}{B^{(1)\prime\prime}(-\zeta_1(\gamma_b))\lambda_1(1-\gamma_b)^2} > 0,$$

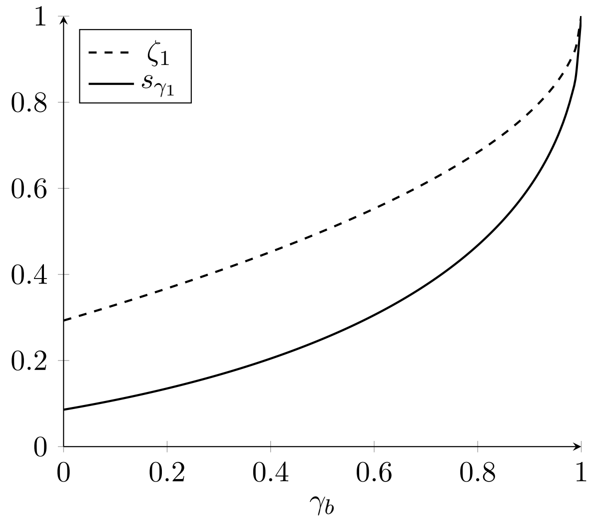


Figure 1: ζ_1 and s_{γ_1} as functions of γ_b for $\lambda_1 = 0.5$ and exponential service times of class-1 with rate $\mu_1 = 1$.

with the inequality following from the fact that all factors in the denominator of the RHS are positive.

From the relation between $s_{\gamma_1}(\gamma_b)$ and $\zeta_1(\gamma_b)$, it follows that $\zeta_1(\gamma_b) \geq s_{\gamma_1}(\gamma_b)$ with the equality only valid for $\gamma_b = 1$.

In Figure 1, we demonstrate the monotonic behavior of $s_{\gamma_1}(\gamma_b)$ and $\zeta_1(\gamma_b)$, that $s_{\gamma_1}(\gamma_b) \geq \zeta_1(\gamma_b)$, and that $\lim_{\gamma_b \to 1} s'_{\gamma_1}(\gamma_b) = +\infty$ for $\lambda_1 = 0.5$ and exponential service times of class-1 with rate $\mu_1 = 1$.

For future reference, we prove that $\tilde{\Gamma}^{(1)}(s; \gamma_b)$ decreases with γ_b for $-s_{\gamma_1} < s < 0$. First, we take the partial derivative of $\tilde{\Gamma}^{(1)}(s; \gamma_b)$ in γ_b:

$$\frac{\partial \tilde{\Gamma}^{(1)}(s; \gamma_b)}{\partial \gamma_b} = \frac{-\lambda_1 (1 - \tilde{\Gamma}^{(1)}(s; \gamma_b)) \tilde{B}^{(1)\prime}(s + \lambda_1 (1 - \gamma_b)(1 - \tilde{\Gamma}^{(1)}(s; \gamma_b)))}{1 + \lambda_1 (1 - \gamma_b) \tilde{B}^{(1)\prime}(s + \lambda_1 (1 - \gamma_b)(1 - \tilde{\Gamma}^{(1)}(s; \gamma_b)))} < 0.$$

The inequality follows from the fact that all three factors in the numerator are negative, while the denominator is positive for $-s_{\gamma_1} < s < 0$. Therefore, $\tilde{\Gamma}^{(1)}(s; \gamma_b)$ decreases with γ_b when $-s_{\gamma_1} < s < 0$.

3.5 The Special Case $\gamma_b = 1$

For $\gamma_b = 1$, the asymptotics of the probability density function corresponding to $\tilde{\Gamma}_1(s)$ are different, since they are in this case given by the asymptotics of the distribution of the class-1 service times.

The other formulas for $\gamma_b < 1$ are also correct for $\gamma_b = 1$.

4 Accreditation Periods Class 2

The function $\tilde{\Gamma}^{(2)}(s)$ as in (4) for $i = 2$ is the LST of an accreditation period started by a class-2 service time. The LST is in fact equal to the LST of a busy period in the M/G/1 queue with arrival rate $\lambda_1(1-\gamma_b)$, service times with LST $\tilde{B}^{(1)}(s)$ and exceptional first service time with LST $\tilde{B}^{(2)}(s)$, cf. [16], p. 24, formula (2.19b).

We first identify the *potential* rightmost singularities of $\tilde{\Gamma}^{(2)}(s)$ and study their behavior as a function of γ_b. We then investigate how γ_b affects which singularity is dominant. Afterwards, we write down the asymptotics of the density function and prove a further property for future reference. Since the case $\gamma_b = 1$ is again somewhat specific, we assume $\gamma_b < 1$ throughout the most part of this section and treat $\gamma_b = 1$ at the end.

4.1 Potential Rightmost Singularities

If $\gamma_b < 1$, the rightmost singularity of $\tilde{\Gamma}^{(2)}(s)$ is either the rightmost singularity $-s_{\gamma_1}$ of $\tilde{\Gamma}_1(s;\gamma_b)$ or the singularity $-\eta_2$ that makes the argument of the function $\tilde{B}^{(2)}$ in (4) equal to the rightmost singularity $-s_{B^{(2)}}$ of $\tilde{B}^{(2)}(s)$. For the latter, we have

$$\eta_2 = s_{B^{(2)}} + \lambda_1(1-\gamma_b)(1-\tilde{B}^{(1)}(-s_{B^{(2)}})).$$

Note that we used that $\tilde{\Gamma}^{(1)}(-\eta_2) = \tilde{B}^{(1)}(-s_{B^{(2)}})$.

4.2 Behavior of the Singularities

We already proved that $-s_{\gamma_1}(\gamma_b)$ decreases with γ_b. We now prove that $-\eta_2(\gamma_b)$ does as well. We have

$$\eta_2'(\gamma_b) = -\lambda_1(1-\tilde{B}^{(1)}(-s_{B^{(2)}})) > 0.$$

Note that $\lim_{\gamma_b \to 1} \eta_2(\gamma_b) = s_{B^{(2)}}$.

Since both singularities decrease with γ_b, the rightmost singularity $-s_{\gamma_2}(\gamma_b)$ of $\tilde{\Gamma}^{(2)}(s;\gamma_b)$ decreases with γ_b as well. It can switch from one singularity to the other (from $-s_{\gamma_1}(\gamma_b)$ to $-\eta_2(\gamma_b)$ or vice versa) when both singularities become equal only.

4.3 Which is the Rightmost Singularity?

We now study which of both singularities $-s_{\gamma_1}(\gamma_b)$ and $-\eta_2(\gamma_b)$ is the rightmost singularity and how γ_b affects this.

When $-\eta_2(\gamma_b)$ exists, it is the rightmost singularity. However, it only exists if $s_{B^{(2)}} < \zeta_1(\gamma_b)$, i.e., if the argument $s + \lambda_1(1-\gamma_b)(1-\tilde{\Gamma}^{(1)}(s;\gamma_b))$ in (4) reaches $-s_{B^{(2)}}$ before it reaches $-\zeta_1(\gamma_b)$ if we decrease s starting from 0. As $\zeta_1(\gamma_b)$ is an increasing function we have following division:

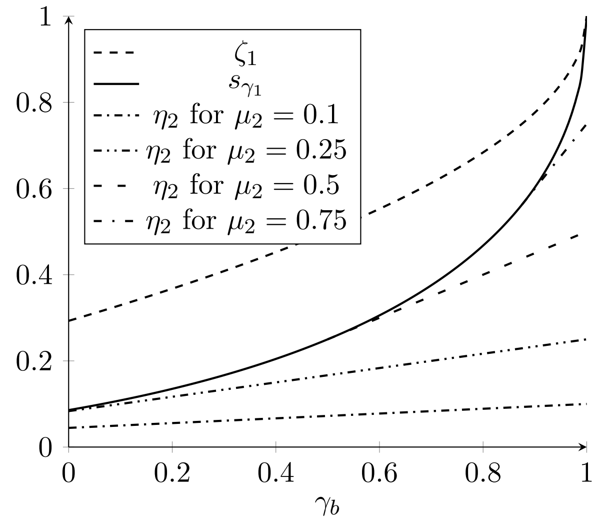


Figure 2: ζ_1, s_{γ_1}, and η_2 as functions of γ_b for $\lambda_1 = 0.5$, exponential service times with rate $\mu_1 = 1$ for class 1 and different rates μ_2 for class 2. The singularity η_2 exists for all γ_b when $\mu_2 = 0.1$ and $\mu_2 = 0.25$, since $\zeta_1(0) > 0.25$ (first case) while it only exists for larger values of γ_b for $\mu_2 = 0.5$ and $\mu_2 = 0.75$ (third case). For $\mu_2 > \zeta_1(1) = 1$, η_2 does not exist for any γ_b (second case).

- If $s_{B^{(2)}} \leq \zeta_1(0)$, $-\eta_2(\gamma_b)$ exists for all γ_b. By definition, $\zeta_1(0)$ is the rightmost solution of $-\tilde{B}^{(1)\prime}(s) = 1/\lambda_1$, cf. (6).
- On the other side of the spectrum, $\zeta_1(1) = s_{B^{(1)}}$. Therefore, if $s_{B^{(2)}} > s_{B^{(1)}}$, $-\eta_2(\gamma_b)$ does not exist for any $\gamma_b < 1$ and $-s_{\gamma_1}(\gamma_b)$ is the rightmost singularity of $\tilde{\Gamma}^{(2)}(s;\gamma_b)$ for all $\gamma_b < 1$.
- If $\zeta_1(0) < s_{B^{(2)}} \leq \zeta_1(1)$, there is a $\gamma_b^* \in [0,1]$ such that $-\eta_2(\gamma_b)$ exists for $\gamma_b \geq \gamma_b^*$ and does not exist for $\gamma_b < \gamma_b^*$. In the former case, $-\eta_2(\gamma_b)$ is the rightmost-singularity, in the latter $-s_{\gamma_1}(\gamma_b)$ is. The threshold γ_b^* is the solution of $\zeta_1(\gamma_b) = s_{B^{(2)}}$ and is given by

  $$\gamma_b^* = 1 + \frac{1}{\lambda_1 \tilde{B}^{(1)\prime}(-s_{B^{(2)}})}.$$

  This threshold increases monotonously with $s_{B^{(2)}}$ from 0 for $s_{B^{(2)}} = \zeta_1(0)$ to 1 for $s_{B^{(2)}} = \zeta_1(1)$.

In Figure 2, we demonstrate the monotonic behavior of $s_{\gamma_1}(\gamma_b)$ and $\eta_1(\gamma_b)$ for $\lambda_1 = 0.5$, exponential service times of class 1 with rate $\mu_1 = 1$, and exponential service times of class 2 with rates $\mu_2 = 0.1$, $\mu_2 = 0.25$, $\mu_2 = 0.5$ and $\mu_2 = 0.75$.

4.4 Asymptotics

We establish the asymptotic behavior of $\tilde{\Gamma}^{(2)}(s)$ in the neighborhood of $-s_{\gamma_1}$ (when it is different from $-\eta_2$):

$$\tilde{\Gamma}^{(2)}(s) \sim \tilde{B}^{(2)}(-\zeta_1) + \tilde{B}^{(2)\prime}(-\zeta_1)\sqrt{\frac{2}{\lambda_1(1-\gamma_b)\tilde{B}^{(1)\prime\prime}(-\zeta_1)}}(s+s_{\gamma_1})^{1/2}, \quad s \to -s_{\gamma_1}.$$

This leads to the following asymptotics for the density function $\gamma^{(2)}(t)$ of the accreditation periods of class 2:

$$\gamma^{(2)}(t) \sim -\tilde{B}^{(2)\prime}(-\zeta_1)\left[2\lambda_1(1-\gamma_b)\tilde{B}^{(1)\prime\prime}(-\zeta_1)\right]^{-1/2}\left(\pi t^3\right)^{-1/2} e^{-s_{\gamma_1} t}.$$

If $-\eta_2$ is the (joint) rightmost singularity, the behavior of $\tilde{\Gamma}^{(2)}(s)$ in the neighborhood of that singularity depends directly on the behavior of $\tilde{B}^{(2)}(s)$ in the neighborhood of $-s_{B^{(2)}}$. As an example, we show the asymptotics in case $-s_{B^{(2)}}$ is a pole of $\tilde{B}^{(2)}(s)$ with multiplicity $n_{B^{(2)}}$, i.e., when

$$\tilde{B}^{(2)}(s) \sim \frac{c_{B^{(2)}}}{(s+s_{B^{(2)}})^{n_{B^{(2)}}}}, \quad s \to -s_{B^{(2)}}, \tag{7}$$

with $c_{B^{(2)}}$ some constant. By combining this expansion of $\tilde{B}^{(2)}(s)$ with the regular expansion of $\tilde{\Gamma}^{(1)}(s)$, cf. [5], p. 411 (the 'supercritical case' of functional composition), we can write

$$\tilde{\Gamma}^{(2)}(s) \sim \frac{c_{B^{(2)}}[1+\lambda_1(1-\gamma_b)\tilde{B}^{(1)\prime}(-s_{B^{(2)}})]^{n_{B^{(2)}}}}{(s+\eta_2)^{n_{B^{(2)}}}}, \quad s \to -\eta_2,$$

where we have also used that

$$\tilde{\Gamma}^{(1)\prime}(-\eta_2) = \frac{\tilde{B}^{(1)\prime}(-s_{B^{(2)}})}{1+\lambda_1(1-\gamma_b)\tilde{B}^{(1)\prime}(-s_{B^{(2)}})}.$$

4.5 Further Properties

For future reference, we prove that $\tilde{\Gamma}^{(2)}(s;\gamma_b)$ decreases with γ_b for $s<0$. We can write

$$\frac{\partial\tilde{\Gamma}^{(2)}(s;\gamma_b)}{\partial\gamma_b} = \frac{\tilde{B}^{(2)\prime}(s+\lambda_1(1-\gamma_b)(1-\tilde{\Gamma}^{(1)}(s;\gamma_b)))}{\tilde{B}^{(1)\prime}(s+\lambda_1(1-\gamma_b)(1-\tilde{\Gamma}^{(1)}(s;\gamma_b)))}\frac{\partial\tilde{\Gamma}^{(1)}(s;\gamma_b)}{\partial\gamma_b} < 0.$$

The inequality follows from the fact that all factors in the RHS are negative.

4.6 The Special Case $\gamma_b = 1$

When $\gamma_b = 1$, $\tilde{\Gamma}^{(1)}(s)$ disappears from the equation and $\tilde{\Gamma}^{(2)}(s;1) = \tilde{B}^{(2)}(s)$, with rightmost singularity $-s_{B^{(2)}}$. Since this is the limit of the singularity $-\eta_2(\gamma_b)$ for $\gamma_b \to 1^-$, we have continuous behavior in cases where $-\eta_2(\gamma_b)$ exists for $\gamma_b \to 1^-$. If $-\eta_2(\gamma_b)$ does not exist for $\gamma_b \to 1^-$, there is a discontinuity of the rightmost singularity and the asymptotic distribution for $\gamma_b = 1$.

5 Waiting Times Class 2

The LST $\tilde{W}^{(2)}(s)$ shows a remarkable resemblance with that of the low-priority waiting time in the priority queue. We can therefore follow a similar reasoning as Abate and Whitt [1].

We first identify the potential rightmost singularities of $\tilde{W}^{(2)}(s)$ and study the behavior of these singularities as functions of γ_b. This helps us in investigating the impact of γ_b on which singularity is dominant. We end with writing down the asymptotics.

5.1 Potential Rightmost Singularities

The rightmost singularity $-s_{w_2}$ of $\tilde{W}^{(2)}(s)$ is one of two possibilities: either the branchpoint $-s_{\gamma_1}$ of $\tilde{\Gamma}^{(1)}(s)$ (and potentially $\tilde{\Gamma}^{(2)}(s)$) or the rightmost zero $-\eta$ (different from 0) of the denominator $s-\lambda_1\gamma_b\left[1-\tilde{\Gamma}^{(1)}(s)\right]-\lambda_2\left[1-\tilde{\Gamma}^{(2)}(s)\right]$ of $\tilde{W}^{(2)}(s)$, cf. (3). Note that when $-\eta_2$ is the rightmost singularity of $\tilde{\Gamma}^{(2)}(s)$, $\eta<\eta_2$ since $\tilde{\Gamma}^{(2)}(s)\to\infty$ for $s\to-\eta_2$.

5.2 Behavior of the Singularities

We already proved that the branchpoint $-s_{\gamma_1}(\gamma_b)$ decreases with γ_b. We now prove that the pole $-\eta(\gamma_b)$ does as well (when it exists). From its definition and the implicit function theorem we find

$$\begin{aligned}\eta'(\gamma_b) &= \frac{\lambda_1(1-\tilde{\Gamma}^{(1)}(-\eta(\gamma_b);\gamma_b))-\lambda_1\gamma_b\dfrac{\partial\tilde{\Gamma}^{(1)}(-\eta(\gamma_b);\gamma_b)}{\partial\gamma_b}-\lambda_2\dfrac{\partial\tilde{\Gamma}^{(2)}(-\eta(\gamma_b);\gamma_b)}{\partial\gamma_b}}{-\left(1+\lambda_1\gamma_b\dfrac{\partial\tilde{\Gamma}^{(1)}(-\eta(\gamma_b);\gamma_b)}{\partial s}+\lambda_2\dfrac{\partial\tilde{\Gamma}^{(2)}(-\eta(\gamma_b);\gamma_b)}{\partial s}\right)}\\ &= -\lambda_1(1-\tilde{\Gamma}^{(1)}(-\eta(\gamma_b);\gamma_b))>0.\end{aligned}$$

The second equality follows from calculating and substituting the partial derivatives of $\tilde{\Gamma}^{(i)}(s;\gamma_b)$.

Since both singularities that are potentially the rightmost singularity $-s_{W^{(2)}}(\gamma_b)$ of $\tilde{W}^{(2)}(s;\gamma_b)$ decrease with γ_b, so will $-s_{W^{(2)}}(\gamma_b)$. Note that this is also intuitive as increasing γ_b means less priority for class 1, hence a more rapidly decaying class-2 waiting time distribution.

5.3 Which is the Rightmost Singularity?

When $-\eta(\gamma_b)$ exists, it is the rightmost singularity. Note that $-\eta(\gamma_b)$ exists for sure if $-\eta_2(\gamma_b)$ exists, since $\eta<\eta_2\le s_{\gamma_1}$ in that case.

Since both singularities decrease with γ_b it is (again) not instantly clear how γ_b influences which singularity is the rightmost one. We distinguish the same cases as for the class-2 accreditation periods:

- If $s_{B^{(2)}}\le\zeta_1(0)$, $-\eta_2$ exists for all γ_b. Therefore, $-\eta$ exists as well and is the rightmost singularity.
- On the other side of the spectrum, if $\zeta_1(1)=s_{B^{(1)}}<s_{B^{(2)}}$, $-\eta_2$ does not exist for any γ_b and the rightmost singularity of the $\tilde{\Gamma}^{(i)}(s)$ is $-s_{\gamma_1}$. We then find that if λ_2 is larger than some threshold, say λ_2^*, $-\eta$ exists and is larger than $-s_{\gamma_1}$. If $\lambda_2<\lambda_2^*$, $-\eta$ does not exist and $-s_{\gamma_1}$ is the rightmost singularity. For $\lambda_2=\lambda_2^*$, $-\eta=-s_{\gamma_1}$.

The threshold value λ_2^* is given by

$$\lambda_2^* = \frac{\zeta_1 - \lambda_1(\tilde{B}^{(1)}(-\zeta_1) - 1)}{\tilde{B}^{(2)}(-\zeta_1) - 1}.$$

This is found as follows: if $-\eta$ exists, the denominator $s - \lambda_1\gamma_b(1-\tilde{\Gamma}^{(1)}(s)) - \lambda_2(1-\tilde{\Gamma}^{(2)}(s))$ of $\tilde{W}^{(2)}(s)$ is negative for $s \in]-\eta, 0[$ and non-negative for $s \in]-s_{\gamma_1}, -\eta]$. Therefore, if plugging $-s_{\gamma_1}$ in this denominator results in a non-negative number, $-\eta$ exists, and vice versa. This condition leads to $\lambda_2 \geq \lambda_2^*$.

Note that λ_2^* is not necessarily positive. In the remainder, we prove the following statements: (i) the numerator of $\lambda_2^*(\gamma_b)$ has one zero $\hat{\gamma}_b \in [0,1]$, (ii) if $\gamma_b \geq \hat{\gamma}_b$, $\lambda_2^* \leq 0$; (iii) for $\gamma_b \in [0, \hat{\gamma}_b[$, $\lambda_2^*(\gamma_b)$ is a positive decreasing function. All this means that $-\eta$ is the rightmost singularity if $\gamma_b \geq \hat{\gamma}_b$ or, if $\gamma_b < \hat{\gamma}_b$ and $\lambda_2 \geq \lambda_2^*$.

We prove this statement by studying the behavior of $\lambda_2^*(\gamma_b)$ and its numerator for all γ_b. We start with the numerator

$$f(\gamma_b) = \zeta_1(\gamma_b) - \lambda_1(\tilde{B}^{(1)}(-\zeta_1(\gamma_b)) - 1).$$

Its first derivative is negative for $\gamma_b > 0$:

$$\begin{aligned} f'(\gamma_b) &= \zeta_1'(\gamma_b)[1 + \lambda_1 \tilde{B}^{(1)\prime}(-\zeta_1(\gamma_b))] \\ &= \zeta_1'(\gamma_b)\frac{-\gamma_b}{1-\gamma_b} < 0, \end{aligned}$$

where we used the definition of ζ_1 in the second step. From the fact that $\tilde{B}^{(1)}(s)$ and $-\tilde{B}^{(1)\prime}(s)$ are decreasing functions for $s < 0$, and from the definition of $-\zeta_1$, it follows that $f(0) > 0$ and $\lim_{\gamma_b \to 1} f(\gamma_b) = -\infty$. Therefore, $f(\gamma_b)$ has a unique zero $\hat{\gamma}_b$ in $[0,1]$ which proves the first part. Note that this zero is independent of $s_{B^{(2)}}$.

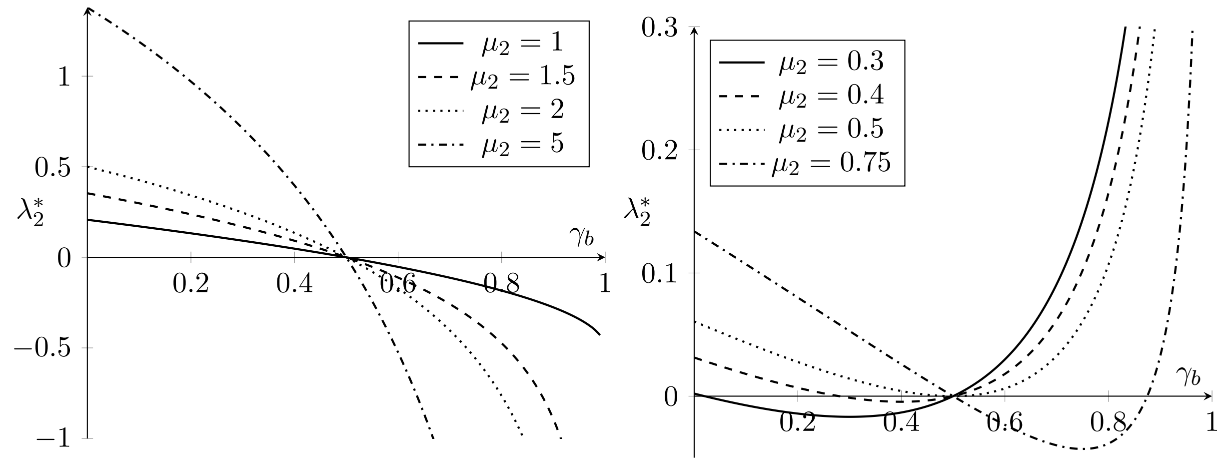
Next, we calculate the derivative of $\lambda_2^*(\gamma_b)$:

$$\lambda_2^{*\prime}(\gamma_b) = \zeta_1'(\gamma_b)\frac{-\dfrac{\gamma_b}{1-\gamma_b} + \lambda_2^*(\gamma_b)\tilde{B}^{(2)\prime}(-\zeta_1(\gamma_b))}{\tilde{B}^{(2)}(-\zeta_1(\gamma_b)) - 1}.$$

Since the first factor and the denominator of the second factor are positive and the numerator of the second factor is negative as long as $\lambda_2^*(\gamma_b)$ is positive, $\lambda_2^*(\gamma_b)$ decreases with γ_b at least until it hits 0. This zero is $\hat{\gamma}_b$ since this is also the only zero: the denominator stays finite ($\zeta_1(\gamma_b) < s_{B^{(2)}}$ for all γ_b in this case). This proves the third part. As a side result, $\lambda_2^*(\gamma_b) < 0$ for $\gamma_b \in]\tilde{\gamma}_b, 1[$, which is the second part of the statement.

We demonstrate these findings in Figure 3(a) where we show λ_2^* as function of γ_b for $\lambda_1 = 0.5$, exponential service times of class 1 with rate 1 and for several rates $\mu_2 \geq 1$ of class 2.

- If $\zeta_1(0) < s_{B^{(2)}} \leq \zeta_1(1) = s_{B^{(1)}}$, we distinguish two subcases:
  - If $\gamma_b \geq \gamma_b^*$, $-\eta_2$ exists. Therefore, $-\eta$ exists as well and is the rightmost singularity.
  - If $\gamma_b < \gamma_b^*$, $-\eta_2$ does not exist. As in the previous case, it then depends on λ_2 and λ_2^*, whether $-\eta$ exists. In this case, we find (see further) that (i) $\lambda_2^*(\gamma_b)$ is a



(a) $\mu_2 \geq 1$ (b) $\mu_2 < 1$

Figure 3: λ_2^* as function of γ_b for $\lambda_1 = 0.5$, exponential service times of class 1 with rate 1 and for several rates μ_2 of class 2.

positive decreasing function for $\gamma_b \in [0, \min(\hat{\gamma}_b, \gamma_b^*)[$, and (ii) if $\hat{\gamma}_b \leq \gamma_b^*$, $\lambda_2^*(\gamma_b) \leq 0$ for $\gamma_b \in [\hat{\gamma}_b, \gamma_b^*]$.

The proof follows the same lines as in the previous case. The difference is that $\lambda_2^*(\gamma_b)$ has two zeroes in the current case, namely the zero $\hat{\gamma}_b$ of the numerator and the pole γ_b^* of the denominator ($\zeta_1(0) < s_{B^{(2)}} < \zeta_1(1) = s_{B^{(1)}}$ in this case). As before, $\lambda_2^*(\gamma_b)$ decreases with γ_b at least until it hits 0 for the first time. This proves the first part of the statement. It then stays negative before hitting the second zero, which is only of interest when the first zero is $\hat{\gamma}_b$ (we are only interested in $\gamma_b < \gamma_b^*$ here). This proves the second statement.

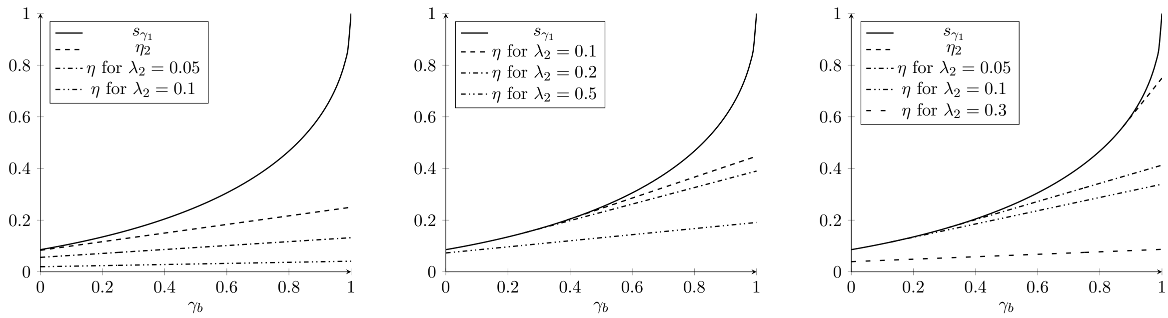
We again demonstrate these findings in Figure 3(b) where we show λ_2^* as function of γ_b for $\lambda_1 = 0.5$, exponential service times of class 1 with rate 1 and for several rates $\mu_2 < 1$ of class 2.

In summary, if the rightmost singularity $-s_{B^{(2)}}$ of $\tilde{B}^{(2)}(s)$ is smaller than a certain threshold, if subsequently γ_b is smaller than a threshold (which depends on $s_{B^{(2)}}$) *and* if λ_2 is smaller than a threshold (which depends on $s_{B^{(2)}}$ and γ_b), the rightmost singularity of $\tilde{W}^{(2)}(s)$ is the branchpoint $-s_{\gamma_1}$ of the LST of the accreditation periods of class 1. In all other cases, the rightmost singularity is the pole $-\eta$.

We demonstrate these conclusions in Figure 4 where we show s_{γ_1}, η_2 and η as functions of γ_b for $\lambda_1 = 0.5$, exponential service times with rate $\mu_1 = 1$ for class 1 and different arrival and service rates λ_2 and μ_2 for class 2.

5.4 Asymptotics

We can distinguish two different cases and one border case for the rightmost singularity of $\tilde{W}^{(2)}(s)$ and corresponding asymptotics of the class-2 waiting time density function $w_2(t)$:



(a) $\mu_2 = 0.25$ (b) $\mu_2 = 1.5$ (c) $\mu_2 = 0.75$

Figure 4: s_{γ_1}, η_2 and η as functions of γ_b for $\lambda_1 = 0.5$, exponential service times with rate $\mu_1 = 1$ for class 1 and different arrival and service rates λ_2 and μ_2 for class 2. The singularity η exists for all γ_b when $\mu_2 = 0.25$ (first case). For $\mu_2 = 0.75$ and $\mu_2 = 1.5$, it exists for all γ_b for λ_2 small enough. For higher λ_2 values, it only exists for large enough values of γ_b (second and third cases).

- $-\eta$ exists and is larger than $-s_{\gamma_1}$. We then have that $-s_{w_2} = -\eta$ and we can write

$$\tilde{W}^{(2)}(s) \sim \frac{(1-\rho)\dfrac{-\eta + \lambda_1(1-\gamma_b)(1-\tilde{\Gamma}^{(1)}(-\eta))}{1+\lambda_1\gamma_b\tilde{\Gamma}^{(1)\prime}(-\eta)+\lambda_2\tilde{\Gamma}^{(2)\prime}(-\eta)}}{s+\eta} \quad \text{as } s \to -\eta.$$

  Asymptotic inversion leads to

$$w^{(2)}(t) \sim (1-\rho)\frac{-\eta + \lambda_1(1-\gamma_b)(1-\tilde{\Gamma}^{(1)}(-\eta))}{1+\lambda_1\gamma_b\tilde{\Gamma}^{(1)\prime}(-\eta)+\lambda_2\tilde{\Gamma}^{(2)\prime}(-\eta)} e^{-\eta t},$$

  in this case.

- $-\eta$ does not exist. We have $-s_{w_2} = -s_{\gamma_1}$, and

$$\tilde{W}^{(2)}(s) \sim \tilde{W}_2(-s_{\gamma_1}) - (1-\rho)\frac{\dfrac{\zeta_1}{1-\gamma_b} + \sum_{i=1}^2 \lambda_i(1-\tilde{B}^{(i)}(-\zeta_1)) - \lambda_2\zeta_1\tilde{B}^{(2)\prime}(-\zeta_1)}{\left[\zeta_1 + \sum_{i=1}^2 \lambda_i(1-\tilde{B}^{(i)}(-\zeta_1))\right]^2} \cdot \sqrt{\frac{2}{\lambda_1(1-\gamma_b)\tilde{B}^{(1)\prime\prime}(-\zeta_1)}}(s+s_{\gamma_1})^{1/2}, \quad s \to -s_{\gamma_1}.$$

  Asymptotic inversion leads to

$$w^{(2)}(t) \sim (1-\rho)\frac{\dfrac{\zeta_1}{1-\gamma_b} + \sum_{i=1}^2 \lambda_i(1-\tilde{B}^{(i)}(-\zeta_1)) - \lambda_2\zeta_1\tilde{B}^{(2)\prime}(-\zeta_1)}{\left[\zeta_1 + \sum_{i=1}^2 \lambda_i(1-\tilde{B}^{(i)}(-\zeta_1))\right]^2}\sqrt{\frac{1}{2\lambda_1(1-\gamma_b)\tilde{B}^{(1)\prime\prime}(-\zeta_1)}} \cdot \left(\pi t^3\right)^{-1/2} e^{-s_{\gamma_1} t}.$$

- $-\eta$ exists and is equal to $-s_{\gamma_1}$. We have $-s_{w_2} = -\eta = -s_{\gamma_1}$, and

$$\tilde{W}^{(2)}(s) \sim (1-\rho) \frac{\zeta_1}{\left[\frac{\gamma_b}{1-\gamma_b} - \lambda_2 \tilde{B}^{(2)\prime}(-\zeta_1)\right] \sqrt{\frac{2}{\lambda_1(1-\gamma_b)\tilde{B}^{(1)\prime\prime}(-\zeta_1)}}(s+s_{\gamma_1})^{1/2}}, \quad s \to -s_{\gamma_1}.$$

Asymptotic inversion leads to

$$w^{(2)}(t) \sim (1-\rho) \frac{\zeta_1}{\left[\frac{\gamma_b}{1-\gamma_b} - \lambda_2 \tilde{B}^{(2)\prime}(-\zeta_1)\right] \sqrt{\frac{2}{\lambda_1(1-\gamma_b)\tilde{B}^{(1)\prime\prime}(-\zeta_1)}}}(\pi t)^{-1/2} e^{-s_{\gamma_1} t}.$$

6 Waiting Times Class 1

The expression (2) of $\tilde{W}^{(1)}(s)$ is more complex than that of $\tilde{W}^{(2)}(s)$ leading to more involved singularity and asymptotic analysis. We first identify the potential rightmost singularities and study their behavior as functions of γ_b. We then characterize which is the rightmost singularity and how this changes with γ_b. Finally, we write down the asymptotics of the density function $w^{(1)}(t)$ of the class-1 waiting time for all cases.

6.1 Potential Rightmost Singularities

The candidate rightmost singularities are:

- the branchpoint $-s_{\gamma_1}/\gamma_b$ of $\tilde{\Gamma}_1(\gamma_b s)$,
- the rightmost zero (< 0) $-\eta/\gamma_b$ of the denominator $\gamma_b s - \lambda_1 \gamma_b (1 - \tilde{\Gamma}^{(1)}(\gamma_b s)) - \lambda_2 (1 - \tilde{\Gamma}^{(2)}(\gamma_b s))$ of $\tilde{W}_2(\gamma_b s)$,
- the rightmost zero (< 0) $-\eta_1$ of the factor $s - \lambda_1(1 - \tilde{B}^{(1)}(s))$ in the denominator of $\tilde{W}^{(1)}(s)$,
- and/or the rightmost singularity $-s_{B^{(2)}}$ of $\tilde{B}^{(2)}(s)$.

6.2 Behavior of the Singularities

Two of the singularities do not depend on γ_b: $-\eta_1$ and $-s_{B^{(2)}}$. Note that $-\eta_1$ is also the dominant singularity of the LST of the waiting time in the regular (class-1) M/G/1 queue, cf. also (5). The other two candidate rightmost singularities $-s_{\gamma_1}(\gamma_b)/\gamma_b$ and $-\eta(\gamma_b)/\gamma_b$ do depend on γ_b.

We start research of the behavior of $-s_{\gamma_1}(\gamma_b)/\gamma_b$ by calculating its first derivative

$$\frac{d}{d\gamma_b}\left(\frac{s_{\gamma_1}(\gamma_b)}{\gamma_b}\right) = \frac{s'_{\gamma_1}(\gamma_b)\gamma_b - s_{\gamma_1}(\gamma_b)}{\gamma_b^2}$$

$$= -\frac{\zeta_1(\gamma_b) + \lambda_1(1 - \tilde{B}^{(1)}(-\zeta_1(\gamma_b)))}{\gamma_b^2}.$$

The numerator is equal to $f(\gamma_b)$ defined earlier and we proved it is a strictly decreasing function. We further have

$$\zeta_1(0) + \lambda_1(1 - \tilde{B}^{(1)}(-\zeta_1(0))) = s_{\gamma_1}(0) > 0,$$
$$\zeta_1(1) + \lambda_1(1 - \tilde{B}^{(1)}(-\zeta_1(1))) = -\infty,$$

which means that $s_{\gamma_1}(\gamma_b)/\gamma_b$ first decreases and then increases again. Note that the γ_b for which the minimum is reached is $\hat{\gamma}_b$, the previously defined unique zero of $f(\gamma_b)$. From the definitions of s_{γ_1} and η_1, it follows that

$$\zeta_1(\hat{\gamma}_b) = \eta_1 = \frac{s_{\gamma_1}(\hat{\gamma}_b)}{\hat{\gamma}_b}.$$

Note that this means that (i) $s_{\gamma_1}(\gamma_b)/\gamma_b \geq \eta_1$ for all γ_b with the equality valid for $\gamma_b = \hat{\gamma}_b$ only, and (ii) $\zeta_1(0) < \eta_1 < \zeta_1(1)$.

Next we research the behavior of $-\eta(\gamma_b)/\gamma_b$ by calculating the first derivative

$$\frac{d}{d\gamma_b}\left(\frac{\eta(\gamma_b)}{\gamma_b}\right) = \frac{\eta'(\gamma_b)\gamma_b - \eta(\gamma_b)}{\gamma_b^2}$$
$$= \frac{\lambda_2(1 - \tilde{\Gamma}^{(2)}(-\eta(\gamma_b);\gamma_b))}{\gamma_b^2} < 0.$$

Thus, the singularity $-\eta(\gamma_b)/\gamma_b$ increases monotonously for $\gamma_b \in [0, 1]$.

6.3 Which is the Rightmost Singularity?

We first have that $-s_{\gamma_1}(\gamma_b)/\gamma_b \leq -\eta_1$ with the equality for $\gamma_b = \hat{\gamma}_b$. Furthermore, we have that $-\eta(\hat{\gamma}_b)$ exists always and $\eta(\hat{\gamma}_b) < s_{\gamma_1}(\hat{\gamma}_b)$. This means that $-s_{\gamma_1}/\gamma_b$ can *never* be the rightmost singularity of $\tilde{W}^{(1)}(s)$ and $s_{W^{(1)}} < s_{\gamma_1}/\gamma_b$.

For $\gamma_b = 0$, $-\eta/\gamma_b$ does not play a role either (or, equivalently, is equal to $-\infty$). The rightmost singularity in that case is therefore either $-\eta_1$ or $-s_{B^{(2)}}$. The locations of these two singularities depend on different parameters (or input distributions), so either might be the rightmost.

For increasing γ_b, $-\eta/\gamma_b$ comes into play. We distinguish following cases:

- If $s_{B^{(2)}} \leq \zeta_1(0)$, $-\eta_2(\gamma_b)$ exists. Therefore, $-\eta(\gamma_b)$ exists as well. In this case it turns out that there is a $\tilde{\gamma}_b$ such that $-s_{B^{(2)}}$ is the rightmost singularity for $\gamma_b < \tilde{\gamma}_b$, $-s_{B^{(2)}} = -\eta(\tilde{\gamma}_b)/\tilde{\gamma}_b$ is the rightmost singularity for $\gamma_b = \tilde{\gamma}_b$, and $-\eta(\gamma_b)/\gamma_b$ is the rightmost singularity for $\gamma_b > \tilde{\gamma}_b$. We prove this claim in the remainder.

  First, $\zeta_1(0) < \eta_1$. This follows from the definitions of ζ_1 and η_1 and from the fact that $\tilde{B}^{(1)}(s)$ is a decreasing and $\tilde{B}^{(1)\prime}(s)$ an increasing function for $-s_{B^{(1)}} < s < 0$. As a result, $s_{B^{(2)}} < \eta_1$ and η_1 is never the rightmost singularity in this case.

Second, since $\eta(\gamma_b) < \eta_2(\gamma_b)$, since $\eta(\gamma_b)/\gamma_b$ decreases with γ_b and since $\eta_2(1) = s_{B^{(2)}}$, there is a unique point $\tilde{\gamma}_b \in]0,1[$ such that $\eta(\tilde{\gamma}_b)/\tilde{\gamma}_b = s_{B^{(2)}}$. For $\gamma_b \in [0, \tilde{\gamma}_b[$, $s_{B^{(2)}} < \eta(\gamma_b)/\gamma_b$; for $\gamma_b \in]\tilde{\gamma}_b, 1]$, $\eta(\gamma_b)/\gamma_b < s_{B^{(2)}}$.

- If $\zeta_1(0) < s_{B^{(2)}} < \eta_1$, we again prove that a $\tilde{\gamma}_b$ exists such that $-s_{B^{(2)}}$ is the rightmost singularity for $\gamma_b < \tilde{\gamma}_b$, $-s_{B^{(2)}} = -\eta(\tilde{\gamma}_b)/\tilde{\gamma}_b$ is the rightmost singularity for $\gamma_b = \tilde{\gamma}_b$, and $-\eta(\gamma_b)/\gamma_b$ is the rightmost singularity for $\gamma_b > \tilde{\gamma}_b$.

  We first have that $-\eta(1)$ exists since we proved before that $-\eta(\gamma_b)$ definitely exists for $\gamma_b \in [\hat{\gamma}_b, 1]$ for a $\hat{\gamma}_b < 1$. Furthermore, just like in the previous case, $\eta(1) < s_B^{(2)}$.

  If we decrease γ_b starting from $\gamma_b = 1$, $-\eta(\gamma_b)/\gamma_b$ decreases until either (i) $-\eta(\gamma_b)$ ceases to exist for a certain γ_b or (ii) until $\gamma_b = 0$. In the latter case, we can repeat the same reasoning as before. In the former case, $-\eta(\gamma_b)/\gamma_b$ hits $-s_{B^{(2)}}$ before $-\eta(\gamma_b)$ ceases to exist, since $-\eta(\gamma_b) = -s_{\gamma_1}(\gamma_b)$ for that γ_b where $-\eta(\gamma_b)$ ceases to exist and since $s_{\gamma_1}/\gamma_b \geq \eta_1 > s_{B^{(2)}}$.

- If $\eta_1 < s_{B^{(2)}}$, we prove that a $\bar{\gamma}_b$ exists such that $-\eta_1$ is the rightmost singularity for $\gamma_b < \hat{\gamma}_b$, $-\eta_1 = -\eta(\bar{\gamma}_b)/\bar{\gamma}_b$ is the rightmost singularity for $\gamma_b = \bar{\gamma}_b$, and $-\eta(\gamma_b)/\gamma_b$ is the rightmost singularity for $\gamma_b > \bar{\gamma}_b$.

  It is clear that, since in this case $\eta_1 < s_{B^{(2)}}$, $-s_{B^{(2)}}$ cannot be the rightmost singularity.

  We know $-\eta(1)$ exists. We also have that $\eta(1) < \eta_1$ and, as a result, $-\eta$ is the rightmost singularity for $\gamma_b = 1$. The inequality follows from the fact that if we plug $-\eta_1$ in the denominator with zero $-\eta(1)$, we have a positive result:

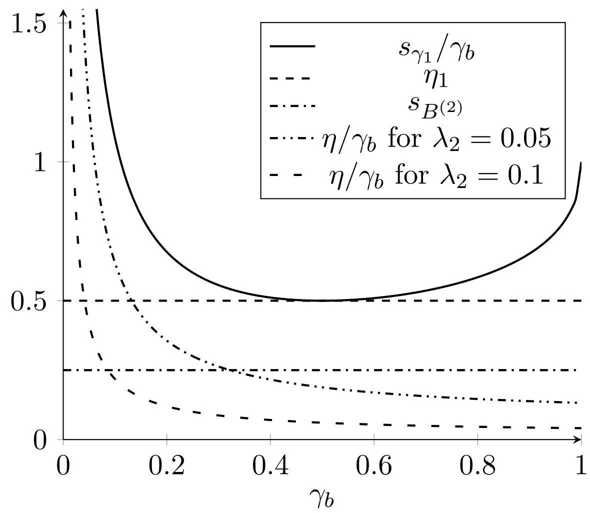
$$\begin{aligned}-\eta_1 - \lambda_1 \left[1 - \tilde{\Gamma}^{(1)}(-\eta_1)\right] & -\lambda_2 \left[1 - \tilde{\Gamma}^{(2)}(-\eta_1)\right] \\ &> -\eta_1 - \lambda_1 \left[1 - \tilde{B}^{(1)}(-\eta_1)\right] - \lambda_2 \left[1 - \tilde{\Gamma}^{(2)}(-\eta_1)\right] \\ &= -\lambda_2 \left[1 - \tilde{\Gamma}^{(2)}(-\eta_1)\right] \\ &> 0.\end{aligned}$$

  The first inequality follows from the definition of $\tilde{\Gamma}_1(s)$ and the equality from the definition of $-\eta_1$.

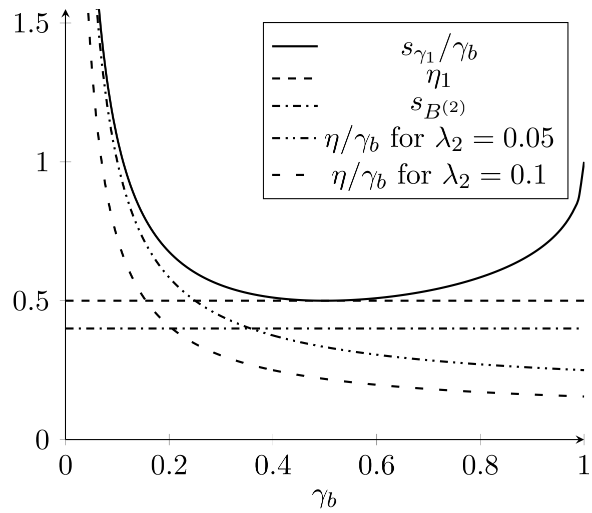
  When we decrease γ_b starting from 1, $-\eta(\gamma_b)/\gamma_b$ decreases, as before, until either (i) $-\eta(\gamma_b)$ ceases to exist for a certain γ_b or (ii) until $\gamma_b = 0$. In the latter case, we can repeat the same reasoning as before. In the former case, $-\eta(\gamma_b)/\gamma_b$ hits $-\eta_1$ before $-\eta(\gamma_b)$ ceases to exist, since $-\eta(\gamma_b) = -s_{\gamma_1}(\gamma_b)$ for that γ_b where $-\eta(\gamma_b)$ ceases to exist and since $s_{\gamma_b}/\gamma_b \geq \eta_1$ for all γ_b.

We conclude that the rightmost singularity of $\tilde{W}^{(1)}(s)$ for small γ_b is $-s_{B^{(2)}}$ or $-\eta_1$, whichever is largest. The values of these singularities do not change with γ_b. When γ_b increases, this singularity remains the rightmost one until a certain γ_b is reached. From this γ_b onwards, the rightmost singularity is $-\eta/\gamma_b$ which increases with γ_b.

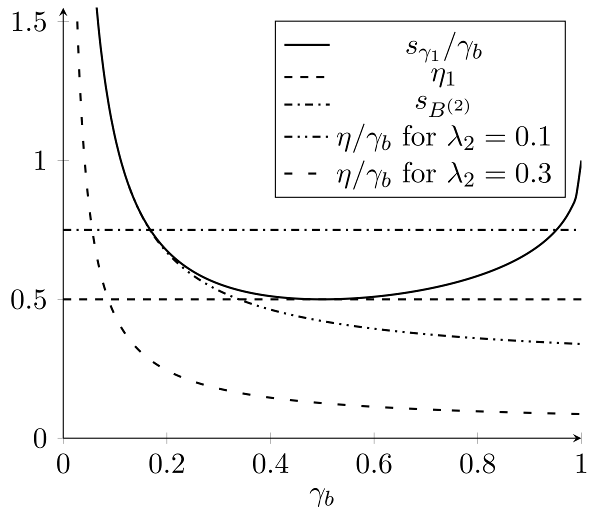
We demonstrate these conclusions in Figure 5 where we show s_{γ_1}/γ_b, η_1, $s_B^{(2)}$, and η/γ_b as functions of γ_b for $\lambda_1 = 0.5$, exponential service times with rate $\mu_1 = 1$ for class 1 and different arrival and service rates λ_2 and μ_2 for class 2.



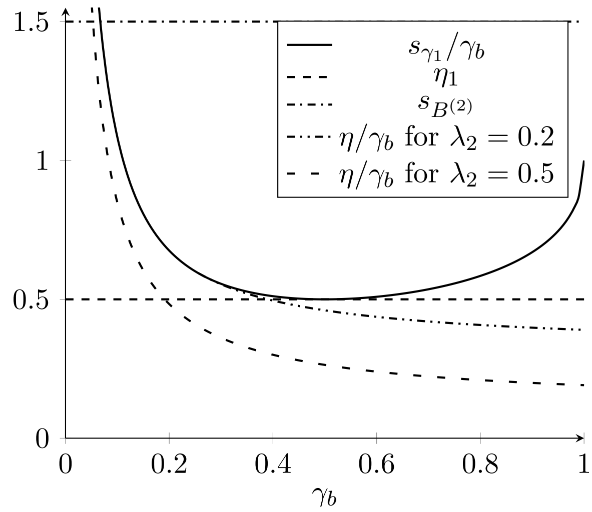
(a) $\mu_2 = 0.25$



(b) $\mu_2 = 0.4$



(c) $\mu_2 = 0.75$



(d) $\mu_2 = 1.5$

Figure 5: s_{γ_1}/γ_b, η_1, $s_{B^{(2)}}$, and η/γ_b as functions of γ_b for $\lambda_1 = 0.5$, exponential service times with rate $\mu_1 = 1$ for class 1 and different arrival and service rates λ_2 and μ_2 for class 2. For $\mu_2 = 0.25$ and $\mu_2 = 0.4$, $s_{B^{(2)}} < \eta_1$ and therefore $s_{B^{(2)}}$ is the rightmost singularity for small γ_b. For $\mu_2 = 0.75$ and $\mu_2 = 1.5$, η_1 is the rightmost singularity for small γ_b. In all cases, η/γ_b is the rightmost singularity for γ_b sufficiently large.

6.4 Asymptotics

We can distinguish the following cases:

- If $-s_{B^{(2)}}$ is the rightmost singularity, the asymptotics depend on the class-2 service times distribution asymptotics. If we again take the example of a dominant pole for $\tilde{B}^{(2)}(s)$, i.e., if we assume (7) for $\tilde{B}^{(2)}(s)$, we have

$$W^{(1)}(s) \sim \frac{c}{(s+s_{B^{(2)}})^{n_{B^{(2)}}}}, \quad s \to -s_{B^{(2)}},$$

with

$$c = \frac{(1-\rho)\lambda_2 c_{B^{(2)}} \left(\gamma_b s_{B^{(2)}} - (1-\gamma_b)\lambda_1 \left[1-\tilde{\Gamma}^{(1)}(-\gamma_b s_{B^{(2)}})\right]\right)}{\left\{s_{B^{(2)}} + \lambda_1\left[1-\tilde{B}^{(1)}(-s_{B^{(2)}})\right]\right\}\left\{\gamma_b s_{B^{(2)}} + \lambda_1\gamma_b\left[1-\tilde{\Gamma}^{(1)}(-\gamma_b s_{B^{(2)}})\right] + \lambda_2\left[1-\tilde{\Gamma}^{(2)}(-\gamma_b s_{B^{(2)}})\right]\right\}}.$$

Asymptotic inversion leads to

$$w^{(1)}(t) \sim \frac{c t^{n_{B^{(2)}}-1}}{(n_{B^{(2)}}-1)!} e^{-s_{B^{(2)}} t}.$$

- If $-\eta_1$ is the rightmost singularity, we can write

$$W^{(1)}(s) \sim \frac{c}{s+\eta_1}, \quad s \to -\eta_1,$$

with

$$\begin{aligned} c = &\frac{1-\rho}{\left[1+\lambda_1\tilde{B}^{(1)\prime}(-\eta_1)\right]\left\{-\gamma_b\eta_1 - \lambda_1\gamma_b\left[1-\tilde{\Gamma}^{(1)}(-\gamma_b\eta_1)\right] - \lambda_2\left[1-\tilde{\Gamma}^{(2)}(-\gamma_b\eta_1)\right]\right\}} \\ &\cdot \left(\left\{\eta_1 + \lambda_1\left[1-\tilde{\Gamma}^{(1)}(-\gamma_b\eta_1)\right]\right\}\left\{\gamma_b\eta_1 + \lambda_2(1-\gamma_b)\left[1-\tilde{B}^{(2)}(-\eta_1)\right]\right\}\right. \\ &\left. +\lambda_2\eta_1\left[\tilde{B}^{(2)}(-\eta_1) - \tilde{\Gamma}^{(2)}(-\gamma_b\eta_1)\right]\right). \end{aligned}$$

Asymptotic inversion leads to

$$w^{(1)}(t) \sim c e^{-\eta_1 t}.$$

- If $-\eta/\gamma_b$ is the rightmost singularity, we can write

$$W^{(1)}(s) \sim \frac{c}{s+\eta/\gamma_b}, \quad s \to -\eta/\gamma_b,$$

with

$$\begin{aligned} c = &\frac{(1-\rho)\lambda_2}{\gamma_b\left\{\eta + \gamma_b\lambda_1\left[1-\tilde{B}^{(1)}(-\eta_2)\right]\right\}\left[1+\lambda_1\gamma_b\tilde{\Gamma}^{(1)\prime}(-\eta) + \lambda_2\tilde{\Gamma}^{(2)\prime}(-\eta)\right]} \\ &\cdot \left(\left[1-\tilde{\Gamma}^{(2)}(-\eta)\right]\left\{\eta + \lambda_2(1-\gamma_b)\left[1-\tilde{B}^{(2)}(\eta/\gamma_b)\right]\right\} - \eta\left[\tilde{B}^{(2)}(-\eta/\gamma_b) - \tilde{\Gamma}^{(2)}(-\eta)\right]\right). \end{aligned}$$

Asymptotic inversion leads to

$$w^{(1)}(t) \sim c e^{-\eta t/\gamma_b}.$$

- The above singularities can also be (co)-dominant. If for instance $-\eta_1 = -\eta/\gamma_b$ is the rightmost singularity, we can write

$$W^{(1)}(s) \sim \frac{c}{(s+\eta_1)^2}, \quad s \to -\eta_1,$$

with

$$c = \frac{1-\rho}{\gamma_b \left[1+\lambda_1 \tilde{B}^{(1)\prime}(-\eta_1)\right] \left[1+\lambda_1\gamma_b \tilde{\Gamma}^{(1)\prime}(-\gamma_b\eta_1) + \lambda_2 \tilde{\Gamma}^{(2)\prime}(-\gamma_b\eta_1)\right]}$$
$$\cdot \left(\left\{\eta_1 + \lambda_1\left[1-\tilde{\Gamma}^{(1)}(-\gamma_b\eta_1)\right]\right\}\left\{\gamma_b\eta_1 + \lambda_2(1-\gamma_b)\left[1-\tilde{B}^{(2)}(-\eta_1)\right]\right\}\right.$$
$$\left. +\lambda_2\eta_1\left[\tilde{B}^{(2)}(-\eta_1) - \tilde{\Gamma}^{(2)}(-\gamma_b\eta_1)\right]\right).$$

Asymptotic inversion leads to

$$w^{(1)}(t) \sim c t e^{-\eta_1 t}.$$

7 Conclusions and Future Work

We have found that the class-2 waiting time distribution asymptotics in the accumulating priority queue resemble very much those of the low-priority waiting time distribution asymptotics of the low-priority class in the regular priority queue. For high class-2 service rates and low class-2 arrival rates, these asymptotics behave like those of the busy period distribution, and otherwise they are purely exponentially decaying. The higher γ_b, (i) the faster the class-2 waiting time distribution asymptotics decay, and (ii) the more the purely exponential asymptotics establish themselves (from a certain γ_b onwards, the other type never occurs).

The class-1 waiting time distribution asymptotics are more intriguing. It always has purely exponentially decaying asymptotics (except for some border case). In other words, the busy period distribution-type of asymptotics is not observed, although the LST of such busy periods (or accreditation periods) does appear in the expression of the LST of the class-1 waiting time distribution. The (mathematical) reason is that the singularity of that LST is never the rightmost singularity of the LST of the class-1 waiting time distribution. Furthermore, also the impact of γ_b is interesting. For $\gamma_b = 1$, we have a pure FIFO scheduling and the class-1 customers do not observe any form of priority. When γ_b decreases, the decay of the class-1 waiting time distribution increases at first, as expected (the class-1 customers accumulate relatively more priority when γ_b decreases). However, this increase stops from a certain γ_b onwards. In other words, if we decrease γ_b even more, the decay of the class-1 waiting time distribution asymptotics stays constant. One might be inclined to propose that this γ_b might be a good choice for the relative priority accumulation between both classes, since you obtain the minimum decay for the class-1 waiting time asymptotics while giving even more priority to class-1 customers would only punish the class-2 customers more.

As future work, we discuss two potential research directions. The first continues on our discussion in the previous paragraph about the optimal γ_b. Much research has been devoted on finding optimal scheduling disciplines (and parameters in these disciplines) to minimize weighted delay costs in multi-class systems. Our asymptotic results could be used to (approximately) minimize such functions [8], especially when high delays are heavily punished. In previous research, the generalized $c\mu$-rule has been proved to be heavy-traffic optimal for convex cost functions [18], which would be a reference point for such a study. This is also related to [15], where the authors used large deviations to find the asymptotic exponential decay rates of the waiting times and some asymptotic optimality of the accumulating priority queue.

The second direction of future research we wish to touch upon is the generalization to more than two customer classes. Next to the two-class system, Stanford et al. [14] also established an algorithm to obtain the LSTs of the waiting times in an accumulating priority queue with a general number of classes, recursively, starting from that of the lowest priority class (the one with lowest accumulating rate). The LST of the waiting time of the lowest class is very similar to that of class-2 in the two-class system, and thus we expect similar results as in this paper for the asymptotics of the density function of that waiting time. Among others, the similarity with the low-priority class in a static priority queue is still present. The other LSTs are not explicitly calculated in Stanford et al. [14], so there would be two options here to proceed: (i) calculate the LSTs explicitly and then invert them asymptotically like we have done in this paper, or (ii) recursively calculate the asymptotics from the recursive equations for the LSTs. Concerning the results, we expect the highest priority class (the one with the highest accumulating rate) to behave like the asymptotics of the class-1 waiting times in the two-class system, i.e., (usually) purely exponentially decaying. The intermediate classes are more interesting. Here, we expect that the asymptotics will not be purely exponentially decaying (like class 2 in the two-class system), since these classes still have lower priority than other classes, but we also expect that some candidate dominant singularities will never be dominant (like for class 1 in the two-class system).

Acknowledgment The authors wish to thank the two anonymous referees and associate editor for valuable comments that improved the paper.

References

[1] J. Abate and W. Whitt. Asymptotics for M/G/1 low-priority waiting-time tail probabilities. *Queueing Systems*, 25(1-4):173–233, 1997.

[2] E. Altman, K. Avrachenkov, and U. Ayesta. A survey on discriminatory processor sharing. *Queueing Systems*, 53:53–63, 2006.

[3] A. Banchs and X. Perez. Distributed weighted fair queuing in 802.11 wireless LAN. In *Proceedings of the 2002 IEEE International Conference on Communications (ICC)*, pages 3121–3127, New York, 2002.

[4] S. De Vuyst, S. Wittevrongel, and H. Bruneel. Place reservation: Delay analysis of a novel scheduling mechanism. *Computers & Operations Research*, 35(8):2447–2462, 2008.

[5] P. Flajolet and R. Sedgewick. *Analytic Combinatorics*. Cambridge University Press, 2009.

[6] L. Kleinrock. *Queueing systems volume II: Computer applications*. John Wiley & Sons, New York, 1976.

[7] C. Knessl, C. Tier, and D. Choi. A dynamic priority queue model for simultaneous service of two traffic types. *SIAM Journal on Applied Mathematics*, 63(2):398–422, 2003.

[8] N. Li, D. Stanford, A. Sharif, R. Caron, and A. Pardhan. Optimising key performance indicator adherence with application to emergency department congestion. *European Journal of Operational Research*, 272:313–323, 2019.

[9] N. Li, D. Stanford, P. Taylor, and I. Ziedins. Nonlinear accumulating priority queues with equivalent linear proxies. *Operations Research*, 65(6):1712–1721, 2017.

[10] Y. Lim and J. Kobza. Analysis of a delay-dependent priority discipline in an integrated multiclass traffic fast packet switch. *IEEE Transactions on Communications*, 38(5):659–685, 1990.

[11] T. Maertens, J. Walraevens, and H. Bruneel. On priority queues with priority jumps. *Performance Evaluation*, 63(12):1235–1252, 2006.

[12] M. Maheswaran, S. Ali, H. Siegel, D. Hensgen, and R. Freund. Dynamic mapping of a class of independent tasks onto heterogeneous computing systems. *Journal of Parallel and Distributed Computing*, 59:107–131, 1999.

[13] A. L. Parekh and R. G. Gallager. A generalized processor sharing approach to flow control in integrated services networks: the multiple node case. *IEEE/ACM Transactions on Networking*, 2(2):137–150, 1994.

[14] D. Stanford, P. Taylor, and I. Ziedins. Waiting time distributions in the accumulating priority queue. *Queueing Systems*, 77:297–330, 2014.

[15] A. Stolyar and K. Ramanan. Largest weighted delay first scheduling: large deviations and optimality. *The Annals of Applied Probability*, 11(1):1–48, 2001.

[16] H. Takagi. *Queueing analysis: a foundation of performance evaluation, volume 1: vacation and priority systems, part 1*. North-Holland, 1991.

[17] T. Takine. The nonpreemptive priority MAP/G/1 queue. *Operations Research*, 47(6):917–927, 1999.

[18] J. Van Mieghem. Dynamic scheduling with convex delay costs: the generalized $c\mu$ rule. *The Annals of Applied Probability*, 5(3):809–833, 1995.

[19] D. Vasiliadis, G. Rizos, and C. Vassilakis. Class-based weighted fair queuing scheduling on quad-priority Delta Networks. *International Journal of Parallel, Emergent and Distributed Systems*, 27(5):435–457, 2012.

[20] J. Walraevens, B. Steyaert, and H. Bruneel. Delay characteristics in discrete-time GI-G-1 queues with non-preemptive priority queueing discipline. *Performance Evaluation*, 50(1):53–75, 2002.

[21] J. Walraevens, J. S. H. van Leeuwaarden, and O. J. Boxma. Power series approximations for two-class generalized processor sharing systems. *Queueing Systems*, 66(2):107–130, 2010.

[22] Y. Wang, K. Huang, and F. Wang. Scheduling online mixed-parallel workflows of rigid tasks in heterogeneous multi-cluster environments. *Future Generation Computer Systems*, 60:35–47, 2016.