



HAL
open science

Covert Cycle Stealing in a Single FIFO Server

Bo Jiang, Philippe Nain, Don Towsley

► **To cite this version:**

Bo Jiang, Philippe Nain, Don Towsley. Covert Cycle Stealing in a Single FIFO Server. ACM Transactions on Modeling and Performance Evaluation of Computing Systems, ACM, 2021, 6 (2), pp.1-33. 10.1145/3462774 . hal-03216762

HAL Id: hal-03216762

<https://hal.inria.fr/hal-03216762>

Submitted on 4 May 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Covert Cycle Stealing in a Single FIFO Server

Bo Jiang ^{*,1}, Philippe Nain ^{†,2}, and Don Towsley ^{‡,3}

¹John Hopcroft Center for Computer Science, Shanghai Jiao Tong University, Shanghai, China

²Inria, Ecole Normale Supérieure de Lyon, LIP, 46 allée d'Italie, Lyon, 69364, France

³College of Information and Computer Sciences, University of Massachusetts, Amherst, MA, 01003, USA.

May 4, 2021

Abstract

Consider a setting where Willie generates a Poisson stream of jobs and routes them to a single server that follows the first-in first-out discipline. Suppose there is an adversary Alice, who desires to receive service without being detected. We ask the question: what is the number of jobs that she can receive covertly, i.e. without being detected by Willie? In the case where both Willie and Alice jobs have exponential service times with respective rates μ_1 and μ_2 , we demonstrate a phase-transition when Alice adopts the strategy of inserting a single job probabilistically when the server idles : over n busy periods, she can achieve a covert throughput, measured by the expected number of jobs covertly inserted, of $\mathcal{O}(\sqrt{n})$ when $\mu_1 < 2\mu_2$, $\mathcal{O}(\sqrt{n}/\log n)$ when $\mu_1 = 2\mu_2$, and $\mathcal{O}(n^{\mu_2/\mu_1})$ when $\mu_1 > 2\mu_2$. When both Willie and Alice jobs have general service times we establish an upper bound for the number of jobs Alice can execute covertly. This bound is related to the Fisher information. More general insertion policies are also discussed.

keywords: Cycle stealing; Covert communication; Queue.

1 Introduction

This paper considers the following problem. Willie has a sequence of jobs that arrive at a first-in first-out (FIFO) queue with a single server, whose processing rate is known to Willie. There exists another actor, Alice, who wants to sneak jobs into the queue for the purpose of stealing processing cycles from Willie. This paper asks the following question: can Alice process her jobs without Willie being able to determine this occurrence beyond making a random guess and, if she can, what is her achievable job processing rate? Answers to this question may apply to several scenarios. Alice could administer a data center, contract to provide Willie with a server with a guaranteed performance, and then resell some of the processing cycles [8]. Similar considerations apply to network contracts. Willie could own a home computer and Alice could install malware for the purpose of stealing computational resources.

In order to address this question of *covert cycle stealing*, we adopt the following model. Willie's jobs arrive according to a Poisson process to a FIFO queue served by a single server with a specified processing rate.

*bjiang@sjtu.edu.cn

†philippe.nain@inria.fr

‡towsley@cs.umass.edu. The research of this author was supported by the US NSF grant CNS-1564067 and the LABEX MILYON (ANR-10-LABX-0070) of Universit de Lyon, within the program "Investissements d'Avenir" (ANR-11-IDEX- 0007) operated by the French National Research Agency (ANR).

Service times of Willie’s jobs are assumed to be independent and identically distributed (iid) according to a general distribution. Alice can insert jobs as she wishes. Her service times are also iid coming from a general distribution that may differ from that of Willie’s. Once an Alice job starts service, it must remain in service until completion; this can interfere with the processing of Willie’s jobs. Last, both Willie and Alice know their own and the other party’s service time distributions and can observe the arrival and departure times of Willie’s jobs.

We formulate the problem as a statistical hypothesis testing problem where Willie’s task is to determine whether or not Alice is stealing cycles, based on observed arrivals and departures. We study the *Insert-at-End-of-Busy-Period* (IEBP) policy, where Alice (probabilistically) inserts a single job each time a Willie busy period (to be defined) ends. We obtain several results, of which the most interesting hold for exponential services for both Willie (rate μ_1) and Alice (rate μ_2) and establish that over n busy periods, Alice can achieve a covert throughput – defined as the expected number of covertly inserted jobs – of $\mathcal{O}(\sqrt{n})$ when $\mu_1 < 2\mu_2$, $\mathcal{O}(\sqrt{n/\log n})$ when $\mu_1 = 2\mu_2$, and $\mathcal{O}(n^{\mu_2/\mu_1})$ when $\mu_1 > 2\mu_2$. This is interesting in part because of the phase transition at $\mu_1 = 2\mu_2$; earlier studies of covert communications and in steganography focused on establishing $\mathcal{O}(\sqrt{n})$ behavior through the control of Alice’s parameters, avoiding regions in the parameter space where this behavior might not hold.

In addition to the above results for the IEBP policy when service times are exponentially distributed, we show that IEBP can also achieve a covert throughput of $\mathcal{O}(\sqrt{n})$ when Willie jobs have general service times and Alice jobs have (hyper-)exponential service times, under some constraints on the service rates.

The rest of the paper is organized as follows. Section 2 discusses related work. Section 3 introduces the model and needed background on hypothesis testing. Section 4 introduces the IEBP policy. Section 5 lists the main results, and some preliminary results are established in Section 6. Sections 7-9 contain the proofs of the main results. Section 10 discusses alternative policies to the IEBP policy. Concluding remarks are given in Section 11.

A word on the notations. For any $a \in [0, 1]$, let $\bar{a} := 1 - a$. We denote the convolution of f and g by $f * g$ and the n -fold tensor product of f with itself is denoted by $f^{\otimes n}$; recall that $f^{\otimes n}(x_1, \dots, x_n) = \prod_{i=1}^n f(x_i)$ with $x_i \in \mathbb{R}^d$ ($d \geq 1$). Throughout we use the shorthand notations $t_{i:j}$ for t_i, \dots, t_j for $i < j$, $a_n \sim_n b_n$ for $\lim_{n \rightarrow \infty} a_n/b_n = 1$, and $\lim_n a_n$ (resp. $\liminf_n a_n$, $\limsup_n a_n$) for $\lim_{n \rightarrow \infty} a_n$ (resp. $\liminf_{n \rightarrow \infty} a_n$, $\limsup_{n \rightarrow \infty} a_n$).

2 Related work

Cycle stealing has been analyzed in the queueing literature in the context of task assignment in multi-server systems. The goal is to allow servers to borrow cycles from other servers while they are idle so as to reduce backlogs and latencies and prevent servers from being under-utilized [10, 17, 18]. These papers focus on the performance analysis of such systems, in particular, mean response times with or without the presence of switching costs. There is no attempt to hide or cover up the theft of cycles.

This paper focuses on the ability of an unknown user to steal cycles without the owner of the server detecting this. Thus it is an instance of a much broader set of techniques used in digital steganography and covert communications. Steganography is the discipline of hiding data in objects such as digital images. A steganographic system modifies fixed-size finite-alphabet coverttext objects into stegotext containing hidden information. A fundamental result of steganography is the *square root law (SRL)*, $\mathcal{O}(\sqrt{n})$ symbols of an n symbol coverttext may safely be altered to hide an $\mathcal{O}(\sqrt{n} \log n)$ -bit message [9]. Covert communications is concerned with the transfer of information in a way that cannot be detected, even by an optimal detector. Here, there exists a similar SRL: suppose Alice may want to communicate to Bob in the face of a third party, Willie, without being detected by Willie. When communication takes place over a channel characterized by

additive white Gaussian noise, it has been established that Alice can transmit $O(\sqrt{n})$ bits of information in n channel uses [3]. This result has been extended to optical (Poisson noise) channels [2], binary channels [6], and many others [5, 21]. It has also been extended to include the presence of jammers [20], and to network settings [19]. Like our work, both steganography and covert communications rely on the use of statistical hypothesis testing. One difference from covert communications is that in our setting Alice hides her jobs in exponentially distributed noise (Willie’s service times) and sometimes generally distributed noise, whereas Alice hides in zero mean Gaussian noise in covert communications. In addition, in the communications context, Alice has control over the power that she transmits at whereas in our context, Alice does not control the size of her jobs, only the rate at which they are introduced.

This work also has ties to the detection of service level agreement (SLA) violation problem. Detecting SLA violation in today’s complex computing infrastructures, such as clouds infrastructures, presents challenging research issues [8]. However no careful analysis of this problem has been conducted. Our work may provide an avenue to doing such.

During the review process a paper related to ours appeared [22]. In [22] Alice’s jobs arrive continuously according to a Poisson process with rate λ_b . It is shown that if Willie knows the number of his jobs successively served in a busy period and λ_b lies below a certain threshold, then the expected number of jobs that Alice can covertly insert over n busy periods is $\mathcal{O}(\sqrt{n})$. If instead Willie knows the length of each busy period instead, the expected number of jobs that Alice can covertly insert is only $\mathcal{O}(1)$.

3 Model and Background

This section gives details about the model we use in the present work and the needed background on hypothesis testing. As mentioned in the introduction, there is a legitimate user, Willie, who sends a sequence of jobs to a single server with known service rate. There is also an illegitimate user, Alice, who wants to introduce a sequence of jobs to be serviced. The questions that we address are the following: can Alice covertly introduce her stream of jobs, i.e. without Willie being able to tell with confidence whether she has introduced the stream or not, and if so, at what rate can she introduce her jobs? We answer these questions under the following assumptions:

1. Willie jobs arrive at the server according to a Poisson process with rate $\lambda \in (0, \infty)$;
2. the service times of all jobs are independent;
3. the service time distributions are known to both parties;
4. the server serves all jobs in a FIFO manner;
5. once in service, Alice jobs cannot be preempted;
6. Willie observes only his arrivals and departures;
7. Alice observes Willie’s and her own arrivals and departures.

The first four assumptions are made mainly for tractability. If Alice jobs can be preempted whenever a Willie job arrives, then Alice can hide her jobs during Willie’s idle periods without affecting his jobs. Consequently, we make the fifth assumption to make the problem interesting. Note that allowing Alice to also observe Willie arrivals and departures gives her the capability to identify idle periods within which to hide her jobs.

Assumption 6 implies that Willie does not know the state of the server. If he does then Alice can only transmit during busy periods. However, if the scheduling policy is FIFO (Assumption 4) Alice cannot

know if an inserted job of hers will not be the last one of the busy period, in which case it will be detected by Willie; relaxing the FIFO assumption appears to be very challenging.

Assume that the system is empty at time 0. Denote by A_i and D_i the arrival and departure times of Willie's i -th job, respectively, for $i \geq 1$. We assume that $D_0 = 0$. Note that $0 < A_i < A_{i+1}$ and $0 < D_i < D_{i+1}$. Let $A_{1:m} = \{A_1, \dots, A_m\}$, $D_{1:m} = \{D_1, \dots, D_m\}$. Let $S_{1:m} = (S_1, \dots, S_m)$ denote the *reconstructed service times* of the first m jobs, which satisfy the following recurrence relation,

$$S_i = \begin{cases} D_1 - A_1, & i = 1, \\ D_i - \max\{A_i, D_{i-1}\}, & i \geq 2. \end{cases} \quad (1)$$

These are the service times perceived by Willie. Note that $(A_{1:m}, S_{1:m})$ and $(A_{1:m}, D_{1:m})$ contain the same information, as they uniquely determine each other. It is also all the information available to Willie in our model.

We define a *Willie Busy Period* (W-BP) to be the time interval between the arrival of a Willie job that finds no other Willie job in the system and the first subsequent departure of a Willie job that leaves no other Willie jobs in the system. Let M_j denote the number of Willie's jobs served in the first j W-BPs, which can be defined recursively by $M_0 = 0$ and

$$M_j = \min \{i > M_{j-1} : A_{i+1} > D_i\}, \quad j \geq 1. \quad (2)$$

Let $N_j = M_j - M_{j-1}$ denote the number of Willie jobs served in the j -th W-BP.

Willie's observation is $W_{1:n}$, where

$$W_j := \left(M_j = m_j, A_{(m_{j-1}+1):m_j} = a_{(m_{j-1}+1):m_j}, S_{(m_{j-1}+1):m_j} = s_{(m_{j-1}+1):m_j} \right), \quad j = 1, \dots, n. \quad (3)$$

In words, Willie observes n W-BPs and for each W-BP records the number of his jobs that have been served, their arrival times and reconstructed service times.

The null hypothesis H_0 is that Alice does not insert jobs and the alternative hypothesis H_1 is that Alice inserts jobs. Willie's test may incorrectly accuse Alice when she does not insert jobs, i.e. he rejects H_0 when it is true. This is known as type I error or false alarm, and, its probability is denoted by P_{FA} [15]. On the other hand, Willie's test may fail to detect insertions of Alice's jobs, i.e. he accepts H_0 when it is false. This is known as type II error or missed detection, and its probability is denoted by P_{MD} . Assume that Willie uses classical hypothesis testing with equal prior probabilities of each hypothesis being true. Then, the lower bound on the sum $P_E = P_{FA} + P_{MD}$ characterizes the necessary trade-off between false alarms and missed detections in the design of a hypothesis test. If prior probabilities are not equal, $\mathbb{P}(H_0) = \pi_0$ and $\mathbb{P}(H_1) = \pi_1$, then, $P_E \geq \min(\pi_0, \pi_1)(P_{FA} + P_{MD})$ [3, Sec. V.B]. Hence scaling results obtained for equal priors apply to the case of non-equal priors and we focus on the former in the remainder of the paper.

4 Insert-at-End-of-Busy-Period Policy

In this section, we consider the strategy that Alice inserts a job probabilistically at the end of each W-BP, which we call the *Insert-at-End-of-Busy-Period (IEBP) Policy*. Note that there may be an Alice job in the system at the start of a W-BP. This occurs when Alice has inserted a job at the end of a W-BP and that her job has not completed service by the time the next Willie job arrives.

Throughout the section, we assume that Willie jobs have service time distribution G_1 with continuous pdf g_1 and finite mean $1/\mu_1 > 0$, and that Alice jobs have service time distribution G_2 with continuous pdf g_2

General	
λ	arrival rate Willie jobs
G_1, g_1	cdf, pdf Willie job service time
G_2, g_2	cdf, pdf Alice job service time
$1/\mu_1$	Willie job expected service time
$1/\mu_2$	Alice job expected service time
H_0	null hypothesis (Alice does not insert jobs)
H_1	alternate hypothesis (Alice inserts jobs)
$T_V(u_0, u_1)$	variation distance between pdfs u_0 and u_1
$H(u_0, u_1)$	Hellinger distance between pdfs u_0 and u_1
X	random variable (rv) with pdf g_1
IEBP policy	
W-BP	Willie Busy Period
Y	Willie first job reconstructed service time
V	Willie idle period duration (exp. rate λ)
f_i	pdf of (Y, V) under $H_i, i = 0, 1$ ($f_0(y, v) = g_1(y)\lambda e^{-\lambda v}$)
q	probability Alice inserts a job ($\bar{q} = 1 - q$); depends on n , the number of W-BPs
$Z(q, y, v)$	$f_1(y, v)/f_0(y, v)$
C_0	Fisher information at origin ($=\mathbb{E}[\rho(X, V)^2]$, with $\rho(x, v)$ defined in (21))
\tilde{f}_i	pdf of Y under $H_i, i = 0, 1$ ($\tilde{f}_0(y) = g_1(y)$)
$\tilde{Z}(q, y)$	$\tilde{f}_1(y)/\tilde{f}_0(y)$
μr	μ_1
μ	μ_2
X_r	exponential rv, rate μr
N_j	nb. Willie jobs served in j -th W-BP
M_j	nb. Willie jobs served in first j W-BPs ($M_j = \sum_{l=1}^j N_l$)
$T(n)$	expected nb. Alice jobs inserted in n W-BPs ($T(n) = nq$)
$T_W(n)$	expected nb. of Willie jobs served in n W-BPs

Figure 1: Glossary of main notations

and finite mean $1/\mu_2 > 0$. Denote by $G_i^*(s) = \int_0^\infty e^{-sx} g_i(x) dx$ the Laplace Stieltjes transform (LST) of g_i for $i = 1, 2$. We also assume $\lambda/\mu_1 < 1$ so that the system is stable under H_0 .

4.1 Introducing the IEBP Policy

To motivate the IEBP policy, we first find the minimum probability that an Alice job interferes with Willie's jobs. Suppose an Alice job is inserted at time t , with service time $\sigma_2 \sim G_2$. Let $U_t \geq 0$ be the unfinished work (of both Alice and Willie jobs) in the system just before time t . The newly inserted Alice job will affect Willie if he sends a job in the interval $(t, t + U_t + \sigma_2)$, the probability of which is

$$\begin{aligned} & \mathbb{P}(\text{at least one Willie job arrives in } (t, t + U_t + \sigma_2)) \\ & \geq \mathbb{P}(\text{at least one Willie job arrives in } (t, t + \sigma_2)) \\ & = \int_0^\infty \mathbb{P}(\text{at least one Willie job arrives in } (t, t + x)) g_2(x) dx \\ & = \int_0^\infty (1 - e^{-\lambda x}) g_2(x) dx = 1 - G_2^*(\lambda) := p. \end{aligned} \tag{4}$$

Thus if Alice is to insert a single job then she should insert it when the system is idle so as to minimize the probability of interfering with a Willie job. Motivated by this observation, we introduce the IEBP policy below.

Alice's strategy. Alice inserts a job with probability q at the end of each W-BP. We refer to this as the *Insert-at-End-of-Busy-Period* (IEBP) policy. Given that Alice does insert a job, the probability that it interferes with a Willie job is given by p in (4). Thus pq is the probability that an interference occurs in a given W-BP.

4.2 Willie's Detector

It is not easy to work directly with the observation process $W_{1:n}$ defined in (3) (Section 3). Instead, we will work with the statistic $(Y_{1:n}, V_{1:n})$, with Y_j denoting the reconstructed service time of the first Willie job in the j -th W-BP and V_j the length of the idle period preceding it. These quantities are given by

$$Y_j = S_{M_{j-1}+1}, \tag{5}$$

$$V_j = A_{M_{j-1}+1} - D_{M_{j-1}}. \tag{6}$$

Denote by $f_i(y, v)$ the joint pdf of (Y, V) at $(Y = y, V = v)$ under H_i for $i = 0, 1$. Also, let $\tilde{f}_i(y) := \int_0^\infty f_i(y, v) dv$ denote the pdf of Y at $Y = y$ under H_i for $i = 0, 1$. Under H_0 the system is a standard $M/G/1$ queue; in particular, the random variables (rvs) Y and V are independent with pdf $g_1(y)$ and $\lambda e^{-\lambda v}$, respectively, yielding

$$f_0(y, v) = g_1(y) \lambda e^{-\lambda v}, \tag{7}$$

$$\tilde{f}_0(y) = g_1(y). \tag{8}$$

Under H_1 , Y_j is the sum of the remaining service time of Alice's job, if any, when a Willie job initiates the j -th W-BP, and of the service time of this Willie's job. Therefore, under H_1 the system is an $M/G/1$ queue with arrival rate λ , exceptional first service time in a busy period with pdf \tilde{f}_1 , and all other service times (the ordinary customers) in a busy period with pdf g_1 . In Proposition 5.5 (Section 5) we derive some performance metrics of interest for this queueing system.

An important feature of the process $(Y_{1:n}, V_{1:n})$ is that $(Y_1, V_1), \dots, (Y_n, V_n)$ form iid rvs due to the Poisson nature of Willie job arrivals and the assumptions that Willie's and Alice's service times are mutually independent processes, further independent of the arrival process. This is the main benefit from using the statistic $(Y_{1:n}, V_{1:n})$ instead of the statistic W_n . From now on (Y, V) denotes a generic (Y_j, V_j) .

The independence of the rvs $(Y_1, V_1), \dots, (Y_n, V_n)$ under H_i ($i = 0, 1$) implies that their joint pdf is given by $f_i^{\otimes n}$, the n -th fold tensor product of f_i with itself.

The following lemma shows that $(Y_{1:n}, V_{1:n})$ is a *sufficient statistic* (e.g. see [15, Chapter 1.9]), that is, Willie does not lose any information by considering the statistic $(Y_{1:n}, V_{1:n})$ instead of the statistic $W_{1:n}$ in order to detect the presence of Alice. The proof is given in Appendix A.

Lemma 4.1. *For every $n \geq 1$, $(Y_{1:n}, V_{1:n})$ is a sufficient statistic.*

Theorem 13.1.1 in [15] is established in the case where a simple hypothesis P_0 is tested against a simple alternative P_1 . In our setting, $q = 0$ is the simple hypothesis against the simple alternative $q = q(n)$, and we are asking how we can scale $q(n)$ down to 0 so that we can not differentiate $q = 0$ and $q = q(n)$, with n the number of observed W-BPs.

Theorem 1. ([15, Theorem 13.1.1])

Using the observed values $(y_{1:n}, v_{1:n})$ of $(Y_{1:n}, V_{1:n})$, any test accepting H_0 if $\prod_{i=1}^n f_0(y_j, v_j) > \prod_{i=1}^n f_1(y_j, v_j)$ and rejecting H_0 if $\prod_{i=1}^n f_0(y_j, v_j) < \prod_{i=1}^n f_1(y_j, v_j)$ minimizes P_E . Furthermore, the minimum P_E is given by

$$P_E^* = 1 - T_V(f_0^{\otimes n}, f_1^{\otimes n}),$$

where

$$T_V(u_0, u_1) = \frac{1}{2} \int |u_0(x) - u_1(x)| dx \quad (9)$$

is the total variation distance between two distributions with densities u_0 and u_1 , respectively.

We will henceforth assume that for a given $(Y_{1:n}, V_{1:n})$ Willie uses the above optimal test.

We say that Alice's insertions are covert provided that, for any $\epsilon > 0$, she has an insertion strategy for each n such that

$$\liminf_n P_E^* \geq 1 - \epsilon, \quad (10)$$

or equivalently from Theorem 1, if for any $\epsilon > 0$,

$$\limsup_n T_V(f_0^{\otimes n}, f_1^{\otimes n}) \leq \epsilon. \quad (11)$$

Note that a sufficient condition for Alice's insertions not being covert is that for some $\delta \in (0, 1)$ there exists a detector such that

$$\limsup_n P_E < \delta. \quad (12)$$

Here the limit is taken over the number of busy periods that Willie observes. This covertness criterion was proposed in the context of low probability of detection (LPD) communications in [3].

Theorem 1 suggests using the total variation distance to analyze Willie's detectors. However, the total variation distance is often unwieldy even for products of pdfs, like $f_0^{\otimes n}$ and $f_1^{\otimes n}$. To overcome this drawback, it is common (e.g. see [3]) to use the following Pinsker's inequality (Lemma 11.6.1 in [7])

$$T_V(u_0, u_1) \leq KL(u_0 \| u_1), \quad (13)$$

where $KL(u_0||u_1) := \int_{\mathbb{R}^d} u_0(x) \ln \frac{u_0(x)}{u_1(x)} dx$ is the *Kullback-Leibler (KL) divergence* between the probability distributions with pdf u_0 and u_1 , respectively.

However, we will work with the Hellinger distance, which has the advantage of offering both lower and upper bounds on the total variation distance. The Hellinger distance between two probability distributions with pdf u_0 and u_1 respectively, denoted $H(u_0, u_1)$, is defined by

$$H(u_0, u_1) = \frac{1}{2} \int_{\mathbb{R}^d} \left(\sqrt{u_0(x)} - \sqrt{u_1(x)} \right)^2 dx. \quad (14)$$

Note that

$$H(u_0, u_1) = 1 - \int_{\mathbb{R}^d} \sqrt{u_0(x)u_1(x)} dx, \quad (15)$$

and $0 \leq H(u_0, u_1) \leq 1$. It is known [13, Lemma 4.1] that

$$H(u_0, u_1) \leq T_V(u_0, u_1) \leq \sqrt{2H(u_0, u_1)}. \quad (16)$$

The upper bound (resp. lower bound) in (16) will be used to establish covert (resp. non-covert) results.

We will also use the following well-known property of the Hellinger distance between pdfs $u_0^{\otimes n}$ and $u_1^{\otimes n}$ [16, Eq. (1.4)]:

$$\begin{aligned} H(u_0^{\otimes n}, u_1^{\otimes n}) &= 1 - \int_{\mathbb{R}^d} \prod_{j=1}^n \sqrt{u_0(x_j)u_1(x_j)} \prod_{j=1}^n dx_j \quad \text{by (15)} \\ &= 1 - \left(\int_{\mathbb{R}^d} \sqrt{u_0(x)u_1(x)} dx \right)^n. \end{aligned} \quad (17)$$

5 Main Results

Let $T(n)$ denote the expected number of jobs that Alice inserts in n W-BPs. Under the IEBP policy,

$$T(n) = nq. \quad (18)$$

This section presents the main results that characterize $T(n)$ under various conditions as n becomes large. Implicit in all asymptotic results as $n \rightarrow \infty$ is that q is a function of n .

Recall that $f_i(y, v)$ is the joint pdf of (Y, V) at $(Y = y, V = v)$ under H_i for $i = 0, 1$, with f_0 given in (7). The likelihood ratio

$$Z(q, y, v) := \frac{f_1(y, v)}{f_0(y, v)}, \quad (19)$$

plays an important role in determining how many jobs Alice can insert covertly. It is shown in Lemma B.1 in Appendix B that Z has the following form,

$$Z(q, y, v) = 1 + q\rho(y, v), \quad (20)$$

where

$$\rho(y, v) := \frac{1}{g_1(y)} \int_0^y g_1(u)g_2(v + y - u)du - \bar{G}_2(v). \quad (21)$$

Since $\rho(y, v)$ does not depend on the insertion probability q , this shows that the likelihood ratio $Z(q, y, v)$ depends linearly on q . Define

$$C_0 := \mathbb{E}[\rho(X, V)^2], \quad (22)$$

where (X, V) has pdf $f_0(x, v)$ at (x, v) .

It is worth noting that $C_0 = J(0)$, with $J(q) := \mathbb{E} \left[\left(\frac{d}{dq} \log f_1(X, V) \right)^2 \right]$ the Fisher information of (Y, V) about the parameter q . Indeed, since $f_1(x, v) = f_0(x, v)(1 + q\rho(x, v))$ by (19)-(20), we have

$$J(q) = \int \frac{1}{f_1(x, v)^2} \left(\frac{d}{dq} f_1(x, v) \right)^2 f_0(x, v) dx dv = \int \frac{\rho(x, v)^2}{(1 + q\rho(x, v))^2} f_0(x, v) dx dv,$$

and therefore $J(0) = \mathbb{E}[\rho(X, V)^2]$. The Fisher information evaluates the amount of information that a random variable carries about an unknown parameter [14].

Proposition 5.1 gives the covert throughput for general service time distributions. Its proof is given in Section 7.

Proposition 5.1 (Covert throughput for general service time distr. and finite C_0). *Assume $C_0 < \infty$. Under the IEBP policy, the number of jobs Alice can insert covertly is $T(n) = \mathcal{O}(\sqrt{n})$ if $\mathbb{E}[\rho(X, V)] = 0$, and $T(n) = \mathcal{O}(1)$ if $\mathbb{E}[\rho(X, V)] \neq 0$.*

Remark 1. $\mathbb{E}[\rho(X, V)] = 0$ for any pdf g_1 if g_2 is the pdf of an exponential or an hyper-exponential rv. Indeed, when $g_2(x) = \mu_2 e^{-\mu_2 x}$, $\rho(x, v)$ in (21) writes

$$\rho(x, v) = e^{-\mu_2 v} \left(\frac{(g_1 * g_2)(x)}{g_1(x)} - 1 \right). \quad (23)$$

By the independence of X and V and the fact that $g_1 * g_2$ is a pdf,

$$\mathbb{E}[\rho(X, V)] = \mathbb{E}[e^{-\mu_2 V}] \cdot \mathbb{E} \left[\left(\frac{(g_1 * g_2)(X)}{g_1(X)} - 1 \right) \right] = 0. \quad (24)$$

The proof when g_2 is the pdf of an hyper-exponential rv is a simple generalization. Note that $\mathbb{E}[\rho(X, V)]$ does not always vanish. In particular, $\mathbb{E}[\rho(X, V)] \neq 0$ when G_2 is an Erlang distribution. Indeed, when $g_2(x) = \frac{(k\mu_2)^k}{(k-1)!} x^{k-1} e^{-k\mu_2 x}$, $k \geq 1$ (Alice service times follow a k -Erlang distribution with mean $1/\mu_2$), it is easy to show that for any pdf g_1 , $\mathbb{E}[\rho(X, V)] = (1 - G_2^*(\lambda)) \left(\frac{(k\mu_2)^k}{\mu_2} - 1 \right) \neq 0$ for all $k > 1$.

The next lemma gives conditions for $C_0 < \infty$ under various distributional assumptions. Its proof is found in Appendix C.

Lemma 5.2 (Finiteness of C_0).

1. Suppose both Alice and Willie have exponential service times, i.e. $g_i(x) = \mu_i e^{-\mu_i x}$ for $i = 1, 2$. Then $C_0 < \infty$ if and only if $\mu_1 < 2\mu_2$.
2. Suppose both Alice and Willie have hyper-exponential service times, i.e.

$$g_i(x) = \sum_{l=1}^{K_i} p_{i,l} \mu_{i,l} e^{-\mu_{i,l} x},$$

where $\sum_{l=1}^{K_i} p_{i,l} = 1$, for $i = 1, 2$. Then $C_0 < \infty$ if and only if

$$\max_{1 \leq l \leq K_1} \mu_{1,l} \leq 2 \min_{1 \leq m \leq K_2} \mu_{2,m}.$$

3. Suppose Willie has Erlang service times and Alice has hyper-exponential service times, i.e.

$$g_1(x) = \frac{\nu_1^{K_1}}{(K_1 - 1)!} x^{K_1-1} e^{-\nu_1 x}$$

and

$$g_2(x) = \sum_{l=1}^{K_2} p_{2,l} \mu_{2,l} e^{-\mu_{2,l} x},$$

where $\sum_{l=1}^{K_2} p_{2,l} = 1$. Then $C_0 < \infty$ if and only if

$$\nu_1 < 2 \min_{1 \leq l \leq K_2} \mu_{2,l}.$$

Proposition 5.1 gives sufficient conditions for Alice to be covert. This raises the following questions:

Q1: When $C_0 < \infty$, can Alice insert covertly more than $\mathcal{O}(\sqrt{n})$ jobs on average during n W-BPs?

Q2: When $C_0 = \infty$, what is the maximum number of jobs that Alice can insert covertly on average during n W-BPs?

We do not have full answers to the above questions. Proposition 5.3 first gives a necessary condition for Alice to be covert under IEBP. Proposition 5.4 then provides a partial answer for the IEBP policy when both Alice and Willie have exponential service times. Proofs are found in Sections 8 and 9.

Proposition 5.3 (Necessary condition for covertness).

Under IEBP Alice cannot be covert if $\limsup_{n \rightarrow \infty} q > 0$.

Consider now the situation when $\lim_{n \rightarrow \infty} q = 0$. The result below is the main result of the paper.

Proposition 5.4 (Covert throughput and converse for exponential service time distr.). *Assume that $g_i(x) = \mu_i e^{-\mu_i x}$ for $i = 1, 2$, and Alice uses the IEBP policy with $\lim_{n \rightarrow \infty} q = 0$. She can be covert if*

$$T(n) = \begin{cases} \mathcal{O}(\sqrt{n}), & \text{if } \mu_1 < 2\mu_2, \\ \mathcal{O}(\sqrt{n/\log n}), & \text{if } \mu_1 = 2\mu_2, \\ \mathcal{O}(n^{\mu_2/\mu_1}), & \text{if } \mu_1 > 2\mu_2. \end{cases} \quad (25)$$

She cannot be covert if

$$T(n) = \begin{cases} \omega(\sqrt{n}), & \text{if } \mu_1 < 2\mu_2, \\ \omega(\sqrt{n/\log n}), & \text{if } \mu_1 = 2\mu_2, \\ \omega(n^{\mu_2/\mu_1}), & \text{if } \mu_1 > 2\mu_2. \end{cases} \quad (26)$$

The above results are in terms of $T(n)$, the expected number of jobs inserted by Alice over n successive W-BPs. It is interesting to determine also the expected number of Willie jobs served during these n W-BPs under the IEBP policy. Let $T_W(n)$ be this number. When $q = 0$, the system behaves like a standard M/G/1 queue with traffic intensity $\lambda/\mu_1 < 1$ and it is known that the expected number of jobs served in a busy period is $(1 - \lambda/\mu_1)^{-1}$ [12], yielding $T_W(n) = n(1 - \lambda/\mu_1)^{-1}$. Proposition 5.5 below shows that the IEBP policy increases each W-BP by a constant factor. The proof is in Appendix H.

Proposition 5.5. *Under IEBP, $T_W(n) = \Theta(n)$. More precisely*

$$T_W(n) = \frac{n}{1 - \lambda/\mu_1} + \frac{\lambda q n}{1 - \lambda/\mu_1} \int_0^\infty t \hat{g}_2(t) dt, \quad (27)$$

which gives the two-sided inequality

$$n \leq T_W(n) \leq n \left(\frac{1 + q\lambda/\mu_2}{1 - \lambda/\mu_1} \right). \quad (28)$$

If $g_2(x) = \mu_2 e^{-\mu_2 x}$, i.e. Alice job service times are exponentially distributed, then

$$T_W(n) = n \left(\frac{1 + pq\lambda/\mu_1}{1 - \lambda/\mu_2} \right). \quad (29)$$

Remark 2. Recall that the Fisher information regarding q is infinity at $q = 0$ when $\mu_1 \geq \mu_2$. It is interesting to speculate that this translates into facilitating Willie's detection task, which appears in the form of the phase transition in equations (25) and (26). Note that one implication of this transition is that Alice should select job sizes with mean size $1/\mu_2 < 2/\mu_1$ to increase throughput without being detected.

6 Preliminary results

We first specialize the two-sided inequality in (16) to the case where $u_0 = f_0^{\otimes n}$ and $u_1 = f_1^{\otimes n}$ and then develop a covert (resp. non-covert) criterion for Alice.

Lemma 6.1. *The Hellinger distance between $f_0^{\otimes n}$ and $f_1^{\otimes n}$ is given by*

$$H(f_0^{\otimes n}, f_1^{\otimes n}) = 1 - \left(\mathbb{E} \left[\sqrt{Z(q, X, V)} \right] \right)^n. \quad (30)$$

Proof. By (17) and (19),

$$\begin{aligned} H(f_0^{\otimes n}, f_1^{\otimes n}) &= 1 - \left(\int_{[0, \infty)^2} \sqrt{f_0(y, v) f_1(y, v)} dy dv \right)^n \\ &= 1 - \left(\int f_0(y, v) \sqrt{Z(q, y, v)} dy dv \right)^n \\ &= 1 - \left(\mathbb{E} \left[\sqrt{Z(q, X, V)} \right] \right)^n. \end{aligned}$$

■

Specializing (16) to the value of $H(f_0^{\otimes n}, f_1^{\otimes n})$ found in Lemma 6.1 gives,

Lemma 6.2 (Lower & upper bounds on total variation distance for statistic $\{Y_j, V_j\}_j$). *For every $n \geq 1$*

$$1 - \left(\mathbb{E} \left[\sqrt{Z(q, X, V)} \right] \right)^n \leq T_V(f_0^{\otimes n}, f_1^{\otimes n}) \leq \sqrt{2} \left(1 - \left(\mathbb{E} \left[\sqrt{Z(q, X, V)} \right] \right)^n \right). \quad (31)$$

Combining Lemma 6.2, the covert criterion (11), and the non-covert criterion (12) yields the following covert/non-covert criterion for Alice:

Corollary 6.1 (Covert/non-covert criteria for the IEPB policy).

Assume that Willie uses an optimal detector for the sufficient statistic $(Y_{1:n}, V_{1:n})$. Alice's insertions are covert if for any $\epsilon > 0$,

$$\liminf_n \left(\mathbb{E} \left[\sqrt{Z(q, X, V)} \right] \right)^n \geq 1 - \epsilon, \quad (32)$$

and Alice's insertions are not covert if for any $\delta > 0$

$$\limsup_n \left(\mathbb{E} \left[\sqrt{Z(q, X, V)} \right] \right)^n < \delta. \quad (33)$$

The lower bound (32) is used to derive the covert throughput (25) in Proposition 5.4.

We now state and prove a non-covert criterion for the IEPB policy. We do so by proposing and analyzing a detector that relies on the (non-sufficient) statistic $\{Y_j\}_j$. Recall that rvs Y_1, \dots, Y_n are iid with common pdf $\tilde{f}_i(y)$ under H_i , for $i = 0, 1$. The non-covert criterion is obtained by applying Theorem 13.1.1 in [15] to the statistic $Y_{1:n}$, which yields the minimum P_E to be $1 - T_V \left(\tilde{f}_0^{\otimes n}, \tilde{f}_1^{\otimes n} \right)$.

The following lemmas are the analog of Lemmas 6.1 and 6.1 for the statistic $\{Y_j\}$.

Lemma 6.3. The Hellinger distance between $\tilde{f}_0^{\otimes n}$ and $\tilde{f}_1^{\otimes n}$ is given by

$$H \left(\tilde{f}_0^{\otimes n}, \tilde{f}_1^{\otimes n} \right) = 1 - \left(\mathbb{E} \left[\sqrt{\tilde{Z}(q, X)} \right] \right)^n, \quad (34)$$

with

$$\tilde{Z}(q, x) := \frac{\tilde{f}_1(x)}{g_1(x)}. \quad (35)$$

Lemma 6.4 (Lower bound on total variation distance for statistic $Y_{1:n}$). For every $n \geq 1$

$$1 - \left(\mathbb{E} \left[\sqrt{\tilde{Z}(q, X)} \right] \right)^n \leq T_V \left(\tilde{f}_0^{\otimes n}, \tilde{f}_1^{\otimes n} \right). \quad (36)$$

The proof of Lemma 6.3 mimics that of Lemma 6.1 and is omitted. Lemma 6.4 follows from Lemma 6.3 and the lower bound in (16).

The non-covert criterion for the statistic $\{Y_j\}_j$ announced earlier is given below. Its proof follows from (12) and (36).

Corollary 6.2 (Non-covert criterion for the IEPB policy).

Assume that Willie uses an optimal detector for the statistic $\{Y_j\}_j$. Alice's insertions are not covert if for any $\delta > 0$

$$\limsup_n \left(\mathbb{E} \left[\sqrt{\tilde{Z}(q, X)} \right] \right)^n < \delta. \quad (37)$$

Corollary 6.2 is used in the proofs of Proposition 5.3 and of the converse (26) in Proposition 5.4.

Remark 3. In direct analogy with the upper bound in Lemma 6.2 it is worth noting that $T_V \left(\tilde{f}_0^{\otimes n}, \tilde{f}_1^{\otimes n} \right)$ in Lemma 6.4 is upper bounded by $\sqrt{2 \left(1 - \left(\mathbb{E} \left[\sqrt{\tilde{Z}(q, X)} \right] \right)^n \right)}$. However, this bound is not useful for establishing a covert result since it does not use the sufficient statistic $(Y_{1:n}, V_{1:n})$.

7 Proof of Proposition 5.1

The proof of Proposition 5.1 relies on an upper bound on the total variation distance between $f_0^{\otimes n}$ and $f_1^{\otimes n}$, given in Lemma 7.1 below. Recall the definition of $\rho(X, V)$ given in (21).

Lemma 7.1. *Let (X, V) be rvs with density f_0 . Assume that $\mathbb{E}[\rho(X, V)] = 0$. Then, for every $n \geq 1$,*

$$T_V(f_0^{\otimes n}, f_1^{\otimes n}) \leq \frac{1}{2} \sqrt{(1 + q^2 C_0)^n - 1}, \quad (38)$$

where C_0 is defined in (22).

Proof. Let $\{(X_j, V_j)\}_j$ be iid rvs with pdf f_0 . By (9),

$$\begin{aligned} 2T_V(f_0^{\otimes n}, f_1^{\otimes n}) &= \int_{[0, \infty)^{2n}} \left| \prod_{j=1}^n f_0(y_j, v_j) - \prod_{j=1}^n f_1(y_j, v_j) \right| \prod_{j=1}^n dy_j dv_j \\ &= \int_{[0, \infty)^{2n}} \prod_{j=1}^n f_0(y_j, v_j) \left| 1 - \prod_{j=1}^n Z(q, y_j, v_j) \right| \prod_{j=1}^n dy_j dv_j \\ &= \mathbb{E} \left[\left| 1 - \prod_{j=1}^n Z(q, X_j, V_j) \right| \right], \end{aligned} \quad (39)$$

where the second equality follows from (19).

Using the inequality $\mathbb{E}|U| \leq \sqrt{\mathbb{E}[U^2]}$ in (39) yields

$$\begin{aligned} (2T_V(f_0^{\otimes n}, f_1^{\otimes n}))^2 &\leq 1 - 2\mathbb{E} \left[\prod_{j=1}^n Z(q, X_j, V_j) \right] + \mathbb{E} \left[\left(\prod_{j=1}^n Z(q, X_j, V_j) \right)^2 \right] \\ &= 1 - 2[\mathbb{E}Z(q, X, V)]^n + [\mathbb{E}[Z(q, X, V)^2]]^n \\ &= -1 + (1 + 2q\mathbb{E}[\rho(X, V)] + q^2\mathbb{E}[\rho(X, V)^2])^n \\ &= -1 + (1 + q^2\mathbb{E}[\rho(X, V)^2])^n, \end{aligned} \quad (40)$$

where we have used the value of $Z(q, x, v)$ obtained in (80) in Appendix B and the assumption that $\mathbb{E}[\rho(X, V)] = 0$ to establish the last two identities. ■

We are now in position to prove Proposition 5.1.

Proof of Proposition 5.1. In order to compute the covert throughput, assume that Willie uses an optimal detector for the sufficient statistic $\{(Y_j, V_j), j = 1, \dots, n\}$. Let

$$q = \frac{\delta}{\phi(n)},$$

with $\delta \in (0, 1]$ and $\phi : \{1, 2, \dots\} \rightarrow [1, \infty)$, so that by (18)

$$T(n) = \frac{\delta n}{\phi(n)}. \quad (41)$$

First consider the case $\mathbb{E}[\rho(X, V)] = 0$. Note that $T(n) = \mathcal{O}(\sqrt{n})$ implies

$$\limsup_n \frac{\sqrt{n}}{\phi(n)} < \infty. \quad (42)$$

By Lemma 7.1

$$\sup_{k \geq n} T_V(f_0^{\otimes k}, f_1^{\otimes k}) \leq \sup_{k \geq n} \frac{1}{2} \sqrt{\left(1 + \frac{\delta^2 C_0}{\phi(k)^2}\right)^k - 1} \sim \frac{\delta \sqrt{C_0}}{2} \sup_{k \geq n} \frac{k}{\phi(k)^2}, \quad \text{as } n \rightarrow \infty$$

as $\lim_n \phi(n) = \infty$ by Lemma D.1 in Appendix D. Therefore,

$$\limsup_n T_V(f_0^{\otimes n}, f_1^{\otimes n}) \leq \frac{\delta \sqrt{C_0}}{2} \left(\limsup_n \frac{\sqrt{n}}{\phi(n)} \right)^2. \quad (43)$$

By making δ small enough, $\limsup_n T_V(f_0^{\otimes n}, f_1^{\otimes n})$ can be made arbitrarily small. We then conclude from (10) that Alice is covert when $T(n) = \mathcal{O}(\sqrt{n})$, which completes the proof for the case $\mathbb{E}[\rho(X, V)] = 0$.

Now consider the case $\mathbb{E}[\rho(X, V)] \neq 0$. Note that $T(n) = \frac{\delta n}{\phi(n)} = \mathcal{O}(1)$ implies there exist $k > 0$ and n_0 such that for all $n \geq n_0$, $0 \leq \frac{n}{\phi(n)} \leq k$. Using inequality (40) and the definition of $Z(q, y, v)$ in (20) gives

$$\begin{aligned} & (2T_V(f_0^{\otimes n}, f_1^{\otimes n}))^2 \\ & \leq 1 - 2e^{n \log\left(1 + \frac{\delta}{\phi(n)} \mathbb{E}[\rho(X, V)]\right)} + e^{n \log\left(1 + \frac{2\delta}{\phi(n)} \mathbb{E}[\rho(X, V)] + \frac{\delta^2}{\phi(n)^2} C_0^2\right)} \\ & \sim 1 - 2e^{\mathbb{E}[\rho(X, V)] \frac{\delta n}{\phi(n)}} + e^{2\mathbb{E}[\rho(X, V)] \frac{\delta n}{\phi(n)} + \delta^2 C_0^2 \frac{n}{\phi(n)^2}}, \end{aligned} \quad (44)$$

as $n \rightarrow \infty$. Since $\frac{n}{\phi(n)}$ and $\frac{n}{\phi(n)^2}$ are bounded away from infinity as $n \rightarrow \infty$, we see that the r.h.s. of (44) can be made arbitrarily close to 0 by letting $\delta \rightarrow 0$. We then conclude from (10) that Alice is covert when $T(n) = \mathcal{O}(1)$, which completes the proof. \blacksquare

8 Proof of Proposition 5.3

The proof uses Corollary 6.2. Take $q = \frac{1}{\phi(n)}$ with $\limsup_n q > 0$ or, equivalently, $\liminf_n \phi(n) < \infty$.

Assume that Willie uses an optimal detector for the statistic $\{Y_j\}_j$ so that Corollary 6.2 applies. Recall (cf. Section 6) that $\tilde{f}_1(y)$ (resp. $g_1(y)$) is the pdf of Y under H_1 (resp. H_0), with $\tilde{Z}(q, y) = \frac{\tilde{f}_1(y)}{g_1(y)}$ being the associated likelihood ratio.

We first derive properties of $\tilde{Z}(q, y)$, to be used in this proof and in the proof of Proposition 5.4.

We claim that

$$\tilde{Z}(q, y) = \int_0^\infty \lambda e^{-\lambda v} Z(q, y, v) dv. \quad (45)$$

Indeed by (19) and (7)

$$\int_0^\infty \lambda e^{-\lambda v} Z(q, y, v) dv = \frac{1}{g_1(y)} \int_0^\infty f_1(y, v) dv = \frac{\tilde{f}_1(y)}{g_1(y)} = \tilde{Z}(q, y),$$

from the definition of $\tilde{Z}(q, y)$ in (35). Hence by (20)

$$\tilde{Z}(q, y) = 1 + q\tilde{\rho}(y), \quad (46)$$

with

$$\begin{aligned}
\tilde{\rho}(y) &:= \int_0^\infty \lambda e^{-\lambda v} \rho(y, v) dv \\
&= \frac{1}{g_1(y)} \int_0^\infty \lambda e^{-\lambda v} \int_0^y g_1(u) g_2(v + y - u) du dv - \int_0^\infty \lambda e^{-\lambda v} \bar{G}_2(v) dv \quad \text{by (21)} \\
&= \frac{(g_1 * \hat{g}_2)(y)}{g_1(y)} - p,
\end{aligned} \tag{47}$$

by using the definition of p in (4), and where

$$\hat{g}_2(t) := \int_0^\infty \lambda e^{-\lambda v} g_2(v + t) dv. \tag{48}$$

We are now ready to prove Proposition 5.3. Recall that X is a rv with pdf g_1 . We have

$$\mathbb{E} \left[\sqrt{\tilde{Z}(q, X)} \right] = \int \sqrt{\prod_{i=0}^1 \tilde{f}_i(x) dx} \leq \sqrt{\prod_{i=0}^1 \int \tilde{f}_i(x) dx} = 1, \tag{49}$$

by Cauchy-Schwarz inequality. Equality holds in (49) if and only if (see e.g. [1, p. 14]) $\tilde{f}_1(x) = c\tilde{f}_0(x)$ for some constant $c > 0$. Since both \tilde{f}_0 and \tilde{f}_1 are densities, integrating over $[0, \infty)$ yields $c = 1$, which is equivalent to $q = 0$ from (35) and (46). This shows that $\mathbb{E} \left[\sqrt{\tilde{Z}(q, X)} \right] < 1$ if and only if $0 < q \leq 1$.

Since $\liminf_n \phi(n) := d < \infty$ by assumption, there exists a subsequence of $\{\phi(n)\}_n$, say $\{\phi(k_n)\}_n$, such that $\phi(k_n) \geq d$ with $\lim_n \phi(k_n) = d$.

Let $M := \sup_{1/d \leq q \leq 1} \mathbb{E} \left[\sqrt{\tilde{Z}(q, X)} \right]$. Note $\sqrt{\tilde{Z}(q, X)} = \sqrt{1 + q\tilde{\rho}(X)} \leq \sqrt{1 + |\tilde{\rho}(X)|} \leq 1 + |\tilde{\rho}(X)|$. By (47), $\mathbb{E}[|\tilde{\rho}(X)|] \leq \int (g_1 * \hat{g}_2)(t) dt + p$ and $g_1 * \hat{g}_2$ is integrable as both g_1 and \hat{g}_2 are integrable. The Dominated Convergence Theorem then guarantees that the function $q \mapsto \mathbb{E} \left[\sqrt{\tilde{Z}(q, X)} \right]$ is continuous.

Since $\mathbb{E} \left[\sqrt{\tilde{Z}(q, X)} \right] < 1$ for all $q \in [1/d, 1]$ as shown above, we have $M < 1$. Therefore,

$$\mathbb{E} \left[\sqrt{\tilde{Z}(1/\phi(k_n), X)} \right] \leq M < 1$$

for all n . As a result

$$\lim_n \left(\mathbb{E} \left[\sqrt{\tilde{Z}(1/\phi(k_n), X)} \right] \right)^n = 0,$$

which implies from Corollary 6.2 that Alice's insertions are not covert when $\liminf_n \phi(n) < 0$, or equivalently when $\limsup_n q = \infty$.

9 Proof of Proposition 5.4

Throughout this section, we assume that Alice and Willie job service times are exponentially distributed with rate μ_2 and μ_1 , respectively, namely, $g_i(x) = \mu_i e^{-\mu_i x}$ for $i = 1, 2$.

The proof of Proposition 5.4 relies on Corollaries 6.1 and 6.2, and Lemma 9.1 below. Before stating the latter, let us introduce some notation.

Let $\mu_1 = r\mu$ and $\mu_2 = \mu$. For $r \neq 1$, define $\beta = \frac{r}{r-1}$ and note that $r = \frac{\beta}{\beta-1}$ and $1 - \beta = \frac{1}{r-1}$. Let X_r denote an exponential rv with rate μr .

For $\theta \in [0, 1]$, $x \geq 0$, define

$$\Xi(\theta, x) = \begin{cases} 1 + \theta(\mu x - 1) & \text{if } r = 1 \\ 1 + \theta \left(\frac{e^{(r-1)\mu x}}{r-1} - \beta \right) & \text{if } r \neq 1. \end{cases} \quad (50)$$

By specializing $Z(q, x, v)$ in (20) to the case where $g_i(x) = \mu_i e^{-\mu_i x}$ for $i = 1, 2$, we obtain from (23) that

$$Z(q, x, v) = \Xi(qe^{-\mu v}, x), \quad \forall q \in [0, 1], x \geq 0, v \geq 0. \quad (51)$$

On the other hand, by (45) and the fact that $\Xi(\theta, x)$ is linear in θ ,

$$\tilde{Z}(q, x) = \int_0^\infty \lambda e^{-\lambda v} \Xi(qe^{-\mu v}, x) dv = \Xi\left(\int_0^\infty q \lambda e^{-(\lambda+\mu)v} dv, x\right) = \Xi(pq, x), \quad \forall q \in [0, 1], x \geq 0, \quad (52)$$

where we have used that $p = \lambda/(\mu_2 + \lambda)$ (see (4)) when Alice job service times are exponentially distributed.

For $\theta \in [0, 1]$, define

$$\xi_r(\theta) := \frac{(\beta - 1)\theta}{1 - \beta\theta}, \quad (53)$$

for $r \geq 2$ (i.e. $1 < \beta \leq 2$), and

$$I_\beta := \beta \int_0^\infty \frac{1 + \frac{1}{2}t - \sqrt{t+1}}{t^{\beta+1}} dt, \quad (54)$$

for $r > 2$. Since $\beta \in (1, 2)$ when $r > 2$, the generalized integral I_β is finite and positive.

Lemma 9.1. For $\theta \in [0, 1]$, define

$$F_r(\theta) := \begin{cases} \frac{1-r}{4(r-2)}\theta + o(\theta)^2 & \text{if } 0 < r < 1 \\ \frac{1}{4}\xi_2^2(\theta) \log \xi_2(\theta) + \Delta_2(\xi_2(\theta)) & \text{if } r = 2 \\ -I_\beta \xi_r^\beta(\theta) + \Delta_r(\xi_r(\theta)) & \text{if } r > 2, \end{cases} \quad (55)$$

where, for $t > 0$,

$$\Delta_r(t) := \begin{cases} o(t^2 \log t) & \text{if } r = 2 \\ o(t^\beta) & \text{if } r > 2. \end{cases} \quad (56)$$

Then, for $r \in (0, 1) \cup [2, \infty)$,

$$\mathbb{E} \left[\sqrt{\Xi(\theta, X_r)} \right] = 1 + F_r(\theta), \quad 0 \leq \theta \leq 1. \quad (57)$$

The proof of Lemma 9.1 is given in Appendix E.

Since $\xi_r(\theta)$ defined in (53) will only be evaluated at $\theta = qe^{-\mu v}$ with $q = \delta/\phi(n)$ and $\delta \in [0, 1]$, thereby yielding $\xi_r(\theta) = \frac{(\beta-1)\delta}{e^{\mu v} \phi(n) - \beta\delta}$, we omit the argument θ in $\xi_r(\theta)$ to simplify notation.

We are now ready to prove Proposition 5.4. Recall that at the beginning of an idle period, Alice inserts a job with probability $q(n) = \frac{\delta}{\phi(n)}$, $\delta \in (0, 1]$, with

$$\lim_n \phi(n) = \infty.$$

Henceforth we drop the argument n in $q(n)$. Furthermore

$$T(n) = nq = \frac{\delta n}{\phi(n)} \quad (58)$$

is the expected number of Alice's insertions in n W-BPs.

9.1 Proof of (25)

We assume that Willie uses an optimal detector for the sufficient statistic $\{(Y_j, V_j)\}_j$, which allows us to apply Corollary 6.1.

9.1.1 Case $\mu_1 < 2\mu_2$

The proof follows from Proposition 5.1 since $\mathbb{E}[\rho(X, V)] = 0$ when Alice and Willie job service times are exponentially distributed (cf. Remark 1) and since $C_0 < \infty$ when $\mu_1 < 2\mu_2$, as shown in Lemma 5.2-(1).

9.1.2 Case $\mu_1 = 2\mu_2$

Without loss of generality we assume in this section that

$$\phi(n) \geq 8, \quad \forall n \geq 1. \quad (59)$$

This assumption is motivated by the need to have $\log \phi(n) > 2$ (for the proof of Lemma F.2 in Appendix F).

Recall that X_2 denotes an exponential rv with rate 2μ . By Lemma 9.1,

$$\begin{aligned} \left(\mathbb{E} \left[\sqrt{Z(\delta/\phi(n), X_2, V)} \right] \right)^n &= \left(1 + \int_0^\infty \lambda e^{-\lambda v} \left(\frac{1}{4} \xi_2^2 \log \xi_2 + \Delta_2(\xi_2) \right) dv \right)^n \\ &= e^{n \log \left(1 + \int_0^\infty \lambda e^{-\lambda v} \xi_2^2 \log \xi_2 \times \left(\frac{1}{4} + \frac{\Delta_2(\xi_2)}{\xi_2^2 \log \xi_2} \right) dv \right)}, \end{aligned} \quad (60)$$

with $\Delta_2(z) = o(z^2 \log z)$ and $\xi_2 = \frac{\delta e^{-\mu v}}{\phi(n) - 2\delta e^{-\mu v}} > 0$ for all $n \geq 1$ and for all $v \geq 0$ thanks to (59). For $v \geq 0$ notice that $\xi_2 > 0$ for all n and $\xi_2 \rightarrow 0$ as $n \rightarrow \infty$. Define

$$D_n := \int_0^\infty \lambda e^{-(\lambda+2\mu)v} \frac{\log \xi_2}{(\phi(n) - 2\delta e^{-\mu v})^2} \left(\frac{1}{4} + \frac{\Delta_2(\xi_2)}{\xi_2^2 \log \xi_2} \right) dv, \quad (61)$$

so that (60) rewrites

$$\left(\mathbb{E} \left[\sqrt{Z(\delta/\phi(n), X_2, V)} \right] \right)^n = e^{n \log(1 + \delta^2 D_n)}. \quad (62)$$

The proof of (25) for $\mu_1 = 2\mu_2$ consists in showing that, as $n \rightarrow \infty$, the r.h.s. of (62) can be made arbitrary close to one by selecting δ small enough, and to apply (32) in Corollary 6.1.

The first step is to show that $D_n \rightarrow 0$ as $n \rightarrow \infty$. This result is shown in Lemma F.1 in Appendix F. Hence,

$$\left(\mathbb{E} \left[\sqrt{Z(\delta/\phi(n), X_2, V)} \right] \right)^n \sim_n e^{\delta^2 n D_n} \quad (63)$$

from (62). The second step is to show that nD_n is bounded as $n \rightarrow \infty$. This result is shown in Lemma F.2 in Appendix F under the condition that $T(n) = \mathcal{O}(\sqrt{n/\log n})$.

The proof is concluded as follows: when $T(n) = \mathcal{O}(\sqrt{n/\log n})$, by (63) and Lemma F.2, as $n \rightarrow \infty$, $\left(\mathbb{E} \left[\sqrt{Z(\delta/\phi(n), X_2, V)} \right]\right)^n$ can be made arbitrarily close to one by taking δ small enough. The proof of (25) for $\mu_1 = 2\mu_2$ then follows from (32) in Corollary 6.1.

9.1.3 Case $\mu_1 > 2\mu_2$

Fix $r > 2$ so that $1 < \beta < 2$. By Lemma 9.1

$$\begin{aligned} \left(\mathbb{E} \left[\sqrt{Z(\delta/\phi(n), X_r, V)} \right]\right)^n &= e^{n \log \left(1 + \int_0^\infty \lambda e^{-\lambda v} \left(-I_\beta \xi_r^\beta + \Delta_r(\xi_r) \right) dv \right)} \\ &= e^{n \log(1 + \delta^\beta (\beta-1)^\beta E_n)}, \end{aligned} \quad (64)$$

with $\Delta_r(z) = o(z^\beta)$, $\xi_r = \frac{\delta(\beta-1)}{e^{\mu\beta\phi(n)} - \beta\delta}$, and

$$E_n := - \int_0^\infty \lambda e^{-\lambda v} \frac{I_\beta - \Delta_r(\xi_r)/\xi_r^\beta}{(e^{\mu\beta\phi(n)} - \beta\delta)^\beta} dv. \quad (65)$$

Lemma F.3 in Appendix F states that $E_n \rightarrow 0$ as $n \rightarrow \infty$ and Lemma F.4 in Appendix F states that nE_n is bounded as $n \rightarrow \infty$ when $T(n) = \mathcal{O}(n^{\mu_2/\mu_1})$. Therefore, cf (64),

$$\left(\mathbb{E} \left[\sqrt{Z(\delta/\phi(n), X_r, V)} \right]\right)^n \sim_n e^{\delta^\beta (\beta-1)^\beta n E_n}, \quad (66)$$

and when $T(n) = \mathcal{O}(n^{\mu_2/\mu_1})$ as $n \rightarrow \infty$ the r.h.s. of (66) can be made arbitrarily close to one by selecting δ small enough. The proof of (25) for $\mu_1 > 2\mu_2$ then follows from (32) in Corollary 6.1.

9.2 Proof of (26)

We assume that Willie uses an optimal detector for the statistic $\{Y_j\}_j$, which will allow us to use the non-covert criterion in Corollary 6.2. Since the proofs in Sections 9.2.1-9.2.3 will not depend on $\delta \in (0, 1]$, we assume without loss of generality that $\delta = 1$, yielding $q = \frac{1}{\phi(n)}$ and $T(n) = \frac{n}{\phi(n)}$.

9.2.1 Case $\mu_1 < \mu_2$

Assume that $T(n) = \omega(\sqrt{n})$, or equivalently

$$\lim_{n \rightarrow \infty} \frac{n}{\phi(n)^2} = \infty. \quad (67)$$

Assume first that $0 < r < 1$. From (52) and Lemma 9.1 we obtain

$$\begin{aligned} \left(\mathbb{E} \left[\sqrt{\tilde{Z}(p/\phi(n), X_r)} \right]\right)^n &= e^{n \log \left(1 + \frac{p^2}{4} \left(\frac{1-r}{r-2} \right) / \phi(n)^2 + o(1/\phi(n)^2) \right)} \\ &\sim_n e^{\frac{n}{\phi(n)^2} \frac{p^2}{4} \left(\frac{1-r}{r-2} \right)} \quad \text{as } \phi(n) \rightarrow \infty \text{ as } n \rightarrow \infty \\ &\sim_n 0, \end{aligned} \quad (68)$$

where the latter follows from (67) together with $\frac{1-r}{r-2} < 0$ when $0 < r < 1$. We invoke Corollary 6.2 to conclude that Alice is not covert when $T(n) = \omega(\sqrt{n})$ and $0 < r < 1$.

It remains to show that Alice is not covert for $1 \leq r < 2$ when $T(n) = \omega(\sqrt{n})$ with $\lim_n n/\phi(n)^2 = \infty$. Without any additional effort, we will prove a stronger result (to be used in the proof of the case $\mu_1 = 2\mu_2$ of (26)) that Alice is not covert when $T(n) = \omega(\sqrt{n})$ and $r \geq 1$. By applying Lemma G.1 in Appendix G to (52), we obtain

$$\mathbb{E} \left[\sqrt{\tilde{Z}(p/\phi(n), X_{r'})} \right] \leq \mathbb{E} \left[\sqrt{\tilde{Z}(p/\phi(n), X_r)} \right] \quad (69)$$

for any $r' \geq r$. Combining now (69) and (68) readily yields

$$\lim_{n \rightarrow \infty} \left(\mathbb{E} \left[\sqrt{\tilde{Z}(p/\phi(n), X_{r'})} \right] \right)^n = 0$$

for any $r' \geq 1$. Similarly to the case $0 < r < 1$ we then conclude from Corollary 6.2 that Alice is not covert $T(n) = \omega(\sqrt{n})$ and $r \geq 1$. In summary, we have shown that Alice is not covert for all $r > 0$ when $T(n) = \omega(\sqrt{n})$.

9.2.2 Case $\mu_1 = 2\mu_2$

Assume that $T(n) = \omega(\sqrt{n/\log n})$ or, equivalently,

$$\lim_n \frac{\phi(n)}{\sqrt{n \log n}} = 0. \quad (70)$$

From (52) and Lemma 9.1

$$\left(\mathbb{E} \left[\sqrt{\tilde{Z}(p/\phi(n), X_2)} \right] \right)^n = e^{n \log(1 + \frac{1}{4} \xi_2^2 \log \xi_2 + o(\xi_2^2 \log \xi_2))}, \quad (71)$$

with $\xi_2 = \frac{p}{\phi(n) - 2p}$. Since $\xi_2 \sim_n 0$ when $\lim_n \phi(n) = \infty$, we have

$$\xi_2 \log \xi_2 \rightarrow 0 \quad \text{as } n \rightarrow \infty.$$

Therefore, from (71),

$$\left(\mathbb{E} \left[\sqrt{\tilde{Z}(p/\phi(n), X_2)} \right] \right)^n \sim_n e^{\frac{n}{4} \xi_2^2 \log \xi_2}. \quad (72)$$

We have proved in the case $\mu_1 < 2\mu_2$ of (26) that Alice is not covert for all $r > 0$ when $T(n) = \omega(\sqrt{n})$. As a result, it suffices to focus on $T(n)$ satisfying (70) when $T(n) \neq \omega(\sqrt{n})$. The latter is equivalent to $\phi(n) = \Omega(\sqrt{n})$, that is,

$$\liminf_n \frac{\phi(n)}{\sqrt{n}} > 0. \quad (73)$$

We have

$$\begin{aligned} n \xi_2^2 \log \xi_2 &= \frac{p^2}{\left(\frac{\phi(n)}{\sqrt{n \log n}} - \frac{2p}{\sqrt{n \log n}} \right)^2} \left(\frac{\log p}{\log n} - \frac{\log(\phi(n) - 2p)}{\log n} \right) \\ &\sim_n \frac{-p^2}{\left(\frac{\phi(n)}{\sqrt{n \log n}} - \frac{2p}{\sqrt{n \log n}} \right)^2} \times \frac{\log(\phi(n) - 2p)}{\log n}. \end{aligned} \quad (74)$$

By (70) the first factor in the r.h.s. of (74) converges to $-\infty$ as $n \rightarrow \infty$. Let us focus on the second factor. We have

$$\frac{\log(\phi(n) - 2p)}{\log n} = \frac{1}{2} + \frac{\log\left(\frac{\phi(n)-2p}{\sqrt{n}}\right)}{\log n} \sim_n \frac{1}{2} + \frac{\log\left(\frac{\phi(n)}{\sqrt{n}}\right)}{\log n}.$$

Assumption (73) ensures that $\frac{\log\left(\frac{\phi(n)}{\sqrt{n}}\right)}{\log n} \rightarrow 0$ as $n \rightarrow \infty$ and

$$\frac{\log(\phi(n) - 2p)}{\log n} \rightarrow \frac{1}{2} \quad \text{as } n \rightarrow \infty.$$

In summary, we have shown that $n\xi_2^2 \log \xi_2 \rightarrow -\infty$ as $n \rightarrow \infty$ which, in turn, implies from (72) that $\lim_n \left(\mathbb{E} \left[\sqrt{\tilde{Z}(p/\phi(n), X_2)} \right] \right)^n = 0$. We conclude from Corollary 6.2 that Alice is not covert if $r = 2$ and $T(n) = \omega(\sqrt{n/\log n})$.

9.2.3 Case $\mu_1 > 2\mu_2$

Assume that $T(n) = n/\phi(n) = \omega(n^{\mu_2/\mu_1})$ or, equivalently,

$$\lim_n \frac{\phi(n)}{n^\beta} = 0. \quad (75)$$

Let $r > 2$ so that $\beta \in (1, 2)$. From (57),

$$\begin{aligned} \left(\mathbb{E} \left[\sqrt{\tilde{Z}(p/\phi(n), X_r)} \right] \right)^n &= e^{n \log(1 - I_\beta \xi_r^\beta + o(\xi_r^\beta))} \\ &\sim_n e^{-I_\beta n \xi_r^\beta}, \end{aligned} \quad (76)$$

since $\xi_r = \frac{(\beta-1)p}{\phi(n) - \beta p} \rightarrow 0$ when $\lim_n \phi(n) = \infty$. We have

$$n \xi_r^\beta = \frac{(\beta-1)p}{\left(\frac{\phi(n)}{n^\beta} - \frac{\beta p}{n^\beta}\right)^\beta} \rightarrow +\infty \quad \text{as } n \rightarrow \infty.$$

Introducing the above limit in (76) and using the finiteness and positiveness of I_β for $\beta \in (1, 2)$, gives $\lim_n \left(\mathbb{E} \left[\sqrt{\tilde{Z}(p/\phi(n), X_r)} \right] \right)^n = 0$, which shows by using again Corollary 6.2 that Alice is not covert if $r > 2$ and $T(n) = \omega(n^{\mu_2/\mu_1})$.

This concludes the proof of Proposition 5.4.

10 Other policies

The first policy – called the *Insert-at-Idle* (II) policy – is a variant of the IEBP policy and works as follows: Alice inserts a job with probability q each time the server idles, and stop inserting with probability \bar{q} (before she tries again at the end of the next W-BP). The difference between the IEBP and II policies is

that under the former Alice inserts at most one job between the end of a W-BP and the start of the next W-BP, whereas under the II policy she may insert more than one job during this time period.

It is shown in [11, Section 5] that when Alice job service times are exponentially distributed all covert/non-covert results obtained under the IEBP policy hold under the II policy. The intuition behind this is that when Alice job service times are exponentially distributed, Willie sees "the same system behavior" under either policy; indeed, under either policy a job of his can interfere with at most one Alice job in a W-BP, whose remaining service time is exponentially distributed.

We have observed at the beginning of Section 4.1 that Alice should preferably inserts jobs at idle times; this was the motivation for introducing and investigating the IEBP policy in Section 4 and its variant, the II policy. But can Alice submit more jobs covertly if she also inserts jobs at other times than at idle times, typically, just after an arrival /departure of a Willie job? Note that, because of the FIFO assumption, Alice cannot benefit from inserting a job at a time $t+$ if time t is neither an arrival time nor a departure time of a Willie job.

This is the motivation for introducing the class of the so-called *Insert-at-Idle-and-at-Arrivals* (II-A) policies, in which

- each time the server idles Alice inserts one job with probability q and does not insert a job with probability \bar{q} ;
- after the arrival of each Willie job, Alice inserts a batch of $s \geq 0$ jobs with probability $qQ(s)$ and with probability $1 - q$ she does not insert any job.

Notice that the II-A policy reduces to the II policy when $Q_B(0) = 1$ (no job inserted at arrival times). Only non-covert results are obtained in [11, Section 6]. More specifically, for exponential service times for both Alice and Willies, Alice is non-covert if $T(n) = \omega(\sqrt{n})$ when the support of Q is finite. A variant of the II-A policy is when Alice inserts a batch of jobs that is geometrically distributed at times the server becomes idle and immediately after the arrival of a Willie job, both with probability q . Non-covert results are also obtained for this policy. The obtained results indicate that batching may be harmful and that Alice should insert only one job at a time.

11 Concluding remarks

In this paper we have studied covert cycle stealing in an M/G/1 queue. We have obtained a phase transition result on the expected number of jobs that Alice can covertly insert in n busy periods when both Alice and Willie's jobs have exponential service times and established partial covert results for arbitrary service times. Several research directions present themselves. We conjecture that Proposition 5.4 holds for a more general class of distributions; it would be interesting to verify this. It would be useful to weaken the assumption that Willie's detectors rely on observations being independent and identically distributed random variables; this would lead to consideration of a larger class of policies on Alice's behalf. Another direction would be to allow Alice to control her job sizes and study what benefit this would provide her. Yet another is to consider other hypothesis testing techniques including generalized likelihood ratio test (GLRT), sequential detection, etc. GLRT could lead to relaxing the need for Willie to know Alice's parameters whereas sequential detection could lead to more timely detection of Alice.

References

- [1] G. Bachmann and L. Narici. *Fourier and Wavelet Analysis*. Springer Science & Business Media, 2012.
- [2] B.A. Bash, A.H. Gheorghie, M. Patel, J.L. Habif, D. Goeckel, D. Towsley, and S. Guha. Quantum-secure covert communication on bosonic channels. In *Nature Communications*, volume 6, 2015.
- [3] B.A. Bash, D. Goeckel, and D. Towsley. Square root law for communication with low probability of detection on awgn channels. In *Proc. IEEE Int. Symp. Inform. Theory (ISIT)*, pages 448–452, Cambridge, MA, USA, July 2012.
- [4] U.N. Bhat. On the busy period of a single server bulk queue with modified service mechanism. *Calcutta Statist. Assoc. Bull.*, 13:163–171, 1964.
- [5] M.R. Bloch. Covert communication over noisy channels: A resolvability perspective. *IEEE Trans. Inf. Theory*, 62(5):2334–2354, 2016.
- [6] P.H. Che, M. Bakshi, C. Chan, and S. Jaggi. Reliable deniable communication with channel uncertainty. In *Proc. IEEE Int. Symp. Inform. Theory Workshop*, pages 448–452, July 2013.
- [7] T.M. Cover and J.A. Thomas. *Elements of Information Theory*. John Wiley & Sons, 2nd edition, 2002.
- [8] V. Emeakaroha, M. Netto T. Ferreto, I. Brandic, and C. De Rose. Casvid: Application level monitoring for SLA violation detection in clouds. In *IEEE Annual Computer Software and Applications Conference (COMPSAC)*, pages 499–508, Cambridge, MA, USA, July 2012.
- [9] J. Fridrich. *Steganography in Digital Media: Principles, Algorithms, and Applications*. Cambridge University Press, 2009.
- [10] M. Harchol-Balter, C. Li, T. Osogami, A. Scheller-Wolf, and M. Squillante. Analysis of task assignment with cycle stealing under central queue. In *Proc. Fifteenth ACM Annual Symposium on Parallel Algorithms and Architectures*, pages 274–285, San Diego, June 2003.
- [11] B. Jiang, P. Nain, and D. Towsley. Covert cycle stealing in a single fifo server (extended version). <https://arxiv.org/abs/2003.05135>, 2021.
- [12] L. Kleinrock. *Queueing Systems, Vol. 1: Theory*. Wiley-Interscience, 1975.
- [13] C.H. Kraft. *Some Conditions for Consistency and Uniform Consistency of Statistical Procedures*, volume 2, pages 125–142. Univ. California Publications in Statist., 1955.
- [14] S. Kullback. *Information Theory and Statistics*. Dover, 1968.
- [15] E. Lehmann and J. Romano. *Testing Statistical Hypotheses*. New York: Springer, 3rd edition, 2005.
- [16] J. Oosterhoff and W. R. Van Zwet. A note on contiguity and hellinger distance. In *Mathematische Statistiek*, volume 36. Stichting Mathematisch Centrum, 1975.
- [17] T. Osogami, M. Harchol-Balter, and A. Scheller-Wolf. Analysis of cycle stealing with switching times and thresholds. In *ACM Sigmetrics*, pages 184–195, San Diego, CA, Junw 2003.
- [18] T. Osogami, M. Harchol-Balter, and A. Scheller-Wolf. Analysis of cycle stealing with switching times and thresholds. *Performance Evaluation*, 61(4):347–369, 2005.

- [19] A. Sheikholeslami, M. Ghaderi, D. Towsley, B. Bash, S. Guha, and D. Goeckel. Multi-hop routing in covert wireless networks. *IEEE Trans. Wireless Communications*, 17(6):3656–3669, 2018.
- [20] T.V. Sobers, B.A. Bash, S.K. Guha, D. Towsley, and D. Goeckel. Covert communication in the presence of an uninformed jammer. *IEEE Trans. Wireless Communications*, 16(9):6193–6206, 2017.
- [21] L. Wang, G.W. Wornell, and L. Zheng. Fundamental limits of communication with low probability of detection. *IEEE Trans. on Information Theory*, 62(6):3493–3503, 2016.
- [22] A. Yardi and T. Bodas. A Covert Queueing Problem with Busy Period Statistic. *IEEE Communication Letters*, 2020.

A Appendix: Proof of Lemma 4.1

Let $p_{i,n}(w_{1:n})$ be the pdf of $W_{1:n} = (W_1, \dots, W_n)$ at $w_{1:n} = (w_1, \dots, w_n)$ under H_i for $i = 0, 1$. Also let $\tilde{p}_{i,j}(w_j)$ be the pdf¹ of W_j at w_j under H_i for $i = 0, 1$. Note that $p_{i,n}$ (resp. $\tilde{p}_{i,j}$) is a generalized pdf since $W_{1:n}$ (resp. W_j) contains integer and continuous components.

By the general multiplicative formula,

$$\tilde{p}_{i,n}(w_{1:n}) = \tilde{p}_{i,1}(w_1) \times \prod_{j=2}^n \tilde{p}_{i,j}(w_j | W_{1:j-1} = w_{1:j-1}). \quad (77)$$

Let $w_j = (m_j, a_{(m_{j-1}+1):m_j}, s_{(m_{j-1}+1):m_j})$, $y_j = s_{m_{j-1}+1}$, and $v_j = a_{m_{j-1}+1} - d_{m_{j-1}}$. We have

$$\tilde{p}_{i,j}(w_j | W_{1:j-1} = w_{1:j-1}) = \tilde{p}_{i,j}(w_j | A_{m_{j-1}} = a_{m_{j-1}}, D_{m_{j-1}} = d_{m_{j-1}}, A_{m_{j-1}+1} > d_{m_{j-1}}),$$

since the probability distribution of the number of customers served in a busy period in an M/G/1 queue is entirely determined once we know the duration of the first service time in this busy period [12, Chapter 5.9]. Hence,

$$\begin{aligned} \tilde{p}_{i,j}(w_j | W_{1:j-1} = w_{1:j-1}) &= \mathbf{1}(a_{m_{j-1}+1} > d_{m_{j-1}}) f_i(y_j, v_j) \\ &\times p\left(M_j = m_j, A_{(m_{j-1}+2):m_j} = a_{(m_{j-1}+2):m_j}, S_{(m_{j-1}+2):m_j} = s_{(m_{j-1}+2):m_j} | Y_j = y_j, V_j = v_j\right), \end{aligned} \quad (78)$$

where the latter density is independent of H_0 and H_1 . The pdf $\tilde{p}_{i,1}(w_1)$ is given by the r.h.s. of (78) by letting $j = 1$. Putting (77) and (78) together yields the factorization result

$$p_{i,n}(w_{1:n}) = \prod_{j=1}^n f_j(y_j, v_j) \times \text{other factors independent of } H_0 \text{ and } H_1, \quad (79)$$

which proves that $(Y_{1:n}, V_{1:n})$ is a sufficient statistic [15, Chapter 1.9].

B Appendix

Recall that $Z(q, y, v) = \frac{f_1(y, v)}{f_0(y, v)}$, with $f_i(y, v)$ the pdf of (Y, V) at (y, v) under H_i for $i = 0, 1$.

¹Clearly $\tilde{p}_{i,j}(w_j) = \int_{\mathbb{R}^{n-1}} p_{i,n}(w_{1:n}) dw_1 \cdots dw_{j-1} dw_{j+1} \cdots dw_n$.

Lemma B.1.

$$Z(q, y, v) = 1 + q\rho(y, v), \quad (80)$$

where

$$\rho(y, v) := \frac{1}{g_1(y)} \int_0^x g_1(u)g_2(v + y - u)du - \bar{G}_2(v). \quad (81)$$

Proof. Consider a generic W-BP. Let σ_1 (resp. σ_2) denote a generic service time of a Willie (resp. Alice) job. Let A be the event that Alice inserts a job at the end of the W-BP. Then

$$Y = \sigma_1 + \mathbb{1}\{A\} \cdot (\sigma_2 - V)^+,$$

where $(z)^+ = \max\{z, 0\}$. We first compute the conditional density $f_1(y | v)$ of Y given V . Given A^c , $Y = \sigma_1$, so

$$f_1(y | v, A^c) = g_1(y).$$

Given A and $V = v$, we have $Y = \sigma_1 + (\sigma_2 - v)^+$, so that

$$f_1(y | v, A) = g_1(y)G_2(v) + \int_0^y g_1(u)g_2(y + v - u)du.$$

Recall the probability of A under H_1 is q , so

$$\begin{aligned} f_1(y | v) &= qf_1(y | v, A) + \bar{q}f_1(y | v, A^c) \\ &= g_1(y) + q \left[\int_0^y g_1(u)g_2(y + v - u)du - g_1(y)\bar{G}_2(v) \right] \\ &= g_1(y)[1 + q\rho(y, v)], \end{aligned} \quad (82)$$

by using the definition of $\rho(y, v)$ in (81). Therefore,

$$Z(q, y, v) = \frac{f_1(y | v)\lambda e^{-\lambda v}}{f_0(y, v)} = 1 + q\rho(y, v),$$

by using (7), which concludes the proof. ■

C Appendix: Proof of Lemma 5.2

Let $g_2(x) = \sum_{l=1}^{K_2} p_{2,l}g_{2,l}(x)$ with $g_{2,l}(x) := \mu_{2,l}e^{-\mu_{2,l}x}$, $p_{2,l} \geq 0$ for all l and $\sum_{l=1}^{K_2} p_{2,l} = 1$, namely, Alice job service times follow an hyper-exponential distribution with mean $1/\mu_2 = \sum_{l=1}^{K_2} 1/\mu_{2,l}$. Denote by $G_1^*(s) = \int_0^\infty e^{-sx}g_1(x)dx$ the Laplace transform of Willie job service times.

By using (21), we find

$$\rho(x, v) = \frac{1}{g_1(x)} \sum_{l=1}^{K_2} p_{2,l}e^{-\mu_{2,l}v}(g_1 * g_{2,l})(x) - \sum_{l=1}^{K_2} p_{2,l}e^{-\mu_{2,l}v},$$

so that

$$\mathbb{E} [\rho(X, V)^2] = \alpha_1 - 2\alpha_2 + \alpha_3$$

with

$$\alpha_1 := \int_{[0, \infty)^2} \frac{\lambda e^{-\lambda v}}{g_1(x)} \left[\sum_{l=1}^{K_2} p_{2,l}e^{-\mu_{2,l}v}(g_1 * g_{2,l})(x) \right]^2 dv dx$$

$$\begin{aligned}
&= \sum_{\substack{l=1 \\ m=1}}^{K_2} \frac{\lambda p_{2,l} p_{2,m}}{\mu_{2,l} + \mu_{2,m}} \int_0^\infty \frac{(g_1 * g_{2,l})(x) \times (g_1 * g_{2,m})(x)}{g_1(x)} dx; \\
\alpha_2 &:= \sum_{\substack{l=1 \\ m=1}}^{K_2} p_{2,l} p_{2,m} \int_{[0,\infty)^2} \lambda e^{-(\lambda+\mu_{2,l})v} (g_1 * g_{2,l})(x) e^{\mu_{2,m}x} dv dx \\
&= \sum_{\substack{l=1 \\ m=1}}^{K_2} p_{2,l} p_{2,m} \frac{\lambda \mu_{2,l} G_1^*(\mu_{2,m})}{(\lambda + \mu_{2,l})(\mu_{2,l} + \mu_{2,m})} \leq 1; \\
\alpha_3 &:= \sum_{\substack{l=1 \\ m=1}}^{K_2} p_{2,l} p_{2,m} \int_{[0,\infty)^2} \lambda e^{-\lambda v} g_1(x) e^{-(\mu_{2,l}+\mu_{2,m})x} dv dx \\
&= \sum_{\substack{l=1 \\ m=1}}^{K_2} p_{2,l} p_{2,m} G_1^*(\mu_{2,l} + \mu_{2,m}) \leq 1.
\end{aligned}$$

We conclude from the above that $\mathbb{E}[\rho(X, V)^2] < \infty$ if and only if

$$\beta_{l,m} := \int_0^\infty \frac{(g_1 * g_{2,l})(x) \times (g_1 * g_{2,m})(x)}{g_1(x)} dx < \infty \quad (83)$$

for all $l, m = 1, \dots, K_2$.

Case 1: $g_1(x) = \sum_{i=1}^{K_1} p_{1,i} \mu_{1,i} e^{-\mu_{1,i}x}$, $p_{1,i} \geq 0$ for all i and $\sum_{i=1}^{K_1} p_{1,i} = 1$, namely, Willie job service times follow an hyper-exponential distribution with mean $1/\mu_1 = \sum_{i=1}^{K_1} 1/\mu_{1,i}$.

We have

$$(g_1 * g_{2,l})(x) = \sum_{i=1}^{K_1} p_{1,i} \mu_{1,i} \mu_{2,l} \left[x e^{-\mu_{2,l}x} \mathbf{1}(\mu_{1,i} = \mu_{2,l}) + \frac{e^{-\mu_{2,l}x} - e^{-\mu_{1,i}x}}{\mu_{1,i} - \mu_{2,l}} \mathbf{1}(\mu_{1,i} \neq \mu_{2,l}) \right]$$

for $l = 1, \dots, K_2$, so that

$$\begin{aligned}
(g_1 * g_{2,l})(x) \times (g_1 * g_{2,m})(x) &= \sum_{\substack{i=1 \\ j=1}}^{K_1} p_{1,i} p_{1,j} \mu_{1,i} \mu_{1,j} \left[P_{i,l}(x) P_{j,m}(x) e^{-(\mu_{2,l}+\mu_{2,m})x} \right. \\
&\quad \left. - a_{i,l} b_{j,m} x e^{-(\mu_{1,i}+\mu_{2,m})x} - a_{j,m} b_{i,l} x e^{-(\mu_{1,j}+\mu_{2,l})x} + b_{i,l} b_{j,m} e^{-(\mu_{1,i}+\mu_{1,j})x} \right],
\end{aligned}$$

with

$$\begin{aligned}
a_{i,l} &:= \mu_{2,l} \mathbf{1}(\mu_{1,i} = \mu_{2,l}) \\
b_{i,l} &:= \frac{\mu_{2,l}}{\mu_{1,i} - \mu_{2,l}} \mathbf{1}(\mu_{1,i} \neq \mu_{2,l}) \\
P_{i,l}(x) &:= a_{i,l} x + b_{i,l},
\end{aligned}$$

for $i = 1, \dots, K_1$, $l = 1, \dots, K_2$. Define $\mu_1^* = \max_{1 \leq i \leq K_1} \mu_{1,i}$. Then,

$$\beta_{l,m} = \sum_{\substack{i=1 \\ j=1}}^{K_1} p_{1,i} p_{1,j} \mu_{1,i} \mu_{1,j} \int_0^\infty \frac{P_{i,l}(x) P_{j,m}(x) e^{-(\mu_{2,l}+\mu_{2,m}-\mu_1^*)x}}{\sum_{i=1}^{K_1} p_{1,i} \mu_{1,i} e^{(\mu_1^*-\mu_{1,i})x}} dx$$

$$\begin{aligned}
& - \sum_{j=1}^{K_1} p_{1,j} \mu_{1,j} b_{j,m} \int_0^\infty \frac{\sum_{i=1}^{K_1} p_{1,i} \mu_{1,i} a_{i,l} e^{-\mu_{1,i} x}}{\sum_{i=1}^{K_1} p_{1,i} \mu_{1,i} e^{-\mu_{1,i} x}} x e^{-\mu_{2,m} x} dx \\
& - \sum_{j=1}^{K_1} p_{1,j} \mu_{1,j} b_{j,l} \int_0^\infty \frac{\sum_{i=1}^{K_1} p_{1,i} \mu_{1,i} a_{i,m} e^{-\mu_{1,i} x}}{\sum_{i=1}^{K_1} p_{1,i} \mu_{1,i} e^{-\mu_{1,i} x}} x e^{-\mu_{2,l} x} dx \\
& + \sum_{j=1}^{K_1} p_{1,j} \mu_{1,j} b_{j,m} \int_0^\infty \frac{\sum_{i=1}^{K_1} p_{1,i} \mu_{1,i} b_{i,l} e^{-\mu_{1,i} x}}{\sum_{i=1}^{K_1} p_{1,i} \mu_{1,i} e^{-\mu_{1,i} x}} e^{-\mu_{1,j} x} dx.
\end{aligned} \tag{84}$$

The second, third, and fourth integrals in the r.h.s. of (84) are finite since $\lim_{x \rightarrow \infty} \frac{\sum_{i=1}^{K_1} p_{1,i} \mu_{1,i} a_{i,l} e^{-\mu_{1,i} x}}{\sum_{i=1}^{K_1} p_{1,i} \mu_{1,i} e^{-\mu_{1,i} x}}$ and $\lim_{x \rightarrow \infty} \frac{\sum_{i=1}^{K_1} p_{1,i} \mu_{1,i} b_{i,l} e^{-\mu_{1,i} x}}{\sum_{i=1}^{K_1} p_{1,i} \mu_{1,i} e^{-\mu_{1,i} x}}$ are finite for any $l = 1, \dots, K_2$. The first integral is finite if and only if

$$\mu_1^* = \max_{1 \leq i \leq K_1} \mu_{1,i} \leq 2 \min_{1 \leq l \leq K_2} \mu_{2,l}. \tag{85}$$

This shows that $C_0 < \infty$ when (85) holds.

In particular, when $K_1 = K_2 = 1$ (exponential service times for both Alice and Willie jobs) then $C_0 < \infty$ if and only if $\mu_1 < 2\mu_2$. For further reference, note that

$$\rho(x, v) = \begin{cases} e^{-\mu_1 v} (\mu_1 x - 1) & \text{if } \mu_1 = \mu_2 \\ e^{-\mu_2 v} \left(\frac{\mu_2 e^{-(\mu_2 - \mu_1)x} - \mu_1}{\mu_1 - \mu_2} \right) & \text{if } \mu_1 \neq \mu_2 \end{cases} \tag{86}$$

when $g_i(x) = \mu_i e^{-\mu_i x}$, $i = 1, 2$.

Case 2: $g_1(x) = \nu_1^{K_1} x^{K_1-1} e^{-\nu_1 x} / (K_1 - 1)!$ with $1/\mu_1 = K_1/\nu_1$ (Willie job service times follow a K_1 -stage Erlang pdf with mean $1/\mu_1$).

We have, with $\eta_l := \frac{\nu_1^{K_1} \mu_{2,l}}{(K_1 - 1)!}$,

$$(g_1 * g_{2,l})(x) = \begin{cases} \eta_l \frac{x^{K_1}}{K_1} e^{-\mu_{2,l} x} & \text{if } \nu_1 = \mu_{2,l} \\ \eta_l e^{-\mu_{2,l} x} \int_0^x u^{K_1-1} e^{-(\nu_1 - \mu_{2,l})u} du & \text{if } \nu_1 \neq \mu_{2,l}, \end{cases}$$

for $l = 1, \dots, K_2$.

Define $\xi_l(k) = \int_0^x u^{k-1} e^{-(\nu_1 - \mu_{2,l})u} du$ for $k \geq 1$. Integrating by part gives

$$\xi_l(k) = \frac{-x^{k-1} e^{-(\nu_1 - \mu_{2,l})x}}{\nu_1 - \mu_{2,l}} + \frac{k-1}{\nu_1 - \mu_{2,l}} \xi_l(k-1), \quad k \geq 2,$$

which yields (use that $\xi_l(1) = (1 - e^{-(\nu_1 - \mu_{2,l})x}) / (\nu_1 - \mu_{2,l})$)

$$\xi_l(k) = -e^{-(\nu_1 - \mu_{2,l})x} \sum_{i=1}^k \frac{(k-1)!}{(k-i)!} \frac{x^{k-i}}{(\nu_1 - \mu_{2,l})^i} + \frac{(k-1)!}{(\nu_1 - \mu_{2,l})^k}.$$

Therefore,

$$(g_1 * g_{2,l})(x) = Q_{1,l}(x) e^{-\mu_{2,l} x} - Q_{2,l}(x) e^{-\nu_1 x}, \tag{87}$$

with

$$Q_{1,l}(x) := \eta_l \frac{x^{K_1}}{K_1} \mathbf{1}(\nu_1 = \mu_{2,l}) + \eta_l \frac{(K_1 - 1)!}{(\nu_1 - \mu_{2,l})^{K_1}} \mathbf{1}(\nu_1 \neq \mu_{2,l})$$

(88)

$$Q_{2,l}(x) := \eta \sum_{i=1}^{K_1} \frac{(K_1 - 1)!}{(K_1 - i)!} \frac{x^{K_1 - i}}{(\nu_1 - \mu_{2,l})^i} \mathbf{1}(\nu_1 \neq \mu_{2,l}).$$

(89)

When $\nu_1 = \mu_{2,l}$ or $\nu_1 = \mu_{2,m}$ it is easily seen from (87)-(89) that $\beta(l, m) < \infty$ if and only if $\nu_1 < \mu_{2,l} + \mu_{2,m}$.

Let us investigate the (less trivial) remaining case when $\nu_1 \neq \mu_{2,l}$ and $\nu_1 \neq \mu_{2,m}$. In this case we have, from (87)-(89),

$$\begin{aligned} & \frac{(g_1 * g_{2,l})(x)(g_1 * g_{2,m})(x)}{g_1(x)} \\ &= \frac{\eta_l \eta_m}{\nu_1^{K_1} x^{K_1} e^{-\nu_1 x}} \left[\frac{(K_1 - 1)!}{(\nu_1 - \mu_{2,l})^{K_1}} e^{-\mu_{2,l} x} - \sum_{i=1}^{K_1} \frac{(K_1 - 1)!}{(K_1 - i)!} \frac{x^{K_1 - i}}{(\nu_1 - \mu_{2,l})^i} e^{-\nu_1 x} \right] \\ & \quad \times \left[\frac{(K_1 - 1)!}{(\nu_1 - \mu_{2,m})^{K_1}} e^{-\mu_{2,m} x} - \sum_{i=1}^{K_1} \frac{(K_1 - 1)!}{(K_1 - i)!} \frac{x^{K_1 - i}}{(\nu_1 - \mu_{2,m})^i} e^{-\nu_1 x} \right] \\ &= \frac{\eta_l \eta_m ((K_1 - 1)!)^2}{\nu_1^{K_1} (\nu_1 - \mu_{2,l})^{K_1} (\nu_1 - \mu_{2,m})^{K_1}} \\ & \quad \times \left[e^{(\nu_1 - \mu_{2,l})x} - \sum_{j=0}^{K_1 - 1} \frac{(x(\nu_1 - \mu_{2,l}))^j}{j!} \right] \times \frac{1}{x^{K_1}} \left[e^{-\mu_{2,m} x} - e^{-\nu_1 x} \sum_{j=0}^{K_1 - 1} \frac{(x(\nu_1 - \mu_{2,m}))^j}{j!} \right], \end{aligned}$$

which shows that $(g_1 * g_{2,l})(x)(g_1 * g_{2,m})(x)/g_1(x)$ is well-defined when $x \rightarrow 0$ and is $[0, \infty)$ -integrable if and only if $\nu_1 < \mu_{2,l} + \mu_{2,m}$.

In summary, $C_0 < \infty$ if and only if $\nu_1 < 2 \min_{1 \leq l \leq K_2} \mu_{2,l}$ or, equivalently, if and only if $\mu_1 < \frac{2}{K_1} \min_{1 \leq l \leq K_2} \mu_{2,l}$.

D Appendix

Lemma D.1. *Let $f, g : \mathbb{N} \rightarrow [0, \infty)$. If $\limsup_n \frac{g(n)}{f(n)} < \infty$ with $\lim_n g(n) = \infty$ then $\lim_n f(n) = \infty$.*

Proof. If $\lim_n \frac{g(n)}{f(n)} = 0$ then clearly $\lim_n f(n) = \infty$. Assume now that there exist $0 < L < \infty$ and n_0 such that for all $n > n_0$

$$\sup_{k \geq n} \frac{g(k)}{f(k)} < L.$$

Since $\sup_{k \geq n} \frac{g(k)}{f(k)} \geq \frac{g(n)}{f(n)}$, $f(n) > L^{-1}g(n)$ for $n > n_0$, which proves the lemma since $\lim_n g(n) = \infty$. \blacksquare

E Appendix: Proof of Lemma 9.1

Proof. Assume that $0 < r < 1$ (i.e. $\beta < 0$). Recalling that X_r is an exponential rv with rate μr , definition (53) yields

$$\begin{aligned} & \mathbb{P} \left(\sqrt{\Xi(\theta, X_r)} > z \right) \\ &= \begin{cases} 0 & \text{if } z > \sqrt{1 - \theta\beta} \\ \left(\frac{1 - \theta\beta - z^2}{\theta(1 - \beta)} \right)^{-\beta} & \text{if } \sqrt{1 - \theta} \leq z \leq \sqrt{1 - \theta\beta} \\ 1 & \text{if } z < \sqrt{1 - \theta}, \end{cases} \end{aligned}$$

which gives

$$\begin{aligned}
\mathbb{E} \left[\sqrt{\Xi(\theta, X_r)} \right] &= \int_0^\infty \mathbb{P} \left(\sqrt{\Xi(\theta, X_r)} > z \right) dz \\
&= \sqrt{1-\theta} + (\theta(1-\beta))^\beta \int_{\sqrt{1-\theta}}^{\sqrt{1-\theta\beta}} (-y^2 + 1 - \theta\beta)^{-\beta} dy \\
&= \sqrt{1-\theta} + \left(\frac{\theta(1-\beta)}{1-\theta\beta} \right)^\beta (1-\theta\beta)^{1/2} \int_{\sqrt{\frac{1-\theta}{1-\theta\beta}}}^1 (1-y^2)^{-\beta} dy.
\end{aligned} \tag{90}$$

Recall that $\xi_r(\theta) = (1-\beta)\theta/(1-\theta\beta)$, so that $\theta = \xi_r(\theta)/(1-\beta + \beta\xi_r(\theta))$. Substitution into (90) yields (with $\xi_r \equiv \xi_r(\theta)$ with a slight abuse of notation)

$$\begin{aligned}
\mathbb{E} \left[\sqrt{\Xi(\theta, X_r)} \right] &= \sqrt{1 - \frac{\xi_r}{1-\beta + \beta\xi_r}} + \xi_r^\beta \left(1 + \frac{\beta\xi_r}{1-\beta} \right)^{-1/2} \int_{\sqrt{1-\xi_r}}^1 (1-y^2)^{-\beta} dy \\
&= \sqrt{1 - \frac{\xi_r}{1-\beta + \beta\xi_r}} + \frac{1}{2} \xi_r^\beta \left(1 + \frac{\beta\xi_r}{1-\beta} \right)^{-1/2} \int_0^{\xi_r} \frac{x^{-\beta}}{\sqrt{1-x}} dx.
\end{aligned}$$

When x is small, $x^{-\beta}/\sqrt{1-x} \sim x^{-\beta} + x^{1-\beta}/2$, so that $\int_0^{\xi_r} \frac{x^{-\beta}}{\sqrt{1-x}} dx \sim \xi_r^{1-\beta}/(1-\beta) + \xi_r^{2-\beta}/(2(2-\beta))$ as ξ_r is small. With this, we obtain

$$\begin{aligned}
\mathbb{E} \left[\sqrt{\Xi(\theta, X_r)} \right] &= 1 + \frac{1-4\beta+2\beta^2}{4(2-\beta)(1-\beta)^2} \xi_r^2 + o(\xi_r^2) \\
&= 1 + \frac{1-4\beta+2\beta^2}{4(2-\beta)} \theta^2 + o(\theta^2).
\end{aligned} \tag{91}$$

Since $\frac{1-4\beta+2\beta^2}{4(2-\beta)} = \frac{1-r}{4(r-2)}$, this proves the lemma when $r < 1$.

Consider now the case where $r \geq 2$. Notice that $1 < \beta \leq 2$ when $r \geq 2$. It is easily seen from (50) that, for $r > 1$,

$$\frac{d}{dz} \mathbb{P}(\Xi(\theta, X_r) < z) = \begin{cases} \frac{\beta(\theta(\beta-1))^\beta}{(z-1+\theta\beta)^{\beta+1}} & \text{if } z \geq 1-\theta \\ 0 & \text{if } z < 1-\theta, \end{cases}$$

which yields

$$\begin{aligned}
\mathbb{E} \left[\sqrt{\Xi(\theta, X_r)} \right] &= \beta\theta^\beta(\beta-1)^\beta \int_{1-\theta}^\infty \frac{\sqrt{z}}{(z-1+\theta\beta)^{\beta+1}} dz \\
&= \frac{2\beta\theta^\beta(\beta-1)^\beta}{(1-\theta\beta)^{\beta-1/2}} \int_{\sqrt{\frac{1-\theta}{1-\theta\beta}}}^\infty \frac{t^2}{(t^2-1)^{\beta+1}} dt.
\end{aligned} \tag{92}$$

We are now ready to address the case when $r \geq 2$.

Assume first that $r = 2$, so that $\beta = 2$, $\xi_2 \equiv \xi_2(\theta) = \frac{\theta}{1-2\theta}$, and $\theta = \frac{\xi}{2\xi+1}$. By (92) we have

$$\begin{aligned}
\mathbb{E} \left[\sqrt{\Xi(\theta, X_2)} \right] &= 4(1+2\xi_2)^{-1/2} \xi_2^2 \int_{\sqrt{\xi_2+1}}^\infty \frac{t^2}{(t^2-1)^3} dt \\
&= 2(1+2\xi_2)^{-1/2} \xi_2^2 \int_{\xi_2}^\infty \frac{\sqrt{y+1}}{y^3} dy.
\end{aligned}$$

Let $h(y) = \frac{1 + \frac{1}{2}y - \sqrt{y+1}}{y^3}$.

We have (Hint: use L'Hôpital's rule to get the 2nd equality)

$$\lim_{\xi_2 \rightarrow 0} \frac{2\xi_2^2 \int_{\xi_2}^{\infty} \frac{\sqrt{y+1}}{y^3} dy - 1 - \xi_2}{2\xi_2^2 \log \xi_2} = \lim_{\xi_2 \rightarrow 0} \frac{-\int_{\xi_2}^{\infty} h(y) dy}{\log \xi_2} = \lim_{\xi_2 \rightarrow 0} \frac{h(\xi_2)}{\xi_2^{-1}} = \lim_{\xi_2 \rightarrow 0} \frac{1 + \frac{1}{2}\xi_2 - \sqrt{\xi_2 + 1}}{\xi_2^2} = \frac{1}{8},$$

and

$$2\xi_2^2 \int_{\xi_2}^{\infty} \frac{\sqrt{y+1}}{y^3} dy = 1 + \xi_2 + \frac{1}{4}\xi_2^2 \log \xi_2 + o(\xi_2^2 \log \xi_2).$$

It follows that

$$\begin{aligned} \mathbb{E} \left[\sqrt{\Xi(\theta, X_2)} \right] &= (1 + 2\xi_2)^{-1/2} 2\xi_2^2 \int_{\xi_2}^{\infty} \frac{\sqrt{y+1}}{y^3} dy \\ &= (1 - \xi_2 + O(\xi_2^2)) \left(1 + \xi_2 + \frac{1}{4}\xi_2^2 \log \xi_2 + o(\xi_2^2 \log \xi_2) \right) \\ &= 1 + \frac{1}{4}\xi_2^2 \log \xi_2 + o(\xi_2^2 \log \xi_2). \end{aligned} \tag{93}$$

This proves the lemma when $r = 2$.

Finally, assume that $r > 2$. Recall that $\xi_r \equiv \xi_r(\theta) = \frac{(\beta-1)\theta}{1-\beta\theta}$, so that $\theta = \frac{\xi_r}{\beta\xi_r + \beta - 1}$ and, by (92),

$$\begin{aligned} \mathbb{E} \left[\sqrt{\Xi(\theta, X_r)} \right] &= 2\beta \left(1 + \frac{\beta}{\beta-1}\xi_r \right)^{-1/2} \xi_r^\beta \int_{\sqrt{\xi_r+1}}^{\infty} \frac{x^2}{(x^2-1)^{\beta+1}} dx \\ &= \beta \left(1 + \frac{\beta}{\beta-1}\xi_r \right)^{-1/2} \xi_r^\beta \int_{\xi_r}^{\infty} \frac{\sqrt{y+1}}{y^{\beta+1}} dy. \end{aligned}$$

Let

$$h(y) = \frac{1 + \frac{1}{2}y - \sqrt{y+1}}{y^{\beta+1}}.$$

Note that $h(y) > 0$ for $y > 0$. As $y \rightarrow \infty$, $h(y) \sim \frac{1}{2}y^{-\beta}$. As $y \rightarrow 0$, $h(y) \sim \frac{1}{8}y^{1-\beta}$. Since $\beta \in (1, 2)$ for $r > 2$, the generalized integral $I_\beta := \beta \int_0^\infty h(y) dy$ is finite and positive.

Therefore,

$$\begin{aligned} \lim_{\xi_r \rightarrow 0} \frac{\beta\xi_r^\beta \int_{\xi_r}^{\infty} \frac{\sqrt{y+1}}{y^{\beta+1}} dy - 1 - \frac{\beta}{2(\beta-1)}\xi_r}{\beta\xi_r^\beta} &= \lim_{\xi_r \rightarrow 0} \int_{\xi_r}^{\infty} \frac{\sqrt{y+1}}{y^{\beta+1}} dy - \frac{1}{\beta}\xi_r^{-\beta} - \frac{1}{2(\beta-1)}\xi_r^{1-\beta} \\ &= \lim_{\xi_r \rightarrow 0} -\int_{\xi_r}^{\infty} h(y) dy = -\frac{I_\beta}{\beta}, \end{aligned}$$

and

$$\beta\xi_r^\beta \int_{\xi_r}^{\infty} \frac{\sqrt{y+1}}{y^{\beta+1}} dy = 1 + \frac{\beta}{2(\beta-1)}\xi_r - I_\beta\xi_r^\beta + o(\xi_r^\beta).$$

It follows that

$$\begin{aligned} \mathbb{E} \left[\sqrt{\Xi(\theta, X_r)} \right] &= \left(1 + \frac{\beta}{\beta-1}\xi_r \right)^{-1/2} \beta\xi_r^\beta \int_{\xi_r}^{\infty} \frac{\sqrt{y+1}}{y^{\beta+1}} dy \\ &= \left(1 - \frac{\beta}{2(\beta-1)}\xi_r + O(\xi_r^2) \right) \left(1 + \frac{\beta}{2(\beta-1)}\xi_r - I_\beta\xi_r^\beta + o(\xi_r^\beta) \right) \\ &= 1 - I_\beta\xi_r^\beta + o(\xi_r^\beta). \end{aligned} \tag{94}$$

This proves the lemma when $r > 2$. ■

F Appendix

Lemma F.1.

$$\lim_n D_n = 0, \quad (95)$$

where D_n is defined in (61).

Proof. Fix $\epsilon > 0$. Since $\Delta_2(z) = o(z^2 \log z)$ there exists $z_\epsilon > 0$ such that for all $0 < z < z_\epsilon$, $\left| \frac{\Delta_2(z)}{z^2 \log z} \right| < \epsilon$. Since for all n such that $\frac{\delta}{\phi(n)-2\delta} < z_\epsilon$ we have $\xi_2 = \frac{\delta e^{-\mu v}}{\phi(n)-2\delta e^{-\mu v}} < z_\epsilon$ for all $v \geq 0$ (Hint: the mapping $v \rightarrow \xi_2$ is nonincreasing in $[0, \infty)$ and $\xi_2 = \frac{\delta}{\phi(n)-2\delta}$ when $v = 0$), we conclude that for n large enough,

$$\sup_{v \geq 0} \left| \frac{\Delta_2(\xi_2)}{\xi_2^2 \log \xi_2} \right| < \epsilon. \quad (96)$$

Hence, for n large enough,

$$\begin{aligned} |D_n| &\leq \left(\frac{1}{4} + \epsilon \right) \int_0^\infty \lambda e^{-(\lambda+2\mu)v} \left| \frac{\log \xi_2}{(\phi(n) - 2\delta e^{-\mu v})^2} \right| dv \\ &= \left(\frac{1}{4} + \epsilon \right) \int_0^\infty \lambda e^{-(\lambda+2\mu)v} \left| \frac{\log(\phi(n) - 2\delta e^{-\mu v}) + \mu v - \log \delta}{(\phi(n) - 2\delta e^{-\mu v})^2} \right| dv. \end{aligned} \quad (97)$$

For n large enough

$$\begin{aligned} a_n(v) &:= \left| \frac{\log(\phi(n) - 2\delta e^{-\mu v}) + \mu v - \log \delta}{(\phi(n) - 2\delta e^{-\mu v})^2} \right| \\ &\leq \mu v - \log \delta + \frac{|\log(\phi(n) - 2\delta e^{-\mu v})|}{(\phi(n) - 2\delta e^{-\mu v})^2} \end{aligned} \quad (98)$$

for all $v \geq 0$. It is easy to check that for all n such that $\phi(n) > 2\delta + \sqrt{e}$, the mapping $v \rightarrow \frac{\log(\phi(n) - 2\delta e^{-\lambda v})}{(\phi(n) - 2\delta e^{-\lambda v})^2}$ is non-decreasing in $[0, \infty)$. Therefore, for all n such that $\phi(n) > 2\delta + \sqrt{e}$,

$$\frac{\log(\phi(n) - 2\delta e^{-\mu v})}{(\phi(n) - 2\delta e^{-\mu v})^2} \leq \frac{\log(\phi(n) - 2\delta)}{(\phi(n) - 2\delta)^2} \quad \text{for all } v \geq 0.$$

This shows that for all n such that $\phi(n) > 2\delta + \sqrt{e}$ [Hint: $\phi(n) - 2\delta e^{-\lambda v} > 1$ for all $v \geq 0$ when $\phi(n) > 2\delta + \sqrt{e}$]

$$\begin{aligned} \frac{|\log(\phi(n) - 2\delta e^{-\mu v})|}{(\phi(n) - 2\delta e^{-\mu v})^2} &= \frac{\log(\phi(n) - 2\delta e^{-\mu v})}{(\phi(n) - 2\delta e^{-\mu v})^2} \\ &\leq \frac{\log(\phi(n) - 2\delta)}{(\phi(n) - 2\delta)^2} \quad \forall v \geq 0. \end{aligned} \quad (99)$$

We conclude from (98) and (99) that for n large enough [Hint: for n large enough, $\log(\phi(n) - 2\delta)/(\phi(n) - 2\delta)^2 < 1$ since $\log t/t^2 \rightarrow 0$ as $t \rightarrow \infty$ and $\phi(n) \rightarrow \infty$ as $n \rightarrow \infty$]

$$0 \leq a_n(v) \leq \mu v + 1 - \log \delta \quad \text{for all } v \geq 0.$$

Since for every $v \geq 0$, $a_n(v) \rightarrow 0$ as $n \rightarrow \infty$ (cf. (98)), and

$$\int_0^\infty \lambda e^{-(\lambda+2\mu)v} (\mu v + 1 - \log \delta) dv = \frac{\lambda \mu}{(\lambda + 2\mu)^2} + \frac{\lambda(1 - \log \delta)}{\lambda + 2\mu} < \infty,$$

we may apply the Bounded Convergence Theorem to the sequence $\{a_n(v)\}_n$, to get from (97) that

$$\lim_n |D_n| \leq \left(\frac{1}{4} + \epsilon\right) \lim_n \int_0^\infty \lambda e^{-(\lambda+2\mu)v} a_n(v) dv = \left(\frac{1}{4} + \epsilon\right) \int_0^\infty \lambda e^{-(\lambda+2\mu)v} \lim_n a_n(v) dv = 0.$$

This concludes the proof of the lemma. ■

Lemma F.2. *Assume that $T(n) = \mathcal{O}(\sqrt{n/\log n})$. Then nD_n is bounded as $n \rightarrow \infty$.*

Proof. Define

$$f_n(v) := -\frac{n \log \xi_2}{(\phi(n) - 2\delta e^{-\mu v})^2} \left(\frac{1}{4} + \frac{\Delta_2(\xi_2)}{\xi_2^2 \log \xi_2}\right),$$

so that (cf. (61))

$$nD_n = -\int_0^\infty \lambda e^{-(\lambda+2\mu)v} f_n(v) dv. \quad (100)$$

Since $\xi_2 = \frac{\delta e^{-\mu v}}{\phi(n) - 2\delta e^{-\mu v}}$, $f_n(v)$ rewrites

$$f_n(v) = \frac{n(\log(\phi(n) - 2\delta e^{-\mu v}) - \log \delta + \mu v)}{(\phi(n) - 2\delta e^{-\mu v})^2} \left(\frac{1}{4} + \frac{\Delta_2(\xi_2)}{\xi_2^2 \log \xi_2}\right). \quad (101)$$

For any $v \geq 0$ and $\delta \in (0, 1)$, we see from (101) that

$$f_n(v) \geq \frac{n(\log \phi(n) - 2)}{(\phi(n) - 2\delta e^{-\mu v})^2} \left(\frac{1}{4} + \frac{\Delta_2(\xi_2)}{\xi_2^2 \log \xi_2}\right). \quad (102)$$

Thanks to assumption (59), $\log \phi(n) > 2$ for n large enough. Also, by (96) $\sup_{v \geq 0} \frac{\Delta_2(\xi_2)}{\xi_2^2 \log \xi_2}$ can be made arbitrarily small by letting $n \rightarrow \infty$. These two properties combined show from (102) that $f_n(v) \geq 0$ for n large enough.

On the other hand, from (96) and (101) we see that, for n large enough,

$$f_n(v) \leq \left(\frac{1}{4} + \epsilon\right) \frac{n \log \phi(n) - \log \delta + \mu v}{(\phi(n) - 2)^2}, \quad \forall v \geq 0.$$

Therefore, for n large enough ,

$$\begin{aligned} 0 &\leq \int_0^\infty \lambda e^{-(\lambda+2\mu)v} f_n(v) dv \\ &\leq \left(\frac{1}{4} + \epsilon\right) \int_0^\infty \lambda e^{-(\lambda+2\mu)v} \frac{n \log \phi(n) - \log \delta + \mu v}{(\phi(n) - 2)^2} dv \\ &= \left(\frac{1}{4} + \epsilon\right) \left(\frac{\lambda}{\lambda + 2\mu}\right) \frac{n \log \phi(n) - \log \delta + \mu/(\lambda + 2\mu)}{(\phi(n) - 2)^2} \\ &\sim_n \left(\frac{1}{4} + \epsilon\right) \left(\frac{\lambda}{\lambda + 2\mu}\right) \frac{n \log \phi(n)}{\phi(n)^2}, \end{aligned} \quad (103)$$

by using that $\lim_n \phi(n) = \infty$. We are left with finding ϕ such that $\frac{n \log \phi(n)}{\phi(n)^2} = \mathcal{O}(1)$.

To this end let $T(n) = \mathcal{O}(\sqrt{n/\log n})$ or, equivalently by (58), $\liminf_n \frac{\phi(n)}{\sqrt{n \log n}} = a$ for some $a > 0$. Let us write $n \log \phi(n)/\phi(n)^2$ as

$$\frac{n \log \phi(n)}{\phi(n)^2} = \frac{1}{(\phi(n)/\sqrt{n \log n})^2} \left(\frac{\log(\phi(n)/\sqrt{n \log n})}{\log n} + \frac{\log(\log n)}{\log n} + \frac{1}{2}\right). \quad (104)$$

Assume that $\limsup_n \frac{\phi(n)}{\sqrt{n \log n}} = b < \infty$. Then,

$$\limsup_{n \rightarrow \infty} \frac{n \log \phi(n)}{\phi(n)^2} \leq \frac{1}{a^2} \limsup_{n \rightarrow \infty} \left(\frac{\log(\phi(n)/\sqrt{n \log n})}{\log n} + \frac{\log(\log n)}{\log n} + \frac{1}{2} \right) = \frac{1}{2a^2} < \infty$$

by using $\lim_{x \rightarrow \infty} \frac{\log x}{x} = 0$. Assume now that $\limsup_n \frac{\phi(n)}{\sqrt{n \log n}} = \infty$. By rewriting (104) as

$$\frac{n \log \phi(n)}{\phi(n)^2} = \frac{1}{\log n} \cdot \frac{\log(\phi(n)/\sqrt{n \log n})}{(\phi(n)/\sqrt{n \log n})^2} + \frac{1}{(\phi(n)/\sqrt{n \log n})^2} \left(\frac{\log(\log n)}{\log n} + \frac{1}{2} \right),$$

we immediately conclude that $\lim_n \frac{n \log \phi(n)}{\phi(n)^2} = 0$ thanks again to $\lim_{x \rightarrow \infty} \frac{\log x}{x} = 0$. This shows that $\frac{n \log \phi(n)}{\phi(n)^2} = \mathcal{O}(1)$ when $T(n) = \mathcal{O}(\sqrt{n/\log n})$, and therefore by (103) and (100), that

$$nD_n = \mathcal{O}(1), \tag{105}$$

when $T(n) = \mathcal{O}(\sqrt{n/\log n})$, which completes the proof. \blacksquare

Lemma F.3.

$$\lim_n E_n = 0, \tag{106}$$

where E_n is defined in (65).

Proof. Fix $\epsilon > 0$. Since $\Delta_r(z) = o(z^\beta)$ when $r > 2$, there exists $z_\epsilon > 0$ such that for all $0 < z < z_\epsilon$, $\left| \frac{\Delta_r(z)}{z^\beta} \right| < \epsilon$. Since for all n such that $\frac{\delta(\beta-1)}{\phi(n)-\delta\beta} < z_\epsilon$ we have $\xi_r = \frac{\delta(\beta-1)}{e^{\mu v} \phi(n) - \delta\beta} < z_\epsilon$ for all $v \geq 0$ (Hint: the mapping $v \rightarrow \xi_r$ is nonincreasing in $[0, \infty)$ and $\xi_r = \frac{\delta(\beta-1)}{\phi(n)-\delta\beta}$ when $v = 0$), we conclude that for n large enough,

$$\sup_{v \geq 0} \left| \frac{\Delta_r(\xi_r)}{\xi_r^\beta} \right| < \epsilon. \tag{107}$$

Hence, for n large enough,

$$\begin{aligned} |E_n| &\leq (I_\beta + \epsilon) \int_0^\infty \frac{\lambda e^{-\lambda v}}{(\phi(n)e^{\mu v} - \delta\beta)^\beta} dv \\ &\leq \frac{I_\beta + \epsilon}{(\phi(n) - \delta\beta)^\beta} \int_0^\infty \lambda e^{-\lambda v} dv \\ &= \frac{I_\beta + \epsilon}{(\phi(n) - \delta\beta)^\beta} \rightarrow 0 \quad \text{as } n \rightarrow \infty. \end{aligned}$$

Lemma F.4. Assume that $T(n) = \mathcal{O}(n^{\mu_2/\mu_1})$ with $\mu_1 > 2\mu_2$. Then E_n defined in (65) is bounded as $n \rightarrow \infty$.

Proof. Define

$$k_n(v) = n \frac{I_\beta - \Delta_r(\xi_r)/\xi_r^\beta}{(e^{\mu\beta v} \phi(n) - \delta\beta)^\beta}, \tag{108}$$

where I_β is defined in (54). With this new function we can rewrite nE_n (cf. (65)) as

$$nE_n = \int_0^\infty \lambda e^{-\lambda v} k_n(v) dv. \tag{109}$$

Notice that

$$\begin{aligned} k_n(v) &= \frac{I_\beta - \Delta_r(\xi_r)/\xi_r^\beta}{(e^{\mu\beta v}\phi(n)/n^{1/\beta} - \delta\beta/n^{1/\beta})^\beta} \\ &\leq \frac{I_\beta - \Delta_r(\xi_r)/\xi_r^\beta}{(\phi(n)/n^{1/\beta} - \delta\beta/n^{1/\beta})^\beta}, \end{aligned} \quad (110)$$

for all $n \geq 1$ and $v \geq 0$. Recall that $I_\beta > 0$. Let $\epsilon < I_\beta$ in (107). From (110) we see that for n large enough

$$0 \leq k_n(v) \leq \frac{I_\beta + \epsilon}{(\phi(n)/n^{1/\beta} - \delta\beta/n^{1/\beta})^\beta} \quad \text{for all } v \geq 0. \quad (111)$$

Recall that $\beta = \frac{r}{r-1}$ yielding $r = \frac{\beta}{\beta-1}$. Assume that $T(n) = \frac{\delta}{n\phi(n)} = \mathcal{O}(n^{1/r})$ for $\delta \in (0, 1)$, or equivalently

$$\liminf_n \frac{\phi(n)}{n^{1/\beta}} = b$$

for some $b > 0$. From (111) we obtain

$$\begin{aligned} 0 &\leq \limsup_n \int_0^\infty \lambda e^{-\lambda v} k_n(v) dv \\ &\leq \limsup_n \frac{I_\beta + \epsilon}{(\phi(n)/n^{1/\beta} - \delta\beta/n^{1/\beta})^\beta} \\ &= \frac{I_\beta + \epsilon}{(\liminf_n \phi(n)/n^{1/\beta})^\beta} = \frac{I_\beta + \epsilon}{b^\beta}. \end{aligned} \quad (112)$$

This shows that $nE_n \in \mathcal{O}(1)$. ■

G Appendix

Lemma G.1. For any $\theta \in [0, 1]$, $r' \geq r$,

$$\mathbb{E} \left[\sqrt{\Xi(\theta, X_{r'})} \right] \leq \mathbb{E} \left[\sqrt{\Xi(\theta, X_r)} \right].$$

Proof. Fix $\theta \in [0, 1]$. When $r' \geq r$ then $X_{r'} \leq_{st} X_r$, which in turn implies that $\Xi(\theta, X_{r'}) \leq_{st} \Xi(\theta, X_r)$ as the mapping $x \rightarrow \Xi(\theta, x)$ in (50) is nondecreasing in $[0, \infty)$,

Therefore,

$$\mathbb{E} \left[\sqrt{\Xi(\theta, X_{r'})} \right] \leq \mathbb{E} \left[\sqrt{\Xi(\theta, X_r)} \right],$$

as the mapping $x \rightarrow \sqrt{x}$ is nondecreasing in $[0, \infty)$. ■

H Appendix: Proof of Proposition 5.5

Recall that under the IEBP policy the system behaves as an M/G/1 queue with an exceptional first job in each busy period. The service times of first jobs in busy periods have pdf \tilde{f}_1 and the service times of the other jobs have pdf g_1 . The numbers of jobs served in different busy periods are iid rvs, characterized by the random variable M , so that the expected number of Willie jobs served during n W-BPs is $T_W(n) = n\mathbb{E}[M]$.

Let us calculate $\mathbb{E}[M]$. To this end, introduce $\mathcal{G}_M(z) = \mathbb{E}[z^M]$, $|z| \leq 1$, the generating function of the number of jobs served in a busy period. Recall (Section 4) that the reconstructed service times of the first job served in different busy periods are iid rvs, and let Y be a generic reconstructed service time. Let $\tau^*(s) = \mathbb{E}[e^{-Ys}] = \int_0^\infty e^{-sx} \tilde{f}_1(x) dx$ be the LST of the reconstructed service time. Since the LTS of the service times all the other Willie jobs in a W-BP is $G_1^*(s)$, we get from [4]

$$\mathcal{G}_M(z) = z\tau^*(\lambda(1-d(z))), \quad |z| \leq 1, \quad (113)$$

where $d(z)$ is the root with the smallest modulus of the equation $t = zG_1^*(\lambda(1-t))$.

Noting that $d(1) = 1$ and $\frac{d}{d(z)}|_{z=1} = \frac{1}{1-\lambda/\mu_1}$, we obtain from (113)

$$\mathbb{E}[M] = \frac{1 - \lambda/\mu_1 + \lambda\mathbb{E}[Y]}{1 - \lambda/\mu_1}, \quad (114)$$

provided that the stability condition $\lambda/\mu_1 < 1$ holds. It remains to find $\mathbb{E}[Y]$. For that, we will use the identity $\mathbb{E}[Y] = -\frac{d\tau^*(s)}{ds}|_{s=0}$. But before that we need to calculate $\tau^*(s)$.

When IEBP is enforced (or, equivalently, under H_1) we know that Y has pdf \tilde{f}_1 (see Section 4.2). Multiplying both sides of (46) by $g_1(x)$ and using the definition of $\tilde{Z}(q, x)$ in (35) along with (47) gives

$$\tilde{f}_1(x) = (1 - qp)g_1(x) + q(g_1 * \hat{g}_2)(x),$$

where $\hat{g}_2(x)$ is defined in (48). Therefore,

$$\begin{aligned} \tau^*(s) &= \int_0^\infty e^{-sx} \tilde{f}_1(x) dx = \int_0^\infty e^{-sx} [(1 - pq)g_1(x) + q(g_1 * \hat{g}_2)(x)] dx \\ &= G_1^*(s) \left(1 - pq + q \int_0^\infty e^{-st} \hat{g}_2(t) dt \right). \end{aligned} \quad (115)$$

Differentiating (115) with respect to s at $s = 0$ and using the identity² $\int_0^\infty \hat{g}_2(t) dt = p$ yields

$$\begin{aligned} \mathbb{E}[Y] &= \frac{1 - pq}{\mu_1} + \frac{q}{\mu_1} \int_0^\infty \hat{g}_2(t) dt + q \int_0^\infty t \hat{g}_2(t) dt \\ &= \frac{1}{\mu_1} + q \int_0^\infty t \hat{g}_2(t) dt, \end{aligned}$$

By (114),

$$\mathbb{E}[M] = \frac{1}{1 - \lambda/\mu_1} + \frac{\lambda q}{1 - \lambda/\mu_1} \int_0^\infty t \hat{g}_2(t) dt, \quad (116)$$

and

$$T_W(n) = n\mathbb{E}[M] = \frac{n}{1 - \lambda/\mu_1} + \frac{\lambda q n}{1 - \lambda/\mu_1} \int_0^\infty t \hat{g}_2(t) dt. \quad (117)$$

Now we upper bound the integral in (116) and (117). We have

$$\begin{aligned} \int_0^\infty t \hat{g}_2(t) dt &= \int_0^\infty \lambda e^{-\lambda v} \int_0^\infty t g_2(v+t) dv dt \\ &\leq \int_0^\infty \lambda e^{-\lambda v} \int_0^\infty (t+v) g_2(v+t) dv dt \\ &= \int_0^\infty \lambda e^{-\lambda v} \int_v^\infty u g_2(u) du dv \end{aligned}$$

² $\int_0^\infty \hat{g}_2(t) dt = \int_0^\infty \lambda e^{-\lambda v} \int_v^\infty g_2(t) dt dv = \int_0^\infty \lambda e^{-\lambda v} (1 - G_2(v)) dv = 1 - G_2^*(\lambda) = p$.

$$\leq \int_0^\infty \lambda e^{-\lambda v} \int_0^\infty u g_2(u) du dv = \frac{1}{\mu_2}.$$

Hence, $\mathbb{E}[M] \leq \frac{1+q\lambda/\mu_2}{1-\lambda/\mu_1}$ and $T_W(n) \leq n \left(\frac{1+q\lambda/\mu_2}{1-\lambda/\mu_1} \right)$. This shows the upper bound in (28). The lower bound is trivial.

If $g_2(x) = \mu_2 e^{-\mu_2 x}$ then from (48) and (4) we find $p = \frac{\lambda}{\mu_2 + \lambda}$ and $\hat{g}_2(x) = \frac{\lambda \mu_2}{(\lambda + \mu_2)} e^{-\mu_2 x} = p \mu_2 e^{-\mu_2 x}$, which yields (29).