CAMBRIDGE
UNIVERSITY PRESS

RESEARCH ARTICLE

Completion times of jobs on two-state service processes and their asymptotic behavior

Melike Baykal-Gürsoy[1] (iD), Marcelo Figueroa-Candia[2] and Zhe Duan[3]

[1]Industrial and Systems Engineering Department, RUTCOR and CAIT, Rutgers University, Piscataway, 08855, NJ
[2]Industrial and Systems Engineering Department, Rutgers University, Piscataway, NJ
[3]Department of Management Science, School of Management, Xi'An Jiaotong University, China

Corresponding author: Melike Baykal-Gürsoy; Email: gursoy@soe.rutgers.edu

We consider the task completion time of a repairable server system in which a server experiences randomly occurring service interruptions during which the server works slowly. Every service-state change preempts the task that is being processed. The server may then *resume* the interrupted task, it may *replace* the task with a different one, or it may *restart* the same task from the beginning, under the new service-state. The total time that the server takes to complete a task of random size including interruptions is called *completion time*. We study the completion time of a task under the last two cases as a function of the task size distribution, the service interruption frequency/severity, and the repair frequency. We derive closed form expressions for the completion time distribution in Laplace domain under *replace* and *restart* recovery disciplines and present their asymptotic behavior. In general, the heavy tailed behavior of completion times arises due to the heavy tailedness of the task time. However, in the preempt-restart service discipline, even in the case that the server still serves during interruptions albeit at a slower rate, completion times may demonstrate power tail behavior for exponential tail task time distributions. Furthermore, we present an $M/G/\infty$ queue with exponential service time and Markovian service interruptions. Our results reveal that the stationary first order moments, that is, expected system time and expected number in the system are insensitive to the way the service modulation affects the servers; system-wide modulation affecting every server simultaneously vs identical modulation affecting each server independently.

1. Introduction

For applications in computer science and telecommunications, it is of interest to study the time it takes for a processor to complete a task of random size when the service can be interrupted or slowed down by the occurrence of random failures or higher priority task arrivals to processor sharing systems. This line of research on interrupted service times that leads to our study started in the sixties with the work of Gaver [18] that investigates the completion time distribution of a task experiencing complete service breakdowns during which times the server stops processing the task. Either the interruption can be postponed until the current task is completed or the current task is preempted whenever an interruption occurs and then the server decides on a *preemption scheme/discipline*. The preemption scheme determines the course of action for the server when a task has been interrupted. Gaver [18] is the first to consider postponable interruptions in addition to the following types of preemption strategies: (1) Preempt-*RESUME*, once the interruption is cleared the server continues with the unfinished task from where it left. (2) Preempt-*REPLACE* (repeat different), the server discards the unfinished task to revisit

later and selects a different task, assuming that such similar tasks are always available. (3) Preempt-*RESTART* (repeat identical), the server starts the unfinished task from the beginning. The author derives the Laplace–Stieltjes transform (Laplace–Stieltjes transform) of the completion time distribution under each interruption/recovery type by counting the number of interruptions until task completion. In [18], the service times are generally distributed while the interruptions arrive as a Poisson process with generally distributed down times.

Such service completion systems have found many opportunities in queuing applications. Researchers investigate the Laplace–Stieltjes transform of the stationary number of customers in the system of an $M/M/1$ queue under server breakdowns [4, 38], and of an $M/M/c$ queue under independent server breakdowns [27]. Coffman *et al.* [11] study a processor-sharing $M/M/1$ queue. The authors obtain the Laplace–Stieltjes transform and the first two moments of the stationary waiting time distribution. Others extend this research to the partial failure case [13, 33, 39]. Such service systems are called Markovian service process (MSP). Nicola [30] considers a mixture of interruption types affecting a single server as Poisson arrivals. He obtains the Laplace–Stieltjes transform of task completion time distribution under various types of interruptions. Kulkarni *et al.* [25] investigate a server affected by a Markov modulated environment in which the service rate in each environmental state is different, that is, the service deteriorates. The authors derive the Laplace–Stieltjes transform of the completion time distribution under each type of recovery schemes using renewal arguments. Furthermore the authors assume that associated to each environmental state a recovery scheme is fixed as either preempt-resume or preempt-restart discipline. They generalize their results to the semi-Markovian environment in [24]. Furthermore, the authors apply their results for preempt-resume service discipline to a processor sharing $M/M/1$ queue. Nicola *et al.* [31] analyze a single server queue with Markov modulated service process, and demonstrate that the queue has a block M/G/∞ structure and provide a procedure to evaluate the moment generating function for the stationary distribution of the number of jobs in the system. Again under preempt-resume service discipline, Boxma and Kurkova [10] consider an $M/M/1$ queue served under alternating speed and generally distributed low-speed times, and investigate the tail behavior of the workload distribution. Baykal-Gürsoy and Xiao [7] study an infinite server queue with two-state Markovian arrival and service processes. Using transform methods, they show that the stationary distribution of the number of jobs in the system is a mixture of two randomized Poisson distributions. Thus show the validity of the stochastic decomposition property. However, this property is not valid for $M/M/c$ queues under two-state service system [8, 28, 29]. Following Neuts [28], others [23, 32] use matrix-analytic methods to solve for the stationary distribution of multi-dimensional queues.

For the complete service breakdown case, Gaver [18] is the first to notice that under the restart strategy the first two moments of the task completion time may not always exist. Later Fiorini *et al.* [16] and Sheahan *et al.* [34] show that under the restart strategy and the same exponential up time assumption, the total time it takes to execute a task not including failures follows a power tailed distribution even when the task service time has exponential tail. Asmussen *et al.* [2] further extend the asymptotic analysis of the restart case in Sheahan *et al.* [34] to more general up time and task time distributions. They notice that the relationship between the up time and task time distributions play an important role impacting the distribution of completion times. In [3], Asmussen, Lipsky, and Thompson show that task completion time is heavy-tailed if the task time has unbounded support. Jelenković and Tan [21] independently study the same restart strategy and approach the analysis by first proving that the number of restarts is power tailed, and then using large deviation theory show that the completion times also have power law distribution irrespective of how heavy or light the distributions of task times and up times may be. Jelenković and Tan [22] extend these results to analyze further how a certain functional relationship between the tail distributions of the up time and task time distributions impacts the distributions of the number of restarts and the completion times.

In this paper, we study the completion time distribution of a task processed by a server experiencing service interruptions. The processor starts working on a task at a random time, and during the interruptions the server works at a lower service rate. Firstly, we derive the Laplace–Stieltjes transform

of the completion time distributions under both replace and restart service disciplines using counting arguments. The approach that we present here yields more detailed results than Kulkarni *et al.* [24] and Asmussen *et al.* [3] for our specific cases. Secondly, we show the asymptotic behavior of these distributions and prove that under the RESTART service discipline the completion time has power tail more directly than Asmussen *et al.* [3]. Finally, we apply our results to an infinite server queue in which the servers experience Markovian partial failures. Using Little's law, we compare the stationary system size and system time distributions of customers in a two server state M/MSP/∞ queue, and in a M/G/∞ queue in which each server experiencing Markovian service interruptions independently of the others.

Section 2 presents the service model for both cases in detail, introducing the corresponding notation and giving the main analytical results for general and specific task-size distributions. Section 3 deals with the asymptotic classification of the resulting completion-time distributions for both the *replace* and *restart* cases. It demonstrates that for the *restart* case not all moments may exists. Section 4 considers the application of our results to an infinite server queue with Poisson arrivals, and service interruptions, and exhibits the insensitivity of the stationary first moments of the system size and system time random variables to system-wide or independent interruptions. Finally, in Section 5, we draw the conclusions of the study and discuss some directions to extend the research.

2. Service model

We consider an unreliable server which from time to time experiences partial failures that reduce the service speed. Upon completion of a repair the server resumes its normal operation speed. We call the periods that the server works with normal speed as *up periods*, and the periods that the server works with low speed as *down periods*. It is so that the server state follows an alternating renewal process of up and down periods.

In general, for a continuous random variable C, we denote its probability density function (pdf) $f_C(t)$, cumulative distribution function (cdf) by $F_C(t)$, tail distribution by $\overline{F_C(t)}$, and its Laplace–Stieltjes transform by $L_C(s)$.

We denote by $F_S(t)$ the cdf for the *task size*, with pdf $f_S(t)$, tail distribution $\overline{F_S(t)}$, Laplace–Stieltjes transform $L_S(s)$, and mean task size $1/\mu$, $\mu > 0$. The task size random variable denoted by S is generally distributed. In literature, this quantity is also called as *service time requirement*.

The up period duration random variable U is exponentially distributed, except noted otherwise, with mean $1/f$, $f \geq 0$. However, down period duration D is generally distributed with $f_D(t)$, $F_D(t)$, $\overline{F_D(t)}$, $L_D(s)$, and mean duration $1/r$, $r \geq 0$. Notice that f and r can be interpreted as failure and repair rates of the server, respectively.

When technical reasons require its definition we will denote the remaining down time for a customer who arrives during down time by Y. It is clear that Y is generally distributed unless D is an exponential random variable.

Finally, without loss of generality the server's normal service speed is considered to be 1, and when a failure happens, the service speed drops to α with $0 < \alpha < 1$. As α does not necessarily take the value of zero this kind of service interruption is called *partial failure*. Note that in this system, a task may be finished during an up period or a down period.

The objective is to derive closed form expressions for the distribution of the task completion time random variable, which is denoted by T for both preempt-replace and preempt-restart service disciplines. Although, for the replace case, a general procedure to calculate the distribution of T in the frequency domain can be found in the works by Kulkarni *et al.* [24, 25], we use a much simpler counting argument for our particular class of service systems yielding more detailed results. In fact, the same argument will also be utilized to analyze the restart service discipline.

If work on a task starts at a random instance, the completion time can be obtained by conditioning it on the instance the work starts on a task. Let us call G^1, the event in which the work starts during an up period, and G^2, the event in which the work starts during a down period. In particular, by renewal arguments it holds that:
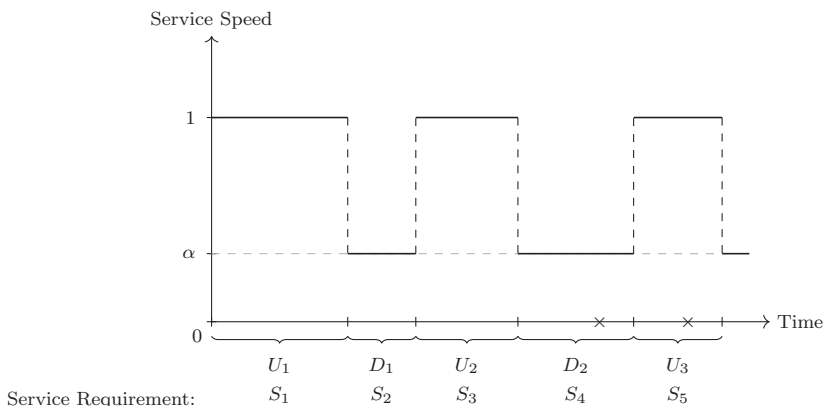
Figure 1. *Sample path of the service system under* G^1.

$$P\left\{G^1\right\} = \frac{r}{f+r}, \qquad P\left\{G^2\right\} = \frac{f}{f+r}. \tag{2.1}$$

The Laplace–Stieltjes transform of the conditional completion time is calculated separately for each one of these events, and then the Laplace–Stieltjes transform of the unconditional completion time is derived. Note that, one can assume that work always starts when the server is working under normal speed, then the conditional completion time under G^1 gives the full task completion time information.

We start by studying the completion time of a task under G^1.

Consider that the work on a task starts at an up period, and denote the subsequent up and down periods as U_i and D_i, respectively, for $i = 1, 2, 3, \ldots$. The service requirement on each period is denoted as S_i, $i = 1, 2, 3, \ldots$. Figure 1 shows a sample path of how the system may evolve. Since up period durations are exponentially distributed, one can assume that the work starts at time zero, and we denote with crosses on the time-axis some possible departure times (realizations of the completion time).

Hence, the conditional completion time given that the work starts during up period, $\{T|G^1\}$, can be written as

$$\{T|G^1\} = \begin{cases} \sum_{i=1}^{n} (U_i + D_i) + S_{2n+1}, & \text{event } A_n, \\ \sum_{i=1}^{n+1} U_i + \sum_{i=1}^{n} D_i + \frac{1}{\alpha} S_{2n+2}, & \text{event } E_n, \end{cases} \quad n = 0, 1, 2, \ldots, \tag{2.2}$$

in terms of events A_n and E_n. Given that the works starts during an up period, A_n denotes the case that the completion time is composed of n complete up and down periods with an incomplete up period finishing the task, and event E_n denotes the case with $n+1$ up periods and n down periods with an incomplete down period finishing the task. In general, A denotes the case that the work on a task starts and ends at the same regime, while E denotes the case that the work starts and ends at different regimes.

Similarly, the conditional completion time given that the work starts at a down period, $\{T|G^2\}$, will be:

$$\{T|G^2\} = \begin{cases} \frac{1}{\alpha} S_1, & \text{event } A_0, \\ Y + \sum_{i=2}^{n} D_i + \sum_{i=1}^{n} U_i + \frac{1}{\alpha} S_{2n+1}, & \text{event } A_n, \quad n = 1, 2, \ldots \\ Y + \sum_{i=2}^{n+1} D_i + \sum_{i=1}^{n} U_i + S_{2n+2}, & \text{event } E_n, \quad n = 0, 1, 2, \ldots \end{cases} \tag{2.3}$$

with event A_n representing the case that the completion time is composed of one residual down time denoted by Y, and $n-1$ complete down and n complete up periods with an incomplete down period finishing the task, and E_n representing the case that the completion time is composed of one residual

down period, and n complete down and up periods with an incomplete up period finishing the task. Next, we will first analyze preempt-replace and then study preempt-restart service disciplines.

2.1. *Preempt-*REPLACE *service discipline*

In the case that the service times (task size) are generally distributed, we assume that any state change reinitiates the service for the current task with the same service requirement distribution, that is, all S_n's for $n = 1, 2, \ldots$ are i.i.d. This last consideration corresponds to the preemption discipline called *repeat different* or *replace*, since the service requirement is resampled at each service-speed change. This is also equivalent to replacing the task with a similar one with the same task size distribution.

Proposition 2.1 shows the resulting completion time pdf in the frequency domain.

Proposition 2.1. *Consider a two-service-speed server as described above. Then, the Laplace–Stieltjes transform of the completion time random variable T is given in Eqs. (2.4)–(2.5).*

$$
\begin{aligned}
L_T(s) &= E[e^{-sT}] \\
&= \frac{1}{1 - V(s)} \left\{ \left(L_S(s+f) + \frac{f}{s+f}[1 - L_S(s+f)] \left[L_S(s/\alpha) - \int_0^\infty e^{-(s/\alpha)t} F_D(t/\alpha) dF_S(t) \right] \right) \right. \\
&\quad \left. \cdot \left(\frac{r}{r+f} + \frac{f}{r+f} \left[L_Y(s) - \int_0^\infty e^{-st} F_S(\alpha t) dF_Y(t) \right] \right) \right\} + \frac{f}{r+f} \left[L_S(s/\alpha) \right. \\
&\quad \left. - \int_0^\infty e^{-(s/\alpha)t} F_Y(t/\alpha) dF_S(t) \right],
\end{aligned}
\tag{2.4}
$$

where

$$
V(s) = \frac{f}{s+f}[1 - L_S(s+f)] \cdot \left[L_D(s) - \int_0^\infty e^{-st} F_S(\alpha t) dF_D(t) \right].
\tag{2.5}
$$

The proof of Proposition 2.1 is constructive and can be found in detail in Appendix A. A special case of interest is detailed in Corollary 2.2, which is stated without proof.

Corollary 2.2. (Exponential service time) *Assume that the service requirement is exponentially distributed with mean $1/\mu$. Then, the Laplace–Stieltjes transform of the completion time distribution given in (2.4)–(2.5) reduces to (2.6)–(2.7).*

$$
\begin{aligned}
L_T(s) &= E\left[e^{-sT}\right] \\
&= \frac{f}{(f+r)} \frac{\mu\alpha \left[s + \mu\alpha - r(1 - L_D(s+\mu\alpha))\right]}{(s+\mu\alpha)^2} \\
&\quad + \frac{1}{1-V(s)} \left\{ \frac{r}{(f+r)} \frac{\mu \left[s + \mu\alpha + f(1+\alpha)(1 - L_D(s+\mu\alpha))\right]}{(s+f+\mu)(s+\mu\alpha)} \right. \\
&\quad \left. + \frac{f}{(f+r)} \frac{fr\mu\alpha \left[1 - L_D(s+\mu\alpha)\right]^2}{(s+f+\mu)(s+\mu\alpha)^2} \right\}
\end{aligned}
\tag{2.6}
$$

where

$$
V(s) = \frac{f}{s+f+\mu} \cdot L_D(s+\mu\alpha).
\tag{2.7}
$$

The following specialized result for exponential up and down durations was proved by Baykal-Gursoy *et al.* [5].

Corollary 2.3. (Exponential down periods) *Assume that the down period duration D is exponentially distributed with mean $1/r$. Then, the Laplace–Stieltjes transform of the completion time distribution given in (2.4)–(2.5) reduces to (2.8)–(2.9).*

$$
\begin{aligned}
L_T(s) &= E[e^{-sT}] \\
&= E[e^{-sT}|G^1]P\{G^1\} + E[e^{-sT}|G^2]P\{G^2\} \\
&= \frac{r}{f+r} \cdot \frac{1}{1-V(s)} L_S(s+f) \left(1 + \frac{f}{s+r}\left[1 - L_S\left(\frac{s+r}{\alpha}\right)\right]\right) \\
&\quad + \frac{f}{f+r} \cdot \frac{1}{1-V(s)} L_S\left(\frac{s+r}{\alpha}\right) \left(1 + \frac{r}{s+f}\left[1 - L_S(s+f)\right]\right),
\end{aligned}
\tag{2.8}
$$

where

$$
V(s) = \frac{rf\left[1 - L_S(s+f)\right]\left[1 - L_S\left(\frac{s+r}{\alpha}\right)\right]}{(s+f)(s+r)},
\tag{2.9}
$$

and with mean

$$
E[T] = \frac{\dfrac{1}{r}(1 - L_S(r/\alpha))\left[1 - \dfrac{r}{r+f}L_S(f)\right] + \dfrac{1}{f}(1 - L_S(f))\left[1 - \dfrac{f}{r+f}L_S(r/\alpha)\right]}{L_S(r/\alpha) + L_S(f) - L_S(f)L_S(r/\alpha)}.
\tag{2.10}
$$

Under very special circumstances the completion time distribution Laplace–Stieltjes transform given in (2.8)–(2.9) can be inverted analytically. In general, it is necessary to resort to numerical Laplace inversion methods (see, e.g., [1], [37], [36]). We have experience with De Hoog's algorithm [12] and found it to be efficient when implemented with accelerated convergence for the continued fraction expansion that is developed by Hollenbeck [20]. It is worth mentioning that no single method for Laplace transform inversion is guaranteed to give good results as this depends greatly on the specific application [14].

The case in which all random variables, that is, up and down time durations, and service time requirement, are exponential was also discussed in [5].

Corollary 2.4. (Exponential down periods and service times) *Consider the server described in Section 2, assume that the down period duration D is exponentially distributed with mean $1/r$, and that the service requirement is exponentially distributed with mean $1/\mu$. Then, the Laplace–Stieltjes transform of completion time r.v. is given as*

$$
L_T(s) = \frac{\mu}{f+r} \cdot \frac{r(s+f+r+\mu\alpha) + f\alpha(s+f+r+\mu)}{(s+f+\mu)(s+r+\mu\alpha) - fr}.
\tag{2.11}
$$

Using Eq. (2.11), the completion time T distribution can be written as shown below by defining two random variables, T_1 and T_2, which are independent, denoting the completion time of customers arriving during up and down periods, respectively,

$$
T = \frac{r}{f+r}T_1 + \frac{f}{f+r}T_2.
\tag{2.12}
$$

The density functions of T_1 and T_2 can be obtained in closed form by inverting the Laplace–Stieltjes transform as shown in (2.13)–(2.14),

$$f_{T_1}(t) = \frac{\mu(s_1 + f + r + \mu\alpha)}{s_1 - s_2} e^{s_1 t} - \frac{\mu(s_2 + f + r + \mu\alpha)}{s_1 - s_2} e^{s_2 t}, \tag{2.13}$$

$$f_{T_2}(t) = \frac{\mu\alpha(s_1 + f + r + \mu)}{s_1 - s_2} e^{s_1 t} - \frac{\mu\alpha(s_2 + f + r + \mu)}{s_1 - s_2} e^{s_2 t}, \tag{2.14}$$

where s_1 and s_2 are the solutions of Eq. (2.15),

$$(s + f + \mu)(s + r + \mu\alpha) - fr = 0, \tag{2.15}$$

and are given in Eqs. (2.16) and (2.17):

$$s_1 = \frac{-(f + r + \mu(1 + \alpha)) + \sqrt{(f + r + \mu(1 + \alpha))^2 - 4\mu(\alpha\mu + f\alpha + r)}}{2}, \tag{2.16}$$

$$s_2 = \frac{-(f + r + \mu(1 + \alpha)) - \sqrt{(f + r + \mu(1 + \alpha))^2 - 4\mu(\alpha\mu + f\alpha + r)}}{2}. \tag{2.17}$$

It holds that s_1 and s_2 are both negative real numbers and it is clear that $s_1 \geq s_2$.

The expected completion time for this case is:

$$E[T] = \frac{1}{\mu}\left(1 + \frac{1 - \alpha}{\alpha} \cdot \frac{f}{f + r} \cdot \frac{f + r + \mu}{f + r/\alpha + \mu}\right), \tag{2.18}$$

which is higher than the mean service time, $1/\mu$.

2.1.1. Approximating mean completion time using renewal arguments

As discussed before, renewal arguments provide the probability of server being up in steady state as $r/(r + f)$, and the probability of server being down as $f/(r + f)$. Then, since the mean completion time in an up period is the mean service time requirement, $1/\mu$, and correspondingly $1/(\alpha\mu)$ for down periods, one might approximate the mean completion time in steady state as the mixture:

$$\text{SSMM} := \frac{r}{r + f}\frac{1}{\mu} + \frac{f}{r + f}\frac{1}{\alpha\mu} \tag{2.19}$$

with SSMM standing for steady-state mixture mean.

However, we can show that in the Markovian case in which all random periods are exponential, the steady state mixture mean over-estimates the expected completion time. In fact, we can calculate the difference explicitly as in Eq. (2.20) [5]

$$\text{SSMM} - E[T] = \frac{fr(1 - \alpha)^2}{\alpha\mu(r + f)(f\alpha + r + \mu\alpha)} \geq 0. \tag{2.20}$$

It is worth mentioning that the generating function of the completion time for the complete breakdown case (which can be obtained by taking the limit $\alpha \to 0$) coincides with Gaver's preemptive-repeat-different interruption case [18].
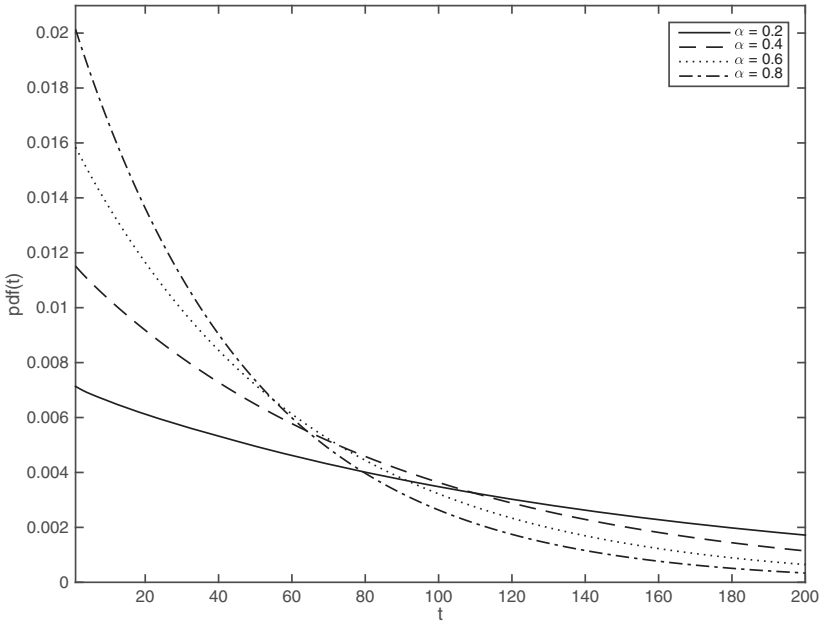
Figure 2. *Completion time distribution under exponential service requirement.*

Table 1. *Expected completion time and SSMM for exponential* S.

α:	0.2	0.4	0.6	0.8
$E[T]$:	141.59	86.38	62.22	48.67
SSMM:	182.22	93.33	63.70	48.89

Figure 2 shows an example of the time domain completion time distribution for exponential service time given in Eqs. (2.13)–(2.14). Here we consider $1/f = 5 < 40 = 1/r$ time units, which corresponds to a system in which failures occur more frequently than repairs. The mean service requirement is $1/\mu = 40$, and we vary α to 0.2, 0.4, 0.6, and 0.8.

It is clear that as α decreases the mass of the distribution is increasingly shifted toward the right-tail. In particular, Table 1 shows the expected completion times for different values of α calculated using Eq. (2.18), in time units. The expected completion time increases as the server capacity drops to a smaller proportion, α. We also give the SSMM for comparison.

2.2. *Preempt-**RESTART** service discipline*

In this case, the service requirement, S, is sampled once, and every time a state-change occurs the same service requirement is repeated, that is, $S_n = S$ for $n = 1, 2, \ldots$. Thus, given that the service requirement, $S = t$, the completion time given in Eqs. (2.2) and (A.11) are rewritten for an up period, and a down period as shown in Eqs. (2.21) and (2.22), respectively,

$$\{T | S \approx t, G^1\} = \begin{cases} t, & \text{event } A_0 \\ X + \sum_{i=2}^{n} U_i + \sum_{i=1}^{n} D_i + t, & \text{event } A_n, \quad n = 1, 2, \ldots \\ X + \sum_{i=2}^{n+1} U_i + \sum_{i=1}^{n} D_i + \frac{t}{\alpha}, & \text{event } E_n, \quad n = 0, 1, 2, \ldots, \end{cases} \quad (2.21)$$

$$\{T|S \approx t, G^2\} = \begin{cases} \frac{t}{\alpha}, & \text{event } A_0 \\ Y + \sum_{i=2}^n D_i + \sum_{i=1}^n U_i + \frac{t}{\alpha}, & \text{event } A_n, \quad n = 1, 2, \ldots \\ Y + \sum_{i=2}^{n+1} D_i + \sum_{i=1}^n U_i + t, & \text{event } E_n, \quad n = 0, 1, 2, \ldots \end{cases} \tag{2.22}$$

where X denoting the remaining up time with cdf $F_X(t)$, tail distribution $\overline{F_X(t)}$, conditional mean $E(X|t) = E[X|X < t$, and conditional Laplace–Stieltjes transform $L_X[s|t] = E[e^{-sX}|X < t]$.

Theorems 2.5–2.8 generalize the results in [16, 34] for the complete breakdown case to the partial breakdown case under the *restart* preemption strategy. Theorem 2.5 presents the Laplace–Stieltjes transform of the completion time conditioned on the task size, and Theorem 2.8 gives the expected completion time conditioned on the task size. The results are stated for general up and down durations with CDFs $F_U(t)$ and $F_D(t)$, tail distributions $\overline{F_U(t)}$ and $\overline{F_D(\frac{t}{\alpha})}$, and Laplace–Stieltjes transform $L_U(s)$ and $L_D(s)$, respectively.

Theorem 2.5. *The conditional Laplace–Stieltjes transform of the completion time distribution for a task of duration t is*

$$L_T(s|S \approx t) = \frac{1}{(r+f)}\Big[re^{-st}\overline{F_X(t)} + fe^{-s\frac{t}{\alpha}}\overline{F_Y(\frac{t}{\alpha})}\Big]$$
$$+ \frac{e^{-st}\overline{F_U(t)}\big[rF_X(t)L_X(s|t)F_D(\frac{t}{\alpha})L_D(s|\frac{t}{\alpha}) + fF_Y(\frac{t}{\alpha})L_Y(s|\frac{t}{\alpha})\big]}{(r+f)\big[1 - F_U(t)L_U(s|t)F_D(\frac{t}{\alpha})L_D(s|\frac{t}{\alpha})\big]}$$
$$+ \frac{e^{-s\frac{t}{\alpha}}\overline{F_D(\frac{t}{\alpha})}\big[rF_X(t)L_X(s|t) + fF_Y(\frac{t}{\alpha})L_Y(s|\frac{t}{\alpha})F_U(t)L_U(s|t)\big]}{(r+f)\big[1 - F_U(t)L_U(s|t)F_D(\frac{t}{\alpha})L_D(s|\frac{t}{\alpha})\big]}, \tag{2.23}$$

with $L_U(s|t) \equiv E[e^{-sU}|U < t]$ and $L_D(s|\frac{t}{\alpha}) \equiv E[e^{-sD}|D < \frac{t}{\alpha}]$ denoting the conditional Laplace–Stieltjes transform of the up time and down time distributions, respectively, given that the up and down times end before the completion of a task of duration t.

The proof of Theorem 2.5 can be found in Appendix B.

Remark 2.6. Note that if the up and down durations are exponential, then the conditional Laplace–Stieltjes transform of the completion time distribution for a task of duration t is

$$E[e^{-sT}|S \approx t] = \frac{e^{-st}\overline{F_U(t)}\big[r + fL_D(s|\frac{t}{\alpha})F_D(\frac{t}{\alpha})\big] + e^{-s\frac{t}{\alpha}}\overline{F_D(\frac{t}{\alpha})}\big[f + rL_U(s|t)F_U(t)\big]}{(r+f)\big[1 - L_U(s|t)L_D(s|\frac{t}{\alpha})F_U(t)F_D(\frac{t}{\alpha})\big]}. \tag{2.24}$$

Remark 2.7. Note that the above result also matches the result of Gaver [18] when the server does not work during down times, that is, $\alpha = 0$.

Theorem 2.8. *The expected completion time for a task of length t is*

$$E[T|S \approx t] = \frac{1}{r+f}\Big[r\big(t\overline{F_X(t)} + E[X|t]F_X(t)\big) + f\big(\frac{t}{\alpha}\overline{F_Y(\frac{t}{\alpha})} + E[Y|\frac{t}{\alpha}]F_Y(\frac{t}{\alpha})\big)\Big]$$
$$+ \frac{\big[rF_X(t)F_D(\frac{t}{\alpha}) + fF_Y(\frac{t}{\alpha})\big]\big[t\overline{F_U(t)} + E[U|t]F_U(t)\big] + \big[rF_X(t) + fF_Y(\frac{t}{\alpha})F_U(t)\big]\big[\frac{t}{\alpha}\overline{F_D(\frac{t}{\alpha})} + E[D|\frac{t}{\alpha}]F_D(\frac{t}{\alpha})\big]}{(r+f)\big(1 - F_U(t)F_D(\frac{t}{\alpha})\big)} \tag{2.25}$$

with $E[U|t] \equiv E[U|U < t]$, $E[X|t] \equiv E[X|X < t]$, $E[D|\frac{t}{\alpha}] \equiv E[D|D < \frac{t}{\alpha}]$, and $E[Y|\frac{t}{\alpha}] \equiv E[Y|D < \frac{t}{\alpha}]$ denoting the conditional mean of the up time, remaining up time, down time, and remaining down time distributions, respectively, given that the up and down times end before the completion of a task of duration t.

The proof of Theorem 2.8 can be found in Appendix C.

Remark 2.9. Note that if up and down durations are exponential, then the conditional expectation of the completion for a task of length t is

$$E[T|S \approx t] = \frac{\left(t\overline{F_U(t)} + E[U|t]F_U(t)\right)\left(r + fF_D(\frac{t}{\alpha})\right) + \left(\frac{t}{\alpha}\overline{F_D(\frac{t}{\alpha})} + E[D|\frac{t}{\alpha}]F_D(\frac{t}{\alpha})\right)\left(f + rF_U(t)\right)}{(r+f)\left(1 - F_U(t)F_D(\frac{t}{\alpha})\right)}.$$

(2.26)

The Laplace–Stieltjes transform and mean completion time r.v., T, can be derived by unconditioning the respective equations (2.23 and 2.25), with respect to the task size, S, that is,

$$L_T(s) = \int_{t=0}^{\infty} L_T(s|S \approx t)dF_S(t),$$

$$E[T] = \int_{t=0}^{\infty} E[T|S \approx t]dF_S(t).$$

3. Tail behavior of the completion time

We now present some asymptotic properties of the completion time. This amounts to studying the characteristics of the distribution's *right-tail*. The right-tail defines the probability of observing abnormally large completion times, and it is calculated as the complementary cumulative distribution, as shown in (3.1)

$$\bar{F}_T(t) = \mathbb{P}(T > t) = 1 - F_T(t) = 1 - \int_0^t f_T(u)du.$$

(3.1)

We again separate the *REPLACE* and *RESTART* cases. We briefly discuss the first case and give some general insights of the tail behavior. For the second case, we derive more specific results on the asymptotic classification.

3.1. Repeat different/**REPLACE**

For the exponential down periods case, the Laplace transform of the tail can be given explicitly as shown in Eq. (3.2)

$$\mathcal{L}[\bar{F}_T](s) = \frac{r}{(r+f)} \frac{(1 - L_S(s+f))\left[s + r + f\left(1 - L_S((s+r)/\alpha)\right)\right]}{\left[(s+f)(s+r) - rf(1 - L_S(s+f))\left(1 - L_S((s+r)/\alpha)\right)\right]}$$
$$+ \frac{f}{(r+f)} \frac{\left(1 - L_S((s+r)/\alpha)\right)\left[s + f + r(1 - L_S(s+f))\right]}{\left[(s+f)(s+r) - rf(1 - L_S(s+f))\left(1 - L_S((s+r)/\alpha)\right)\right]},$$

(3.2)

where $L_S(\cdot)$ is the Laplace–Stieltjes transform of the service time requirement. In general, the Laplace transform in (3.2) has to be obtained numerically.

More generally, it can be shown via the counting argument in the proof of Proposition 2.1 (see Eqs. (A.3) and (A.11)), that T is a stochastic sum of the up and down period random variables (U, D) and the service time requirement random variable (S). Then, it suffices for the D or S to be heavy tailed for T to be heavy tailed [17].

3.2. Repeat identical/RESTART

In order to analyze the asymptotic behavior of the completion time, one needs to consider the following function for an r.v. C with support contained in $[0, +\infty)$ and distribution $F_C(t)$:

$$\phi(\theta; C) := \int_0^\infty e^{\theta t} dF_C(t). \tag{3.3}$$

Let $\theta_{\min}(C) := \sup\{\theta \mid \phi(\theta; C) < \infty\}$. Using this definition, we can classify any distribution function on having finite range, a light tail, an exponential tail, or a heavy tail. This corresponds to the cases of finite support, $\theta_{\min}(C) = \infty$, $0 < \theta_{\min}(C) < \infty$, and $\theta_{\min}(C) = 0$, respectively [34]. We write $\theta_{\min}$ when there is no ambiguity on the r.v. C.

Theorem 3.1 shows conditions under which the completion time distribution is *power tailed*.

Theorem 3.1. *Let U and D be exponentially distributed with failure rate f and repair rate r, and call $\Delta = \min\{f, r/\alpha\}$. Assume that the service-time requirement (task size) has an exponential tail, that is: $0 < \theta_{min}(S) = \mu < \infty$, and denote $\varepsilon := \theta_{min}(S)/\Delta = \mu/\min\{f, r/\alpha\}$. Then,*

$$E[T^m] = \infty, \quad \forall m \geq \varepsilon, \tag{3.4}$$

and the completion time has power tail,

$$\bar{F}_T(t) \sim \frac{c}{t^\varepsilon}. \tag{3.5}$$

The proof of Theorem 3.1 is shown in Appendix D.

We note that the tail behavior of completion times depends on the ratio of the maximum of the means of up time and the speed adjusted down time, with the mean task time. If the mean task time is close to the larger mean then all moments of the completion time are undefined. The details of this derivation can be found in the proof of Theorem 3.1 in Appendix D.

Although the preemptive-replace and preemptive-repeat service disciplines seem similar, their tail behaviors are quite different. This happens because in the preemptive-replace case service interruptions may decrease the completion time over what it would be in the preemptive-repeat case, since each service change brings a new task that may have shorter time requirement. In a sense, these interruptions may terminate tasks that require longer time requirement, thus shortening the completion time [18].

4. An application: insensitivity of the stationary first order moments for infinite server queues with two-state MSP

Several different queueing systems with service degradation can be defined by using the completion time model described above as a generally distributed service time. Next we discuss a model having Poisson arrivals with parameter λ and infinitely many servers of $M/G/\infty$ type.

With exponential up and down periods, the process controlling the service rate is a two-state continuous-time Markov chain, independent of the arrival process, and is considered as the external environment. In the case of multiple servers, if all servers are controlled simultaneously, which means that interruptions occur system-wide then the queue is said to have an MSP. On the other hand, similar

interruption processes may affect each server independently. These interruptions are assumed to be of preemptive-replace type.

In the case of finitely many servers, system-wide partial failures, and exponential task times, the system becomes the M/MSP/c queue analyzed in [6] with two service states. The independent server breakdown case is studied by Mitrany and Avi-Itzhak [27].

In the case of infinitely many servers, system-wide partial failures, and exponential task time distribution, the system coincides with the M/MSP/∞ queue considered in [7]. Baykal-Gürsoy and Xiao [7] show that the steady-state number in the system, *N*, is the sum of two independent random variables: a Poisson r.v. representing the stationary number of customers in an uninterrupted M/G/∞ system, and a randomized Poisson r.v. representing the extra customers accumulated during interruptions. Then, the mean steady-state number of customers in the system is derived as

$$E[N] = \frac{\lambda}{\mu} + \frac{\lambda f(1-\alpha)}{f+r} \cdot \frac{f+r+\mu}{\mu(f\alpha+r+\mu\alpha)}. \tag{4.1}$$

On the other hand, assuming that there are no jockeying between the servers, one can analyze the independent server interruption case similar to an M/G/∞ system as will be discussed below.

When Markovian service interruptions arrive independently to each of the infinitely many servers, a job joining the system experiences exactly the same task completion time derived in Corollary 2.3. Each task completion time is independent and identically distributed. Hence, this system becomes an M/G/∞ system with the service time equal to the completion time, *T*, that is given in the frequency domain by Eqs. (2.8) and (2.9).

Clearly, the stationary number of customers in the M/G/∞ queue is Poisson distributed with parameter $\rho = \lambda \cdot E[T]$. Hence, the expected number of customers in the system, *N*, using the expected completion time from Eq. (2.18) is given as in (4.1) (see Eq. (3.6) in [7]).

The expected system time for this system, that is, the job completion time, in turn coincides with the mean system time for the M/MSP/∞ queue obtained via Little's law from Eq. (4.1) as $E[T] = E[N]/\lambda$. For the M/MSP/∞ queue in [7], however, the steady-state variance of the number in the system is not equal to its expected value, which has to be the case in the M/G/∞ setting with independent server failures and preempt-replace service discipline. Thus, except the first order moments, the higher moments do not match in the case of system wide interruptions with the interruptions affecting each server independently.

5. Conclusion

We study the task completion time of a server experiencing randomly occurring service deterioration. Focusing on preempt-replace and preempt-repeat service recovery disciplines following each service interruption, we derive the Laplace–Stieltjes transform of the task completion times using counting arguments. We observe that in general the resulting distributions are difficult to obtain explicitly in the time domain and one has to resort to numerical inversion of the transforms. For the specific case of exponential down period and exponential task time case we can determine the exact form of the completion time. Furthermore, we show how the steady-state mixture can be compared against the expected completion time, which can be useful for applications in which a simpler steady state model is preferred.

The analysis of the tail distribution demonstrates that in the preempt-repeat service discipline even when the task time distribution has exponential tail, the completion time distribution may have power tail. Our results provide conditions under which the moments of the completion time exist. Moreover, the connection between the expected task completion time presented in this study and the expected system time for the M/MSP/∞ queue studied by Baykal-Gürsoy and Xiao [7], reveals that the first order moments at an infinite server queue in random environment are insensitive to how the random environment affects the servers.

References

[1] Abate, J. & Whitt, W. (1995). Numerical inversion of Laplace transforms of probability distributions. *ORSA Journal on Computing* 7(1): 36–43.

[2] Asmussen, S., Fiorini, P., Lipsky, L., Rolski, T., & Sheahan, R. (2008). Asymptotic behavior of total times for jobs that must start over if a failure occurs. *Mathematics of Operations Research* 33(4): 932–944. 10.1287/moor.1080.0329, http://mor.journal.informs.org/content/33/4/932.abstract, http://mor.journal.informs.org/content/33/4/932.full.pdf+html.

[3] Asmussen, S., Lipsky, L., & Thompson, S. (2016). Markov renewal methods in restart problems in complex systems. *The Fascination of Probability, Statistics and Their Applications: In Honour of Ole E Barndorff-Nielsen*: 501–527.

[4] Avi-Itzhak, B. & Naor, P. (1963). Some queueing problems with the service station subject to breakdown. *Operations Research* 11(3): 303–320.

[5] Baykal-Gürsoy, M., Benton, A.R., Gerum, P.C.L., & Candia, M.F. (2022). How random incidents affect travel-time distributions. *IEEE Transactions on Intelligent Transportation Systems* 23(8): 13000–13010, 10.1109/TITS.2021.3119024

[6] Baykal-Gürsoy, M. & Duan, Z. (2006). M/M/C queues with Markov modulated service processes. In *First International Conference on Performance Evaluation Methodologies and Tools-Valuetools*. Pisa, IT.

[7] Baykal-Gürsoy, M. & Xiao, W. (2004). Stochastic decomposition in M/M/∞ queues with Markov-modulated service rates. *Queueing Systems* 48: 75–88.

[8] Baykal-Gürsoy, M., Xiao, W., & Ozbay, K.M.A. (2009). Modeling traffic flow interrupted by incidents. *European Journal of Operational Research* 195(1): 127–138.

[9] Bell, E.T. (1927). Partition polynomials. *Annals of Mathematics* 29(1/4): 38–46.

[10] Boxma, O. & Kurkova, I. (2000). The M/M/1 queue in a heavy-tailed random environment. *Statistica Neerlandica* 54(2): 221–236.

[11] Coffman, E.G., Muntz, R.R., & Trotter, H. (1970). Waiting time distributions for processor-sharing systems. *Journal of the Association for Computing Machinery (JACM)* 17(1): 123–130.

[12] de Hoog, F. (1987). A new algorithm for solving Toeplitz systems of equations. *Linear Algebra and its Applications* 88: 123–138.

[13] Eisen, M. & Tainiter, M. (1963). Stochastic variations in queuing processes. *Operations Research* 11(6): 922–927.

[14] Epstein, C.L. & Schotland, J. (2008). The bad truth about Laplace's transform. *Society for Industrial and Applied Mathematics (SIAM) Review* 50(3): 504–520.

[15] Faa di Bruno, F. (1855). Sullo sviluppo delle funzioni. *Annali di Scienze Matematiche e Fisiche* 6: 479–480.

[16] Fiorini, P.M., Sheahan, R., & Lipsky, L. (2005). On unreliable computing systems when heavy-tails appear as a result of the recovery procedure. *ACM SIGMETRICS Performance Evaluation Review* 33(2): 15–17.

[17] Foss, S., Korshunov, D., & Zachary, S. (2013). *An introduction to heavy-tailed and subexponential distributions.*, New York: Springer Science & Business Media.

[18] Gaver, D.P. (1962). A waiting line with interrupted service, including priorities. *Journal of the Royal Statistical Society Series B (Methodological)* 24(1): 73–90.

[19] Greiner, M., Jobmann, M., & Lipsky, L. (1999). The importance of power-tail distributions for modeling queueing systems. *Operations Research* 47(2): 313–326.

[20] Hollenbeck, K.J. (1998). INVLAP. M: A matlab function for numerical inversion of Laplace transforms by the de Hoog algorithm.

[21] Jelenković, P.R. & Tan, J. (2007). Can retransmissions of superexponential documents cause subexponential delays? In *INFOCOM 2007. 26th IEEE International Conference on Computer Communications*. IEEE, pp. 892–900, 10.1109/INFCOM.2007.109

[22] Jelenković, P.R. & Tan, J. (2013). Characterizing heavy-tailed distributions induced by retransmissions. *Advances in Applied Probability* 45(1): 106–138.

[23] Katehakis, M.N., Smit, L.C., & Spieksma, F.M. (2015). DES and RES processes and their explicit solutions. *Probability in the Engineering and Informational Sciences* 29(2): 191–217, 10.1017/S0269964814000291

[24] Kulkarni, V.G., Nicola, V.F., & Trivedi, K.S. (1987). The completion time of a job on multimode systems. *Advances in Applied Probability* 19(4): 932–954.

[25] Kulkarni, VG., Nicola, VF., & Trivedi, KS. (1986) On modeling the performance and reliability of multimode computer systems. *Tech. Rep.*, Duke University.

[26] Mihoubi, M. (2008). Bell polynomials and binomial type sequences. *Discrete Mathematics* 308(12): 2450–2459.

[27] Mitrany, I. & Avi-Itzhak, B. (1968). A many-server queue with service interruptions. *Operations Research* 16(3): 628–638.

[28] Neuts, M. (1978). Further results on the M/M/1 queue with randomly varying rates. *Operations Research Society of India (OPSEARCH)* 15(4): 158–168.

[29] Neuts, M. (1981). *Matrix-geometric solutions in stochastic models: An algorithmic approach.* The John Hopkins University Press.

[30] Nicola, V.F. (1986). A single server queue with mixed types of interruptions. *Acta Informatica* 23(4): 465–486.

[31] Nicola, V.F., Kulkarni, V.G., & Trivedi, K.S. (1987). Queueing analysis of fault-tolerant computer systems. *Software Engineering*. IEEE Transactions on 3: 363–375.

[32] O'Cinneide, C. & Purdue, P. (1986). The M/M/∞ queue in a random environment. *Journal of Applied Probability* 23: 175–184.

[33] Purdue, P. (1973). The M/M/1 queue in a Markovian environment. *Operations Research* 22(3): 562–569.

[34] Sheahan, R., Lipsky, L., Fiorini, P.M., & Asmussen, S. (2006). On the completion time distribution for tasks that must restart from the beginning if a failure occurs. *SIGMETRICS Performance Evaluation Review* 34(3): 24–26, http://doi.acm.org/10.1145/1215956.1215967, 10.1145/1215956.1215967

[35] Sloane, N.J.A., editor (2017). The On-Line Encyclopedia of Integer Sequences, Sequence A028246. https://oeis.org/A028246, Accessed June 9.

[36] Talbot, A. (1979). The accurate numerical inversion of Laplace transforms. *Institute of Mathematics and its Applications (IMA) Journal of Applied Mathematics* 23(1): 97–120.

[37] Weeks, W.T. (1966). Numerical inversion of Laplace transforms using Laguerre functions. *Journal of Association for Computing Machinery* 13(3): 419–429, https://doi.org/10.1145/321341.321351, 10.1145/321341.321351.

[38] White, H. & Christie, L. (1958). Queuing with preemptive priorities or with breakdown. *Operations Research* 6(1): 79–95.

[39] Yechiali, U. & Naor, P. (1971). Queueing problems with heterogeneous arrivals and service. *Operations Research* 19(3): 722–734.

Appendix A. Proof of Proposition 2.1

As mentioned before, we condition the completion time on the epoch that work starts on a task. Let G^1 denote the event in which work starts during an up period, and G^2 the event in which work starts during a down period. Then the occurrence probabilities of events G^1 and G^2 are given by Eq. (2.1). The Laplace–Stieltjes transform of conditional completion time is calculated separately for each one of these events and then unconditioned.

Consider the case that work starts during an up period, and denote the subsequent up and down periods as U_i and D_i, respectively, for $i = 1, 2, 3, \ldots$. The task time at each period is denoted as S_i, $i = 1, 2, 3, \ldots$. Refer to Figure 1 for a sample path of how the system could evolve. Define the following events:

- A_n, for $n = 0, 1, 2, \ldots$. There are n complete up and down periods, and another incomplete up period in the completion time. In general, A denotes the case that the work on a task starts and ends at the same regime. In G^1, work starts while the server is up. The second cross in Figure 1 is an example of event A_2. Notice that the occurrence of event A_n implies that necessarily:

  (1) $S_{2i-1} > U_i$, for all $i = 1, 2, \ldots, n$,
  (2) $\frac{1}{\alpha} S_{2i} > D_i$, for all $i = 1, 2, \ldots, n$,
  (3) $S_{2n+1} < U_{n+1}$.

- E_n, for $n = 0, 1, 2, \ldots$. There are $n + 1$ complete up and n complete down periods, and another incomplete down period in the completion time. In general, E denotes the case that the work on a task starts and ends at different regimes. The first cross in Figure 1 is an example of event E_1. Here, necessarily:

  (1) $S_{2i-1} > U_i$, for all $i = 1, 2, \ldots, n + 1$,
  (2) $\frac{1}{\alpha} S_{2i} > D_i$, for all $i = 1, 2, \ldots, n$,
  (3) $\frac{1}{\alpha} S_{2n+2} < D_{n+1}$.

Since clearly events S_i, U_i, and D_i, are independent and respectively identically distributed for all i, then for all $n = 0, 1, 2, \ldots$

$$P\left\{A_n | G^1\right\} = P^n\{S > U\} \cdot P^n\{S > \alpha D\} \cdot P\{S < U\}, \tag{A.1}$$

$$P\{E_n|G^1\} = P^{n+1}\{S > U\} \cdot P^n\{(1/\alpha)S > D\} \cdot P\{(1/\alpha)S < D\}. \tag{A.2}$$

Notice that the probabilities defined in (A.1)–(A.2) were obtained by simple enumeration of the number of state transitions. Then, as already stated in Eq. (2.2), the conditional completion time $\{T|G^1\}$ under each event stated above is

$$\{T|G^1\} = \begin{cases} \sum_{i=1}^{n}(U_i + D_i) + S_{2n+1}, & \text{event } A_n, \\ \sum_{i=1}^{n+1} U_i + \sum_{i=1}^{n} D_i + \frac{1}{\alpha}S_{2n+2}, & \text{event } E_n, \end{cases} \quad n = 0, 1, 2, \dots. \tag{A.3}$$

In order to specify the conditional completion time Laplace–Stieltjes transform under some event H, we use the indicator variable for H defined as

$$\mathbf{1}\{H\} = \begin{cases} 1, & \text{event } H, \\ 0, & \text{otherwise} \end{cases}$$

then, for all $n = 0, 1, 2, \dots$

$$\begin{aligned} E[e^{-sT}\mathbf{1}\{A_n\}|G^1] &= E[e^{-sT}|A_n, G^1] \cdot P\{A_n|G^1\} \\ &= \left(E[e^{-sU}|S > U]P\{S > U\}\right)^n \left(E[e^{-sD}|S > \alpha D]P\{S > \alpha D\}\right)^n \\ &\quad \times E[e^{-sS}|S < U]P\{S < U\}, \end{aligned} \tag{A.4}$$

where s is a complex number with positive real part, $U \sim \text{Exp}(f)$, S denotes the generally distributed task time random variable under normal conditions, and D is the generally distributed down period random variable.

The Laplace–Stieltjes transform of the up period for up times lasting less than the service requirement is derived below as

$$\begin{aligned} E[e^{-sU}|S > U]P\{S > U\} &= E[e^{-sU}\mathbf{1}\{S > U\}] = \int_{x=0}^{\infty} \int_{u=0}^{x} e^{-su} f e^{-fu} du\, dF_S(x) \\ &= \frac{f}{s+f} - \int_{u=0}^{\infty} f e^{-(s+f)u} F_S(u)\, du = \frac{f}{s+f}[1 - L_S(s+f)]. \end{aligned} \tag{A.5}$$

The last equality follows directly from the properties of Laplace–Stieltjes transform or by application of integration by parts to the integral term. Similarly, the following holds

$$E[e^{-sD}|S > \alpha D]P\{S > \alpha D\} = L_D(s) - \int_0^{\infty} e^{-st} F_S(\alpha t) dF_D(t), \tag{A.6}$$

$$E[e^{-sS}|S < U]P\{S < U\} = L_S(s+f). \tag{A.7}$$

By substituting Eqs. (A.5)–(A.7) into (A.4), the conditional Laplace–Stieltjes transform can be written for all $n = 0, 1, 2, \dots$ as

$$E[e^{-sT}\mathbf{1}\{A_n\}|G^1] = \left\{\frac{f}{s+f}[1 - L_S(s+f)]\right\}^n \cdot \left\{L_D(s) - \int_0^{\infty} e^{-st} F_S(\alpha t) dF_D(t)\right\}^n L_S(s+f). \tag{A.8}$$

Similarly for events E_n, we have for all $n = 0, 1, 2, \ldots$

$$E[e^{-sT}\mathbf{1}\{E_n\}|G^1] = E[e^{-sT}|E_n, G^1] \cdot P\{E_n|G^1\}$$
$$= \left(E[e^{-sU}|S > U]P\{S > U\}\right)^{n+1} \left(E[e^{-sD}|S > \alpha D]P\{S > \alpha D\}\right)^n E[e^{-s\frac{1}{\alpha}S}|S < \alpha D]P\{S < \alpha D\},$$

where the last term can be derived as

$$E[e^{-s\frac{1}{\alpha}S}|S < \alpha D]P\{S < \alpha D\} = E[e^{-s\frac{1}{\alpha}S}\mathbf{1}\{S < \alpha D\}]$$
$$= \int_{t=0}^{\infty}\int_{x=0}^{\alpha t} e^{-\frac{s}{\alpha}x}dF_S(x)dF_D(t) = \int_{x=0}^{\infty}\left\{\int_{t=\frac{x}{\alpha}}^{\infty}dF_D(t)\right\}e^{-\frac{s}{\alpha}x}dF_S(x)$$
$$= \int_{x=0}^{\infty} e^{-\frac{s}{\alpha}x}\left(1 - F_D\left(\frac{x}{\alpha}\right)\right)dF_S(x) = L_S\left(\frac{s}{\alpha}\right) - \int_{x=0}^{\infty} e^{-\frac{s}{\alpha}x}F_D\left(\frac{x}{\alpha}\right)dF_S(x)$$

giving for all $n = 0, 1, 2, \ldots$

$$E[e^{-sT}\mathbf{1}\{E_n\}|G^1] = \left\{\frac{f}{s+f}[1 - L_S(s+f)]\right\}^{n+1} \cdot \left\{L_D(s) - \int_0^{\infty} e^{-st}F_S(\alpha t)dF_D(t)\right\}^n$$
$$\cdot \left[L_S\left(\frac{s}{\alpha}\right) - \int_{x=0}^{\infty} e^{-\frac{s}{\alpha}x}F_D\left(\frac{x}{\alpha}\right)dF_S(x)\right]. \tag{A.9}$$

Then, from (A.8) and (A.9), we obtain the Laplace–Stieltjes transform of the conditional completion time, given that the work starts during up period as shown in (A.10)

$$E[e^{-sT}|G^1] = \sum_{n=0}^{\infty} E\left[e^{-sT}|A_n, G^1\right]P\{A_n|G^1\} + \sum_{n=0}^{\infty} E\left[e^{-sC}|E_n, G^1\right]P\{E_n|G^1\}$$
$$= \frac{1}{1 - V(s)} \cdot \left\{L_S(s+f) + \frac{f}{s+f}\cdot[1 - L_S(s+f)]\cdot\left[L_S\left(\frac{s}{\alpha}\right) - \int_{x=0}^{\infty} e^{-\frac{s}{\alpha}x}F_D\left(\frac{x}{\alpha}\right)dF_S(x)\right]\right\}, \tag{A.10}$$

where

$$V(s) = \frac{f}{s+f}[1 - L_S(s+f)] \cdot \left[L_D(s) - \int_0^{\infty} e^{-st}F_S(\alpha t)dF_D(t)\right].$$

Consider now that work starts during down period, that is, under G^2. Figure A1 shows a sample path of a run for the system when work starts during down period.

Remember that we denote by Y the remaining down time after work starts. Let, again, A_n denote the case that the work on a task starts and ends while the server is down, and there are n up and n down periods with one of them being the remaining down time, and one incomplete down time. Similarly, let E_n denote the case that the work starts while the server is down and ends while the server is up, and there are $(n + 1)$ down periods and n up period with an incomplete up period. Then, the conditional completion time given that work starts during a down period will be

$$\{T|G^2\} = \begin{cases} \frac{1}{\alpha}S_1, & \text{event } A_0 \\ Y + \sum_{i=2}^{n} D_i + \sum_{i=1}^{n} U_i + \frac{1}{\alpha}S_{2n+1}, & \text{event } A_n, \quad n = 1, 2, \ldots \\ Y + \sum_{i=2}^{n+1} D_i + \sum_{i=1}^{n} U_i + S_{2n+2}, & \text{event } E_n, \quad n = 0, 1, 2, \ldots. \end{cases} \tag{A.11}$$
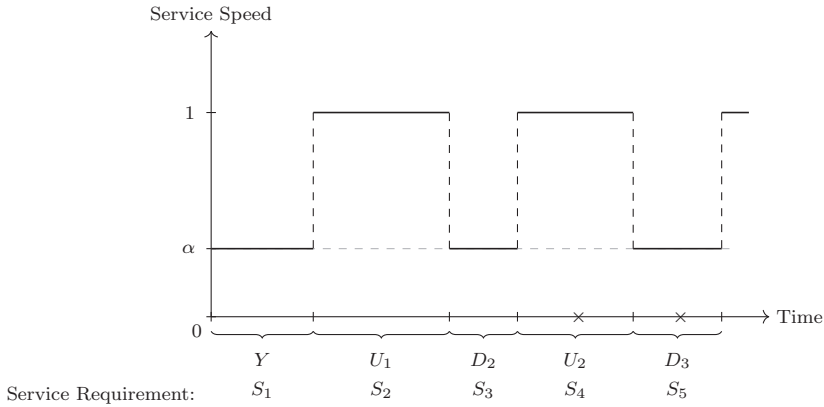
Figure A1. *Sample path of service system under* G^2.

The pdf of the remaining down time Y is

$$f_Y(t) = r[1 - F_D(t)], \qquad t > 0, \tag{A.12}$$

and its Laplace–Stieltjes transform is

$$L_Y(s) = \frac{r}{s}[1 - L_D(s)]. \tag{A.13}$$

We write the Laplace–Stieltjes transform of the conditional completion time under each event as

$$E[e^{-sT}\mathbf{1}\{A_0\}|G^2] = E[e^{-sT}|A_0, G^2] \cdot P\{A_0|G^2\} = L_S(s/\alpha) - \int_0^\infty e^{-\frac{s}{\alpha}t} F_Y(t/\alpha)\, dF_S(t), \tag{A.14}$$

$$E[e^{-sT}\mathbf{1}\{E_0\}|G^2] = E[e^{-sT}|E_0, G^2] \cdot P\{E_0|G^2\} = L_S(s+f) \cdot \left[L_Y(s) - \int_0^\infty e^{-st} F_S(\alpha t) dF_Y(t)\right], \tag{A.15}$$

for all $n = 1, 2, \ldots$

$$\begin{aligned}
E[e^{-sT}\mathbf{1}\{A_n\}|G^2] &= E[e^{-sT}|A_n, G^2] \cdot P\{A_n|G^2\} \\
&= \left[L_Y(s) - \int_0^\infty e^{-st} F_S(\alpha t) dF_Y(t)\right] \cdot \left[L_D(s) - \int_0^\infty e^{-st} F_S(\alpha t) dF_D(t)\right]^{n-1} \cdot \\
&\quad \left[\frac{f}{s+f}(1 - L_S(s+f))\right]^n \left[L_S\left(\frac{s}{\alpha}\right) - \int_{x=0}^\infty e^{-\frac{s}{\alpha}x} F_D\left(\frac{x}{\alpha}\right) dF_S(x)\right], \tag{A.16}
\end{aligned}$$

$$\begin{aligned}
E[e^{-sT}\mathbf{1}\{E_n\}|G^2] &= E[e^{-sT}|E_n, G^2] \cdot P\{E_n|G^2\} \\
&= \left[L_Y(s) - \int_0^\infty e^{-st} F_S(\alpha t) dF_Y(t)\right] \cdot \left[L_D(s) - \int_0^\infty e^{-st} F_S(\alpha t) dF_D(t)\right]^n \cdot \\
&\quad \left[\frac{f}{s+f}(1 - L_S(s+f))\right]^n \cdot L_S(s+f). \tag{A.17}
\end{aligned}$$

Using (A.14)–(A.17), we obtain the Laplace–Stieltjes transform of the conditional completion times as

$$
E[e^{-sT}|G^2] = \frac{1}{1 - V(s)} \cdot \left[L_Y(s) - \int_0^\infty e^{-st} F_S(\alpha t) dF_Y(t) \right] \cdot
$$
$$
\left\{ \frac{f}{s+f} [1 - L_S(s+f)] \cdot \left[L_S\left(\frac{s}{\alpha}\right) - \int_{x=0}^\infty e^{-\frac{s}{\alpha}x} F_D\left(\frac{x}{\alpha}\right) dF_S(x) \right] + L_S(s+f) \right\} +
$$
$$
\left[L_S\left(\frac{s}{\alpha}\right) - \int_0^\infty e^{-\frac{s}{\alpha}t} F_Y\left(\frac{t}{\alpha}\right) dF_S(t) \right]. \tag{A.18}
$$

Finally, by combining the conditional completion times (A.10) and (A.18) using the corresponding probabilities $P\{G^1\}$ and $P\{G^2\}$, the unconditional completion time Laplace–Stieltjes transform is obtained as shown in (2.4)–(2.5).

Appendix B. Proof of Theorem 2.5

We aim to obtain the conditional Laplace–Stieltjes transform of the completion time distribution for a service time requirement of size t, so we set $S \approx t$. From here, the completion time for work starting during an up period is as given in Eq. (2.21) with probabilities

$$
P\{A_n|S \approx t, G^1\} = \begin{cases} \overline{F_X(t)}, & n = 0 \\ F_X(t) \left(F_U(t)F_D(\frac{t}{\alpha}) \right)^{n-1} F_D(\frac{t}{\alpha})\overline{F_U(t)}, & n = 1, 2, \ldots \end{cases} \tag{B.1}
$$

$$
P\{E_n|S \approx t, G^1\} = F_X(t) \left(F_U(t)F_D(\frac{t}{\alpha}) \right)^n \overline{F_D(\frac{t}{\alpha})}, \qquad n = 0, 1, 2, \ldots. \tag{B.2}
$$

Then, we have that the conditional Laplace–Stieltjes transform of the completion time for A_0, under events G^1 and $S = t$ as

$$
E[e^{-sT}\mathbf{1}\{A_0\}|S \approx t, G^1] = e^{-st}\overline{F_X(t)},
$$

and for A_n as

$$
E[e^{-sT}\mathbf{1}\{A_n\}|S \approx t, G^1] = E\left[e^{-sT}|S = t, G^1, A_n \right] P\{A_n|G^1, S = t\}
$$
$$
= E\left[e^{-s(X + \sum_{i=2}^n U_i + \sum_{i=1}^n D_i + t)}|S = t, G^1, A_n \right] F_X(t) \left(F_U(t)F_D(\frac{t}{\alpha}) \right)^{n-1}
$$
$$
\times F_D(\frac{t}{\alpha})\overline{F_U(t)}
$$
$$
= e^{-st}F_X(t)L_X(s|t) \left(F_U(t)L_U(s|t)F_D(\frac{t}{\alpha})L_D(s|\frac{t}{\alpha}) \right)^{n-1}
$$
$$
\times F_D(\frac{t}{\alpha})L_D(s|\frac{t}{\alpha})\overline{F_U(t)}, \forall n = 1, 2, \ldots.
$$

For the union of events A_n, $n = 0, 1, 2, \ldots$,

$$
E\left[e^{-sT}\mathbf{1}\{\cup_{n=0}^\infty A_n\}|S \approx t, G^1 \right] = e^{-st} \left[\overline{F_X(t)} + \frac{F_X(t)L_X(s|t)F_D(\frac{t}{\alpha})L_D(s|\frac{t}{\alpha})\overline{F_U(t)}}{1 - F_U(t)L_U(s|t)F_D(\frac{t}{\alpha})L_D(s|\frac{t}{\alpha})} \right]. \tag{B.3}
$$

Similarly,

$$
E[e^{-sT}\mathbf{1}\{E_n\}|S \approx t, G^1] = e^{-s\frac{t}{\alpha}}F_X(t)L_X(s|t) \left(F_U(t)L_U(s|t)F_D(\frac{t}{\alpha})L_D(s|\frac{t}{\alpha}) \right)^n \overline{F_D(t/\alpha)}
$$

and

$$E\left[e^{-sT}\mathbf{1}\{\cup_{n=0}^{\infty}E_n\}|S\approx t, G^1\right] = \frac{e^{-st/\alpha}F_X(t)L_X(s|t)\overline{F_D(\frac{t}{\alpha})}}{1 - F_U(t)L_U(s|t)F_D(\frac{t}{\alpha})L_D(s|\frac{t}{\alpha})}.\tag{B.4}$$

Summing up (B.3) and (B.4), we obtain the conditional Laplace–Stieltjes transform of the completion time for customers arriving during an up period:

$$L_T\left(s|S\approx t, G^1\right) = e^{-st}\overline{F_X(t)} + \frac{F_X(t)L_X(s|t)\left(e^{-st}F_D(\frac{t}{\alpha})L_D(s|\frac{t}{\alpha})\overline{F_U(t)} + e^{-s\frac{t}{\alpha}}\overline{F_D(\frac{t}{\alpha})}\right)}{1 - F_U(t)L_U(s|t)F_D(\frac{t}{\alpha})L_D(s|\frac{t}{\alpha})}.\tag{B.5}$$

For work starting during a down period, we use the interpretation of events $A_n, E_n, n = 0, 1, \ldots$, as in the proof of Proposition 1 under G^2. The conditional completion time from Eq. (A.11) becomes

$$\{T|S\approx t, G^2\} = \begin{cases} \frac{t}{\alpha}, & \text{event } A_0, \\ Y + \sum_{i=2}^{n}D_i + \sum_{i=1}^{n}U_i + \frac{t}{\alpha}, & \text{event } A_n, \quad n = 0, 1, 2, \ldots \\ Y + \sum_{i=2}^{n+1}D_i + \sum_{i=1}^{n}U_i + t, & \text{event } E_n, \quad n = 0, 1, 2, \ldots \end{cases}\tag{B.6}$$

with probabilities,

$$P\{A_0|S\approx t, G^2\} = \overline{F_Y(\frac{t}{\alpha})},\tag{B.7}$$

$$P\{A_n|S\approx t, G^2\} = \left(F_U(t)F_D(\frac{t}{\alpha})\right)^{(n-1)}F_Y(\frac{t}{\alpha})F_U(t)\overline{F_D(\frac{t}{\alpha})}, \qquad n = 1, 2, \ldots\tag{B.8}$$

$$P\{E_n|S\approx t, G^2\} = \left(F_U(t)F_D(\frac{t}{\alpha})\right)^{n}F_Y(\frac{t}{\alpha})\overline{F_U(t)}, \qquad n = 0, 1, \ldots.\tag{B.9}$$

From here,

$$E[e^{-sT}\mathbf{1}\{A_n\}|S\approx t, G^2]$$
$$= \begin{cases} e^{-st/\alpha}\overline{F_Y(\frac{t}{\alpha})}, & n = 0, \\ e^{-st/\alpha}\left(F_U(t)L_U(s|t)F_D(\frac{t}{\alpha})L_D(s|\frac{t}{\alpha})\right)^{n-1}F_Y(\frac{t}{\alpha})L_Y(s|\frac{t}{\alpha})F_U(t)L_U(s|t)\overline{F_D(\frac{t}{\alpha})}, & n = 1, 2, \ldots \end{cases}$$

and

$$E[e^{-sT}\mathbf{1}\{E_n\}|S\approx t, G^2] = e^{-st}\left(F_U(t)L_U(s|t)F_D(\frac{t}{\alpha})L_D(s|\frac{t}{\alpha})\right)^{n}F_Y(\frac{t}{\alpha})L_Y(s|\frac{t}{\alpha})\overline{F_U(t)}, \quad n = 0, 1, 2, \ldots.$$

Taking the union of events,

$$E\left[e^{-sT}\mathbf{1}\{\cup_{n=0}^{\infty}A_n\}|S\approx t, G^2\right] = e^{-s\frac{t}{\alpha}}\left[\overline{F_Y(\frac{t}{\alpha})} + \frac{F_Y(\frac{t}{\alpha})L_Y(s|\frac{t}{\alpha})F_U(t)L_U(s|t)\overline{F_D(\frac{t}{\alpha})}}{1 - F_U(t)L_U(s|t)F_D(\frac{t}{\alpha})L_D(s|\frac{t}{\alpha})}\right]$$

and

$$E\left[e^{-sT}\mathbf{1}\{\cup_{n=0}^{\infty}E_n\}|S\approx t, G^2\right] = \frac{e^{-st}F_Y(\frac{t}{\alpha})L_Y(s|\frac{t}{\alpha})\overline{F_U(t)}}{1 - F_U(t)L_U(s|t)F_D(\frac{t}{\alpha})L_D(s|\frac{t}{\alpha})},$$

and combining all events under G^2,

$$L_T\left(s|S\approx t, G^2\right) = e^{-st\frac{t}{\alpha}}\overline{F_Y(\frac{t}{\alpha})} + \frac{F_Y(\frac{t}{\alpha})L_Y(s|\frac{t}{\alpha})\left(e^{-st}\overline{F_U(t)} + e^{-s\frac{t}{\alpha}}F_U(t)L_U(s|t)\overline{F_D(\frac{t}{\alpha})}\right)}{1 - F_U(t)L_U(s|t)F_D(\frac{t}{\alpha})L_D(s|\frac{t}{\alpha})}. \quad \text{(B.10)}$$

Finally combining Laplace–Stieltjes transform for events G^1 and G^2 from (B.5) and (B.10), and considering the probabilities in (2.1), we obtain

$$\begin{aligned}
L_T\left(s|S\approx t\right) = \ &\frac{r}{(r+f)}\Big[e^{-st}\overline{F_X(t)} + \frac{F_X(t)L_X(s|t)\left(e^{-st}F_D(\frac{t}{\alpha})L_D(s|\frac{t}{\alpha})\overline{F_U(t)} + e^{-st/\alpha}\overline{F_D(\frac{t}{\alpha})}\right)}{1 - F_U(t)L_U(s|t)F_D(\frac{t}{\alpha})L_D(s|\frac{t}{\alpha})}\Big] \\
&+ \frac{f}{(r+f)}\Big[e^{-s\frac{t}{\alpha}}\overline{F_Y(\frac{t}{\alpha})} + \frac{F_Y(\frac{t}{\alpha})L_Y(s|\frac{t}{\alpha})\left(e^{-st}\overline{F_U(t)} + e^{-s\frac{t}{\alpha}}F_U(t)L_U(s|t)\overline{F_D(\frac{t}{\alpha})}\right)}{1 - F_U(t)L_U(s|t)F_D(\frac{t}{\alpha})L_D(s|\frac{t}{\alpha})}\Big],
\end{aligned}$$

gives the result in Eq. (2.23).

Appendix C. Proof of Theorem 2.8

As in Theorem 2.5, we condition the service time to a task length $S = t$. Recalling the events in Eq. (2.21), and the probabilities in Eqs. (B.1)–(B.2), it can be shown that

$$\begin{aligned}
E[T\cdot\mathbf{1}\{A_n\}|S\approx t, G^1] &= E[T|A_n, S\approx t, G^1]P\{A_n|G^1\} \\
&= \begin{cases} t\overline{F_X(t)}, & n = 0 \\ \big[t + E[X|t] + E[D|\frac{t}{\alpha}] + (n-1)\big(E[U|t] + E[D|\frac{t}{\alpha}]\big)\big]F_X(t)\big(F_U(t)F_D(\frac{t}{\alpha})\big)^{n-1}F_D(\frac{t}{\alpha})\overline{F_U(t)}, & n = 1, 2, \dots \end{cases} \\
E[T\cdot\mathbf{1}\{E_n\}|S\approx t, G^1] &= \big[\frac{t}{\alpha} + E[X|t] + n\big(E[U|t] + E[D|\frac{t}{\alpha}]\big)\big]F_X(t)\big(F_U(t)F_D(\frac{t}{\alpha})\big)^{n}\overline{F_D(\frac{t}{\alpha})}, \qquad n = 0, 1, 2, \dots.
\end{aligned}$$

From here,

$$\begin{aligned}
E[T\cdot\mathbf{1}\{\cup_{i=0}^{\infty}A_n\}|S\approx t, G^1] = \ &\frac{\big[t + E[X|t] + E[D|\frac{t}{\alpha}]\big]F_X(t)F_D(\frac{t}{\alpha})\overline{F_U(t)}}{1 - F_U(t)F_D(\frac{t}{\alpha})} \\
&+ \frac{\big(E[U|t] + E[D|\frac{t}{\alpha}]\big)F_X(t)F_U(t)\big(F_D(\frac{t}{\alpha})\big)^2\overline{F_U(t)}}{(1 - F_U(t)F_D(\frac{t}{\alpha}))^2}, \qquad \text{(C.1)}
\end{aligned}$$

$$\begin{aligned}
E[T\cdot\mathbf{1}\{\cup_{i=0}^{\infty}E_n\}|S\approx t, G^1] = \ &\frac{\big(\frac{t}{\alpha} + E[X|t]\big)F_X(t)\overline{F_D(\frac{t}{\alpha})}}{1 - F_U(t)F_D(\frac{t}{\alpha})} \\
&+ \frac{\big(E[U|t] + E[D|\frac{t}{\alpha}]\big)F_X(t)F_D(\frac{t}{\alpha})F_U(t)\overline{F_D(\frac{t}{\alpha})}}{(1 - F_U(t)F_D(\frac{t}{\alpha}))^2}. \qquad \text{(C.2)}
\end{aligned}$$

Summing up (C.1) and (C.2) and simplifying we obtain

$$\begin{aligned}
E[T|S\approx t, G^1] = \ &t\overline{F_X(t)} + F_X(t)\Big[E[X|t] \\
&+ \frac{\big(tF_D(\frac{t}{\alpha})\overline{F_U(t)} + \frac{t}{\alpha}\overline{F_D(\frac{t}{\alpha})}\big) + F_D(\frac{t}{\alpha})\big(E[U|t]F_U(t) + E[D|\frac{t}{\alpha}]\big)}{1 - F_U(t)F_D(\frac{t}{\alpha})}\Big]. \qquad \text{(C.3)}
\end{aligned}$$

Consider now the events from Eq. (2.22), and the probabilities in (B.8)–(B.9). Then it holds that

$$E[T \cdot \mathbf{1}\{A_0\}|S \approx t, G^2] = \frac{t}{\alpha}\overline{F_Y(\frac{t}{\alpha})},$$

$$E[T \cdot \mathbf{1}\{A_n\}|S \approx t, G^2] = \left[\frac{t}{\alpha} + E[Y|\frac{t}{\alpha}] + (n-1)\left(E[U|U<t] + E[D|\frac{t}{\alpha}]\right)\right.$$
$$\left. + E[U|t]\right]F_Y(\frac{t}{\alpha})F_U(t)\left(F_U(t)F_D(\frac{t}{\alpha})\right)^{n-1}\overline{F_D(\frac{t}{\alpha})},$$

$$E[T \cdot \mathbf{1}\{E_n\}|S \approx t, G^2] = \left[t + E[Y|\frac{t}{\alpha}] + n\left(E[U|t] + E[D|\frac{t}{\alpha}]\right)\right]F_Y(\frac{t}{\alpha})\left(F_U(t)F_D(\frac{t}{\alpha})\right)^n\overline{F_U(t)}.$$

Taking the union of events, we get

$$E[T \cdot \mathbf{1}\{\cup_{i=0}^{\infty}A_n\}|S \approx t, G^2] = \frac{t}{\alpha}\overline{F_Y(\frac{t}{\alpha})} + \frac{\frac{t}{\alpha}F_Y(\frac{t}{\alpha})F_U(t)\overline{F_D(\frac{t}{\alpha})}}{1 - F_U(t)F_D(\frac{t}{\alpha})}$$
$$+ \frac{\left(E[Y|\frac{t}{\alpha}] + E[U|t]\right)F_D(\frac{t}{\alpha})F_U(t)\overline{F_D(\frac{t}{\alpha})}}{\left(1 - F_U(t)F_D(\frac{t}{\alpha})\right)}$$
$$+ \frac{\left(E[U|t] + E[D|\frac{t}{\alpha}]\right)F_Y(\frac{t}{\alpha})(F_U(t))^2 F_D(\frac{t}{\alpha})\overline{F_D(\frac{t}{\alpha})}}{\left(1 - F_U(t)F_D(\frac{t}{\alpha})\right)^2}, \qquad (\text{C.4})$$

$$E[T \cdot \mathbf{1}\{\cup_{i=0}^{\infty}E_n\}|S \approx t, G^2] = \frac{\left(t + E[Y|\frac{t}{\alpha}]\right)F_Y(\frac{t}{\alpha})\overline{F_U(t)}}{1 - F_U(t)F_D(\frac{t}{\alpha})}$$
$$+ \frac{\left(E[U|t] + E[D|\frac{t}{\alpha}]\right)F_Y(\frac{t}{\alpha})F_U(t)F_D(\frac{t}{\alpha})\overline{F_U(t)}}{\left(1 - F_U(t)F_D(\frac{t}{\alpha})\right)^2} \qquad (\text{C.5})$$

and summing up (C.4) and (C.5) we obtain,

$$E[T|S \approx t, G^2] = \frac{t}{\alpha}\overline{F_Y(\frac{t}{\alpha})} + F_Y(\frac{t}{\alpha})\left[E[Y|\frac{t}{\alpha}]\right.$$
$$\left. + \frac{\left(\frac{t}{\alpha}F_U(t)\overline{F_D(\frac{t}{\alpha})} + t\overline{F_U(t)}\right) + F_U(t)\left(E[U|t] + E[D|\frac{t}{\alpha}]F_D(\frac{t}{\alpha})\right)}{1 - F_U(t)F_D(\frac{t}{\alpha})}\right]. \qquad (\text{C.6})$$

Finally, combining (C.3) and (C.6) with the probabilities (2.1), we obtain the result in (2.25).

Appendix D. Proof of Theorem 3.1

Proof. As in Fiorini *et al.* [16], we calculate the moments of the completion time distribution to assess the heaviness of the tail. We assume here that the task size has an exponential tail, which amounts to $\theta_{\min}(S) = \mu$ from Eq. (3.3), and denote $\Delta = \min\{f, r/\alpha\} > 0$ and $\Delta^+ = f + r/\alpha - \Delta = \max\{f, r/\alpha\} > 0$. Let U and D be exponentially distributed with failure rate f and repair rate r, respectively. Their distributions are $F_U(t) = 1 - e^{-ft}$, and $F_D(t) = 1 - e^{-rt}$.

Notice that the task-size-conditioned Laplace–Stieltjes transform can be expressed as in (D.1)–(D.2)

$$L_U(s|t) = E[e^{-sU}|t > U] = \int_{x=0}^{t} e^{-sx}dF_U(x|t) = \int_{x=0}^{t} e^{-sx}\frac{fe^{-fx}}{F_U(t)}dx = \frac{f}{s+f}\frac{1 - e^{-(s+f)t}}{1 - e^{-ft}}, \qquad (\text{D.1})$$

$$L_D(s|t/\alpha) = E[e^{-sD}|t > D] = \int_{x=0}^{t/\alpha} e^{-sx} dF_D(x|t/\alpha) = \int_{x=0}^{t/\alpha} e^{-sx} \frac{re^{-rx}}{F_D(t/\alpha)} dx = \frac{r}{s+r} \frac{1 - e^{-(s+r)t/\alpha}}{1 - e^{-rt/\alpha}}.$$
(D.2)

From Theorem 2.5, the conditional Laplace–Stieltjes transform of the completion time in Eq. (2.23) becomes

$$L_T(s|t) = \frac{re^{-(s+f)t}\left[1 + \frac{f}{s+r}(1 - e^{-(s+r)t/\alpha})\right] + fe^{-(s+r)t/\alpha}\left[1 + \frac{r}{s+f}(1 - e^{-(s+f)t})\right]}{(r+f)\left[1 - \frac{rf}{(s+r)(s+f)}(1 - e^{-(s+f)t})(1 - e^{-(s+r)t/\alpha})\right]},$$

where we call

$$h_r(s,t) := \frac{f}{s+r}(1 - e^{-(s+r)t/\alpha}),$$
$$h_f(s,t) := \frac{r}{s+f}(1 - e^{-(s+f)t}),$$

obtaining

$$L_T(s|t) = \frac{1}{r+f} \cdot \frac{re^{-(s+f)t}(1 + h_r(s,t)) + fe^{-(s+r)t/\alpha}(1 + h_f(s,t))}{1 - h_r(s,t)h_f(s,t)}.$$
(D.3)

To calculate the mth moment $E[T^m]$ of the completion time T, we resort to the derivatives of the Laplace–Stieltjes transform of the completion time distribution evaluated at $s = 0$ as shown in Eq. (D.4)

$$E[T^m] = (-1)^m \frac{d^m}{ds^m} L_T(s)\Big|_{s=0}.$$
(D.4)

In particular, however, we calculate the mth partial derivative with respect to s of the *conditioned* Laplace–Stieltjes transform for the completion time in (D.3), evaluate at $s = 0$, and uncondition with respect to the task size whenever is possible.

Consider the following function definitions and notation for the partial derivatives of the conditional Laplace Stieljest Transform (LSTm):

(1) Main term for the nth derivative:

$$L_T(s;n|t) := \frac{1}{r+f} \cdot \frac{re^{-(s+f)t}(1 + h_r(s,t)) + \frac{f}{\alpha^n}e^{-(s+r)t/\alpha}(1 + h_f(s,t))}{1 - h_r(s,t)h_f(s,t)}, \quad n = 0, 1, 2, \ldots.$$

Notice $L_T(s;0|t) := L_T(s|t)$.

(2) First recurrent factor of higher-order term for nth derivative:

$$v_n(s,t) := \frac{re^{-(s+f)t} + \frac{f}{\alpha^{n-1}}e^{-(s+r)t/\alpha}}{r+f}, \quad n = 1, 2, \ldots.$$
(D.5)

It holds that $\frac{\partial^n}{\partial s^n} v_1(s,t) = (-1)^n t^n v_{n+1}(s,t)$.

(3) Recurrent rational exponential functions and terms:

$$\theta_i(s, t; a, b) := \frac{e^{-i(s+a)t/b}}{1 - e^{-i(s+a)t/b}}. \tag{D.6}$$

It can be shown that for $n \geq 1$,

$$\frac{\partial^n}{\partial s^n} \theta_1(s, t; a, b) = (-1)^n \left(\frac{t}{b}\right)^n \sum_{i=1}^{n+1} W(n, i) \frac{e^{-i(s+a)t/b}}{(1 - e^{-(s+a)t/b})^i}$$

$$:= (-1)^n \left(\frac{t}{b}\right)^n \sum_{i=1}^{n+1} W(n, i) \theta_i(s, t; a, b), \tag{D.7}$$

where $W(n, i)$ denotes the nth row and ith column's Worpitzky number from the Worpitzky triangle [35], as defined in (D.8)

$$W(n, i) = \frac{1}{i} \sum_{j=0}^{i} (-1)^{i-j} \binom{i}{j} j^n, \quad n \geq 1, \quad 1 \leq i \leq n. \tag{D.8}$$

In order to simplify the formulas, we also give the following notation:

$$\Theta_1(s, t; a, b) := \frac{t}{b} \cdot \frac{e^{-(s+a)t/b}}{1 - e^{-(s+a)t/b}} - \frac{1}{s + a}$$

$$= \frac{t}{b} \theta_1(s, t; a, b) - \frac{1}{s + a}. \tag{D.9}$$

From here, similarly, for $n \geq 0$,

$$\frac{\partial^n}{\partial s^n} \Theta_1(s, t; a, b) = (-1)^n \left(\frac{t^{n+1}}{b^{n+1}} \sum_{i=1}^{n+1} W(n, i) \theta_i(s, t; a, b) - \frac{n!}{(s + a)^{n+1}}\right)$$

$$:= \Theta_{n+1}(s, t; a, b). \tag{D.10}$$

Finally, we call

$$\omega_n(s, t) := \Theta_n(s, t; f, 1) + \Theta_n(s, t; r, \alpha), \quad n = 1, 2, \dots \tag{D.11}$$

and by definition, $\frac{\partial^n}{\partial s^n} \omega_1(s, t) = \omega_{n+1}(s, t)$.

Now we calculate de first derivative of the conditioned Laplace–Stieltjes transform in Eq. (D.3). Henceforth, functional dependencies of $h_r(s, t)$ and $h_f(s, t)$ may be omitted for cleaner notation

$$\frac{\partial}{\partial s} L_T(s|t) = \frac{\partial}{\partial s} L_T(s;0|t)$$

$$= -t L_T(s;1|t) + \nu_1(s,t) \frac{\frac{\partial}{\partial s}(h_r h_f)}{(1 - h_r h_f)^2}$$

$$+ \frac{s+f}{r+f} \theta_1(s,t;f,1) \Big(\Theta_1(s,t;r,\alpha) + h_r h_f \Theta_1(s,t;f,1) \Big) \frac{h_r h_f}{(1 - h_r h_f)^2}$$

$$+ \frac{s+r}{r+f} \theta_1(s,t;r,\alpha) \Big(\Theta_1(s,t;f,1) + h_r h_f \Theta_1(s,t;r,\alpha) \Big) \frac{h_r h_f}{(1 - h_r h_f)^2}$$

$$:= -t L_T(s;1|t) + \nu_1(s,t) \frac{\frac{\partial}{\partial s}(h_r h_f)}{(1 - h_r h_f)^2}$$

$$+ \left[\frac{s+f}{r+f} \theta_1(s,t;f,1) A(s,t) + \frac{s+r}{r+f} \theta_1(s,t;r,\alpha) B(s,t) \right] \frac{h_r h_f}{(1 - h_r h_f)^2}, \tag{D.12}$$

with $A(s,t) = \Theta_1(s,t;r,\alpha) + h_r h_f \Theta_1(s,t;f,1)$ and $B(s,t) = \Theta_1(s,t;f,1) + h_r h_f \Theta_1(s,t;r,\alpha)$.

It can be shown that

$$\frac{\partial}{\partial s}(h_r h_f) = h_r h_f \omega_1(s,t). \tag{D.13}$$

Then, calling

$$u_1(s,t) := \frac{h_r h_f}{(1 - h_r h_f)^2}, \tag{D.14}$$

we have from (D.12),

$$\frac{\partial}{\partial s} L_T(s|t) = -t L_T(s;1|t) + u_1(s,t) \left[\nu_1(s,t)\omega_1(s,t) + \frac{s+f}{r+f} \theta_1(s,t;f,1) A(s,t) \right.$$

$$\left. + \frac{s+r}{r+f} \theta_1(s,t;r,\alpha) B(s,t) \right]. \tag{D.15}$$

Moreover,

$$\frac{\partial}{\partial s} L_T(s;n|t) = -t L_T(s;n+1|t) + u_1(s,t) \left[\nu_{n+1}(s,t)\omega_1(s,t) + \frac{s+f}{r+f} \theta_1(s,t;f,1) A(s,t) \right.$$

$$\left. + \frac{s+r}{r+f} \theta_1(s,t;r,\alpha) B(s,t) \right]. \tag{D.16}$$

To calculate the higher order partial derivatives of $L_T(s|t)$ we have to use the general Leibniz rule on the second term in (D.15). For this, we have to study the partial derivatives of the factor $u(s,t)$, and the functions $A(s,t)$ and $B(s,t)$. We start with $u(s,t)$. From (D.13), we observe that

$$h_r(s,t)h_f(s,t) = C(t)e^{\int \omega_1(s,t)ds}, \tag{D.17}$$

where $C(t) := rf\, e^{-t(f+r/\alpha)}$.

Consider now t fixed and the two following one-dimensional functions:

$$g(x(s)) := C(t)e^{\int x(s)ds}, \quad f(y) := \frac{y}{(1-y)^2}. \tag{D.18}$$

Using the functions defined above Eqs. (D.14) and (D.17) can be rewritten as

$$h_r(s,t)h_f(s,t) = g(w_1(s,t)), \tag{D.19}$$

$$u_1(s,t) = f(g(w_1(s,t))), \tag{D.20}$$

thus, $u_1(s,t)$ is a nested composite function.

The derivatives of composite functions can be written using Faà di Bruno's formula [15] as shown in Eq. (D.21):

$$\frac{d^n}{dx^n}f(g(x)) = \sum_{k=1}^{n} f^{(k)}(g(x)) \cdot B_{n,k}\Big(g'(x), g''(x), \ldots, g^{(n-k+1)}(x)\Big), \tag{D.21}$$

where $B_{n,k}$ are the incomplete (or partial) exponential Bell polynomials [9, 26] defined below

$$B_{n,k}(x_1, x_2, \ldots, x_{n-k+1}) = \sum \frac{n!}{j_1! j_2! \cdots j_{n-k+1}!} \left(\frac{x_1}{1!}\right)^{j_1} \left(\frac{x_2}{2!}\right)^{j_2} \cdots \left(\frac{x_{n-k+1}}{(n-k+1)!}\right)^{j_{n-k+1}}, \tag{D.22}$$

where the summation is carried over all sequences $j_1, j_2, j_3, \ldots, j_{n-k+1}$ of non-negative integers such that conditions below are satisfied

$$j_1 + j_2 + \cdots + j_{n-k+1} = k, \tag{D.23}$$

$$j_1 + 2j_2 + 3j_3 + \cdots + (n-k+1)j_{n-k+1} = n. \tag{D.24}$$

The nth complete exponential Bell polynomial is given by the sum of the incomplete polynomials as:

$$B_n(x_1, \ldots, x_n) = \sum_{k=1}^{n} B_{n,k}(x_1, x_2, \ldots, x_{n-k+1}). \tag{D.25}$$

The derivatives of composite functions in Eq. (D.21) are similar to the complete exponential Bell polynomials, of which the variables $x_1, \ldots, x_n$ are replaced by the successive derivatives of the inner function $g(x)$, and the coefficients of the polynomials are factored by the derivatives of the outer function. Since $u(s,t) = f(g(w_1(s,t)))$ is a double composite function, its partial derivatives are given by nested Bell polynomials.

From (D.11), we know that the partial derivatives of the innermost function $\omega_1(s,t)$ with respect to s are

$$\frac{\partial^n}{\partial s^n}\omega_1(s,t) = \omega_{n+1}(s,t). \tag{D.26}$$

The partial derivatives of the first composite function $g(w_1(s,t))$ are calculated using (D.21) where the variable coefficients are the derivatives of $g(x(s)) = C(t)e^{\int x(s)ds}$ evaluated in $\omega_1(s,t)$, and the incomplete Bell polynomials are evaluated on the partial derivatives of $\omega_1(s,t)$ from Eq. (D.26). The general derivative is given below, we omit the functional dependency on $\omega_i(s,t)$ to keep the notation uncluttered

$$G_n(s,t) := \frac{\partial^n}{\partial s^n}\big(h_r(s,t)h_f(s,t)\big) = h_r(s,t)h_f(s,t)B_n(\omega_1, \omega_2, \ldots, \omega_n).$$

The similarities of the derivatives to the complete Bell polynomials are due to the exponential nature of $g(x(s))$, which makes the successive derivatives of the outer function self-similar, and hence it can

be factored out from the coefficients of the partial polynomials. Notice that derivatives of low order for $g(x(s))$ may also be easily calculated recursively following from Eq. (D.13).

The derivatives of the outermost function $f(y)$ are

$$\frac{d^n}{dy^n}f(y) = \frac{d^n}{dy^n}\frac{y}{(1-y)^2} = \frac{n!}{(1-y)^2}\frac{y+n}{(1-y)^n}. \tag{D.27}$$

With this we give below the general form of the partial derivatives of $u(s,t)$ where we drop the functional dependencies of $G_n(s,t)$ to keep the notation uncluttered

$$u_{n+1}(s,t) := \frac{\partial^n}{\partial s^n}u_1(s,t)$$
$$= \sum_{k=1}^{n} k!\frac{(k+h_rh_f)}{(1-h_rh_f)^{k+2}} \cdot B_{n,k}\Big(G_1, G_2, \ldots, G_{n-k+1}\Big), \quad n = 1, 2, \ldots. \tag{D.28}$$

The derivatives of $A(s,t)$ and $B(s,t)$ are also written in terms of Bell polynomials. This is due to the fact that, from Eq. (D.17) and by definition both functions are composite functions:

$$A(s,t) = \Theta_1(s,t;r,\alpha) + C(t)e^{\int(\Theta_1(s,t;f,1)+\Theta_1(s,t;r,\alpha))ds}\Theta_1(s,t;f,1),$$
$$B(s,t) = \Theta_1(s,t;f,1) + C(t)e^{\int(\Theta_1(s,t;f,1)+\Theta_1(s,t;r,\alpha))ds}\Theta_1(s,t;r,\alpha).$$

Denote by $\Theta_{1:k}(s,t;a,b)$ the ordered list $\Theta_1(s,t;a,b), \ldots, \Theta_k(s,t;a,b)$. It can be shown by induction that

$$\frac{\partial^n}{\partial s^n}A(s,t) = \Theta_{n+1}(s,t;r,\alpha) + h_r(s,t)h_f(s,t)\sum_{k=0}^{n}\binom{n}{k}B_{n-k+1}\big(\Theta_{1:(n-k+1)}(s,t;f,1)\big)B_k\big(\Theta_{1:k}(s,t;r,\alpha)\big),$$
$$\tag{D.29}$$

$$\frac{\partial^n}{\partial s^n}B(s,t) = \Theta_{n+1}(s,t;f,1) + h_r(s,t)h_f(s,t)\sum_{k=0}^{n}\binom{n}{k}B_{n-k+1}\big(\Theta_{1:(n-k+1)}(s,t;r,\alpha)\big)B_k\big(\Theta_{1:k}(s,t;f,1)\big),$$
$$\tag{D.30}$$

where the summations are binomial expansions on the indices of $\Theta_k(s,t;a,b)$, with the first index shifted to the right by one. Getting back to Eq. (D.15), we had that

$$\frac{\partial}{\partial s}L_T(s|t) = -tL_T(s;1|t) + u_1(s,t)\left[v_1(s,t)\omega_1(s,t) + \frac{s+f}{r+f}\theta_1(s,t;f,1)A(s,t)\right.$$
$$\left. + \frac{s+r}{r+f}\theta_1(s,t;r,\alpha)B(s,t)\right]. \tag{D.31}$$

It can be shown from successive differentiation of Eq. (D.31) and considering (D.16) that the nth partial derivative of the conditioned Laplace transform of the completion time is given as:

$$
\frac{\partial^n}{\partial s^n} L_T(s|t) = (-t)^n L_T(s; n|t)
$$

$$
+ \sum_{k=1}^{n} (-t)^{n-k} \frac{\partial^{k-1}}{\partial s^{k-1}} \left(u_1(s,t) \left[v_{n-k+1}(s,t) \omega_1(s,t) \right. \right.
$$

$$
\left. \left. + A(s,t) \left(\tfrac{s+f}{r+f} \theta_1(s,t;f,1) \right) + B(s,t) \left(\tfrac{s+r}{r+f} \theta_1(s,t;r,\alpha) \right) \right] \right), \tag{D.32}
$$

where the partial derivatives in the right hand side are calculated with the General Leibniz rule, which is given in (D.33) for the product of three functions $f_1(x), f_2(x), f_3(x)$:

$$
\frac{d^n}{dx^n} \left(f_1(x) f_2(x) f_3(x) \right) = \sum_{k_1 + k_2 + k_3 = n} \frac{n!}{k_1! \, k_2! \, k_3!} \frac{d^{k_1}}{dx^{k_1}} f_1(x) \frac{d^{k_2}}{dx^{k_2}} f_2(x) \frac{d^{k_3}}{dx^{k_3}} f_3(x), \tag{D.33}
$$

with the summation extending over all triplets (k_1, k_2, k_3) of non-negative integers that are such that $\sum_{t=1}^{3} k_t = n$.

Finally, the mth moment of the completion time T is obtained by unconditioning (D.32) with respect to the task size t whenever the improper integral converges

$$
E[T^m] = (-1)^m \int_0^{\infty} \left(\tfrac{\partial^m}{\partial s^m} L_T(s|t) \big|_{s=0} \right) dF_S(t)
$$

$$
= (-1)^m \mu \int_0^{\infty} \left(\tfrac{\partial^m}{\partial s^m} L_T(s|t) \big|_{s=0} \right) e^{-\mu t} dt. \tag{D.34}
$$

To study the convergence of the integral in (D.34) we look at the asymptotic behavior of the different terms and factors in (D.32) evaluated at $s = 0$. We will use the definition of asymptotic equivalence given in (D.35). Recall also that we denote $\Delta = \min\{f, r/\alpha\} > 0$ and $\Delta^+ = f + r/\alpha - \Delta = \max\{f, r/\alpha\} > 0$

$$
\psi_1(t) \sim \psi_2(t) \text{ if and only if } \lim_{t \to \infty} \frac{\psi_1(t)}{\psi_2(t)} = 1. \tag{D.35}
$$

We classify according to the asymptotic order.

(1) Constant factors:
    (a) Main factor of the first term: $L_T(0; n|t)$

$$
L_T(0; n|t) \sim C_1(n) := \begin{cases} 1, & f < r/\alpha, \\ 1/\alpha^n, & f > r/\alpha, \\ \tfrac{1}{2}(1 + 1/\alpha^n), & f = r/\alpha \end{cases} \quad n = 0, 1, 2, \ldots.
$$

(b) Recurrent factor $\Theta_n(0, t; a, b)$:

$$\Theta_1(0, t; a, b) = \frac{t}{b} \frac{e^{-at/b}}{1 - e^{-at/b}} - 1/a \sim -1/a,$$

$$\Theta_n(0, t; a, b) = (-1)^{n-1} \left(\frac{t^n}{b^n} \sum_{i=1}^{n} W(n, i) \frac{e^{-i \cdot at/b}}{(1 - e^{-at/b})^i} - \frac{(n-1)!}{a^n} \right)$$

$$\sim (-1)^n \frac{(n-1)!}{a^n}, \quad n = 2, 3, \ldots.$$

(c) Recurrent factor $\omega_n(0, t)$:

$$\omega_n(0, t) = \Theta_n(0, t; f, 1) + \Theta_n(0, t; r, \alpha) \sim (-1)^n (n-1)! \left(\frac{1}{f^n} + \frac{1}{r^n} \right), \quad n = 1, 2, \ldots.$$

(d) Recurrent factors $A(s, t)$ and $B(s, t)$, and derivatives: from (D.29)–(D.30), and by the continuity of Bell polynomials, we have that for every $n \geq 0$ there exist constants $C_A(n) > 0$ and $C_B(n) > 0$ such that

$$\frac{\partial^n}{\partial s^n} A(s, t) \sim (-1)^{n+1} C_A(n),$$

$$\frac{\partial^n}{\partial s^n} B(s, t) \sim (-1)^{n+1} C_B(n).$$

(2) Polynomial factors with exponential damping:

(a) Recurrent factor $v_n(0|t)$:

$$v_n(0|t) \sim e^{-\Delta t} \cdot C_v(n) := e^{-\Delta t} \cdot \begin{cases} \frac{r}{r+f}, & f < r/\alpha, \\ \frac{f}{(r+f)} \frac{1}{\alpha^{n-1}}, & f > r/\alpha, \\ \frac{1}{r+f} \left(r + \frac{f}{\alpha^{n-1}} \right), & f = r/\alpha \end{cases} \quad n = 1, 2, \ldots.$$

(b) Recurrent factors $\theta_k(0, t; a, b)$ and derivatives of $\theta_1(0, t; a, b)$:

$$\theta_k(s, t; a, b) \sim e^{-k \cdot at/b},$$

$$\frac{\partial^n}{\partial s^n} \theta_1(s, t; a, b) \Big|_{s=0} = (-1)^n \left(\frac{t}{b} \right)^n \sum_{i=1}^{n+1} W(n, i) \theta_i(0, t; a, b) \sim (-1)^n \left(\frac{t}{b} \right)^n e^{-at/b}, \quad n = 1, 2, \ldots.$$

(c) Accompanying factors of $A(0, t)$ and $B(0, t)$, and derivatives:

$$\frac{\partial^n}{\partial s^n} \left(\frac{s+a}{r+f} \theta_1(s, t; a, b) \right) \Big|_{s=0} \sim \frac{1}{r+f} \left[n + a(-1)^{n-1} \frac{t^{n-1}}{b^{n-1}} \right] e^{-at/b}, \quad n = 1, 2, \ldots.$$

(3) Exponential growth factors: recurrent factor $u_n(0, t)$.

Again due to the continuity of the Bell polynomials, for every $n \in \mathbb{N}$ there exists a constant $C_G(n) > 0$ such that $G_n(0, t) \sim (-1)^n C_G(n)$, and since the dominant term in (D.28) is for $k = n$, there exists a constant $C_B(n-1) > 0$ such that $B_{n-1,n-1}(G_1)\big|_{s=0} \sim (-1)^{n-1} C_B(n-1)$. From here, given that

$$\frac{\left(n + h_r(0,t)h_f(0,t)\right)}{\left(1 - h_r(0,t)h_f(0,t)\right)^{n+2}} \sim (n+1)e^{(n+2)\Delta t},$$

we obtain

$$u_n(0,t) \sim (-1)^{n-1}n! \, C_B(n-1) \cdot e^{(n+1)\Delta t}, \quad n = 1, 2, \dots.$$

Notice that $u_n(0,t)$ is the only factor that has an asymptotic behavior that is not decaying nor constant for large t, and also that the asymptotic order increases with n, which is the order of the partial derivative.

It is easy to see from the first derivative in Eq. (D.15) that $\frac{\partial}{\partial s}L_T(s|t)\big|_{s=0} \sim -Ce^{\Delta t}$, for some constant $C > 0$, and that the expected completion time $E[T]$ only exists when $\Delta = \min\{f, r/\alpha\} < \mu$. For higher order derivatives in Eq. (D.32), we only have to observe the last term in the summation, which will yield the higher order terms overall. Additionally, since the partial derivatives of $u_1(s,t)$ dominate the asymptotic behavior, we only have to look at the term coming from the Leibniz rule for which $u_1(s,t)$ has the partial derivative of order $n-1$ and the remaining factors are not differentiated. Then, for $n \geq 1$,

$$\begin{aligned}
\frac{\partial^n}{\partial s^n}L_T(s|t)\bigg|_{s=0} &\sim u_n(0,t)\left[\nu_1(0,t)\omega_1(0,t) + A(0,t)\left(\tfrac{f}{r+f}\,\theta_1(0,t;f,1)\right)\right. \\
&\left. + B(0,t)\left(\tfrac{r}{r+f}\,\theta_1(0,t;r,\alpha)\right)\right] \sim \mathcal{C}(n)e^{n\Delta t},
\end{aligned} \tag{D.36}$$

where $\mathcal{C}(n)$ is a constant that depends on n. Then, from (D.34), the mth moment $E[T^m]$ is not finite whenever $m\Delta - \mu \geq 0$ which is $m \geq \mu/\Delta = \mu/\min\{f, r/\alpha\} = \varepsilon$, and the completion time distribution is power-tailed [19], that is, there exists a positive constant c such that

$$\overline{F_T(t)} \sim \frac{c}{t^\varepsilon}. \tag{D.37}$$

$\square$