**Estimating the input of a Lévy queue
by Poisson sampling of the workload process**

L. Ravner, O. Boxma, M. Mandjes

Estimating the input of a Lévy queue by Poisson sampling of the workload process

Liron Ravner[1,2], Onno Boxma[2], and Michel Mandjes[1]

[1]University of Amsterdam
[2]Eindhoven University of Technology

August 14, 2018

Abstract

This paper aims at semi-parametrically estimating the input process to a Lévy-driven queue by sampling the workload process at Poisson times. We construct a method-of-moments based estimator for the Lévy process' characteristic exponent. This method exploits the known distribution of the workload sampled at an exponential time, thus taking into account the dependence between subsequent samples. Verifiable conditions for consistency and asymptotic normality are provided, along with explicit expressions for the asymptotic variance. The method requires an intermediate estimation step of estimating a constant (related to both the input distribution and the sampling rate); this constant also features in the asymptotic analysis. For subordinator Lévy input, a partial MLE is constructed for the intermediate step and we show that it satisfies the consistency and asymptotic normality conditions. For general spectrally-positive Lévy input a biased estimator is proposed that only uses workload observations above some threshold; the bias can be made arbitrarily small by appropriately choosing the threshold.

1 Introduction

To optimally design and control queueing systems, it is of great importance to have reliable estimates of the model primitives such as the arrival rate and the service-time distribution. In many situations, however, one cannot observe the users' interarrival times and service times; instead, one only has periodic observations of the workload process. This leaves us with the challenging task of inferring the model primitives from such workload observations. A complicating factor is that subsequent observations are dependent. As a consequence, standard estimation techniques (such as maximum-likelihood estimation) are not applicable in almost all queuing systems.

This work focuses on the problem described above in the setting of a single-server queue with work arriving according to a *spectrally-positive Lévy* inputs process, i.e., a Lévy process $X(\cdot)$ with only positive jumps. The resulting model includes the M/G/1 queue as a special case; then the input process is a Compound Poisson (CP) process. The set of spectrally-positive Lévy input processes, however, is substantially broader than CP: it covers a broad range of other relevant processes, such as Brownian motion, the Gamma process and the Inverse Gaussian process. The Lévy input process is uniquely defined by its *characteristic exponent* function

1

$\varphi(\alpha) := \log \mathbb{E} e^{-\alpha X(1)}$, with $\alpha \geq 0$ [26]. We assume that $\varphi(\cdot)$ is unknown and it is therefore the objective to estimate this function. For a detailed account of Lévy-driven queues we refer to [11].

Approach. In our approach we exploit the fact that, given some initial workload, there is an explicit expression for the Laplace-Stieltjes Transform (LST) of the workload after an exponentially distributed time; see, e.g. [25]. When the workload is sampled at random times according to an independent Poisson process (with some specified rate ξ), this LST enables us to deal with the dependence between the subsequent workload observations.In particular, method of moments estimation is carried out by equating the empirical and expected LST, thus obtaining so called *Z-estimators* [33]. Our setting can be viewed as non-parametric in the sense that we do not make specific distributional assumptions about the input process except that it is a spectrally-positive Lévy process. However, we use parametric estimation techniques for pointwise estimation of the exponent function.

There is the complication that, due to the specific form of the above-mentioned LST, it is necessary to first estimate $\varphi^{-1}(\xi)$, i.e., the inverse of φ at the sampling rate ξ. This intermediate step is also crucial for determining the asymptotic performance of the estimator for $\varphi(\alpha)$. In particular, we show that consistency and asymptotic normality in the intermediate step (to estimate $\varphi^{-1}(\xi)$, that is) carry over to consistency and asymptotic normality in the main estimation step (to estimate $\varphi(\alpha)$). For the case of subordinator input (i.e., a Lévy process whose paths are non-decreasing almost surely) we construct a maximum-likelihood procedure for estimating $\varphi^{-1}(\xi)$, which we show to be consistent and asymptotically normal. This result therefore covers the special cases of the driving Lévy process being CP, a Gamma process and an Inverse Gaussian process. For general spectrally-positive Lévy input, i.e., also including a Brownian motion component, we provide an estimation method for $\varphi^{-1}(\xi)$ that uses workload observations only above some threshold; this procedure is inherently biased, but the bias can be controlled by choosing the threshold parameter appropriately.

Later in this introduction we describe in detail the differences with existing estimation approaches. A major difference is that previous papers tend to aim at using workload observations to estimate quantities pertaining to the steady-state workload distribution, from which the model primitives are then inferred, whereas we explicitly use the knowledge of the dependence between two subsequent workload observations.

Main contributions. We now summarize the main contributions of our work.

- Thus far, the literature has mostly focused on the case of Compound Poisson input, whereas we address the larger class of spectrally-positive Lévy processes.

- In our approach we do not rely on estimating quantities pertaining to the steady-state workload distribution, but rather explicitly use the system's correlation structure.

- We provide verifiable conditions for consistency and asymptotic normality of our estimator (including an expression for the asymptotic variance). These conditions are verified for the case of subordinator input, thus implying the asymptotic properties for the special case of CP. When $\varphi(\alpha)$ is estimated for multiple values of α simultaneously, we provide the corresponding asymptotic covariances.

- The idea to use a Z-estimator in this Lévy-driven setting is novel. As mentioned, our approach requires knowledge of $\varphi^{-1}(\xi)$, for which we present an estimation procedure.

- We present simulation examples that explore the accuracy of the procedure. These show that the accuracy is generally good, but degrades for large α (because the asymptotic variance of the error increases with α). We also observe that the accuracy is best for low ξ, which is to be expected: when ξ is high, the workload process is very frequently sampled, such that new observations offer hardly any new information. As mentioned, for non-subordinator input the estimator of $\varphi^{-1}(\xi)$ is biased, but our experiments show that the accuracy is high, even for a relatively low threshold.

Background and related literature. Observing a system at Poisson times is commonly known as *Poisson probing*. In communication networks it has been used to estimate (steady-state) packet delays in a network with unknown traffic input [2, 32]. In a practical implementation of Poisson probing, very small jobs are sent through the system and their delays are recorded. This method relies on the *Poisson Arrivals See Time Averages* (PASTA) property: the stationary distribution coincides with the distribution observed at Poisson instants; theoretical justification can be found in e.g. [14, 34].

In the M/G/1 case the queue is fed by a Poisson arrival process with rate λ, whereas the service times are i.i.d. random variables $G_i \sim G$ with LST G^*. In [8] a non-parametric approach is proposed for estimating λ and the LST of G using observations of the lengths of the busy and idle periods. This procedure overcomes the issue of dependent observations as busy periods are independent, but it may require a very long observation period to get a large sample of busy periods, especially in heavily loaded systems. The estimator exploits the fixed-point equation $B^*(\alpha) = G^*(\alpha + \lambda(1 - B^*(\alpha)))$, where B^* is the LST of the busy period, and uses an estimator $\hat{G}_n(\alpha)$ based on the empirical LST of the busy periods (with n observations). The estimator is consistent and asymptotically normal: $\sqrt{n}(\hat{G}_n(\cdot) - G^*(\cdot))$ converges in distribution to a Gaussian process. Reference [19] provides a nonparametric procedure for estimating the service-time distribution, using an equidistant sample of the workload. It consists of two steps: (1) estimating the residual service-time distribution by inverting the convolution workload formula, $F_V = (1 - \rho) \sum_{k=0}^{\infty} \rho^k G_e^{\star k}$, where F_V is the stationary cdf of the workload and $G_e^{\star k}$ is the k-fold convolution of the residual service time distribution, and (2) then estimating the service-time distribution itself. As the first step relies on the stationary workload distribution, the procedure does not use any knowledge on the dependence between observations. Nevertheless, the estimation of the residual service-time distribution is shown to be consistent and asymptotically normal under some conditions on the system load. For the second step only heuristic methods are suggested. The numerical experiments in [19] show that the procedure performs well for moderately loaded systems but often performs poorly for systems with low or high load.

A general framework for generalized method-of-moments (GMM) estimators for Markov processes sampled according to a Poisson process, that is possibly state and time dependent, is provided in [12]. As it turns out, the fact that in our setup the quantity $\varphi^{-1}(\xi)$ has to be estimated entails that we cannot use this approach. In Section 7 we further discuss the relation of our work with the GMM-based procedure of [12] in the context of our model.

Non-parametric estimation of the stationary waiting time distribution for a GI/G/1 queue was studied in [29]. If the number of arrivals during a service time can be observed, then estimation of the service-time distribution reduces to the problem of decompounding Poisson sums, as studied in [10]. As is often the case, when moving from a single-server queue to its infinite-server counterpart, the analysis greatly simplifies. The object of interest in this line of

3

research is typically the service time distribution, and we refer the reader to [8, 9, 15, 16, 31]. A related approach appeared in [28] where an estimator for the arrival rate is derived by relying on counting the number of arrivals during a busy period. In [30] diffusion approximations are used.

Estimation of the model primitives from queue observations also has important managerial applications, for example in the contexts of demand estimation [4] and dynamic pricing [1]. In [3] the traffic intensity of a non-homogeneous-in-time queue was estimated using a (non-Poisson) probing scheme that depends on the service time of the probes. This list of papers on statistical inference of queueing systems is by no means exhaustive; see [5] for an extensive bibliography of this strand of research.

In the present paper we estimate (the characteristic exponent of) a Lévy process by observing a Lévy-driven queue. Various papers consider the counterpart in which the Lévy process itself is observed. We refer to e.g. [7, 27]; [17] studies the case of a Lévy process without small jumps.

Paper organization. The remainder of the paper is organized as follows. Section 2 introduces the queueing model and sampling scheme. In Section 3 the estimator for $\varphi(\alpha)$ is defined and conditions for consistency and asymptotic normality are provided. The following two sections deal with the intermediate estimation step of $\varphi^{-1}(\xi)$. In Section 4 an MLE is constructed for the subordinator case and it is shown to satisfy the consistency and asymptotic normality conditions. Section 5 presents an approach for the intermediate step that relies only on workload observations above some threshold. Further, a bound is derived for the bias and explicitly computed in the limit when the sample size grows. In Section 6 we numerically analyze the performance of the estimators using simulated data. Section 7 contains conclusions and suggestions for further research. Finally, proofs and further technical details are provided in the appendices.

2 Model and preliminaries

We study a queue with a spectrally-positive Lévy input process $J(t)$. The distribution of the process can be represented by the exponent function,

$$\log \mathrm{E}e^{-\alpha J(1)} = -c\alpha + \frac{1}{2}\sigma^2\alpha^2 - \int_{(0,\infty)} (1 - e^{-\alpha x})\nu(\mathrm{d}x) , \tag{1}$$

where c and σ are non-negative constants and ν is the Lévy jump measure. The output of the system is a unit-rate negative linear drift. This means that the net input process is $X(t) = J(t) - t$, which is also a spectrally-positive Lévy process. Our objective is to estimate the (unknown) exponent function of the net input process, $\varphi(\alpha) := \log \mathrm{E}e^{-\alpha X(1)}$, with $\alpha \geq 0$. Let $(V(t))_{t\in\mathbb{R}}$ denote the workload process, which can be represented as the net input process reflected at zero: $V(t) = X(t) + \max\{V(0), L(t)\}$, where $L(t) := -\inf_{0 \leq s \leq t} X(s)$. If $\mathrm{E}X(1) < 0$, then the stationary distribution of $V := V(\infty)$ is given by the *generalized Pollaczek-Khintchine* (GPK) formula [11, p. 27]:

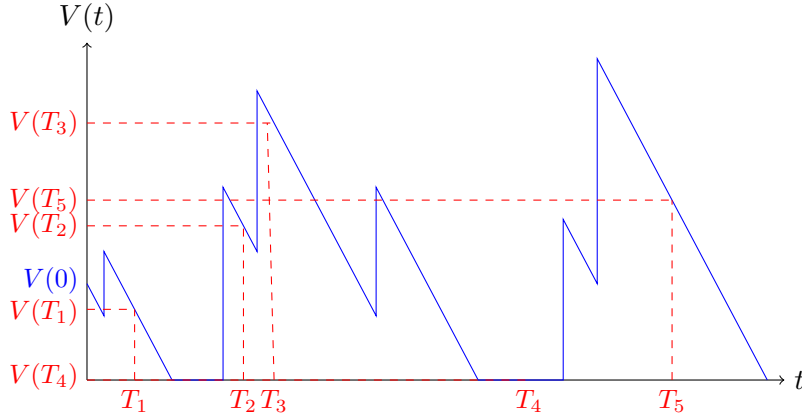$$\mathrm{E}e^{-\alpha V} = \frac{\alpha\varphi'(0)}{\varphi(\alpha)} . \tag{2}$$

4

Figure 1: The workload $V(t)$ of an M/G/1 queue is sampled according to the probing process at times $T_1, T_2, \ldots$. The jump sizes and interarrival times are not observed. Inference is made based on the sample $(V(T_1), V(T_2), \ldots)$.

Furthermore, if $\sigma = 0$ then the input process $J(t)$ is a subordinator (i.e., an almost surely non-decreasing Lévy process), and the stationary probability that the workload is 0 equals

$$\mathrm{P}(V = 0) = \varphi'(0) = -\mathrm{E}X(1) \ . \tag{3}$$

This is the is case if, for example, the input is CP (as in the M/G/1 model). Importantly, however, it also holds for spectrally-positive Lévy input without a Brownian component, i.e., a Lévy jump process with non-negative jumps (covering e.g. the Gamma process and the Inverse Gaussian process). Our transient sampling scheme does not require that $\mathrm{E}X(1) < 0$, but all our asymptotic results naturally do rely on this assumption.

Suppose that the workload is observed at n random times, sampled according to an independent Poisson process with rate ξ. Denote these times by $\{T_i\}_{i=1}^n$ such that $T_i - T_{i-1}$ are i.i.d. samples from an exponential distribution with parameter ξ (putting $T_0 := 0$). The sampling procedure is illustrated for the workload of an M/G/1 queue in Figure 1.

The distribution of the workload at T_i conditional on the workload at sample T_{i-1} is

$$\mathrm{E}\left[e^{-\alpha V(T_i)}|V(T_{i-1})\right] = \frac{\xi}{\xi - \varphi(\alpha)}\left(e^{-\alpha V(T_{i-1})} - \frac{\alpha}{\psi(\xi)}e^{-\psi(\xi)V(T_{i-1})}\right) \ , \tag{4}$$

see, e.g. [25] and [11, Ch. IV]; here ψ is the inverse of φ. For $\alpha = \psi(\xi)$, the expression (4) is not well defined as the denominator $\xi - \varphi(\psi(\xi))$ equals zero, but the LST can still be computed by applying L'Hôpital's rule. The conditional expected workload at observation time T_i is

$$\mathrm{E}\left(V(T_i)|V(T_{i-1})\right) = V(T_{i-1}) + \frac{e^{-\psi(\xi)V(T_{i-1})}}{\psi(\xi)} - \frac{\varphi'(0)}{\xi} \ , \tag{5}$$

and the conditional second moment is

$$\mathrm{E}\left[V(T_i)^2|V(T_{i-1})\right] = V(T_{i-1})^2 - 2\frac{\varphi'(0)}{\xi}\left(V(T_{i-1}) + \frac{e^{-\psi(\xi)V(T_{i-1})}}{\psi(\xi)} - \frac{\varphi'(0)}{\xi}\right) + \frac{\varphi''(0)}{\xi} \ . \tag{6}$$

5

The above holds for any Lévy-driven queue with spectrally-positive input. In the subordinator case, essentially due to the fact that the workload process spends time at zero with positive probability, by sending α to ∞ in (4), it turns out that

$$P(V(T_i) = 0 | V(T_{i-1}) = v) = \frac{\xi}{\psi(\xi)} e^{-\psi(\xi)v} . \tag{7}$$

The above results are in terms of LSTs; the joint distribution of the sample can in principle be evaluated by numerical inversion of the LSTs involved, but this is computationally demanding. Therefore, in the sequel we use a method-of-moments type estimator that requires finding the root of a function equating the empirical LST or the empirical moments to their respective expectations. The term $\psi(\xi) = \varphi^{-1}(\xi)$, which depends on the net input process as well as the sampling rate, appears in the above expressions. Estimating this quantity plays a crucial role in this work.

Notation. Much of this work involves large sample asymptotic analysis, and as such we will frequently make use of the following notations: $\rightarrow_{\mathbb{P}}$ for convergence in probability, $\rightarrow_{as}$ for almost sure convergence, $\rightarrow_d$ for convergence in distribution, and $\approx$ to indicate that two random sequences have the same limit in probability.

3   Semiparametric estimation of φ

This section proposes a Z-estimator for $\varphi(\alpha)$ for a given $\alpha > 0$. We then study the asymptotic properties of the estimator, in that we provide conditions for consistency and asymptotic normality given we know $\psi(\xi)$; the estimation of $\psi(\xi)$ is the topic of the next section.

For brevity, we throughout denote $V_i := V(T_i)$. Given a sample of workload observations at Poisson times, $(V_1, \ldots, V_n)$, let ψ_n denote a series of estimators for $\psi(\xi)$. An estimator for $\varphi(\alpha)$ is derived by equating the conditional LST (4) and the empirical LST and solving

$$\sum_{i=1}^{n} e^{-\alpha V_i} = \sum_{i=1}^{n} \frac{\xi}{\xi - \varphi(\alpha)} \left(e^{-\alpha V_{i-1}} - \frac{\alpha}{\psi_n} e^{-\psi_n V_{i-1}} \right) , \tag{8}$$

Letting

$$Z_i(\alpha) := e^{-\alpha V_i} - \frac{\xi}{\xi - \varphi(\alpha)} \left(e^{-\alpha V_{i-1}} - \frac{\alpha}{\psi_n} e^{-\psi_n V_{i-1}} \right) , \tag{9}$$

the estimator is given by the root of $\frac{1}{n} \sum_{i=1}^{n} Z_i(\alpha)$. We rearrange (8) to

$$\varphi(\alpha) = \xi \left[1 - \frac{1}{n} \sum_{i=1}^{n} \left(e^{-\alpha V_{i-1}} - \frac{\alpha}{\psi_n} e^{-\psi_n V_{i-1}} \right) \middle/ \frac{1}{n} \sum_{i=1}^{n} e^{-\alpha V_i} \right] , \tag{10}$$

yielding the estimator

$$\hat{\varphi}_n(\alpha; \psi_n) = \frac{\xi\alpha}{\psi_n} \left(\frac{\psi_n}{\alpha n} \left(e^{-\alpha V_n} - e^{-\alpha V_0} \right) + \frac{1}{n} \sum_{i=1}^{n} e^{-\psi_n V_{i-1}} \right) \middle/ \frac{1}{n} \sum_{i=1}^{n} e^{-\alpha V_i} . \tag{11}$$

This type of estimator is generally referred to as a Z-estimator [33, Ch. V]. Note that in our case the samples are dependent, thus preventing direct use of classical results on consistency and asymptotic normality. Another justification for this estimator is that it coincides with the estimator minimizing the sum of quadratic deviations of the empirical LST from their respective conditional expectations, i.e. minimizing the conditional mean square error (MSE).

Remark 1. This method can also be applied for parametric estimation of the moments. For example, an estimator for the first moment of the net input process, i.e., $\theta := \mathrm{E}X(1)$, can be derived. This is done by taking the derivative in (11):

$$\theta_n = -\hat{\varphi}'_n(0; \psi_n) = -\lim_{\alpha \downarrow 0} \frac{d}{d\alpha} \hat{\varphi}_n(\alpha) = -\frac{\xi}{n\psi_n} \sum_{i=1}^{n} e^{-\psi_n V_{i-1}} \ . \tag{12}$$

The same procedure can be followed for the (joint) estimation of higher moments. The estimators of k-th moment $\mathrm{E}X(1)^k$ follows from

$$\theta_n^{(k)} = (-1)^k \hat{\varphi}_n^{(k)}(0; \psi_n) \ . \tag{13}$$

3.1 Consistency

The estimator for $\varphi(\alpha)$ does generally not satisfy $\hat{\varphi}(\psi_n) = \xi$. However, Theorem 1 claims that consistency of the estimator ψ_n implies consistency of the estimator $\hat{\varphi}_n(\alpha; \psi_n)$.

Theorem 1. *If $\psi_n \to_{\mathbb{P}} \psi(\xi)$ as $n \to \infty$ then for every $\alpha > 0$ the pointwise estimator $\hat{\varphi}_n(\alpha; \psi_n)$ is consistent: $\hat{\varphi}_n(\alpha; \psi_n) \to_{\mathbb{P}} \varphi(\alpha)$ as $n \to \infty$.*

Proof. If $\mathrm{E}X(1) < 0$, then $(V(t))_{t \in \mathbb{R}}$ has a stationary distribution. By PASTA the limit distribution at Poisson sampling instants equals the stationary distribution of $(V(t))_{t \in \mathbb{R}}$, which we denote by V. Denoting the Poisson sampling process by $N(t)$, with $t_n := \inf\{t : N(t) = n - 1\}$,

$$\frac{1}{n} \sum_{i=1}^{n} e^{-\beta V_{i-1}} = \frac{1}{N(t_n)} \sum_{i=1}^{N(t_n)} e^{-\beta V(t_i)} \ . \tag{14}$$

As $t_n \to \infty$ when $n \to \infty$, Equation (2) implies that for any $\beta \geq 0$, as $n \to \infty$,

$$\frac{1}{n} \sum_{i=1}^{n} e^{-\beta V_{i-1}} = \frac{1}{N(t_n)} \sum_{i=1}^{N(t_n)} e^{-\beta V(t_i)} \to_{\mathbb{P}} \mathrm{E}e^{-\beta V} = \frac{\beta \varphi'(0)}{\varphi(\beta)} \ . \tag{15}$$

We define the right-continuous process

$$H(t) := \sum_{n=1}^{\infty} \mathbf{1}(t_n \leq t < t_{n+1})\psi_n, \ t \geq 0 \ , \tag{16}$$

and observe that $H(t_n) = \psi_n$ for every $n \geq 1$. By Slutsky's lemma [33, p. 11], $\psi_n V(t_n) = H(t_n)V(t_n) \to_{\mathbb{P}} \psi(\xi)V$ as $n \to \infty$. Hence, applying PASTA once more,

$$\frac{1}{N(t_n)} \sum_{i=1}^{N(t_n)} e^{-H(t_n)V(t_i)} \to_{\mathbb{P}} \frac{\psi(\xi)\varphi'(0)}{\varphi(\psi(\xi))} \ . \tag{17}$$

Furthermore, the term $\frac{\psi_n}{\alpha n} \left(e^{-\alpha V_n} - e^{-\alpha V_0} \right)$ converges to zero almost surely if the stability condition holds. Next we can use the fact that $\varphi(\psi(\xi)) = \varphi(\varphi^{-1}(\xi)) = \xi$ and apply the continuous mapping theorem (see [33], p. 7) to (11), thus obtaining

$$\hat{\varphi}_n(\alpha; \psi_n) \to_{\mathbb{P}} \left(\alpha \xi \frac{\varphi^{-1}(\xi)\varphi'(0)}{\varphi(\varphi^{-1}(\xi))} \right) \bigg/ \left(\varphi^{-1}(\xi) \frac{\alpha \varphi'(0)}{\varphi(\alpha)} \right) = \varphi(\alpha) \ . \tag{18}$$

Thus, the pointwise estimator of $\varphi(\alpha)$ is consistent for every $\alpha > 0$. $\qquad\square$

Remark 2. Consider the case of parametric estimation of the (first and higher) moments, as described in Remark 1. Using the same arguments it is straightforward to show that the estimators are consistent if ψ_n is consistent.

3.2   Asymptotic normality of $\hat{\varphi}_n$

In this section we establish, under specific conditions, asymptotic normality of our estimator $\hat{\varphi}_n(\alpha; \psi_n)$ as $n \to \infty$: we show that $\sqrt{n}(\hat{\varphi}_n(\alpha; \psi_n) - \varphi(\alpha))$ converges to a zero-mean normal random variable, for any fixed $\alpha > 0$. We find an expression for the corresponding asymptotic variance, so that one can assess the estimation error for large samples.

We then extend this result to a multivariate setting: we establish the joint normality of the vector $\hat{\mathbf{\Phi}}_n(\boldsymbol{\alpha}) \in \mathbb{R}^p$, where $p \geq 1$, the i-th component of which is defined by $\sqrt{n}\,(\hat{\varphi}_n(\alpha_i; \psi_n) - \varphi(\alpha_i))$ (with $i = 1, \ldots, p$). Each of the estimation equations (8) is solved independently for α_i, but they rely on the same sample and are therefore dependent.

We start the section by stating a Central Limit Theorem (CLT) for martingale difference sums. Lemma 2 below consists of two parts; both are essentially known results, that we customized to our needs. The first part is the univariate result of [18, Thm. 3.2]. The second part is a special case of the multidimensional extension of [24, Thm. 12.6]; here we remark that the result in [24] relates to continuous time, but as we are only interested in the discrete-time case we present a more concise statement without the quadratic variation terms.

All convergence statements in the following are as $n \to \infty$; we therefore omit this specification from now on.

Lemma 2.

(a) *Let M_n be a discrete-time martingale with respect to a filtration $\{\mathcal{F}_n\}_{n \geq 0}$, and denote*

$$Z_{ni} := \frac{M_i - M_{i-1}}{\sqrt{\mathrm{Var}(M_n)}}, \quad i = 1, \ldots, n. \tag{19}$$

If

$$\mathrm{E} \max_{1 \leq i \leq n} Z_{ni}^2 < \infty, \quad n = 1, 2, \ldots, \tag{20}$$

and

$$\max_{1 \leq i \leq n} |Z_{ni}| \to_{\mathbb{P}} 0, \tag{21}$$

then $\sum_{i=1}^n Z_{ni} \to_d N(0, 1)$.

(b) *Let*

$$\boldsymbol{M}_n^\top := \big(M_n^{(1)}, \ldots, M_n^{(p)}\big)$$

be a vector of discrete-time martingales, and let

$$\boldsymbol{Z}_n^\top := \Big(\sum_{i=1}^n Z_{ni}^{(1)}, \ldots, \sum_{i=1}^n Z_{ni}^{(p)} \Big),$$

with $Z_{ni}^{(k)}$, for $k = 1, \ldots, p$, the respective normalized differences. If these objects satisfy all conditions of (a)*, and*

$$\lim_{n \to \infty} \mathrm{E}\left[\boldsymbol{Z}_n \boldsymbol{Z}_n^\top \right] = S, \tag{22}$$

where $S \in \mathbb{R}^{p \times p}$ is a positive-definite matrix with finite elements, then $\boldsymbol{Z}_n \to_d N(0, S)$.

3.2.1 Univariate asymptotic normality

Recall that the estimation procedure also relies on an external estimator for $\psi(\xi)$. As the CLT below shows, the asymptotic distribution of $\sqrt{n}(\psi_n - \psi(\xi))$ and its joint distribution with the estimator for $\varphi(\alpha)$ play a crucial role in the analysis.

Theorem 3. *Suppose that* $\mathrm{E}X(1) < 0$ *and* $\psi_n - \psi(\xi) \approx \frac{1}{n}M_n + R_n$ *such that:*

(i) $\sqrt{n}R_n \to_{\mathbb{P}} 0$ *and* M_n *is a martingale with respect to* $(V_0, \ldots, V_n)$ *that satisfies the conditions of Lemma 2a,*

(ii) $\lim_{n\to\infty} \mathrm{Var}(M_n)/n = \sigma_\xi^2 < \infty,$

(iii) $\lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^n \mathrm{E}\left[Z_i(\alpha)M_n\right] < \infty,$

where $Z_i(\alpha)$ *is given by* (9). *Then* $\sqrt{n}(\psi_n - \psi(\xi)) \to_{\mathrm{d}} \mathrm{N}(0, \sigma_\xi^2)$, *and for every* $\alpha > 0$,

$$\sqrt{n}(\hat{\varphi}_n(\alpha; \psi_n) - \varphi(\alpha)) \to_{\mathrm{d}} \mathrm{N}\left(0, \sigma_{\alpha,\xi}^2\right) \ , \tag{23}$$

where $0 < \sigma_{\alpha,\xi}^2 < \infty$.

Remark 3. The normality assumptions involving the asymptotic estimation error of $\psi(\xi)$ have to be verified for the specific ψ_n used. Condition (i) entails that the estimation error satisfies a martingale CLT, condition (ii) specifies the convergence rate of the variance and can in principle be generalized to other rates, and condition (iii) further demands that the asymptotic covariance between estimation errors is bounded with the appropriate scaling. As it turns out, under some regularity conditions these assumptions are valid for various estimation procedures (see, e.g., [24]) . In the next section we present an MLE for $\psi(\xi)$ for the subordinator case, and show that it satisfies the conditions of Theorem 3.

Proof of Theorem 3. In this proof a number of lemmas are needed. These are stated after the end of the proof, and are proven in Appendix A.

We focus on the marginal asymptotic distribution of the estimation error for a fixed $\alpha > 0$. We denote the estimator by $\varphi_n := \hat{\varphi}_n(\alpha; \psi_n)$ and recall that the true value at α is $\varphi(\alpha)$.

The starting point is the estimation equation (8), which can be written as

$$J_n(\psi, \varphi) := \frac{1}{n}\sum_{i=1}^n \left[e^{-\alpha V_i} - \frac{\xi}{\xi - \varphi}\left(e^{-\alpha V_{i-1}} - \frac{\alpha}{\psi}e^{-\psi V_{i-1}}\right)\right] = 0 \ . \tag{24}$$

The standard *delta method* can be applied for deriving the asymptotic distribution of φ_n; see, e.g., [33, p. 51]. Concretely, taking a Taylor expansion of $J_n(\psi_n, \varphi_n)$, where ψ_n and φ_n are the estimators, around the true parameters $\psi(\xi)$ and $\varphi(\alpha)$, yields

$$J_n(\psi_n, \varphi_n) \approx J_n(\psi(\xi), \varphi(\alpha)) + (\psi_n - \psi(\xi))\frac{\partial}{\partial \psi}J_n(\psi(\xi), \varphi(\alpha)) + (\varphi_n - \varphi(\alpha))\frac{\partial}{\partial \varphi}J_n(\psi(\xi), \varphi(\alpha)) \ , \tag{25}$$

where the remainder will be argued to vanish, in a convergence-in-probability sense, in Lemma 6. As $J_n(\psi_n, \varphi_n) = 0$, we thus obtain

$$\sqrt{n}(\varphi_n - \varphi(\alpha)) \approx \frac{\sqrt{n}J_n(\psi(\xi), \varphi(\alpha)) + \sqrt{n}(\psi_n - \psi(\xi))\frac{\partial}{\partial \psi}J_n(\psi(\xi), \varphi(\alpha))}{-\frac{\partial}{\partial \varphi}J_n(\psi(\xi), \varphi(\alpha))} \ . \tag{26}$$

Lemma 4 states that partial derivatives featuring in the right-hand side of (26) converge almost surely to constants, which we call ∂J_ψ and ∂J_φ. Thus, by applying Slutsky's lemma to (26), $\sqrt{n}(\varphi_n - \varphi(\alpha))$ converges in distribution to the same limit as

$$-\frac{1}{\partial J_\varphi} \left(\sqrt{n}\, J_n(\psi(\xi), \varphi(\alpha)) + \sqrt{n}(\psi_n - \psi(\xi))\, \partial J_\psi \right) \; . \tag{27}$$

By condition (i),

$$\sqrt{n}(\psi_n - \psi(\xi)) \approx \frac{1}{\sqrt{n}} M_n = \sqrt{\frac{\mathrm{Var}(M_n)}{n}} \sum_{i=1}^{n} Z_{ni},$$

where $\sum_{i=1}^{n} Z_{ni}$ is a normalized sum of martingale differences that, due to Lemma 2a, converges to a standard normal random variable. We conclude, by condition (ii), that the limiting distribution of $\sqrt{n}(\psi_n - \psi(\xi))$ is normal. In Lemma 5a we further show that $\sqrt{n} J_n(\psi(\xi), \varphi(\alpha)) \to_d \mathrm{N}\left(0, \sigma_\alpha^2\right)$, where σ_α^2 is given by (33). In order to complete the proof of Theorem 3 it is now sufficient to prove that the limiting joint distribution of $\sqrt{n} J_n(\psi(\xi), \varphi(\alpha))$ and $\sqrt{n}(\psi_n - \psi(\xi))$ is (bivariate) normal, or equivalently that any linear combination of them is asymptotically normal. This is done in Lemma 5b by applying Lemma 2b, which is facilitated by the assumption of condition (iii). Finally, we conclude that the asymptotic variance of $\sqrt{n}(\varphi_n - \varphi(\alpha))$ is

$$\sigma_{\alpha,\xi}^2 := \frac{\sigma_\alpha^2 + 2\partial J_\psi \sigma_{\alpha,\psi}^2 + (\partial J_\psi)^2 \sigma_\xi^2}{(\partial J_\varphi)^2} \; , \tag{28}$$

where

$$\sigma_{\alpha,\psi}^2 := \lim_{n\to\infty} \mathrm{Cov}\left[\sqrt{n} J_n(\psi(\xi), \varphi(\alpha)), \sqrt{n}(\psi_n - \psi(\xi))\right] \; . \tag{29}$$

This proves the claimed asymptotic normality. $\qquad\square$

Remark 4. The explicit form of the term $\sigma_{\alpha,\psi}^2$ depends on the specific estimator ψ_n. In the sequel we derive an explicit expression for the asymptotic variance in the subordinator case when ψ_n is an MLE.

Lemma 4. *As $n \to \infty$,*

$$\frac{\partial}{\partial \psi} J_n(\psi(\xi), \varphi(\alpha)) \to_{\mathrm{as}} -\frac{\alpha \varphi'(0)}{(\xi - \varphi(\alpha))\psi'(\xi)} =: \partial J_\psi \; , \tag{30}$$

and

$$\frac{\partial}{\partial \varphi} J_n(\psi(\xi), \varphi(\alpha)) \to_{\mathrm{as}} -\frac{\alpha \varphi'(0)}{\varphi(\alpha)(\xi - \varphi(\alpha))} =: \partial J_\varphi \; . \tag{31}$$

Lemma 5.   (a) *For any $\alpha > 0$,*

$$\sqrt{n} J_n(\psi(\xi), \varphi(\alpha)) \to_d \mathrm{N}\left(0, \sigma_\alpha^2\right) \; , \tag{32}$$

where

$$\sigma_\alpha^2 := \frac{2\xi^2 \alpha \varphi'(0)}{(\xi - \varphi(\alpha))^2} \left(\frac{\left(\frac{\varphi(\alpha)}{\xi}\right)^2 - \frac{2\varphi(\alpha)}{\xi}}{\varphi(2\alpha)} - \frac{\alpha}{\psi(\xi)\varphi(2\psi(\xi))} + \frac{\alpha + \psi(\xi)}{\psi(\xi)\varphi(\alpha + \psi(\xi))} \right) \; . \tag{33}$$

(b) *Furthermore, under the conditions of Theorem 3,*

$$\sqrt{n} \begin{pmatrix} J_n(\psi(\xi), \varphi(\alpha)) \\ \\ \psi_n - \psi(\xi) \end{pmatrix} \rightarrow_d N \left(\mathbf{0}, \begin{pmatrix} \sigma_\alpha^2 & \sigma_{\alpha,\psi}^2 \\ \\ \sigma_{\alpha,\psi}^2 & \sigma_\xi^2 \end{pmatrix} \right) . \tag{34}$$

Consequently,

$$\sqrt{n} \left(J_n(\psi(\xi), \varphi(\alpha)) + \partial J_\psi(\psi_n - \psi(\xi)) \right) \rightarrow_d N \left(0, \sigma_\alpha^2 + 2\partial J_\psi \, \sigma_{\alpha,\psi}^2 + (\partial J_\psi)^2 \, \sigma_\xi^2 \right) . \tag{35}$$

Lemma 6. *The remainder term of the Taylor expansion in* (25) *converges in probability to zero.*

3.2.2 Multivariate asymptotic normality

We now consider the situation that $\varphi(\alpha)$ is estimated for multiple values of α simultaneously. The main result is that $\hat{\mathbf{\Phi}}_n(\boldsymbol{\alpha})$ converges to a multivariate normal random variable, under the conditions of Theorem 3. The proof combines the lines of reasoning developed for the univariate asymptotic normality with the use of the Cramér-Wold device.

We first introduce some notation. Denote for any $x > 0$,

$$\partial J_{\psi,x} := \lim_{n \to \infty} \frac{\partial}{\partial \psi} J_n(\psi(\xi), \varphi(x)), \quad \partial J_{\varphi,x} := \lim_{n \to \infty} \frac{\partial}{\partial \varphi} J_n(\psi(\xi), \varphi(x)) , \tag{36}$$

as defined in (30) and (31). In addition, let

$$\sigma_{J,\alpha,\beta}^2 := \frac{(\alpha + \beta)\varphi'(0)}{\varphi(\alpha + \beta)} - \frac{\xi^2 \varphi'(0)}{(\xi - \varphi(\alpha))(\xi - \varphi(\beta))} \times$$
$$\left(\frac{\alpha + \beta}{\varphi(\alpha + \beta)} + \frac{2\alpha\beta}{\psi(\xi)\varphi(2\psi(\xi))} - \frac{\beta(\alpha + \psi(\xi))}{\psi(\xi)\varphi(\alpha + \psi(\xi))} - \frac{\alpha(\beta + \psi(\xi))}{\psi(\xi)\varphi(\beta + \psi(\xi))} \right) , \tag{37}$$

and, for every $\alpha \neq \beta$,

$$\sigma_{\alpha,\beta}^2 := \frac{\sigma_{J,\alpha,\beta}^2 + \partial J_{\psi,\beta}\sigma_{\alpha,\xi}^2 + \partial J_{\psi,\alpha}\sigma_{\beta,\xi}^2 + \partial J_{\psi,\alpha}\partial J_{\psi,\beta}\sigma_\xi^2}{\partial J_{\varphi,\alpha}\partial J_{\varphi,\beta}} . \tag{38}$$

Theorem 7. *Suppose that* $EX(1) < 0$. *Let* $\hat{\mathbf{\Phi}}_n(\boldsymbol{\alpha}) := (\hat{\varphi}_n(\alpha_1; \psi_n), \dots, \hat{\varphi}_n(\alpha_p; \psi_n))$, *where* $(\alpha_1, \dots, \alpha_p) \in (0, \infty)^p$. *If* ψ_n *satisfies the conditions of Theorem 3, and in particular (iii) is satisfied for every* α_i, $i = 1, \dots, p$, *then*

$$\hat{\mathbf{\Phi}}_n(\boldsymbol{\alpha}) \rightarrow_d N(0, \Sigma) , \tag{39}$$

where $\Sigma_{ij} = \sigma_{\alpha_i,\xi}^2$ *if* $i = j$ *and* $\sigma_{\alpha_i,\alpha_j}^2$ *otherwise.*

4 Estimating $\psi(\xi)$ for subordinator input

Theorems 1 and 3 have established that, under specific conditions, the Z-estimator for $\varphi(\alpha)$ is consistent (asymptotically normal, respectively) if we have a consistent (asymptotically normal) estimator for $\psi(\xi)$. In this section we propose such an estimator for $\psi(\xi)$ using a likelihood approach for the subordinator input case, i.e., the input process has almost surely non-decreasing paths. An important special case is the CP model: nonnegative jumps (with LST $G^*(\alpha)$) arrive according to a Poisson process with rate λ. We thus have that for this example,

$$\varphi(\alpha) = \lambda\left(G^*(\alpha) - 1\right) + \alpha . \tag{40}$$

In the following an ML-type estimator is constructed for $\psi(x)$. Suppose we have a sample of workload observations $\boldsymbol{V} = (V_0, V_1, \ldots, V_n)$. Let $Y_i := \mathbf{1}(V_i = 0)$, then by (7),

$$\mathrm{P}(Y_i = 1 \mid \boldsymbol{V}) = \frac{\xi}{\psi(\xi)} e^{-\psi(\xi)V_{i-1}}, \ i = 1, \ldots, n . \tag{41}$$

Conditional on $\mathbf{V}$, the sample $\boldsymbol{Y} = (Y_1, \ldots, Y_n)$ is then distributed as a sample of independent, but non-identically distributed, Bernoulli trials, with likelihood

$$\mathrm{P}_{\psi(\xi)}(Y_1, \ldots, Y_n | \mathbf{V}) = \prod_{i=1}^{n} \left[\frac{\xi}{\psi(\xi)} e^{-\psi(\xi)V_{i-1}}\right]^{Y_i} \left[1 - \frac{\xi}{\psi(\xi)} e^{-\psi(\xi)V_{i-1}}\right]^{1-Y_i} . \tag{42}$$

We can thus define the MLE for $\psi(\xi)$: $\hat{\psi}_n = \arg\max_{\psi \in \Psi}\{\mathrm{P}_\psi(Y_1, \ldots, Y_n | \mathbf{V})\}$.

We assume that the parameter space Ψ is compact and that the true value lies within the interior. This assumption is reasonable and can be achieved by setting $\Psi = [\xi, C]$, where C is an arbitrarily large constant. Note that $\psi(\xi) \geq \xi$ by (7). For special cases the parameter space can be specified in a more detailed manner. For example, the exponent function for the CP case is given in (40), and as $G^*(\alpha) \in [0, 1]$, it follows that $\alpha - \lambda \leq \varphi(\alpha) \leq \alpha$. This implies

$$\psi(\xi) - \lambda \leq \varphi(\psi(\xi)) \leq \psi(\xi) . \tag{43}$$

As $\psi(\xi) = \varphi^{-1}(\xi)$, we conclude that the parameter space is $\Psi = [\xi, \xi + \lambda]$. In this definition there is an implicit assumption that λ is known. If it is unknown, one could again follow the pragmatic approach of replacing $\xi + \lambda$ by an arbitrarily large constant.

For any $\psi \in \Psi$, the log-likelihood function, $L_n(\psi) = \log \mathrm{P}_\psi(Y_1, \ldots, Y_n \mid \boldsymbol{V})$, is

$$L_n(\psi) = \sum_{i=1}^{n} \left[Y_i(\log\xi - \log\psi - \psi V_{i-1}) + (1 - Y_i)\log\left(1 - \frac{\xi}{\psi}e^{-\psi V_{i-1}}\right)\right] , \tag{44}$$

and taking derivatives yields

$$L_n'(\psi) = \sum_{i=1}^{n} \frac{\left(\frac{1}{\psi} + V_{i-1}\right)\left(\frac{\xi}{\psi}e^{-\psi V_{i-1}} - Y_i\right)}{1 - \frac{\xi}{\psi}e^{-\psi V_{i-1}}} . \tag{45}$$

The MLE is the solution of the nonlinear optimization problem,

$$\hat{\psi}_n := \arg\max_{\psi \in \Psi} L_n(\psi) . \tag{46}$$

12

The function $L_n(\psi)$ is not always concave with respect to ψ but as the parameter space is compact a maximizer always exists. If $Y_i = 1$ for all $i = 1, \ldots, n$ (i.e., all observations are at idle periods), then by (42) the likelihood $L_n(\psi)$ decreases with ψ, and therefore the MLE is the lower boundary of the parameter space. On the other hand if $Y_i = 0$ for all $i = 1, \ldots, n$ (i.e., no idle periods are observed), then by (42) the likelihood $L_n(\psi)$ increases with ψ, and then the MLE is the upper boundary of the support. Otherwise, the solution may be in the interior or boundary of the parameter space. In the following subsection we will show that as n increases the asymptotic log-likelihood has a unique maximizer, that solves the first order equation $L_n'(\hat\psi_n) = 0$.

4.1 Consistency

In this subsection we establish strong consistency of the MLE $\hat\psi_n$. As $\hat\psi_n$ is consistent (actually even strongly consistent), Theorem 1 thus yields the consistency of $\hat\varphi_n(\alpha; \hat\psi_n)$. The key for determining consistency and asymptotic normality is the asymptotic analysis of the log-likelihood function given in (44) and its first two derivatives. We define $\ell_n(\psi) := n^{-1} L_n(\psi)$, and $\ell_n'(\psi)$ and $\ell_n''(\psi)$ as the respective first two derivatives. There are several issues that prevent us from directly applying standard MLE consistency results. Firstly, the observations are dependent. Additionally, the function $|\ell_n(\psi)|$ may be unbounded at the boundaries of the parameter space: e.g., if there is some i such that $Y_i = 1$ and $V_{i-1} = 0$ then $\ell_n(\psi) \to -\infty$ as $\psi \downarrow \xi$. Therefore, we will need to verify that as n grows the MLE moves away from the boundary almost surely.

In the next theorem, we assume that Ψ is compact and that the true parameter belongs to the interior of the parameter space, Ψ^{o}. In practice, choosing a very high upper bound for Ψ should ensure this condition.

Theorem 8. *If* $\mathrm{E}X(1) < 0$, *and if the parameter space* Ψ *is compact and* $\psi(\xi) \in \Psi^{\mathrm{o}}$, *then the MLE is strongly consistent:* $\hat\psi_n \to_{\mathrm{as}} \psi(\xi)$.

Upon combining Theorems 1 and 8, we thus have the following result.

Corollary 9. *If* $\mathrm{E}X(1) < 0$, *and if the parameter space* Ψ *is compact and* $\psi(\xi) \in \Psi^{\mathrm{o}}$, *then the Z-estimator for* $\varphi(\alpha)$ *defined in* (11) *using the MLE for* $\psi(\xi)$ *is pointwise consistent:* $\hat\varphi_n(\alpha; \hat\psi_n) \to_{\mathbb{P}} \varphi(\alpha)$, *for all* $\alpha > 0$.

The remainder of this subsection focuses on proving Theorem 8. We commence by providing two lemmas, which are proven in Appendix B, featuring properties of the log-likelihood function that will be utilized in the asymptotic analysis. In particular, we assert smoothness and the construction of a strong law of large numbers for the first two derivatives of the log-likelihood.

Lemma 10. *If the parameter space* $\Psi \subset [\xi, \infty)$ *is a compact interval, then the functions* $\ell_n(\psi)$ *and* $\ell_n'(\psi)$ *are bounded and smooth on* Ψ^{o}.

Lemma 11. *If* $\mathrm{E}X(1) < 0$, *then for all* $\psi \in \Psi^{\mathrm{o}}$,

$$\ell_n'(\psi) \to_{\mathrm{as}} \mathrm{E}(\mathrm{E}(\ell_1'(\psi)|V)) = \mathrm{E}\left[\frac{\frac{1}{\psi} + V}{1 - \frac{\xi}{\psi}e^{-\psi V}} \left(\frac{\xi}{\psi}e^{-\psi V} - \frac{\xi}{\psi(\xi)}e^{-\psi(\xi)V} \right) \right], \qquad (47)$$

and

$$\ell_n''(\psi) \to_{\mathrm{as}} \mathrm{E}(\mathrm{E}(\ell_1''(\psi)|V)) . \qquad (48)$$

13

Proof of Theorem 8. The following condition is sufficient for the consistency of an MLE with dependent observations [22]: for every $\psi \neq \psi(\xi)$ (in the interior of Ψ) there exists a $\delta > 0$ such that

$$\lim_{n \to \infty} \sup_{\zeta \in \mathcal{B}_\delta(\psi)} \{\ell_n(\zeta) - \ell_n(\psi(\xi))\} < 0 , \tag{49}$$

almost surely, where $\mathcal{B}_\delta(\psi) := (\psi - \delta, \psi + \delta)$ is an open ball around ψ. Firstly, applying (47) yields

$$\ell'_n(\psi) \to_{\text{as}} \text{E}(\text{E}(\ell'_1(\psi)|V)) \begin{cases} > 0, & \psi < \psi(\xi), \\ = 0, & \psi = \psi(\xi), \\ < 0, & \psi > \psi(\xi). \end{cases} \tag{50}$$

As $\psi(\xi) \in \Psi^\circ$, for $\psi < \psi(\xi)$ we can consider a ball such that $\psi + \delta < \psi(\xi)$, and then for any $\zeta < \psi + \delta < \psi(\xi)$, by Lemma 10 the difference can be written as

$$\ell_n(\zeta) - \ell_n(\psi(\xi)) = - (\ell_n(\psi(\xi)) - \ell_n(\zeta)) = - \int_\zeta^{\psi(\xi)} \ell'_n(x) \, \mathrm{d}x . \tag{51}$$

By (50), we have that as $n \to \infty$, $\ell'_n(\psi) > 0$ almost surely for any $\psi \in [\psi + \delta, \psi(\xi)]$, which is a non-empty interval, and thus

$$\sup_{\zeta \in (\psi - \delta, \psi + \delta)} \{\ell_n(\zeta) - \ell_n(\psi(\xi))\} \leq - \int_{\psi + \delta}^{\psi(\xi)} \ell'_n(x) \mathrm{d}x < 0 \tag{52}$$

almost surely. Similarly, for any $\psi > \psi(\xi)$ we can consider a ball such that $\psi - \delta > \psi(\xi)$ and conclude that

$$\sup_{\zeta \in (\psi - \delta, \psi + \delta)} \{\ell_n(\zeta) - \ell_n(\psi(\xi))\} \leq \int_{\psi(\xi)}^{\psi - \delta} \ell'_n(x) \mathrm{d}x < 0 \tag{53}$$

almost surely. Therefore, for any $\psi \neq \psi(\xi)$ there exists a $\delta < |\psi - \psi(\xi)|$ such that (49) holds, almost surely. As a consequence, (49) is satisfied. $\qquad\square$

4.2 Asymptotic normality

We now discuss the asymptotic distribution of the MLE $\hat{\psi}_n$. In this subsection we show that the estimation term is asymptotically normal with rate $\sqrt{n}$, and that it further satisfies the conditions of Theorem 3. Useful conditions for asymptotic normality of the MLE when the observations are dependent are given in [23]. Again the delta method is relied on; we also make use of arguments used in the proof of Cramér's Theorem [13, p. 121]. Define

$$I_\xi := \text{E} \left[\frac{\frac{\xi}{\psi(\xi)} e^{-\psi(\xi)V} \left(\frac{1}{\psi(\xi)} + V \right)^2}{1 - \frac{\xi}{\psi(\xi)} e^{-\psi(\xi)V}} \right] . \tag{54}$$

Theorem 12. *If* $\text{E}X(1) < 0$, *and if the parameter space* Ψ *is compact and* $\psi(\xi) \in \Psi^\circ$, *then*

$$\sqrt{n}(\hat{\psi}_n - \psi(\xi)) \to_{\text{d}} \text{N} \left(0, \frac{1}{I_\xi} \right) . \tag{55}$$

14

We first outline the two main steps required to establish the asymptotic distribution. This will be followed by a separate lemma verifying each of these steps.

Proof. The smoothness of $\ell_n'(\cdot)$ established in Lemma 10 implies

$$\ell_n'(\hat{\psi}_n) = \ell_n'(\psi(\xi)) + \int_{\psi(\xi)}^{\hat{\psi}_n} \ell_n''(x)\ \mathrm{d}x\ , \tag{56}$$

and by applying a change of variables

$$\begin{aligned}
\ell_n'(\hat{\psi}_n) &= \ell_n'(\psi(\xi)) + (\hat{\psi}_n - \psi(\xi)) \int_0^1 \ell_n''(\psi(\xi) + y(\hat{\psi}_n - \psi(\xi)))\ \mathrm{d}y \\
&= \ell_n'(\psi(\xi)) + (\hat{\psi}_n - \psi(\xi)) \int_0^1 \ell_n''(y\hat{\psi}_n + (1-y)\psi(\xi))\ \mathrm{d}y\ .
\end{aligned} \tag{57}$$

If $\psi(\xi)$ is in the interior of Ψ, then by Lemma 11, as $n \to \infty$, the MLE is given by the root of the first order condition $L_n'(\hat{\psi}_n) = 0$, almost surely. To verify this conclusion one only has to consider the limits of $\ell_n'(\psi)$ at the lower and upper boundary points of the compact interval Ψ, which are almost surely positive and negative, respectively. Therefore, $\sqrt{n}\ell_n'(\hat{\psi}_n) = \frac{1}{\sqrt{n}} L_n'(\hat{\psi}_n) \to_{\mathrm{as}} 0$. Hence, denoting

$$D_n := \int_0^1 \ell_n''(y\hat{\psi}_n + (1-y)\psi(\xi)))\ \mathrm{d}y, \tag{58}$$

we have

$$\sqrt{n}\ell_n'(\psi(\xi)) + \sqrt{n}(\hat{\psi}_n - \psi(\xi))D_n \approx 0\ , \tag{59}$$

and therefore

$$\sqrt{n}(\hat{\psi}_n - \psi(\xi)) \approx \frac{\sqrt{n}\ell_n'(\psi(\xi))}{-D_n}\ . \tag{60}$$

We show that $\sqrt{n}\ell_n'(\psi(\xi)) \to_{\mathrm{d}} N(0, I_\xi)$ and $D_n \to_{\mathrm{as}} -I_\xi$ in Lemmas 13 and 14, respectively. Then, by Slutsky's lemma, (60) implies (55). $\qquad\square$

Lemma 13. *Under the conditions of Theorem 8, $\sqrt{n}\ell_n'(\psi(\xi)) \to_{\mathrm{d}} N(0, I_\xi)$.*

Lemma 14. *Under the conditions of Theorem 8, $D_n \to_{\mathrm{as}} = -I_\xi$.*

For the proofs of these two lemmas we refer to Appendix B.

Remark 5. The asymptotic variance (54) in Theorem 12 is $\sigma_\xi^2 := 1/I_\xi$. In the CP case $\varphi(\alpha) = \lambda(G^*(\alpha) - 1) + \alpha$, and therefore $\psi(\xi)$ is an increasing function of ξ such that $\psi(0) = 0$ and

$$\frac{\xi}{\psi(\xi)} = \frac{\lambda(G^*(\psi(\xi)) - 1) + \psi(\xi)}{\psi(\xi)} = 1 + \lambda \frac{G^*(\psi(\xi)) - 1}{\psi(\xi)} \xrightarrow{\xi\downarrow 0} \varphi'(0). \tag{61}$$

This implies that the asymptotic variance σ_ξ^2 increases with the sampling rate ξ, and that for a decreasing rate the variance disappears: $\sigma_\xi^2 \to 0$ as $\xi \downarrow 0$.

4.3 Asymptotic error of $\hat{\varphi}_n(\alpha; \hat{\psi}_n)$

Theorem 12 has implications also for the asymptotic estimation error of $\varphi(\alpha)$. We verified the martingale CLT conditions of Lemma 2a for $\sqrt{n}(\psi_n - \psi(\xi))$ in Theorem 12, and thus if we further show that the covariance of the estimation errors has a finite limit we can apply Lemma 2b and Theorem 3 to conclude that $\sqrt{n}(\varphi_n(\alpha; \hat{\psi}_n) - \varphi(\alpha))$ is asymptotically normal. This is done in the proof of Theorem 15 in Appendix B.

Theorem 15. *If* $\mathrm{E}X(1) < 0$*,* $\hat{\psi}_n$ *is the MLE for* $\psi(\xi)$*, then*

$$\sqrt{n}(\varphi_n(\alpha; \hat{\psi}_n) - \varphi(\alpha)) \to_d \mathrm{N}\left(0, \sigma_{\alpha,\xi}^2\right) \ , \tag{62}$$

where $\sigma_{\alpha,\xi}^2$ *is given by* (28) *with*

$$\sigma_{\alpha,\psi}^2 := \frac{\xi^2}{\psi(\xi)(\xi - \varphi(\alpha))I_\xi} \mathrm{E}\left[\frac{\left(\frac{1}{\psi(\xi)} + V\right)e^{-\psi(\xi)V}}{1 - \frac{\xi}{\psi(\xi)}e^{-\psi(\xi)V}}\left(e^{-\alpha V} - \frac{\alpha}{\psi(\xi)}e^{-\psi(\xi)V} - \frac{\xi - \varphi(\alpha)}{\xi}\right)\right]. \tag{63}$$

5 Estimating $\psi(\xi)$ for spectrally-positive Lévy input

The Z-estimator $\hat{\varphi}_n(\alpha; \psi_n)$ requires an estimator for $\psi(\xi)$. In Section 4 we considered the subordinator case, utilizing that there is a non-zero probability that the queue is empty. For general spectrally-positive Lévy input with a Brownian motion component this property does not hold. In this section we present an alternative procedure.

A moment based Z-estimator ψ_n can be derived using (5) by solving

$$\sum_{i=1}^{n}(V_i - V_{i-1}) - \sum_{i=1}^{n}\left(\frac{e^{-\psi_n V_{i-1}}}{\psi_n} - \frac{\varphi'(0)}{\xi}\right) = 0 \ . \tag{64}$$

After some rearranging of the terms, we define the estimator ψ_n as the solution of

$$V_n - V_0 + \frac{n\varphi'(0)}{\xi} = \sum_{i=1}^{n}\frac{e^{-\psi_n V_{i-1}}}{\psi_n} \ , \tag{65}$$

when it exists. There is at most one solution as the left-hand side is constant and the right-hand side is decreasing from infinity to zero on $\psi_n \in \mathbb{R}^+$. However, if the constant on the left is negative there is no solution and then we can set $\psi_n = \bar{\psi}$, where $\bar{\psi} > 0$ is some arbitrary large constant. Observe that under the stability condition $\varphi'(0) = -\mathrm{E}X(1) > 0$, the probability of having a negative term on the left vanishes as $n \to \infty$.

The problem, however, is that $\varphi'(0)$ in (65) is also unknown and needs to be somehow estimated. Note that if we naïvely plug $\hat{\varphi}_n'(0; \psi_n)$ into (65) then all terms including ψ_n cancel out because the estimator of the first derivative relies on the very same equation, i.e., (5). In the next subsections we will present an estimation scheme for $\varphi'(0)$ that relies on considering a sub-sample with high initial workloads above some given threshold. We also show how the estimator's inherent bias can be made arbitrarily small.

5.1 Estimating $\varphi'(0)$

We construct an estimator for $\varphi'(0)$ relying on the fact that $\xi \, \mathrm{E} X(T) = -\varphi'(0)$ where $T \sim \exp(\xi)$, and estimating the net input only using observations that are far from zero. If a high workload was observed then the probability of reflection (i.e., the queue hitting zero) before the next observation is low. In particular, we can quantify this probability so as to provide an upper bound for the bias of the estimator.

For some (large) threshold $\tau > 0$ the workload is sampled until there are $m \in \mathbb{N}$ observations V_i such that $V_{i-1} \geq \tau$. Let $M(m, \tau)$ denote the total number of observations, which is now a random variable, and denote the j-th such observation by $i(j) \in \{1, \dots, M(m, \tau)\}$, for $j = 1, \dots, m$. The estimator for $\vartheta = -\varphi'(0) = \xi \mathrm{E} X(T)$ is

$$\hat{\vartheta}_m(\tau) := \frac{\xi}{m} \sum_{j=1}^m (V_{i(j)} - V_{i(j)-1}) \,. \tag{66}$$

Observe that the statistic $\hat{\vartheta}_m(\tau)$ only depends on the increments of the workload process which makes it convenient for analysis due to the independent increments property of Lévy processes. We first derive an upper bound for the bias of $\hat{\vartheta}_m(\tau)$ and in the sequel provide an accurate asymptotic analysis of the bias. The proofs for this section are detailed in Appendix C.

Proposition 16. *Let* $b_m(\hat{\vartheta}; \tau) := \mathrm{E}[\hat{\vartheta}_m(\tau) - \vartheta]$. *Then, for any* $m \in \mathbb{N}$ *and* $\tau > 0$,

$$0 < b_m(\hat{\vartheta}; \tau) \leq \xi \, \mathrm{E}[V(T)|V(0) = 0] \, e^{-\psi(\xi)\tau} \,. \tag{67}$$

The bound (67) for the bias decays to zero as τ grows, but at the same time the probability of observing workloads above τ also diminishes if $\mathrm{E} X(1) < 0$, so that the required number of samples $M(m, \tau)$ also grows. This entails that there is a tradeoff between τ and m in terms of accuracy, information and required sample size. The expectation in the bound, i.e., $\mathrm{E}[V(T)|V(0) = 0]$, can be approximately estimated by using observations $V_i - V_{i-1}$ such that $V_{i-1} \leq \epsilon$ for some low threshold $\epsilon > 0$. The bound (67) also depends on $\psi(\xi)$ which can be estimated using (65) by plugging in the estimator $\hat{\vartheta}_m(\tau)$ instead of $-\varphi'(0)$. There is a (seemingly) circular argument here: the estimated bias bound is also biased because it also relies on the biased estimator $\hat{\vartheta}_m$. In the next proposition, however, we argue that this bias is bounded as well and diminishes with τ.

Proposition 17. *Let* $\tilde{\psi}_m$ *be the Z-estimator given by* (65) *with* $\varphi'(0)$ *replaced with the estimator* $\hat{\vartheta}_m(\tau)$, *then the bias* $\mathrm{E}[\tilde{\psi}_m - \psi(\xi)]$ *diminishes to zero as* $\tau \to \infty$.

5.2 Asymptotic analysis

In this subsection we consider asymptotic properties of $\hat{\vartheta}_m(\tau)$ as $m \to \infty$. For any $\tau < \infty$ there is positive bias due to the possibility of reflection, even though it may be very small, and therefore the estimator cannot be consistent. A possible solution to this is defining an estimator with a dynamic threshold τ_m that grows with the sample size. In this section we analyze the asymptotic properties of the biased estimator based on the static threshold τ. We first establish that the estimator $\hat{\vartheta}_m(\tau)$ converges in probability (as $m \to \infty$, that is) to a constant that depends on τ. This further enables the computation of the exact asymptotic bias.

Proposition 18. *If* $\mathrm{E}X(1) < 0$ *then for any* $\tau > 0$, *as* $m \to \infty$, $\hat{\vartheta}_m(\tau) \to_{\mathbb{P}} \vartheta(\tau)$, *where*

$$\vartheta(\tau) := \frac{\xi \mathrm{E}\left[e^{-\psi(\xi)V}\mathbf{1}(V \geq \tau)\right]}{\psi(\xi)\mathrm{P}(V \geq \tau)} - \varphi'(0) > 0 \;, \tag{68}$$

and V *is the steady-state workload.*

An immediate result of Proposition 18 is that from (68) we can derive a bound on the asymptotic bias that does not depend on τ.

Corollary 19. *For any* $\tau \geq 0$,

$$b\left(\hat{\vartheta};\tau\right) := \lim_{m \to \infty} b_m\left(\hat{\vartheta};\tau\right) = \frac{\xi \mathrm{E}\left[e^{-\psi(\xi)V}\mathbf{1}(V \geq \tau)\right]}{\psi(\xi)\mathrm{P}(V \geq \tau)} \leq \frac{\xi}{\psi(\xi)} \;. \tag{69}$$

As $\varphi(\alpha)$ is a convex and increasing function, the term $\xi/\psi(\xi)$ is an increasing function of the sampling rate ξ. In line with earlier observations, it shows that sampling slowly is to be preferred when there is no, or little, cost associated with the sampling duration.

6   Simulation analysis

In order to numerically test the performance of the estimation procedure we carried out simulations of a queue with Lévy input that consists of three processes: $X(t) = C(t) + W(t) + U(t)$, where

- $\circ$ $C(t)$ is Compound Poisson with rate λ and Gamma jumps-size distribution, $\Gamma(\eta, \mu)$,

- $\circ$ $W(t)$ is Brownian motion with drift d and variance σ^2,

- $\circ$ $U(t)$ is a Gamma process (subordinator) with parameters (β, γ).

The exponent function is then

$$\varphi(\alpha) = \lambda\left[\left(\frac{\mu}{\mu+\alpha}\right)^{\eta} - 1\right] - \alpha d + \frac{1}{2}\alpha^2\sigma^2 + \beta \log\left(\frac{\gamma}{\gamma+\alpha}\right) \;, \tag{70}$$

and the expected input per unit of time is $\mathrm{E}X(1) = -\varphi'(0) = \lambda\eta/\mu + d + \beta/\gamma$. We first examine the performance of the Z-estimator for φ that uses the MLE estimator for $\psi(\xi)$ in the M/M/1 case $(\lambda, \eta, \mu, d, \sigma^2, \beta, \gamma) = (0.8, 1, 1, -1, 0, 0, 0)$. In Figure 2 the real and estimated exponent functions are compared along with the confidence interval based on the asymptotic normal approximation. Even for a moderate sample size of $n = 30$ the estimated function is quite close to the real one. Observe, however, that the estimate is less accurate as α increases. In Figure 3 we illustrate the increase in accuracy when the sample size n increases. Theorem 1 established that the estimator is consistent and here we can see that the bias diminishes very quickly. Increasing the sample size is more important when one wishes to estimate $\varphi(\alpha)$ for high values of α.

In Figure 4 the asymptotic variance $\sigma^2_{\alpha,\xi}$ is plotted as a function of α for different sampling rates. These functions are all increasing and appear to be convex and unbounded. The asymptotic correlation,

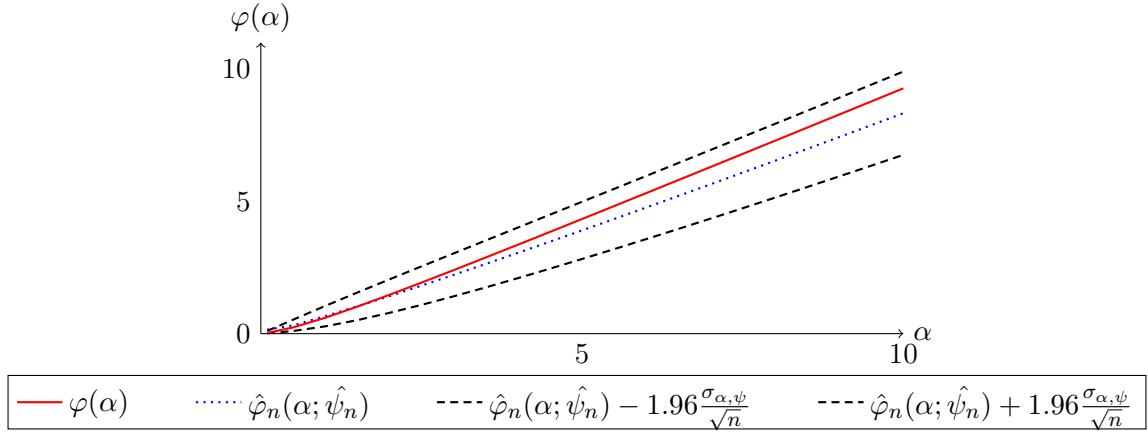$$r(\alpha, 1) := \frac{\sigma^2_{\alpha,1}}{\sqrt{\sigma^2_{\alpha,\xi}\sigma^2_{1,\xi}}} \tag{71}$$

Figure 2: M/M/1 with $\lambda = 0.8$ and $\mu = 1$. The exponent function $\varphi(\alpha)$ (solid red line) and its Z-estimator $\hat{\varphi}_n(\alpha)$ (dotted blue line) from a sample of $n = 30$ observations with sampling rate $\xi = 1$. The intermediate estimator for $\psi(\xi)$ is the MLE of Section 4. Also plotted are the upper and lower bounds of the confidence interval using the normal approximation and asymptotic variance.
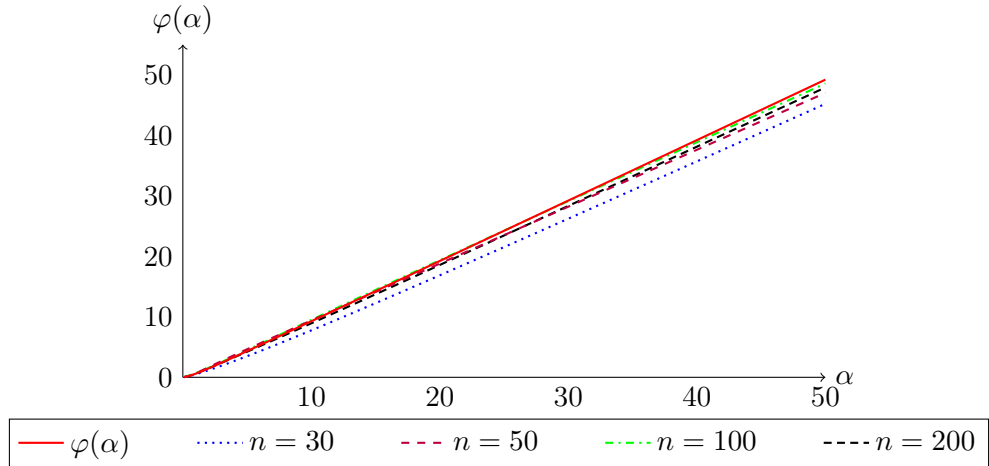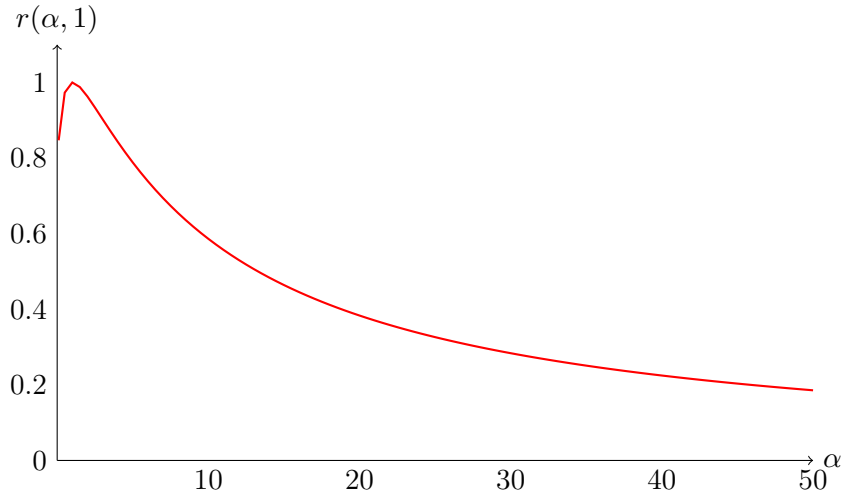

Figure 3: M/M/1 with $\lambda = 0.8$ and $\mu = 1$. The exponent function $\varphi(\alpha)$ (solid red line) and its Z-estimator $\hat{\varphi}_n(\alpha)$ (dotted blue line) based on sample sizes $n \in \{30, 50, 100, 200\}$ with sampling rate $\xi = 1$. The intermediate estimator for $\psi(\xi)$ is the MLE of Section 4. Also plotted are the upper and lower bounds of the confidence interval using the normal approximation and asymptotic variance.

Figure 5: M/M/1 with $\lambda = 0.8$ and $\mu = 1$. The asymptotic covariance $\sigma_{\alpha,1}^2$, between the estimators $\hat{\varphi}_n(\alpha; \hat{\psi}_n)$ and $\hat{\varphi}_n(1; \hat{\psi}_n)$. The sampling rate is $\xi = 1$.

(see Theorem 7), between the estimated function $\hat{\varphi}_n(\alpha; \hat{\psi}_n)$ and $\hat{\varphi}_n(1; \hat{\psi}_n)$ is illustrated in Figure 5. The correlation is a decreasing function of the distance from one, but interestingly does not disappear to zero but to a lower bound (about 0.2 in this example). This should come as no surprise after some thought, as the function is estimated using the same sample of observations for all $\alpha > 0$.
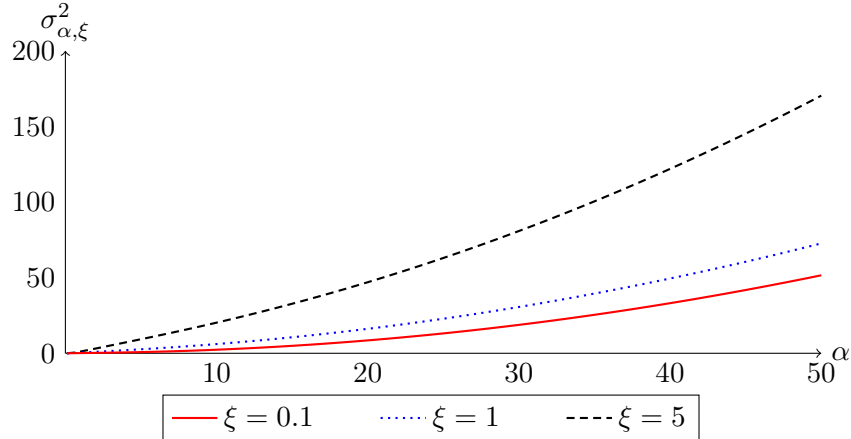

Figure 4: M/M/1 with $\lambda = 0.8$ and $\mu = 1$. The asymptotic variance $\sigma_{\alpha,\xi}^2$ as a function of α for different sampling rates.

We used the two-step estimation method of $\psi(\xi)$ introduced in Section 5 to test the performance of the estimator $\hat{\varphi}_n$ for Lévy input. The data was simulated for an input process with parameter vector $(\lambda, \eta, \mu, d, \sigma, \beta, \gamma) = (0.2, 1.2, 0.5, -1, 0.1, 1, 5)$ and initial workload $V(0) = 0$. In Figure 6 the estimator is based on $m = 200$ observations with a threshold $\tau = 2$ for different sampling rates. As before, accuracy improves with a slow sampling rate and decreases as α grows. For a fixed sampling rate of $\xi = 1$, Figure 7 illustrates the estimated functions
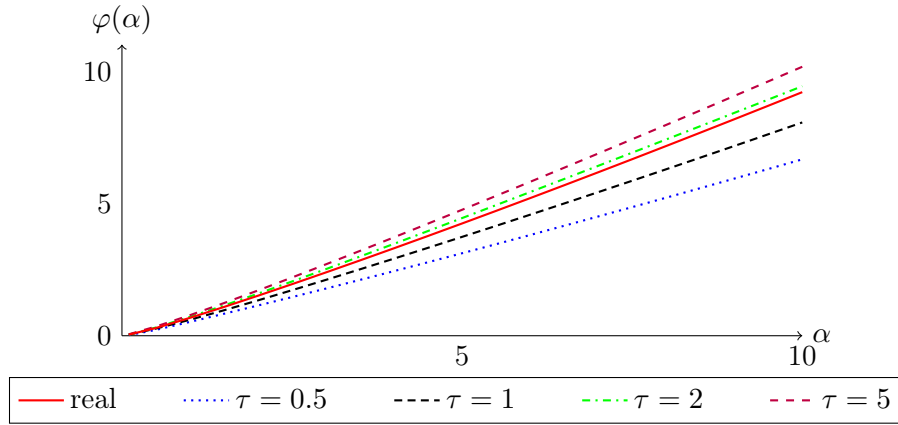
20

Figure 7: Lévy input with $(\lambda, \eta, \mu, d, \sigma, \beta, \gamma) = (0.2, 1.2, 0.5, -1, 0.1, 1, 5)$. The exponent function $\varphi(\alpha)$ (solid red line) and its Z-estimator $\hat{\varphi}_n(\alpha)$ from a sample of $m = 200$ observations for different thresholds τ and sampling rate $\xi = 1$.

for different threshold levels for the biased intermediate estimation step. Both low and high thresholds give less accurate estimators than the intermediate level, $\tau = 2$ in this example. This is because a low threshold has a large bias, whereas a high threshold has a small bias but a smaller sample size and thus a larger variance.
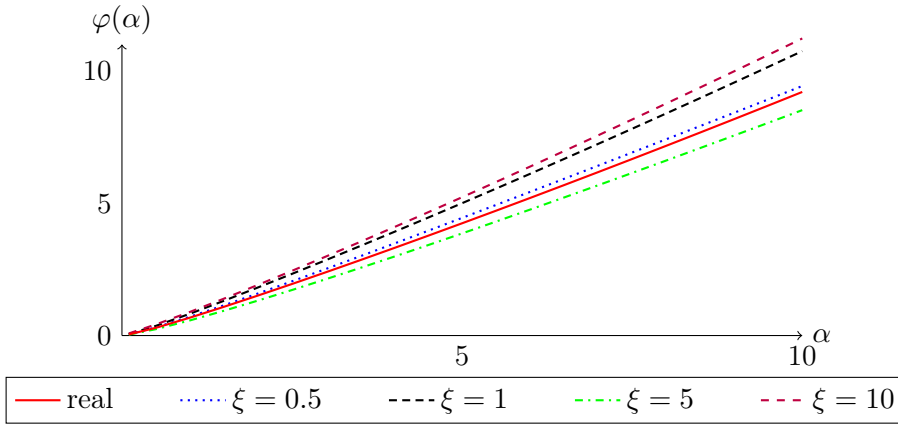

Figure 6: Lévy input with $(\lambda, \eta, \mu, d, \sigma, \beta, \gamma) = (0.2, 1.2, 0.5, -1, 0.1, 1, 5)$. The exponent function $\varphi(\alpha)$ (solid red line) and its Z-estimator $\hat{\varphi}_n(\alpha)$ from a sample of $m = 200$ observations for different sampling rates ξ and threshold $\tau = 2$.

7   Discussion and concluding remarks

In this paper we have studied the problem of statistical inference on the input process to a Lévy-driven queue, where the workload is observed at Poisson times. An estimation method for the characteristic exponent function of the input process has been introduced. The setting is semi-parametric because although the goal is to estimate a function, the methods employed are

parametric (namely, method-of-moments and ML-based). Under specific verifiable conditions, consistency and asymptotic normality are proven.

Relation with previous work. In the introduction we already reflected on the papers [8, 19]; we now focus on the relation with the generalized method-of-moments (GMM) framework of [12]. Recall that our main estimation device is the Z-estimator $\hat{\varphi}_n(\alpha)$ which is derived by solving an estimation equation that equates the empirical LST of the workload to the conditional expected value after an exponential sampling time. In this respect our work is related to [12] that studies GMM estimation of parameters of Markov processes based on data sampled according to a random process which may be time and state dependent. In our setting, a GMM procedure is used to estimate the vector $(\varphi(\alpha_1), \ldots, \varphi(\alpha_p))$ by solving the respective estimation equations (8) for $i = 1, \ldots, p$.

Our estimation equations can be stated in terms of the generator functions used in [12] with a homogeneous Poisson sampling process. However, there is a need to externally estimate the term $\psi(\xi)$; importantly, this cannot be achieved simultaneously using the same estimation equations. This is more than just a technical issue associated with the estimation method, and is inherent to the non-parametric setting of estimating the function φ, as opposed to finite-dimensional parameter estimation. Another issue is that we make inference on the input process without observing it directly but rather from observations of the reflected process. Finally, the asymptotic analysis of [12] only deals with asymptotic normality for the case of a one-dimensional parameter space, where we have derived the multi-dimensional asymptotic distribution of the estimation errors and provided explicit terms for all associated asymptotic (co-)variance terms.

Identifiability for CP input. For the CP case one can in principle identify the λ and $G^*(\alpha)$ separately from the estimates $\hat{\varphi}_n(\alpha; \hat{\psi}_n)$. First one sets, for some large value of α_+,

$$\hat{\lambda}_n = \alpha_+ - \hat{\varphi}_n(\alpha_+; \hat{\psi}_n), \tag{72}$$

and then, for any α,

$$\hat{G}_n^*(\alpha) = 1 + \frac{1}{\hat{\lambda}} \left(\hat{\varphi}_n(\alpha; \hat{\psi}_n) - \alpha \right) . \tag{73}$$

Parametric special cases. Our estimation method is directly applicable to various parametric special cases using the transient moment equations. For example, when the input is Brownian motion one has $\varphi(\alpha) = -d\alpha + \frac{1}{2}\alpha^2\sigma^2$. In this case the first two conditional moment equations, (5) and (6), can be used to construct estimators for d and σ^2. For the M/M/1 queue jump sizes follow an exponential distribution with rate μ. Therefore, $G^*(\alpha) = \mu/(\mu + \alpha)$, which yields

$$\varphi(\alpha) = \frac{\alpha(\mu + \alpha - \lambda)}{\mu + \alpha}, \quad \psi(\xi) = \frac{1}{2} \left(\xi + \lambda - \mu + \sqrt{(\xi + \lambda - \mu)^2 + 4\xi\mu} \right). \tag{74}$$

In Section 4 for the subordinator input case an MLE for $\psi(\xi)$ was derived. This can be used to estimate the two parameters of an M/M/1, or any other two-parameter formulation for that matter, by using only the first moment equation (5). This can be done similarly for other types of parametric subordinator input such as a Gamma process.

Future research. There are several interesting questions that are left open in this work. First of all, the question whether a consistent estimator for $\psi(\xi)$ can be constructed for the general case. Although simulation analysis suggests that in practice the biased estimator performs well, it is

still of interest to find a consistent estimator for theoretical and practical efficiency purposes. Another issue is that the object of estimation is the function φ but all asymptotic results are in terms of the function at a finite collection of points. A natural extension would be to derive functional limit theorems, for example uniform-consistency and a FCLT approximation for the estimation error as a Gaussian process. In Section 6 we observed that the asymptotic variance appears to be increasing and unbounded with α which may lead to technical difficulties, and in particular it is likely that the estimator is not uniformly-consistent.

Applications. We conclude this paper by mentioning a few applications for which the model and methods presented here may be useful.

- *Optimal sampling rate.* As we saw, explicitly in Section 4 and numerically in Section 6, the accuracy of the estimators improves as the sampling rate decreases. This, however, comes at a cost of a longer sampling duration. In applications where there is a cost associated with the duration of the learning period of the system, this leads to the problem of finding an optimal sampling rate that balances this tradeoff. Our results can be used to derive an approximate solution.

- *Stability detection.* Throughout this work we assumed that $\mathrm{E}X(1) < 0$ and consequently the workload process is stable. However when the input is unknown, we do not know *a priori* whether $\mathrm{E}X(1) < 0$. For example if the processing speed can be controlled in an M/G/1 queue but the input is unknown then it is possible that for some processing speeds the queue is unstable. In this case a likelihood ratio test can be constructed using the setup of Section 4. In the general case, the biased estimator $\hat{\vartheta}_n$ for $-\mathrm{E}X(1)$ (Section 5) can be used to set up a threshold rule for detecting instability.

- *Dynamic pricing.* Suppose that the server can charge an admission price with the goal of maximizing social welfare or revenue. The optimal pricing scheme obviously depends on the input distribution [21], which may not be known. This naturally leads to a model with an exploration-exploitation element, and using our framework the exploration step can be carried out for transient observations without relying on any approximations. The extension to the dynamic setting requires a state-dependent input model such as [6].

- *Testing distributional assumptions.* An issue that is related to the two previous points is that often controllable system parameters are chosen with the goal of maximizing specific performance measures. This is typically done under some modeling assumptions. Our framework enables testing whether the input to the system is consistent with the modeling assumptions (as long as the input alternatives are within the class of spectrally positive Lévy processes).

- *Service level monitoring.* Service systems are often obliged to provide performance guarantees. For example, when an Internet Service Provider commits to certain bandwidth, it is important for users and regulators to be able to estimate the performance of the system in order to monitor whether the contract is upheld or not [20, Ch. IX]. Our method allows to estimate performance measures, such as moments and tail probabilities, using a transient sample of observations.

Appendices

A   Proofs for Section 3

Proof of Lemma 4. We first prove (30). From (24),

$$\frac{\partial}{\partial \psi} J_n(\psi, \varphi)\bigg|_{\varphi=\varphi(\alpha), \psi=\psi(\xi)} = -\frac{\xi}{\xi - \varphi(\alpha)} \cdot \frac{1}{n} \sum_{i=1}^{n} \left(\frac{\alpha}{\psi(\xi)^2} e^{-\psi(\xi) V_{i-1}} + \frac{\alpha V_{i-1}}{\psi(\xi)} e^{-\psi(\xi) V_{i-1}} \right) . \quad (75)$$

Applying the GPK formula (2), PASTA and the law of large numbers,

$$\begin{aligned}
\frac{\partial}{\partial \psi} J_n(\psi(\xi), \varphi(\alpha)) \to_{\text{as}} &-\frac{\xi}{\xi - \varphi(\alpha)} \left(\frac{\alpha}{\psi(\xi)^2} \mathrm{E} e^{-\psi(\xi) V_{i-1}} + \frac{\alpha}{\psi(\xi)} \xi \,\mathrm{E}[V e^{-\psi(\xi) V}] \right) \\
&= -\frac{\alpha \varphi'(0)}{(\xi - \varphi(\alpha)) \psi'(\xi)} = \partial J_\psi ,
\end{aligned} \quad (76)$$

as $n \to \infty$, where it is used that

$$\mathrm{E}[V e^{-\alpha V}] = \frac{\varphi'(0)}{(\varphi(\alpha))^2} \big(\alpha \varphi'(\alpha) - \varphi(\alpha) \big) .$$

Now we prove (31). Following the same steps as above,

$$\frac{\partial}{\partial \varphi} J_n(\psi, \varphi)\bigg|_{\varphi=\varphi(\alpha), \psi=\psi(\xi)} \to_{\text{as}} -\frac{\alpha \varphi'(0)}{\varphi(\alpha)(\xi - \varphi(\alpha))} =: \partial J_\varphi , \quad (77)$$

as $n \to \infty$. $\qquad\square$

Proof of Lemma 5.   a. We first compute the asymptotic variance of $\sqrt{n} J_n(\psi(\xi), \varphi(\alpha))$ and then apply the martingale CLT of Lemma 2a to verify that the limiting distribution is normal. Let

$$Z_i := e^{-\alpha V_i} - \frac{\xi}{\xi - \varphi(\alpha)} \left(e^{-\alpha V_{i-1}} - \frac{\alpha}{\psi(\xi)} e^{-\psi(\xi) V_{i-1}} \right) . \quad (78)$$

Then $J_n(\psi(\xi), \varphi(\alpha)) = \frac{1}{n} \sum_{i=1}^{n} Z_i$. By (4), $\mathrm{E}[Z_i | V_{i-1}] = 0$. Therefore,

$$n \mathrm{Var}(J_n(\psi(\xi), \varphi(\alpha))) = \frac{1}{n} \mathrm{E} \left(\sum_{i=1}^{n} Z_i \right)^2 = \frac{1}{n} \left(\sum_{i=1}^{n} \mathrm{E} Z_i^2 + 2 \sum_{i=2}^{n} \sum_{j<i} \mathrm{E} Z_i Z_j \right) , \quad (79)$$

and as

$$\mathrm{E} Z_i Z_j = \mathrm{E}[\mathrm{E}(Z_i Z_j | (V_0, \ldots, V_{i-1}))] = \mathrm{E}[Z_j \mathrm{E}(Z_i | V_{i-1})] = 0 \quad \forall j < i , \quad (80)$$

we have that $n\mathrm{Var}(J_n(\psi(\xi),\varphi(\alpha))) = \frac{1}{n}\sum_{i=1}^{n}\mathrm{E}Z_i^2$. The second moment of Z_i is

$$
\mathrm{E}Z_i^2 = \mathrm{E}\left[e^{-2\alpha V_i} - \frac{2\xi}{\xi-\varphi(\alpha)}e^{-\alpha V_i}\left(e^{-\alpha V_{i-1}} - \frac{\alpha}{\psi(\xi)}e^{-\psi(\xi)V_{i-1}}\right)\right.
$$
$$
\left. + \left(\frac{\xi}{\xi-\varphi(\alpha)}\right)^2\left(e^{-2\alpha V_{i-1}} - \frac{2\alpha}{\psi(\xi)}e^{-(\alpha+\psi(\xi))V_{i-1}} + \left(\frac{\alpha}{\psi(\xi)}\right)^2 e^{-2\psi(\xi)V_{i-1}}\right)\right]
$$
$$
= \mathrm{E}\left[e^{-2\alpha V_i} + \left(\frac{\xi}{\xi-\varphi(\alpha)}\right)^2 e^{-2\alpha V_{i-1}} + \left(\frac{\xi\alpha}{\psi(\xi)(\xi-\varphi(\alpha))}\right)^2 e^{-2\psi(\xi)V_{i-1}}\right.
$$
$$
\left. - \frac{2\xi}{\xi-\varphi(\alpha)}e^{-\alpha V_i - \alpha V_{i-1}} + \frac{2\xi\alpha}{\psi(\xi)(\xi-\varphi(\alpha))}e^{-\alpha V_i - \psi(\xi)V_{i-1}} - \frac{2\xi^2\alpha}{\psi(\xi)(\xi-\varphi(\alpha))^2}e^{-(\alpha+\psi(\xi))V_{i-1}}\right].
$$
$$\tag{81}$$

The limit of $\mathrm{E}e^{-\alpha V_i}$ as $i \to \infty$ is given by PASTA and the GPK formula (2). The joint distribution of (V_i, V_{i-1}) is obtained by conditioning on V_{i-1} and applying (4), for any $\alpha > 0$ and $\beta > 0$,

$$
\mathrm{E}e^{-\alpha V_i - \beta V_{i-1}} = \mathrm{E}\left[e^{-\beta V_{i-1}}\frac{\xi}{\xi-\varphi(\alpha)}\left(e^{-\alpha V_{i-1}} - \frac{\alpha}{\psi(\xi)}e^{-\psi(\xi)V_{i-1}}\right)\right]
$$
$$
= \frac{\xi}{\xi-\varphi(\alpha)}\mathrm{E}\left[e^{-(\alpha+\beta)V_{i-1}} - \frac{\alpha}{\psi(\xi)}e^{-(\psi(\xi)+\beta)V_{i-1}}\right] \tag{82}
$$
$$
\xrightarrow{i\to\infty} \frac{\xi\varphi'(0)}{\xi-\varphi(\alpha)}\left(\frac{\alpha+\beta}{\varphi(\alpha+\beta)} - \frac{\alpha(\psi(\xi)+\beta)}{\psi(\xi)\varphi(\psi(\xi)+\beta)}\right).
$$

Therefore,

$$
\lim_{i\to\infty}\mathrm{E}Z_i^2 = \left(1 + \left(\frac{\xi}{\xi-\varphi(\alpha)}\right)^2\right)\frac{2\alpha\varphi'(0)}{\varphi(2\alpha)} + \left(\frac{\xi\alpha}{\psi(\xi)(\xi-\varphi(\alpha))}\right)^2 \cdot \frac{2\psi(\xi)\varphi'(0)}{\varphi(2\psi(\xi))}
$$
$$
- \frac{2\xi^2}{(\xi-\varphi(\alpha))^2}\left(\frac{2\alpha\varphi'(0)}{\varphi(2\alpha)} - \frac{\alpha(\alpha+\psi(\xi))\varphi'(0)}{\psi(\xi)\varphi(\alpha+\psi(\xi))}\right)
$$
$$
+ \frac{2\xi^2\alpha}{\psi(\xi)(\xi-\varphi(\alpha))^2}\left(\frac{(\alpha+\psi(\xi))\varphi'(0)}{\varphi(\alpha+\psi(\xi))} - \frac{2\alpha\varphi'(0)}{\varphi(2\psi(\xi))}\right) \tag{83}
$$
$$
- \frac{2\xi^2\alpha}{\psi(\xi)(\xi-\varphi(\alpha))^2} \cdot \frac{(\alpha+\psi(\xi))\varphi'(0)}{\varphi(\alpha+\psi(\xi))}.
$$

Rearranging terms we obtain $\lim_{i\to\infty}\mathrm{E}Z_i^2 = \sigma_\alpha^2$, as defined in (33). Therefore, we conclude that

$$
n\mathrm{Var}\left(J_n(\psi(\xi),\varphi(\alpha))\right) = \frac{1}{n}\sum_{i=1}^{n}\mathrm{E}Z_i^2 \xrightarrow{n\to\infty} \sigma_\alpha^2. \tag{84}
$$

We now verify the conditions of Lemma 2a. Let

$$
Z_{ni} := \frac{Z_i}{\sqrt{\mathrm{Var}(nJ_n(\psi(\xi),\varphi(\alpha)))}} = \frac{nJ_i(\psi(\xi),\varphi(\alpha)) - nJ_{i-1}(\psi(\xi),\varphi(\alpha))}{\sqrt{\mathrm{Var}(nJ_n(\psi(\xi),\varphi(\alpha)))}}, \tag{85}
$$

and observe that $nJ_n(\psi(\xi),\varphi(\alpha))$ is a martingale as $\mathrm{E}[Z_i|V_{i-1}] = 0$ for every $i \geq 1$. Furthermore, for any n, $\mathrm{Var}(nJ_n(\psi(\xi),\varphi(\alpha))) > 0$ and $\mathrm{P}(Z_i < \infty) = 1$ as $\mathrm{E}Z_i = 0$

for all $i = 1, \ldots, n$. Hence, $\mathrm{E}\max_{1 \leq i \leq n} Z_{ni}^2 < \infty$ and condition (20) holds. Above, we established that $n\mathrm{Var}(J_n(\psi(\xi), \varphi(\alpha)))$ goes to a constant, and thus

$$\mathrm{Var}(nJ_n(\psi(\xi), \varphi(\alpha))) = n^2 \mathrm{Var}(J_n(\psi(\xi), \varphi(\alpha))) \to \infty , \tag{86}$$

and as $\mathrm{P}(Z_i < \infty) = 1$, we conclude that condition (21) is satisfied. Therefore, the Martingale CLT holds and $\sum_{i=1}^n Z_{ni} \to_d \mathrm{N}(0, 1)$. Finally, applying Slutsky's lemma yields

$$\sqrt{n} J_n(\psi(\xi), \varphi(\alpha)) = \sqrt{n} \sum_{i=1}^n Z_i = \sqrt{n \mathrm{Var}\left(J_n(\psi(\xi), \varphi(\alpha))\right)} \sum_{i=1}^n Z_{ni} \to_d \sqrt{\sigma_\alpha^2} \mathrm{N}(0, 1) . \tag{87}$$

b. In the previous part of the lemma we established that $J_n(\psi(\xi), \varphi(\alpha))$ satisfies the conditions of Lemma 2a. If we further assume that $\psi_n - \psi(\xi) \approx \frac{1}{n} M_n + R_n$ such that M_n satisfies the conditions of Lemma 2a, then

$$\sqrt{n}\left(\psi_n - \psi(\xi)\right) \approx \frac{M_n}{\sqrt{n}} \approx \sqrt{\frac{\mathrm{Var}(M_n)}{n}} \mathrm{N}(0, 1) , \tag{88}$$

where $\sqrt{\mathrm{Var}(M_n)/n} \to \sigma_\xi$ by condition (ii). We can next apply Lemma 2b to conclude joint asymptotic normality:

$$\sqrt{n}\begin{pmatrix} J_n(\psi(\xi), \varphi(\alpha)) \\ \\ \psi_n - \psi(\xi) \end{pmatrix} \approx \sqrt{n}\begin{pmatrix} J_n(\psi(\xi), \varphi(\alpha)) \\ \\ \frac{M_n}{n} \end{pmatrix} \to_d \mathrm{N}(0, S) , \tag{89}$$

if the covariance matrix S has finite terms. The variance terms are $S_{11} = \sigma_\alpha^2$ and

$$S_{22} = \lim_{n\to\infty} \mathrm{E}[M_n^2] = \lim_{n\to\infty} \mathrm{E}[(\psi_n - \psi(\xi))^2] = \sigma_\xi^2 . \tag{90}$$

The asymptotic covariance is finite by condition (iii):

$$S_{12} = S_{21} = \lim_{n\to\infty} \mathrm{E}\left[\sqrt{n} J_n(\psi(\xi), \varphi(\alpha)) \frac{M_n}{\sqrt{n}}\right]$$
$$= \lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^n \mathrm{E}\left[Z_i M_n\right] = \sigma_{\alpha,\psi}^2 < \infty . \tag{91}$$

By the Cramér-Wold device this implies that any linear combination of $J_n(\psi(\xi), \varphi(\alpha))$ and $\psi_n - \psi(\xi)$ is also normal, and thus (35) follows.

$\square$

Proof of Lemma 6. As the proof is identical for all three second order terms, we restrict ourselves to the first of them:

$$\sqrt{n}(\varphi_n - \varphi(\alpha))^2 \partial_\varphi^2 J = (\varphi_n - \varphi(\alpha))\sqrt{n}(\varphi_n - \varphi(\alpha))\partial_\varphi^2 J, \quad \partial_\varphi^2 J := \frac{\partial^2}{\partial\varphi^2} J_n(\psi(\xi), \varphi(\alpha)). \tag{92}$$

By explicit evaluation it is straightforward to verify that $\partial_\varphi^2 J < \infty$, almost surely if $\mathrm{E}X(1) < 0$, i.e., the stationary workload V exists and is bounded almost surely. The consistency assumption $\psi_n \to_{\mathbb{P}} \psi(\xi)$ implies, by Theorem 1, that $\varphi_n - \varphi(\alpha) \to_{\mathbb{P}} 0$. The scaled error term $\sqrt{n}(\varphi_n - \varphi(\alpha))$ is almost surely bounded as the first two moments of the error have finite limits due to Lemma 5. Therefore the error term indeed diminishes to zero, as $n \to \infty$. $\qquad \square$

Proof of Theorem 7. In Lemma 5 we verified the martingale CLT conditions of Lemma 2a for every element in $\boldsymbol{J}_n := (J_n(\psi(\xi), \varphi(\alpha_1)), \dots J_n(\psi(\xi), \varphi(\alpha_p)))$. Along the same lines, it is shown that any linear combination of the entries of $\boldsymbol{J}_n$ converges to a normal random variable, so that by the Cramér-Wold device we conclude joint asymptotic normality. All that remains is to compute for every pair $\alpha \neq \beta$ the covariance

$$\lim_{n \to \infty} \mathrm{Cov}\left[\sqrt{n}(\hat{\varphi}_n(\alpha; \psi_n) - \varphi(\alpha)), \sqrt{n}(\hat{\varphi}_n(\beta; \psi_n) - \varphi(\beta))\right]. \tag{93}$$

By (27), this covariance has the same limit as

$$\frac{n}{\partial J_{\varphi,\alpha}\partial J_{\varphi,\beta}}\mathrm{Cov}\left[J_n(\psi(\xi), \varphi(\alpha)) + \partial J_{\psi,\alpha}(\psi_n - \psi(\xi)), J_n(\psi(\xi), \varphi(\beta)) + \partial J_{\psi,\beta}(\psi_n - \psi(\xi))\right]. \tag{94}$$

The covariance can be computed by considering the covariance of the different pairs,

$$
\begin{aligned}
n\mathrm{Cov}&\left[J_n(\psi(\xi), \varphi(\alpha)) + \partial J_{\psi,\alpha}(\psi_n - \psi(\xi)), J_n(\psi(\xi), \varphi(\beta)) + \partial J_{\psi,\beta}(\psi_n - \psi(\xi))\right] \\
&= n\mathrm{Cov}\left[J_n(\psi(\xi), \varphi(\alpha)), J_n(\psi(\xi), \varphi(\beta))\right] \\
&\quad + n\partial J_{\psi,\beta}\mathrm{Cov}\left[J_n(\psi(\xi), \varphi(\alpha)), \psi_n - \psi(\xi)\right] \\
&\quad + n\partial J_{\psi,\alpha}\mathrm{Cov}\left[J_n(\psi(\xi), \varphi(\beta)), \psi_n - \psi(\xi)\right] \\
&\quad + \partial J_{\psi,\alpha}\partial J_{\psi,\beta}\mathrm{Var}\left(\sqrt{n}(\psi_n - \psi(\xi))\right) .
\end{aligned}
\tag{95}
$$

The limit of the latter three terms are $\sigma_{\alpha,\xi}^2$, $\sigma_{\beta,\xi}^2$ and σ_ξ^2, respectively. As $\mathrm{E}J_n(\psi(\xi), \varphi(\alpha)) = \mathrm{E}J_n(\psi(\xi), \varphi(\beta)) = 0$,

$$\mathrm{Cov}\left(J_n(\psi(\xi), \varphi(\alpha)), J_n(\psi(\xi), \varphi(\beta))\right) = \mathrm{E}\left[J_n(\psi(\xi), \varphi(\alpha))J_n(\psi(\xi), \varphi(\beta))\right] . \tag{96}$$

Similar to the arguments made in the proof of Lemma 5, by conditioning on V_{i-1} and applying (4) we obtain that, for every $i > j$,

$$\mathrm{E}\left[\left(e^{-\alpha V_i} - \frac{\xi}{\xi - \varphi(\alpha)}\left(e^{-\alpha V_{i-1}} - \frac{\alpha}{\psi(\xi)}e^{-\psi(\xi)V_{i-1}}\right)\right)\right. \tag{97}$$

$$\left.\cdot\left(e^{-\beta V_j} - \frac{\xi}{\xi - \varphi(\beta)}\left(e^{-\beta V_{j-1}} - \frac{\beta}{\psi(\xi)}e^{-\psi(\xi)V_{j-1}}\right)\right)\right] = 0 , \tag{98}$$

and similarly by conditioning on V_{j-1} for every $i < j$ we get a zero expectation for the respective terms. Hence, by (24),

$$n\mathrm{E}\left[J_n(\psi(\xi), \varphi(\alpha))J_n(\psi(\xi), \varphi(\beta))\right]$$

$$= \frac{1}{n}\sum_{i=1}^n \mathrm{E}\left[\left(e^{-\alpha V_i} - \frac{\xi}{\xi - \varphi(\alpha)}\left(e^{-\alpha V_{i-1}} - \frac{\alpha}{\psi(\xi)}e^{-\psi(\xi)V_{i-1}}\right)\right)\right.$$

$$\cdot \left(e^{-\beta V_i} - \frac{\xi}{\xi - \varphi(\beta)}\left(e^{-\beta V_{i-1}} - \frac{\beta}{\psi(\xi)}e^{-\psi(\xi)V_{i-1}}\right)\right)\Bigg]. \tag{99}$$

We compute the expectation for every $i \geq 1$,

$$\mathrm{E}\left[\left(e^{-\alpha V_i} - \frac{\xi}{\xi - \varphi(\alpha)}\left(e^{-\alpha V_{i-1}} - \frac{\alpha}{\psi(\xi)}e^{-\psi(\xi)V_{i-1}}\right)\right)\left(e^{-\beta V_i} - \frac{\xi}{\xi - \varphi(\beta)}\left(e^{-\beta V_{i-1}} - \frac{\beta}{\psi(\xi)}e^{-\psi(\xi)V_{i-1}}\right)\right)\right]$$

$$= \mathrm{E}\bigg[e^{-(\alpha+\beta)V_i} + \frac{\xi^2}{(\xi - \varphi(\alpha))(\xi - \varphi(\beta))}e^{-(\alpha+\beta)V_{i-1}} + \frac{\alpha\beta\xi^2}{\psi(\xi)^2(\xi - \varphi(\alpha))(\xi - \varphi(\beta))}e^{-2\psi(\xi)V_{i-1}}$$

$$- \frac{\beta\xi^2}{\psi(\xi)(\xi - \varphi(\alpha))(\xi - \varphi(\beta))}e^{-(\alpha+\psi(\xi))V_{i-1}} - \frac{\xi}{\xi - \varphi(\beta)}e^{-\alpha V_i - \beta V_{i-1}} + \frac{\beta\xi}{\psi(\xi)(\xi - \varphi(\beta))}e^{-\alpha V_i - \psi(\xi)V_{i-1}}$$

$$- \frac{\alpha\xi^2}{\psi(\xi)(\xi - \varphi(\alpha))(\xi - \varphi(\beta))}e^{-(\beta+\psi(\xi))V_{i-1}} - \frac{\xi}{\xi - \varphi(\alpha)}e^{-\beta V_i - \alpha V_{i-1}} + \frac{\alpha\xi}{\psi(\xi)(\xi - \varphi(\alpha))}e^{-\beta V_i - \psi(\xi)V_{i-1}}\bigg], \tag{100}$$

and by applying the GPK formula (2), with the additional conditioning step used before for the joint limiting distribution of (V_{i-1}, V_i), we obtain the limit of the expectation,

$$\frac{(\alpha+\beta)\varphi'(0)}{\varphi(\alpha+\beta)} + \frac{\xi^2\varphi'(0)}{(\xi - \varphi(\alpha))(\xi - \varphi(\beta))}\left(\frac{\alpha+\beta}{\varphi(\alpha+\beta)} + \frac{2\alpha\beta}{\psi(\xi)\varphi(2\psi(\xi))} - \frac{\beta(\alpha+\psi(\xi))}{\psi(\xi)\varphi(\alpha+\psi(\xi))} - \frac{\alpha(\beta+\psi(\xi))}{\psi(\xi)\varphi(\beta+\psi(\xi))}\right.$$

$$\left. - 2\left(\frac{\alpha+\beta}{\varphi(\alpha+\beta)} - \frac{\alpha(\beta+\psi(\xi))}{\psi(\xi)\varphi(\beta+\psi(\xi))} - \frac{\beta(\alpha+\psi(\xi))}{\psi(\xi)\varphi(\alpha+\psi(\xi))} + \frac{2\alpha\beta}{\psi(\xi)\varphi(2\psi(\xi))}\right)\right). \tag{101}$$

Rearranging the terms yields the desired expression for $\sigma_{J,\alpha,\beta}^2$. $\qquad\square$

B Proofs for Section 4

Proof of Lemma 10. The smoothness is due to the fact that $\ell_n(\psi)$, $\ell_n'(\psi)$ and $\ell_n''(\psi)$ are all continuous functions of ψ. Furthermore, for every $\xi < \psi < \infty$ the functions are well defined and finite and therefore bounded on any compact interval. $\qquad\square$

Proof of Lemma 11. The first derivative of the empirical log-likelihood function is

$$\ell_n'(\psi) := \frac{1}{n}L_n'(\psi) = \frac{1}{n}\sum_{i=1}^{n}\frac{\frac{1}{\psi} + V_{i-1}}{1 - \frac{\xi}{\psi}e^{-\psi V_{i-1}}}\left(\frac{\xi}{\psi}e^{-\psi V_{i-1}} - Y_i\right), \tag{102}$$

for any $\psi > \xi$ (where we note that it is otherwise not always well defined). The limiting (time-average) distribution of the pair (Y_n, V_{n-1}) is given by the distribution of a stationary workload observation and the probability to find the system empty in an exponentially distributed time after the observation. Thus, the joint distribution is characterized by $V = V(\infty)$, combined with conditional probability $\mathrm{P}(Y = 1 \,|\, V) = \frac{\xi}{\psi}e^{-\psi V}$. Applying PASTA for the sampling process $N(t)$,

$$\ell_n'(\psi) = \frac{1}{N(t_n)}\ell_{N(t_n)}'(\psi) \to_{\mathrm{as}} \mathrm{E}\ell_1'(\psi) = \mathrm{E}\left[\frac{\frac{1}{\psi} + V}{1 - \frac{\xi}{\psi}e^{-\psi V}}\left(\frac{\xi}{\psi}e^{-\psi V} - Y\right)\right], \tag{103}$$

and by iterating the expectation we obtain (47),

$$\ell'_n(\psi) \to_{\mathrm{as}} \mathrm{E}(\mathrm{E}(\ell'_1(\psi)|V)) = \mathrm{E}\left[\frac{\frac{1}{\psi} + V}{1 - \frac{\xi}{\psi}e^{-\psi V}}\left(\frac{\xi}{\psi}e^{-\psi V} - \frac{\xi}{\psi(\xi)}e^{-\psi(\xi)V}\right)\right] . \tag{104}$$

The same argument yields (48). □

Proof of Lemma 13. From

$$\mathrm{E}\ell'_n(\psi(\xi)) = \frac{1}{n}\sum_{i=1}^{n}\mathrm{E}\left[\frac{\frac{1}{\psi(\xi)} + V_{i-1}}{1 - \frac{\xi}{\psi(\xi)}e^{-\psi(\xi)V_{i-1}}}\left(\frac{\xi}{\psi(\xi)}e^{-\psi(\xi)V_{i-1}} - Y_i\right)\right]$$

$$= \frac{1}{n}\sum_{i=1}^{n}\mathrm{E}\left[\frac{\frac{1}{\psi(\xi)} + V_{i-1}}{1 - \frac{\xi}{\psi(\xi)}e^{-\psi(\xi)V_{i-1}}}\left(\frac{\xi}{\psi(\xi)}e^{-\psi(\xi)V_{i-1}} - \frac{\xi}{\psi(\xi)}e^{-\psi(\xi)V_{i-1}}\right)\right] = 0 , \tag{105}$$

we find

$$\mathrm{Var}(\ell'_n(\psi(\xi))) = \frac{1}{n^2}\sum_{i=1}^{n}\mathrm{E}\left[\frac{\frac{1}{\psi(\xi)} + V_{i-1}}{1 - \frac{\xi}{\psi(\xi)}e^{-\psi(\xi)V_{i-1}}}\left(\frac{\xi}{\psi(\xi)}e^{-\psi(\xi)V_{i-1}} - Y_i\right)\right]^2 + \frac{1}{n^2}\sum_{i=1}^{n}\sum_{j\neq i}s_{ij} , \tag{106}$$

where $s_{ij} :=$

$$\mathrm{E}\left[\frac{\frac{1}{\psi(\xi)} + V_{i-1}}{1 - \frac{\xi}{\psi(\xi)}e^{-\psi(\xi)V_{i-1}}}\left(\frac{\xi}{\psi(\xi)}e^{-\psi(\xi)V_{i-1}} - Y_i\right)\right]\left[\frac{\frac{1}{\psi(\xi)} + V_{j-1}}{1 - \frac{\xi}{\psi(\xi)}e^{-\psi(\xi)V_{j-1}}}\left(\frac{\xi}{\psi(\xi)}e^{-\psi(\xi)V_{j-1}} - Y_j\right)\right] . \tag{107}$$

By conditioning on V_{i-1} and using the fact that $\mathrm{P}(Y_i = 1|V_{i-1}) = \frac{\xi}{\psi(\xi)}e^{-\psi(\xi)V_{j-1}}$, where $\{Y_i|V_{i-1}\}$ is independent of (V_{j-1}, Y_j) for any $j < i$, we conclude that $s_{ij} = 0$ for all $i < j$. Similarly, conditioning on V_{j-1} yields that $s_{ij} = 0$ for all $j > i$. We find that

$$\mathrm{Var}(\ell'_n(\psi(\xi))) = \frac{1}{n^2}\sum_{i=1}^{n}\mathrm{E}\left[\frac{\frac{1}{\psi(\xi)} + V_{i-1}}{1 - \frac{\xi}{\psi(\xi)}e^{-\psi(\xi)V_{i-1}}}\left(\frac{\xi}{\psi(\xi)}e^{-\psi(\xi)V_{i-1}} - Y_i\right)\right]^2 . \tag{108}$$

Observe that, for $i = 1, \ldots, n$,

$$m_i(v) := \mathrm{E}\left[\left(\frac{\frac{1}{\psi(\xi)} + V_{i-1}}{1 - \frac{\xi}{\psi(\xi)}e^{-\psi(\xi)V_{i-1}}}\right)^2\left(\frac{\xi}{\psi(\xi)}e^{-\psi(\xi)V_{i-1}} - Y_i\right)^2\bigg|V_{i-1} = v\right]$$

$$= \left(\frac{\frac{1}{\psi(\xi)} + v}{1 - \frac{\xi}{\psi(\xi)}e^{-\psi(\xi)v}}\right)^2\left(\frac{\xi^2}{\psi(\xi)^2}e^{-2\psi(\xi)v} - 2\frac{\xi}{\psi(\xi)}e^{-\psi(\xi)v}\mathrm{E}[Y_i|V_{i-1} = v] + \mathrm{E}[Y_i^2|V_{i-1} = v]\right)$$

$$= \left(\frac{\frac{1}{\psi(\xi)} + v}{1 - \frac{\xi}{\psi(\xi)}e^{-\psi(\xi)v}}\right)^2\left(\frac{\xi^2}{\psi(\xi)^2}e^{-2\psi(\xi)v} - 2\frac{\xi^2}{\psi(\xi)^2}e^{-2\psi(\xi)v} + \frac{\xi}{\psi(\xi)}e^{-\psi(\xi)v}\right)$$

$$= \frac{\frac{\xi}{\psi(\xi)}e^{-\psi(\xi)v}\left(\frac{1}{\psi(\xi)} + v\right)^2}{1 - \frac{\xi}{\psi(\xi)}e^{-\psi(\xi)v}} . \tag{109}$$

Applying PASTA once more yields

$$n\mathrm{Var}(\ell'_n(\psi(\xi))) = \frac{1}{N(t_n)} \sum_{i=1}^{N(t_n)} \mathrm{E}\, m_i(V(t_{i-1})) \to_{\mathrm{as}} \mathrm{E}\, m_1(V) = I_\xi \; . \tag{110}$$

Next we apply the martingale CLT of Lemma 2, using the same notation as before for the corresponding martingale difference terms. For $i \in \{1, \ldots, n\}$, we denote

$$Z_i := \frac{\left(\frac{1}{\psi(\xi)} + V_{i-1}\right)\left(\frac{\xi}{\psi(\xi)} e^{-\psi(\xi)V_{i-1}} - Y_i\right)}{1 - \frac{\xi}{\psi(\xi)} e^{-\psi(\xi)V_{i-1}}} \; , \quad Z_{ni} := \frac{Z_i}{\sqrt{\mathrm{Var}(L'_n(\psi(\xi)))}}. \tag{111}$$

As $\mathrm{E}Z_i = 0$, $L'_n(\psi(\xi)) = \sum_{i=1}^n Z_i$ is a martingale with respect to $\{V_1, \ldots, V_n\}$. For any n, $\mathrm{Var}(L'_n(\psi(\xi))) > 0$ and $\mathrm{P}(Z_i < \infty) = 1$ as $\mathrm{E}Z_i = 0$ for all $i \in \{1, \ldots, n\}$. Hence $\mathrm{E}\max_{1 \le i \le n} Z_{ni}^2 < \infty$ so that the condition (20) holds. We have established that

$$\frac{1}{n}\mathrm{Var}(L'_n(\psi(\xi))) = n\mathrm{Var}(\ell'_n(\psi(\xi))) \to_{\mathrm{as}} I_\xi < \infty \; , \tag{112}$$

and therefore $\sqrt{\mathrm{Var}(L'_n(\psi(\xi)))} \to_{\mathrm{as}} \infty$. Furthermore, as $\mathrm{P}(Z_i < \infty) = 1$,

$$\max_{1 \le i \le n} |Z_{ni}| = \max_{1 \le i \le n} \frac{|Z_i|}{\sqrt{\mathrm{Var}(L'_n(\psi(\xi)))}} \to_{\mathrm{as}} 0 \tag{113}$$

and thus condition (21) is satisfied as well. Therefore, by Lemma 2,

$$\frac{L'_n(\psi(\xi))}{\sqrt{\mathrm{Var}(L'_n(\psi(\xi)))}} \to_{\mathrm{d}} \mathrm{N}(0,1) \; . \tag{114}$$

Recall that $\ell'_n(\psi(\xi)) = n^{-1} L'_n(\psi(\xi))$, so that

$$\frac{L'_n(\psi(\xi))}{\sqrt{\mathrm{Var}(L'_n(\psi(\xi)))}} = \frac{n\ell'_n(\psi(\xi))}{\sqrt{\mathrm{Var}(n\ell'_n(\psi(\xi)))}} = \frac{\sqrt{n}\ell'_n(\psi(\xi))}{\sqrt{n\mathrm{Var}(\ell'_n(\psi(\xi)))}} \; . \tag{115}$$

Because $\sqrt{n\mathrm{Var}(\ell'_n(\psi(\xi)))} \to_{\mathrm{as}} \sqrt{I_\xi}$, and applying Slutsky's lemma, we obtain the claimed result: $\sqrt{n}\ell'_n(\psi(\xi)) \to_{\mathrm{d}} \sqrt{I_\xi}\mathrm{N}(0,1)$. $\qquad\square$

Proof of Lemma 14. Taking derivative of $\ell'_n(\psi)$, as given in (45), yields

$$\begin{aligned}
\ell''_1(\psi) &= -\frac{1}{\left(1 - \frac{\xi}{\psi(\xi)} e^{-\psi(\xi)V_0}\right)^2}\Bigg[\frac{\xi}{\psi(\xi)} e^{-\psi(\xi)V_0}\left(\frac{1}{\psi(\xi)} + V_0\right)^2 (1 - Y_1) \\
&\quad + \frac{1}{\psi(\xi)^2}\left(1 - \frac{\xi}{\psi(\xi)} e^{-\psi(\xi)V_0}\right)\left(\frac{\xi}{\psi(\xi)} e^{-\psi(\xi)V_0} - Y_1\right)\Bigg] \\
&= -\frac{\frac{\xi}{\psi(\xi)} e^{-\psi(\xi)V_0}\left(\frac{1}{\psi(\xi)} + v\right)^2}{1 - \frac{\xi}{\psi(\xi)} e^{-\psi(\xi)V_0}} \; .
\end{aligned} \tag{116}$$

30

The conditional expectation for any initial workload $V_0 = v$ is

$$\mathrm{E}[\ell_1''(\psi(\xi))|V_0 = v] = -\frac{1}{\left(1 - \frac{\xi}{\psi(\xi)}e^{-\psi(\xi)v}\right)^2}\left[\frac{\xi}{\psi(\xi)}e^{-\psi(\xi)v}\left(\frac{1}{\psi(\xi)} + v\right)^2(1 - \mathrm{P}_{\psi(\xi)}(Y_1|V_0 = v))\right.$$

$$\left. + \frac{1}{\psi(\xi)^2}\left(1 - \frac{\xi}{\psi(\xi)}e^{-\psi(\xi)v}\right)\left(\frac{\xi}{\psi(\xi)}e^{-\psi(\xi)v} - \mathrm{P}_{\psi(\xi)}(Y_1|V_0 = v)\right)\right]$$

$$= -\frac{\frac{\xi}{\psi(\xi)}e^{-\psi(\xi)v}\left(\frac{1}{\psi(\xi)} + v\right)^2}{1 - \frac{\xi}{\psi(\xi)}e^{-\psi(\xi)v}} .$$
(117)

Denote the limit of the second derivative, that is given by (48) in Lemma 11, at value ψ by $k_2(\psi)$. Recall that by (7), $\mathrm{P}_{\psi(\xi)}(Y_1|V_0 = v) = \frac{\xi}{\psi(\xi)}e^{-\psi(\xi)v}$, and thus by taking expectation with respect to the stationary workload we obtain

$$\ell_n''(\psi) \to_{\mathrm{as}} \mathrm{E}\left[\mathrm{E}[\ell_1''(\psi(\xi))|V]\right] = k_2(\psi(\xi)) = -I_\xi .$$
(118)

Now consider the difference between D_n and $-I_\xi$:

$$|D_n - (-I_\xi)| = \left|\int_0^1 \left[\ell_n''(t\hat{\psi}_n + (1-t)\psi(\xi)) + I_\xi\right] \mathrm{d}t\right|$$

$$\leq \int_0^1 \left|\ell_n''(t\hat{\psi}_n + (1-t)\psi(\xi)) + I_\xi\right| \mathrm{d}t$$
(119)

$$\leq \sup_{t\in[0,1]} \left|\ell_n''(t\hat{\psi}_n + (1-t)\psi(\xi)) + I_\xi\right| .$$

By the consistency of the MLE established in Theorem 8, $t\hat{\psi}_n + (1-t)\psi(\xi) \to_{\mathrm{as}} \psi(\xi)$ for any $t \in [0,1]$. Hence, for any $\epsilon > 0$ there exists some N such that $|t\hat{\psi}_n + (1-t)\psi(\xi) - \psi(\xi)| < \epsilon$, for all $n \geq N$ almost surely, and therefore

$$|D_n + I_\xi| \leq \sup_{\psi:\ |\psi-\psi(\xi)|<\epsilon} \left|\ell_n''(\psi) + I_\xi\right|$$

$$\leq \sup_{\psi:\ |\psi-\psi(\xi)|<\epsilon} \left|\ell_n''(\psi) - k_2(\psi)\right| + \sup_{\psi:\ |\psi-\psi(\xi)|<\epsilon} \left|k_2(\psi) + I_\xi\right| .$$
(120)

By Lemma 11, $\ell_n''(\psi) \to_{\mathrm{as}} k_2(\psi)$. Consequently, the first term goes to zero almost surely. Let $\delta(\epsilon) := \sup_{\psi:\ |\psi-\psi(\xi)|<\epsilon} |k_2(\psi) + I_\xi|$, then (as k_2 is continuous and bounded on finite intervals) we have that $\delta(\epsilon) \to 0$ as $\epsilon \to 0$. We conclude that

$$|D_n + I_\xi| \to_{\mathrm{as}} 0 ,$$
(121)

and the result follows. $\square$

Proof of Theorem 15. Let $\Omega_\xi(\alpha) := \xi/(\xi - \varphi(\alpha))$ and

$$\Xi_{\xi,i}(\alpha) := e^{-\alpha V_{i-1}} - \frac{\alpha}{\psi(\xi)}e^{-\psi(\xi)V_{i-1}}.$$
(122)

As $\mathrm{E}\left[J_n(\psi(\xi), \varphi(\alpha))\right] = 0$,

$$\mathrm{Cov}(\sqrt{n}J_n(\psi(\xi), \varphi(\alpha)), \sqrt{n}(\hat{\psi}_n - \psi(\xi))) = n\mathrm{E}\left[J_n(\psi(\xi), \varphi(\alpha))(\hat{\psi}_n - \psi(\xi))\right] . \tag{123}$$

Recall that as $n \to \infty$, we almost surely have $\sqrt{n}(\hat{\psi}_n - \psi(\xi)) \approx \sqrt{n}\ell'_n(\psi(\xi))/I_\xi$, and therefore

$$\lim_{n\to\infty} \mathrm{Cov}(\sqrt{n}J_n(\psi(\xi), \varphi(\alpha)), \sqrt{n}(\hat{\psi}_n - \psi(\xi))) = \frac{1}{I_\xi}\lim_{n\to\infty} n\mathrm{E}\left[J_n(\psi(\xi), \varphi(\alpha))\ell'_n(\psi(\xi))\right] . \tag{124}$$

Direct computations yield that for any finite n the expectation $\mathrm{E}\left[J_n(\psi(\xi), \varphi(\alpha))\,\ell'_n(\psi(\xi))\right]$ equals

$$\frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n}\mathrm{E}\left[\frac{\left(e^{-\alpha V_i} - \Omega_\xi(\alpha)\Xi_{\xi,i}(\alpha)\right)\left(\frac{1}{\psi(\xi)} + V_{j-1}\right)\left(\frac{\xi}{\psi(\xi)}e^{-\psi(\xi)V_{j-1}} - Y_j\right)}{1 - \frac{\xi}{\psi(\xi)}e^{-\psi(\xi)V_{j-1}}}\right] . \tag{125}$$

Conditioning on $(V_0, \ldots, V_{i-1})$, all terms in the above display with $i > j$ can be written as

$$\frac{\left(\frac{1}{\psi(\xi)} + V_{j-1}\right)\left(\frac{\xi}{\psi(\xi)}e^{-\psi(\xi)V_{j-1}} - Y_j\right)}{1 - \frac{\xi}{\psi(\xi)}e^{-\psi(\xi)V_{j-1}}}\mathrm{E}\left[e^{-\alpha V_i} - \Omega_\xi(\alpha)\Xi_{\xi,i}(\alpha)\Big|V_{i-1}\right] = 0 , \tag{126}$$

and for the terms such that $i < j$, conditioning on $(V_0, \ldots, V_{j-1})$ yields

$$\left(e^{-\alpha V_i} - \Omega_\xi(\alpha)\Xi_{\xi,i}(\alpha)\right)\mathrm{E}\left[\frac{\left(\frac{1}{\psi(\xi)} + V_{j-1}\right)\left(\frac{\xi}{\psi(\xi)}e^{-\psi(\xi)V_{j-1}} - Y_j\right)}{1 - \frac{\xi}{\psi(\xi)}e^{-\psi(\xi)V_{j-1}}}\Big|V_{j-1}\right] = 0 . \tag{127}$$

For $j = i$, we compute the expectation conditioned on V_{i-1},

$$\mathrm{E}\left[\frac{\left(e^{-\alpha V_i} - \Omega_\xi(\alpha)\Xi_{\xi,i}(\alpha)\right)\left(\frac{1}{\psi(\xi)} + V_{i-1}\right)\left(\frac{\xi}{\psi(\xi)}e^{-\psi(\xi)V_{i-1}} - Y_i\right)}{1 - \frac{\xi}{\psi(\xi)}e^{-\psi(\xi)V_{i-1}}}\Big|V_{i-1}\right]$$

$$= \mathrm{E}\left[\frac{\frac{1}{\psi(\xi)} + V_{i-1}}{1 - \frac{\xi}{\psi(\xi)}e^{-\psi(\xi)V_{i-1}}}\left(\frac{\xi}{\psi(\xi)}e^{-\psi(\xi)V_{i-1}}\mathrm{E}[e^{-\alpha V_i}|V_{i-1}] - \mathrm{E}[Y_ie^{-\alpha V_i}]\right)\right.$$

$$\left. - \Omega_\xi(\alpha)\left(\frac{\Xi_{\xi,i}(\alpha)\left(\frac{1}{\psi(\xi)} + V_{i-1}\right)\mathrm{E}\left[\frac{\xi}{\psi(\xi)}e^{-\psi(\xi)V_{i-1}} - Y_i\Big|V_{i-1}\right]}{1 - \frac{\xi}{\psi(\xi)}e^{-\psi(\xi)V_{i-1}}}\right)\right]$$

$$= \mathrm{E}\left[\frac{\frac{1}{\psi(\xi)} + V_{i-1}}{1 - \frac{\xi}{\psi(\xi)}e^{-\psi(\xi)V_{i-1}}}\left(\frac{\xi}{\psi(\xi)}e^{-\psi(\xi)V_{i-1}}\mathrm{E}[e^{-\alpha V_i}|V_{i-1}] - \mathrm{P}(Y_i = 1, V_i = 0)\right)\right]$$

$$= \mathrm{E}\left[\frac{\xi^2\left(\frac{1}{\psi(\xi)} + V_{i-1}\right)e^{-\psi(\xi)V_{i-1}}}{\psi(\xi)(\xi - \varphi(\alpha))\left(1 - \frac{\xi}{\psi(\xi)}e^{-\psi(\xi)V_{i-1}}\right)}\left(\Xi_{\xi,i}(\alpha) - \frac{\xi - \varphi(\alpha)}{\xi}\right)\right] , \tag{128}$$

where the second equality follows from the fact that given V_{i-1}, Y_i is a Bernoulli random variable with probability $\frac{\xi}{\psi(\xi)}e^{-\psi(\xi)V_{i-1}}$ and equals one only when $V_{i-1} = 0$. Finally, by applying PASTA we conclude that

$$\lim_{n\to\infty} \mathrm{Cov}(\sqrt{n}J_n(\psi(\xi), \varphi(\alpha), \sqrt{n}(\hat{\psi}_n - \psi(\xi))) = \frac{1}{I_\xi}\lim_{n\to\infty} n\mathrm{E}\left[J_n(\psi(\xi), \varphi(\alpha))\ell'_n(\psi(\xi))\right]$$

32

$$= \frac{\xi^2}{\psi(\xi)(\xi - \varphi(\alpha))I_\xi} \lim_{n \to \infty} \frac{1}{n} \sum_{i=1}^{n} \mathrm{E}\left[\frac{\left(\frac{1}{\psi(\xi)} + V_{i-1}\right) e^{-\psi(\xi)V_{i-1}}}{1 - \frac{\xi}{\psi(\xi)} e^{-\psi(\xi)V_{i-1}}} \left(\Xi_{\xi,i}(\alpha) - \frac{\xi - \varphi(\alpha)}{\xi} \right) \right]$$

$$= \frac{\xi^2}{\psi(\xi)(\xi - \varphi(\alpha))I_\xi} \mathrm{E}\left[\frac{\left(\frac{1}{\psi(\xi)} + V\right) e^{-\psi(\xi)V}}{1 - \frac{\xi}{\psi(\xi)} e^{-\psi(\xi)V}} \left(e^{-\alpha V} - \frac{\alpha}{\psi(\xi)} e^{-\psi(\xi)V} - \frac{\xi - \varphi(\alpha)}{\xi} \right) \right] . \quad (129)$$

This completes the proof. $\qquad\square$

C   Proofs for Section 5

Proof of Proposition 16. Consider the expected workload after an exponential time $T \sim \exp(\xi)$ given that the initial workload was v,

$$\mathrm{E}(V(T)|V(0) = v) = \mathrm{E}\left[V(T)\mathbf{1}\left(\inf_{0 \leq s \leq T} X(s) \geq -v \right) |V(0) = v \right]$$
$$+ \mathrm{E}\left[V(T)\mathbf{1}\left(\inf_{0 \leq s \leq T} X(s) < -v \right) |V(0) = v \right]$$
$$= \mathrm{E}\left[(v + X(T))\mathbf{1}\left(\inf_{0 \leq s \leq T} X(s) \geq -v \right) |V(0) = v \right]$$
$$+ \mathrm{P}\left(\inf_{0 \leq s \leq T} X(s) < -v \right) \mathrm{E}[V(T)|V(0) = 0]$$
$$\leq v + \mathrm{E}X(T) + \mathrm{P}\left(\inf_{0 \leq s \leq T} X(s) < -v \right) \mathrm{E}[V(T)|V(0) = 0] ,$$

yielding, with the last step being due to [11, Lemma 6.2]

$$\mathrm{E}(V(T) - v \,|\, V(0) = v) - \mathrm{E}X(T) \leq \mathrm{P}\left(\inf_{0 \leq s \leq T} X(s) < -v \right) \mathrm{E}[V(T) \,|\, V(0) = 0]$$
$$= e^{-\psi(\xi)v} \mathrm{E}[V(T)|V(0) = 0] . \quad (130)$$

Consequently, for any threshold $\tau > 0$ and all $v \geq \tau$,

$$\mathrm{E}(V(T) - v \,|\, V(0) = v) - \mathrm{E}X(T) \leq \mathrm{E}[V(T) \,|\, V(0) = 0]e^{-\psi(\xi)\tau} . \quad (131)$$

The expected bias of $\hat\vartheta_m(\tau)$ is

$$b_m(\vartheta; \tau) = \mathrm{E}\left[\hat\vartheta_m(\tau) - \vartheta \right] = \mathrm{E}\left[\frac{\xi}{m} \sum_{j=1}^{m} (V_{i(j)} - V_{i(j)-1}) - \left(-\varphi'(0) \right) \right]$$
$$= \frac{\xi}{m} \sum_{j=1}^{m} \mathrm{E}\left[(V_{i(j)} - V_{i(j)-1}) - \mathrm{E}X(T) \right] . \quad (132)$$

For every $j = 1 \ldots, m$, conditioning on $V_{i(j)-1} = v$ (with $v \geq \tau$), applying (131) yields

$$\mathrm{E}\left[(V_{i(j)} - V_{i(j)-1}) - \mathrm{E}X(T) \Big| V_{i(j)-1} = v \right] \leq \mathrm{E}[V(T)|V(0) = 0]e^{-\psi(\xi)\tau} , \quad (133)$$

and the bound (67) follows by total expectation. $\qquad\square$

Proof of Proposition 17. For any sample $\boldsymbol{V} = (V_0, \ldots, V_n)$ such that $n \geq m$, let

$$f(x, n) := \frac{\xi(V_n - V_0)}{n} - \frac{\xi}{n} \sum_{i=1}^{n} \frac{e^{-xV_{i-1}}}{x} , \tag{134}$$

then the estimator $\tilde{\psi}_m$ is given by solving (65), i.e., $\hat{\vartheta}_m(\tau) = f(\tilde{\psi}_m, M(m, \tau))$. By (5), for any $\{M(m, \tau) = n\}$,

$$\mathrm{E}[f(\psi(\xi), n) - f(\tilde{\psi}_m, n)] = \mathrm{E}\left[\hat{\vartheta}_m(\tau) + \varphi'(0)\right] , \tag{135}$$

thus, the upper bound (67) for the bias of $\hat{\varphi}_m(\tau)$ we obtain yields

$$\mathrm{E}[f(\psi(\xi), n) - f(\tilde{\psi}_m, n)] \leq \xi e^{-\psi(\xi)\tau} \mathrm{E}[V(T)|V(0) = 0] . \tag{136}$$

As $f(x, n)$ is continuous and monotone decreasing with x, this also means that the bias of $f(\tilde{\psi}_n, n)$ from $f(\psi(\xi), n)$ is bounded. The bound further holds when taking expectation over the possible values of $M(m, \tau)$. Moreover, if the threshold τ is increased, then the bias vanishes, and by continuity so does the bias of the estimator for $\tilde{\psi}$. $\qquad \square$

Proof of Proposition 18. We define

$$K(t) := \sum_{i=1}^{N(t)} \mathbf{1}(V(t_i) \geq \tau), \tag{137}$$

where $N(t)$ is the Poisson sampling process and t_i are the event times. By PASTA, as $t \to \infty$,

$$\frac{K(t)}{N(t)} = \frac{1}{N(t)} \sum_{i=1}^{N(t)} \mathbf{1}(V(t_i) \geq \tau) \to_{\mathbb{P}} \mathrm{P}(V > \tau) . \tag{138}$$

Observe that $m = K(t_{M(m,\tau)})$, so that for $M(m, \tau) = n$,

$$\begin{aligned}
\hat{\vartheta}_m(\tau) &= \frac{\xi}{K(t_n)} \sum_{i=1}^{N(t_n)} (V(t_i) - V(t_{i-1})) \mathbf{1}(V(t_{i-1}) \geq \tau) \\
&= \xi \frac{N(t_n)}{K(t_n)} \cdot \frac{1}{N(t_n)} \sum_{i=1}^{N(t_n)} (V(t_i) - V(t_{i-1})) \mathbf{1}(V(t_{i-1}) \geq \tau) .
\end{aligned} \tag{139}$$

The term $N(t_n)/K(t_n)$ converges to $(\mathrm{P}(V > \tau))^{-1}$ as $n \to \infty$, and this also holds for $m \to \infty$ as $M(m, \tau) \to \infty$ almost surely in this case. By PASTA, the limiting joint distribution of the indicator and the increment is given by $(V(T) - V, \mathbf{1}(V \geq \tau))$, where $T \sim \exp(\xi)$. Therefore, we can apply Slutsky's lemma to conclude that the limiting distribution of the product is given by the product of the limiting distributions,

$$\hat{\vartheta}_m(\tau) \to_{\mathbb{P}} \xi \frac{\mathrm{E}\left[\mathrm{E}[(V(T) - V)\mathbf{1}(V \geq \tau) \mid V]\right]}{\mathrm{P}(V > \tau)} . \tag{140}$$

By first conditioning on $V = v$ and then plugging in the explicit expectation of $\mathrm{E}[V(T) - v]$ given in (5) we obtain that for $v \geq \tau$

$$\mathrm{E}\left[(V(T) - V)\mathbf{1}(V \geq \tau)\Big|V = v\right] = \frac{e^{-\psi(\xi)v}}{\psi(\xi)} - \frac{\varphi'(0)}{\xi} , \tag{141}$$

and 0 else. Then the asymptotic estimator is

$$\vartheta(\tau) = \frac{\xi}{\mathrm{P}(V > \tau)} \int_\tau^\infty \left[\frac{e^{-\psi(\xi)v}}{\psi(\xi)} - \frac{\varphi'(0)}{\xi} \right] \mathrm{d}F_V(v) = \frac{\xi \mathrm{E}\left[e^{-\psi(\xi)V} \mathbf{1}(V \geq \tau) \right]}{\psi(\xi) \mathrm{P}(V > \tau)} - \varphi'(0) \ . \quad (142)$$

$\square$

Acknowledgments

References

[1] P. Afèche and B. Ata. Bayesian dynamic pricing in queueing systems with unknown delay cost characteristics. *Manufacturing & Service Operations Management*, 15(2):292–304, 2013.

[2] N. Antunes, G. Jacinto, and A. Pacheco. Probing a M/G/1 queue with general Input and service times. *SIGMETRICS Perform. Eval. Rev.*, 41(3):34–36, Jan. 2014.

[3] N. Antunes, G. Jacinto, A. Pacheco, and C. Wichelhaus. Estimation of the traffic intensity in a piecewise-stationary $M_t/G_t/1$ queue with probing. *SIGMETRICS Perform. Eval. Rev.*, 44(2):3–5, Sept. 2016.

[4] M. Armony and E. L. Plambeck. The impact of duplicate orders on demand estimation and capacity investment. *Management Science*, 51(10):1505–1518, 2005.

[5] A. Asanjarani, Y. Nazarathy, and P. K. Pollett. Parameter and state estimation in queues and related stochastic models: A bibliography. *ArXiv:1701.08338*, 2017.

[6] R. Bekker, O. J. Boxma, and J. A. C. Resing. Lévy processes with adaptable exponent. *Advances in Applied Probability*, 41(1):177–205, 2009.

[7] D. Belomestny. Statistical inference for time-changed Lévy processes via composite characteristic function estimation. *The Annals of Statistics*, 39(4):2205–2242, 2011.

[8] N. H. Bingham and S. M. Pitts. Nonparametric inference from M/G/1 busy periods. *Communications in Statistics. Stochastic Models*, 15(2):247–272, 1999.

[9] N. Blanghaps, Y. Nov, and G. Weiss. Sojourn time estimation in an M/G/∞ queue with partial information. *Journal of Applied Probability*, 50(4):1044–1056, 2013.

[10] B. Buchmann and R. Grübel. Decompounding: an estimation problem for Poisson random sums. *The Annals of Statistics*, 31(4):1054–1074, 2003.

[11] K. Dębicki and M. Mandjes. *Queues and Lévy Fluctuation Theory.* Springer, 2015.

[12] D. Duffie and P. Glynn. Estimation of continuous-time Markov processes sampled at random time intervals. *Econometrica*, 72(6):1773–1808, 2004.

[13] T. S. Ferguson. *A Course in Large Sample Theory.* Chapman & Hall, 1996.

[14] P. W. Glynn, B. Melamed, and W. Whitt. Estimating customer and time averages. *Operations Research*, 41(2):400–408, 1993.

[15] A. Goldenshluger. Nonparametric estimation of the service time distribution in the M/G/∞ queue. *Advances in Applied Probability*, 48(4):1117–1138, 2016.

[16] A. Goldenshluger. The M/G/∞ estimation problem revisited. *Bernoulli*, 24(4A):2531–2568, 2018.

[17] S. Gugushvili. Nonparametric estimation of the characteristic triplet of a discretely observed Lévy process. *Journal of Nonparametric Statistics*, 21(3):321–343, 2009.

[18] P. Hall and C. C. Heyde. *Martingale Limit Theory and its Application*. Academic Press, 1980.

[19] M. B. Hansen and S. M. Pitts. Nonparametric inference from the M/G/1 workload. *Bernoulli*, 12:737–759, 2006.

[20] R. Hassin. *Rational Queueing*. CRC Press, 2016.

[21] R. Hassin and M. Haviv. *To Queue or Not to Queue: Equilibrium Behavior in Queueing Systems*. Springer, 2003.

[22] R. D. H. Heijmans and J. R. Magnus. Consistent maximum-likelihood estimation with dependent observations: the general (non-normal) case and the normal case. *Journal of Econometrics*, 32(2):253–285, 1986.

[23] R. D. H. Heijmans and J. R. Magnus. On the first–order efficiency and asymptotic normality of maximum likelihood estimators obtained from dependent observations. *Statistica Neerlandica*, 40(3):169–188, 1986.

[24] C. C. Heyde. *Quasi-Likelihood and its Application: A General Approach to Optimal Parameter Estimation*. Springer Science & Business Media, 1997.

[25] O. Kella, O. Boxma, and M. Mandjes. A Lévy process reflected at a Poisson age process. *Journal of Applied Probability*, 43(1):221–230, 2006.

[26] A. Kyprianou. *Introductory Lectures on Fluctuations of Lévy Processes with Applications*. Springer, 2006.

[27] M. H. Neumann and M. Reiß. Nonparametric estimation for Lévy processes from low-frequency observations. *Bernoulli*, 15(1):223–248, 2009.

[28] A. Novak, P. Taylor, and D. Veitch. The distribution of the number of arrivals in a subinterval of a busy period of a single server queue. *Queueing Systems*, 53(3):105–114, 2006.

[29] S. M. Pitts. Nonparametric estimation of the stationary waiting time distribution function for the GI/G/1 queue. *The Annals of Statistics*, 22(3):1428–1446, 1994.

[30] J. V. Ross, T. Taimre, and P. K. Pollett. Estimation for queues from queue length data. *Queueing Systems*, 55(2):131–138, 2007.

[31] S. Schweer and C. Wichelhaus. Nonparametric estimation of the service time distribution in the discrete-time GI/G/∞ queue with partial information. *Stochastic Processes and their Applications*, 125(1):233–253, 2015.

[32] D. E. Smith and M. Li. Estimating data delays from Poisson probe delays, Aug. 6 2002. US Patent 6,430,160.

[33] A. W. van der Vaart. *Asymptotic Statistics*, volume 3. Cambridge University Press, 1998.

[34] D. D. Yao. On Wolff's PASTA martingale. *Operations Research*, 40(3-supplement-2):S352–S355, 1992.