Bernoulli **24**(4A), 2018, 2531–2568
https://doi.org/10.3150/17-BEJ936

The $M/G/\infty$ estimation problem revisited

ALEXANDER GOLDENSHLUGER

Department of Statistics, University of Haifa, Haifa 31905, Israel. E-mail: goldensh@stat.haifa.ac.il

The subject of this paper is *the $M/G/\infty$ estimation problem*: the goal is to estimate the service time distribution G of the $M/G/\infty$ queue from the arrival–departure observations without identification of customers. We develop estimators of G and derive exact non-asymptotic expressions for their mean squared errors. The problem of estimating the service time expectation is addressed as well. We present some numerical results on comparison of different estimators of the service time distribution.

Keywords: $M/G/\infty$ queue; nonparametric estimation; Poisson point process; rates of convergence

1. Introduction

Background and motivation

The $M/G/\infty$ queueing model postulates that customers come to a system at time instances of a homogeneous Poisson process, obtain service immediately upon arrival, and leave the system after the service completion. The service times are assumed to be independent identically distributed random variables, independent of the arrival process, with common distribution G. The $M/G/\infty$ queue is one of the basic models in queueing theory; it is well understood from the probabilistic point of view and widely used in different applications.

Some problems of statistical inference for the $M/G/\infty$ queues were also considered in the literature. Motivated by a study of low density Poisson traffic streams, [9] studied the problem of estimating the service time distribution G in the $M/G/\infty$ queue from *the arrival–departure data* when observations of the arrival and departure epochs are available without identification of customers. In this setting, [9] developed a sequence of estimators $\{G_n, n \geq 1\}$ with G_n depending on the data up to the nth departure, and proved the consistency, $\sup_x |G_n(x) - G(x)| \overset{\text{a.s.}}{\to} 0$ as $n \to \infty$.

It is well known that the departure (output) process of the $M/G/\infty$ queue is Poisson of the same intensity λ as the arrival (input) process [see, e.g., [12], Chapter VIII, Section 5]; therefore G cannot be inferred from the output process alone. In addition, even under parametric assumptions on G, the likelihood function for such incomplete observations is unavailable in a usable form. Thus, estimation of G is not a trivial task.

The construction of estimators $\{G_n, n \geq 1\}$ in [9] is remarkable. Let $(\tau_j)_{j\in\mathbb{Z}}$ and $(t_j)_{j\in\mathbb{Z}}$ denote the arrival and departure epochs, respectively. Suppose that the output stream is observed starting from some departure, say t_0, and until the nth departure t_n. Associate every departure point t_j, $j = 1, \dots, n$ to the closest arrival point to the left, and consider the corresponding distances z_j, $j = 1, \dots, n$. It is shown that (z_j) is a stationary ergodic process with marginal distribution D which is related to the service time distribution by the simple formula:

$$D(x) = 1 - \big(1 - G(x)\big)e^{-\lambda x}, \tag{1.1}$$

where λ is the intensity of the arrival Poisson process. Then one can estimate D empirically from the data and invert (1.1) for G.

The idea of pairing departure points with arrivals to the left was also exploited in the subsequent work by [4]. The authors use distances from departure points to the rth nearest arrival to the left, and show consistency of the proposed estimators. On the basis of extensive simulations it is argued that for some service time distributions the estimators with properly chosen r can be advantageous. The recent paper, [25] considers a discrete-time $GI/G/\infty$ queue and derives a functional central limit theorem for a Brown-like estimator of the service time distribution.

A closely related problem is that of estimating G in the $M/G/\infty$ queue from observation of *the queue-length* (*number-of-busy-servers*) *process* $\{X(t), t \in \mathbb{R}\}$ over a finite time interval. If $(\tau_j)_{j\in\mathbb{Z}}$, $(t_j)_{j\in\mathbb{Z}}$ are arrival and departure epochs, and $(\sigma_j)_{j\in\mathbb{Z}}$ is the sequence of service times then in the stationary regime

$$X(t) = \sum_{j\in\mathbb{Z}} \mathbf{1}\{\tau_j \le t, \sigma_j > t - \tau_j\}, \qquad t \in \mathbb{R}.$$

The observation of the queue-length process $X(t)$ over a finite time interval is equivalent to observing arrival and departure epochs up to initial state of the queue: arrivals and departures are easily extracted from observations of $X(t)$, and the process $X(t)$ can be reconstructed from arrival and departure epochs, provided that the initial state is known. Thus the queue-length observations contain a bit more information than the arrival–departure data: the initial state of the system.

It is well known that if the service time expectation α is finite, $\alpha := \frac{1}{\mu} := \int_0^\infty [1 - G(u)]\,\mathrm{d}u < \infty$, then $X(t)$ has Poisson distribution with parameter $\rho := \lambda/\mu = \lambda\alpha$ for every t, where ρ is the *traffic intensity*. In addition, the correlation function of $\{X(t), t \ge 0\}$ is given by

$$H(t) = \mathrm{corr}_G\bigl\{X(s+t), X(s)\bigr\} = \mu \int_t^\infty \bigl[1 - G(u)\bigr]\,\mathrm{d}u; \tag{1.2}$$

see, e.g., [1]. This fact suggests that G can be reconstructed from correlation structure of the queue-length process. The work of [3] discusses this approach and provides standard results from the time series literature for estimation of H. The idea of reconstructing the service time distribution from correlation structure of the queue-length process was also exploited by [23] for a discrete-time queue model. Recently [13] constructed a local polynomial estimator of G based on (1.2) and investigated its worst-case accuracy over a suitable class of target distributions.

It is worth mentioning that problems of statistical inference on the service time distribution of the $M/G/\infty$ queue from incomplete data are ubiquitous in practice. They appear in such diverse areas as in studies of cell mobility [16,24], low density traffic [22] and in communication systems [1] and [17]. Additional references can be found in [2] and [4].

Although some estimators of the service time distribution in the $M/G/\infty$ queue were proposed in the literature, very little is known about their accuracy. In the original $M/G/\infty$ estimation problem with the arrival–departure data only consistency results [4,9] were established, and it is not clear what is the achievable accuracy in this problem. As for the setting with the queue-length data, the recent work of [13] derives non-asymptotic bounds on the risk and shows that under a local smoothness assumptions on G the pointwise risk of the proposed estimator of

G converges to zero at the nonparametric rate $T^{-\beta/(2\beta+2)}$, where T is the observation horizon, and β is the smoothness index of G.

Main results

In this paper, we revisit the $M/G/\infty$ estimation problem of [9] and concentrate on construction of estimators with provable theoretical accuracy guarantees. The main contributions of this paper are as follows.

(a) We study properties of a service time distribution estimator which is based on the formulas for covariance measures of the bivariate arrival–departure point process. This approach was briefly discussed by [20] who considered the problem of identifiability of random translations of stationary point processes. We show that the proposed estimator is unbiased and derive exact non-asymptotic expressions for its variance.

(b) We consider the problem of estimating G from the superposed arrival–departure data, when the arrival and departure epochs are registered without knowledge of the epoch type. This setting is particularly relevant in studies of particle mobility when crossings of the observation region can be recorded without knowledge of the crossing direction. The fact that estimation of service time distribution from such data is possible seems surprising. However, in a study of identifiability of random translations of Poisson processes [18] points at this possibility. In this setting, we develop unbiased estimator and derive exact non-asymptotic expressions for its variance.

(c) We present some numerical results in order to compare accuracy of the developed estimator with that of estimators in [9] and [13]. In particular, we study numerically the influence of the arrival rate λ and of tail behavior of G on the accuracy of estimators.

(d) The problem of estimating the service time expectation from arrival–departure data is also addressed. We propose an estimator and derive non-asymptotic bounds on its accuracy.

Our development relies upon tools from the theory of stationary point processes [11], and we use the statistical framework as expounded in [10] and [6]. We study the underlying bivariate arrival–departure and superposed arrival–departure point processes, derive their Laplace functionals and corresponding covariance measures up to the fourth order. The obtained expressions are used in order to construct estimating equations and to analyze properties of the corresponding estimators. Some of these technical results are new and interesting in their own right.

Further related literature

The $M/G/\infty$ estimation setting can be viewed as a particular case of the general point process system identification problem [7]. Here the unknown system translates every time instance of the input process by a random amount, and we would like to recover characteristics of the system from the input–output data. Some related results appear in [5] and [8]. In particular, [5] uses a spectral approach in order to construct an estimator of the service time distribution in the $G/G/\infty$ model, and shows asymptotic normality under suitable conditions. Brillinger [8] discusses the problem of estimating the second-order intensities of a bivariate stationary point process.

The problem of estimating the service time distribution G in the $M/G/\infty$ queue was also considered under other observation schemes. For instance, [14] discusses the setting where observations of durations of busy periods are available. Using a relationship between busy period distribution and the service time distribution, the authors construct an estimator of G and study its accuracy. This setting was also considered in [3].

Organization of the paper

The rest of the paper is structured as follows. In Section 2, we present the formulation of the problem. Section 3 collects some preliminary results which are instrumental for constructing estimators of the service time distribution G. In Section 4, we define the estimators of G based on the arrival–departure and superposed arrival–departure data, and present the results on their accuracy. Numerical experiments on comparison of different estimators are described in Section 5. In Section 6, we discuss the problem of estimating the expected service time from the arrival–departure data; some concluding remarks are brought in Section 7. Proofs of all results are given in Appendices.

2. Problem formulation

Let M be a homogeneous Poisson point process on $\mathbb{R}$ of intensity λ with representation

$$M := \sum_{j\in\mathbb{Z}} \varepsilon_{\tau_j}, \qquad \varepsilon_x(A) = \begin{cases} 1, & x \in A, \\ 0, & x \notin A, \end{cases} \qquad \forall A \in \mathscr{B},$$

where $\mathscr{B}$ is the Borel σ-algebra of $\mathbb{R}$. Suppose that every point of the process M is displaced by a random amount giving rise to the point process

$$N = \sum_{j\in\mathbb{Z}} \varepsilon_{t_j}, \qquad t_j := \tau_j + \sigma_j, \qquad j \in \mathbb{Z}.$$

Here $(\sigma_j)_{j\in\mathbb{Z}}$ is the sequence of independent identically distributed random variables, independent of M, with common distribution G. It is well known that N is also a homogeneous Poisson process of intensity λ.

We regard M and N as the arrival (input) and departure (output) processes, respectively. We also define the superposed point process S with the representation

$$S := M + N = \sum_{j\in\mathbb{Z}} \varepsilon_{\tau_j} + \sum_{j\in\mathbb{Z}} \varepsilon_{t_j} =: \sum_{j\in\mathbb{Z}} \varepsilon_{s_j}.$$

The following two different observation schemes will be considered.

(i) Assume that we observe the bivariate point process (M, N) restricted to a window $\mathcal{T}^2 = \mathcal{T} \times \mathcal{T} \subset \mathbb{R} \times \mathbb{R}$, $(M, N)|_{\mathcal{T}^2}$; thus the available data is

$$\mathcal{D}_{\mathcal{T}} := (M, N)|_{\mathcal{T}^2} = \{(\tau_j : \tau_j \in \mathcal{T}), (t_k : t_k \in \mathcal{T})\}.$$

(ii) Assume that the superposed process S restricted to a set $\mathcal{T} \subset \mathbb{R}$ is observed, so that the available data is

$$\mathcal{D}_{\mathcal{T}} := S|_{\mathcal{T}} = \{s_j : s_j \in \mathcal{T}\}.$$

Using the data $\mathcal{D}_{\mathcal{T}}$ our goal is to estimate the displacement distribution G or a functional thereof. In particular, we will be primarily interested in estimating the service time distribution G at a fixed point x_0, and the expectation of displacements $\alpha := \mathrm{E}_G[\sigma]$. Here and in all what follows E_G denotes the expectation with respect to the probability measure P_G generated by the observations $\mathcal{D}_{\mathcal{T}}$ when the displacement distribution is G.

Another interpretation of the observed data can be given in terms of marked point processes [11], Section 6.4. Specifically, consider the marked point process $\{(s_j, \varkappa_j)\}_{j\in\mathbb{Z}}$, where $(s_j)_{j\in\mathbb{Z}}$ are the locations of the superposed process S, and the marks $\varkappa_j \in \{1, 2\}$ are defined as $\varkappa_j = 1$ if the corresponding location s_j belongs to the input process M, and $\varkappa_j = 2$ if the corresponding location s_j belongs to the output process N. Then in the scenario (i) $\mathcal{D}_{\mathcal{T}} = \{(s_j, \varkappa_j) : s_j \in \mathcal{T}\}$, while in the scenario (ii) $\mathcal{D}_{\mathcal{T}} = \{s_j : s_j \in \mathcal{T}\}$, and the corresponding marks $\{\varkappa_j : s_j \in \mathcal{T}\}$ are not available.

By an estimator of $G(x_0)$ or $\alpha = \mathrm{E}_G[\sigma]$, we mean any measurable function of the available observations $\mathcal{D}_{\mathcal{T}}$. We measure accuracy of estimators by the mean squared error:

$$\mathcal{R}_{x_0}[\hat{G}, G] = \mathrm{E}_G\big|\hat{G}(x_0) - G(x_0)\big|^2, \qquad \mathcal{R}[\hat{\alpha}, \alpha] = \mathrm{E}_G|\hat{\alpha} - \alpha|^2.$$

In what follows, with slight abuse of notation, for any interval $I = (a, b]$ we denote $G(I)$ the probability mass assigned by the distribution G to I, $G(I) := G(b) - G(a)$.

3. Preliminaries

For the ease of reference, in this section we collect supporting preliminary results on the point processes involved; some of these results can be found, for example, in [11].

We start with a statement about the Laplace functional of bivariate process point (M, N).

Proposition 1. *Let $\{A_i\}_{i=1,\dots,m}$ and $\{B_l\}_{l=1,\dots,n}$ be two families of disjoint intervals of the real line; then for any $(\eta_1, \dots, \eta_m) \in \mathbb{R}^m$ and $(\xi_1, \dots, \xi_n) \in \mathbb{R}^n$ one has*

$$\begin{aligned}
&\log \mathrm{E}_G \exp\left\{\sum_{i=1}^m \eta_i M(A_i) + \sum_{l=1}^n \xi_l N(B_l)\right\} \\
&\quad = \lambda \sum_{i=1}^m (e^{\eta_i} - 1)|A_i| + \lambda \sum_{l=1}^n (e^{\xi_l} - 1)|B_l| + \lambda \sum_{i=1}^m \sum_{l=1}^n (e^{\eta_i} - 1)(e^{\xi_l} - 1) Q(A_i, B_l),
\end{aligned} \tag{3.1}$$

where $|\cdot|$ *is the Lebesgue measure on* $\mathbb{R}$, *and*

$$Q(A, B) := \int_A G(B - x)\,\mathrm{d}x. \tag{3.2}$$

The next result for the superposed process S is an immediate consequence of Proposition 1.

Proposition 2. *Let* $\{A_i\}_{i=1,\dots,m}$ *be disjoint intervals of* $\mathbb{R}$; *then for any* $\eta = (\eta_1, \dots, \eta_m) \in \mathbb{R}^m$

$$\log \mathrm{E}_G \exp\left\{\sum_{i=1}^m \eta_i S(A_i)\right\} = 2\lambda \sum_{i=1}^m \big(e^{\eta_i} - 1\big)|A_i| + \lambda \sum_{i=1}^m \sum_{l=1}^m \big(e^{\eta_i} - 1\big)\big(e^{\eta_l} - 1\big) Q(A_i, A_l),$$

where $Q(\cdot,\cdot)$ *is defined in* (3.2).

Remark 1. (a) In the specific case $m = 1$, $n = 1$ the formula (3.1) was obtained by [18] by queueing theoretical considerations. For general case, we refer to a related result given in Example 6.3(e) in [11].

(b) The bivariate point process (M, N) is closely related to the Gauss–Poisson processes introduced in [21] and further studied in [19]. In particular, the bivariate probability generating functional of (M, N) is given by

$$\begin{aligned}
&\mathcal{G}_{(M,N)}(\eta, \xi) \\
&\quad := \mathrm{E}_G \exp\left\{\int \log \eta(\tau)\,\mathrm{d}M(\tau) + \int \log \xi(t)\,\mathrm{d}N(t)\right\} \\
&\quad = \mathrm{E}_G\left[\prod_i \eta(\tau_i) \prod_l \xi(t_l)\right] \\
&\quad = \exp\left\{\lambda \int \big[\eta(\tau) - 1\big]\,\mathrm{d}\tau + \lambda \int \big[\xi(t) - 1\big]\,\mathrm{d}t + \lambda \iint \big[\eta(\tau) - 1\big]\big[\xi(t) - 1\big] Q(\mathrm{d}\tau, \mathrm{d}t)\right\},
\end{aligned}$$

where functions $0 \le \eta \le 1$ and $0 \le \xi \le 1$ are such that $1 - \eta$ and $1 - \xi$ vanish outside a common compact set, and $Q(\mathrm{d}\tau, \mathrm{d}t) = \mathrm{d}G(t - \tau)\,\mathrm{d}\tau$. This is an immediate consequence of Proposition 1.

(c) It is well known that the superposed process $S = M + N$ is the Gauss–Poisson process, and its probability generating functional is

$$\mathcal{G}_S(\eta) = \exp\left\{2\lambda \int \big[\eta(\tau) - 1\big]\,\mathrm{d}\tau + \lambda \iint \big[\eta(\tau) - 1\big]\big[\eta(t) - 1\big] Q(\mathrm{d}\tau, \mathrm{d}t)\right\}.$$

This fact follows immediately from (3.1); see also [19] and [11].

Using Propositions 1 and 2, we can calculate covariance measures of the bivariate process (M, N) and of the superposed process S. The next two statements follow from Propositions 1 and 2, respectively. They are proved in Appendix A along with other results on covariance measures of higher orders.

Corollary 1. *For any two intervals A, B one has*

$$\mathrm{E}_G\big[M(A)N(B)\big] = \lambda^2|A||B| + \lambda Q(A, B).$$

In particular, for the differential increments $\mathrm{d}M(\tau) = M((\tau, \tau + \mathrm{d}\tau])$ *and* $\mathrm{d}N(t) = N((t, t + \mathrm{d}t])$ *we have*

$$\mathrm{E}_G\big[\mathrm{d}M(\tau)\,\mathrm{d}N(t)\big] = \lambda^2\,\mathrm{d}\tau\,\mathrm{d}t + \lambda\,\mathrm{d}G(t-\tau)\,\mathrm{d}\tau.$$

Corollary 2. *Let A_1 and A_2 be disjoint intervals; then*

$$\mathrm{E}_G\big[S(A_1)S(A_2)\big] = 4\lambda^2|A_1||A_2| + \lambda\big[Q(A_1, A_2) + Q(A_2, A_1)\big].$$

In particular, for $A_1 = (\tau, \tau + \mathrm{d}\tau]$ and $A_2 = (t, t + \mathrm{d}t]$ with $\tau \neq t$ one has

$$\mathrm{E}_G\big[\mathrm{d}S(\tau)\,\mathrm{d}S(t)\big] = 4\lambda^2\,\mathrm{d}\tau\,\mathrm{d}t + \lambda\big[\mathrm{d}G(t-\tau)\,\mathrm{d}\tau + \mathrm{d}G(\tau - t)\,\mathrm{d}t\big].$$

Corollaries 1 and 2 provide the basis for constructing estimators of the service time distribution from the arrival–departure and superposed arrival–departure data. The estimators are presented in Section 4.

4. Estimation of service time distribution

Now we turn to the problem of estimating the service time distribution. First, we consider the setting with arrival–departure data.

4.1. Arrival–departure data

Corollary 1 implies that for any function φ satisfying

$$\iint \big|\varphi(\tau, t)\big|\,\mathrm{d}\tau\,\mathrm{d}t < \infty, \qquad \iint \big|\varphi(\tau, t)\big|\,\mathrm{d}G(t-\tau)\,\mathrm{d}\tau < \infty$$

one has

$$\mathrm{E}_G\bigg[\iint \varphi(\tau, t) M(\mathrm{d}\tau) N(\mathrm{d}t)\bigg] = \lambda^2 \iint \varphi(\tau, t)\,\mathrm{d}\tau\,\mathrm{d}t + \lambda \iint \varphi(\tau, t)\,\mathrm{d}G(t-\tau)\,\mathrm{d}\tau. \qquad (4.1)$$

The relationship (4.1) can serve as the estimating equation for constructing estimators of $G(x_0)$.

Suppose that for an interval $I = (a, b]$ we are interested in estimating $\theta_I := G(I) = G(b) - G(a)$, and the available data is $\mathcal{D}_\mathcal{T} = (M, N)|_\mathcal{T}$ where

$$\mathcal{T} = \mathcal{T}_M \times \mathcal{T}_N := [\tau_{\min}, \tau_{\max}] \times [\tau_{\min} + a, \tau_{\max} + b]$$

for some fixed $\tau_{\min} < \tau_{\max}$. Thus,

$$\mathcal{D}_\mathcal{T} := \big\{(\tau_j : \tau_{\min} \le \tau_j \le \tau_{\max}), (t_k : \tau_{\min} + a \le t_k \le \tau_{\max} + b)\big\}.$$

Let $T := \tau_{\max} - \tau_{\min}$, $\varphi_*(\tau, t) := \mathbf{1}_{[\tau_{\min},\tau_{\max}]}(\tau)\mathbf{1}_I(t-\tau)$, and define

$$\hat{\theta}_I = \frac{1}{\lambda T}\iint \varphi_*(\tau,t)\,\mathrm{d}M(\tau)\,\mathrm{d}N(t) - \lambda|I| = \frac{1}{\lambda T}\sum_{j\in\mathbb{Z}}\sum_{k\in\mathbb{Z}}\mathbf{1}_{\mathcal{T}_M}(\tau_j)\mathbf{1}_I(t_k-\tau_j) - \lambda|I|. \quad (4.2)$$

We are interested in accuracy of the estimator $\hat{\theta}_I$.

The formulas (4.1) and (4.2) appeared in [20], who studied identifiability of the random translations of stationary point processes. Mori [20] attributes the estimator (4.2) to [10]; Section 6 of this survey paper contains a discussion of statistical analysis for bivariate point processes. Following [20], we call $\hat{\theta}_I$ *the Cox–Lewis estimator*.

Theorem 1. *For any G, $\hat{\theta}_I$ is an unbiased estimator of θ_I, and*

$$\begin{aligned}\operatorname{var}_G\{\hat{\theta}_I\} &= \frac{2\lambda|I|}{T}\left\{|I| + \int_{-T}^{T} G(I+u)\left(1-\frac{|u|}{T}\right)\mathrm{d}u - \frac{|I|^2}{6T}\right\}\\ &\quad + \frac{|I|}{T} + \frac{2}{T}|I|G(I) + \frac{1}{T}\int_{-T}^{T} G(I+u)G(I-u)\left(1-\frac{|u|}{T}\right)\mathrm{d}u\\ &\quad + \frac{2}{T}\int_0^{|I|}\bigl[G(I)+G(b-u)-G(a+u)\bigr]\left(1-\frac{u}{T}\right)\mathrm{d}u + \frac{G(I)}{\lambda T}.\end{aligned} \quad (4.3)$$

In the context of the $M/G/\infty$ estimation problem, we set $\tau_{\min}=0$, $\tau_{\max}=T$, $I=[0,x_0]$ and assume $G(0)=0$. Then the Cox–Lewis estimator $\hat{G}(x_0)$ of $G(x_0)$ is given by

$$\hat{G}(x_0) = \hat{\theta}_I = \frac{1}{\lambda T}\sum_{j\in\mathbb{Z}}\sum_{k\in\mathbb{Z}}\mathbf{1}_{[0,T]}(\tau_j)\mathbf{1}_{[0,x_0]}(t_k-\tau_j) - \lambda x_0. \quad (4.4)$$

Note that $\hat{G}(x_0)$ is based on the data $\mathcal{D}_{\mathcal{T}} = \{(\tau_j : 0\le\tau_j\le T), (t_k : 0\le t_k\le T+x_0)\}$.

The next statement is an immediate consequence of Theorem 1 for the $M/G/\infty$ setting.

Theorem 2. *If $G(0)=0$ and $x_0\in(0,T)$, then $\hat{G}(x_0)$ is an unbiased estimator of $G(x_0)$, and*

$$\begin{aligned}\operatorname{var}_G\bigl\{\hat{G}(x_0)\bigr\} &= \frac{2\lambda x_0}{T}\left\{x_0 + \int_{-T}^{T}\bigl[G(x_0+u)-G(u)\bigr]\left(1-\frac{|u|}{T}\right)\mathrm{d}u - \frac{x_0^2}{6T}\right\} + \frac{x_0}{T}\\ &\quad + \frac{2}{T}x_0 G(x_0)\\ &\quad + \frac{1}{T}\int_{-T}^{T}\bigl[G(x_0+u)-G(u)\bigr]\bigl[G(x_0-u)-G(-u)\bigr]\left(1-\frac{|u|}{T}\right)\mathrm{d}u\\ &\quad + \frac{2}{T}\int_0^{x_0}\bigl[G(x_0)+G(x_0-u)-G(u)\bigr]\left(1-\frac{u}{T}\right)\mathrm{d}u + \frac{G(x_0)}{\lambda T}.\end{aligned} \quad (4.5)$$

Remark 2. (a) Theorem 2 shows that the service time distribution in the $M/G/\infty$ problem is estimated with the root mean squared error tending to zero at the parametric rate $T^{-1/2}$, $T \to \infty$. The result shows, however, that the Cox–Lewis estimator is less accurate in the heavy-traffic regime when the arrival rate λ is large. The last term on the right-hand side of (4.3) [and (4.5)] also indicates that the accuracy is poor when λ is very small. In addition, the farther point x_0 is from the origin, the worse accuracy of $\hat{G}(x_0)$ is.

(b) The results of Theorems 1 and 2 do not require any conditions on the service time distribution G; for instance, G can have infinite expectation. However, accuracy of the Cox–Lewis estimator depends on tails of G, and this dependence is quantified by (4.3) and (4.5).

(c) In practice the Cox–Lewis estimator $\hat{G}(x_0)$, when considered to be a function of x_0, should be monotonized and confined to the interval [0, 1].

4.2. Superposed arrival–departure data

Now, we turn to the problem of estimating the service time distribution G from the observations of the superposed process $S = \sum_{j\in\mathbb{Z}} \varepsilon_{s_j}$ on the time interval $\mathcal{T} = [0, T]$. In particular, assume that the available data is

$$\mathcal{D}_{\mathcal{T}} = S|_{\mathcal{T}} = \{s_j : 0 \le s_j \le T\}.$$

Similarly to (4.1), Corollary 2 implies that

$$\begin{aligned}\mathrm{E}_G\left[\iint \varphi(\tau,t)\,\mathrm{d}S(\tau)\,\mathrm{d}S(t)\right] &= 4\lambda^2 \iint \varphi(\tau,t)\,\mathrm{d}\tau\,\mathrm{d}t + \lambda \iint \varphi(\tau,t)\,\mathrm{d}G(t-\tau)\,\mathrm{d}\tau \\ &\quad + \lambda \iint \varphi(\tau,t)\,\mathrm{d}G(\tau-t)\,\mathrm{d}t\end{aligned} \tag{4.6}$$

for any function φ for which the right-hand side is well defined.

Suppose that $G(0) = 0$, and let

$$\varphi_*(\tau,t) = \mathbf{1}_{[0,T]}(\tau)\mathbf{1}_{[0,x_0]}(t-\tau)\chi(\tau,t), \qquad \chi(\tau,t) := \begin{cases} 1, & \tau \ne t, \\ 0, & \text{otherwise}. \end{cases}$$

Consider the following estimator of $G(x_0)$:

$$\begin{aligned}\tilde{G}(x_0) &= \frac{1}{\lambda T}\iint \varphi_*(\tau,t)\,\mathrm{d}S(\tau)\,\mathrm{d}S(t) - 4\lambda x_0 \\ &= \frac{1}{\lambda T}\sum_{\substack{j\in\mathbb{Z}\\ j\ne k}}\sum_{k\in\mathbb{Z}} \mathbf{1}_{[0,T]}(s_j)\mathbf{1}_{[0,x_0]}(s_k - s_j) - 4\lambda x_0.\end{aligned}$$

Theorem 3. *If $G(0) = 0$ and $x_0 \in (0, T)$, then $\tilde{G}(x_0)$ is an unbiased estimator of $G(x_0)$, and*

$$\mathrm{var}_G\{\tilde{G}(x_0)\} = \frac{1}{T}R_T^{(1)}(\lambda, x_0; G) + \frac{1}{T^2}R_T^{(2)}(\lambda, x_0; G),$$

where $R_T^{(1)}$ and $R_T^{(2)}$ are positive functions satisfying

$$R_T^{(1)}(\lambda, x_0; G) \le 76\lambda x_0^2 + 36x_0 + \frac{1}{\lambda}G(x_0), \qquad R_T^{(2)}(\lambda, x_0; G) \le 36\lambda x_0^3, \qquad \forall T, \forall G.$$

Remark 3. (a) Exact expressions for functions $R_T^{(1)}(\lambda, x_0; G)$ and $R_T^{(2)}(\lambda, x_0; G)$ are given in the proof of the theorem. As it could be expected, accuracy of the estimator $\tilde{G}(x_0)$ is worse than the accuracy of the Cox–Lewis estimator $\hat{G}(x_0)$ based on the whole arrival–departure data. It should be noticed, however, that qualitatively the behavior is similar: the rate of convergence is parametric, the variance grows linearly with λ as λ increases, and quadratically with x_0 as x_0 grows.

(b) Similarly to Theorem 2, the statement does not require any conditions on the service time distribution G.

5. Numerical experiments

We conducted a small simulation study in order to compare performance of three different estimators of the service time distribution. The following estimators were considered.

The Cox–Lewis estimator

The estimator is given by

$$\hat{G}_{\mathrm{CL}}(x_0) = \frac{1}{\lambda T}\sum_{k\in\mathbb{Z}}\sum_{j\in\mathbb{Z}} \mathbf{1}_{[0,T]}(t_k)\mathbf{1}_{[0,x_0]}(t_k - \tau_j) - \lambda x_0.$$

In our computations, we considered monotonized and truncated version of $\hat{G}_{\mathrm{CL}}(x_0)$,

$$\hat{G}^*_{\mathrm{CL}}(x_0) := \Big\{\max_{u\le x_0}\hat{G}_{\mathrm{CL}}(u)\Big\}_{[0,1]}, \qquad \{x\}_{[0,1]} := \begin{cases} 1, & x > 1, \\ x, & 0 \le x \le 1, \\ 0, & x < 0. \end{cases}$$

Note that $\hat{G}_{\mathrm{CL}}(x_0)$ is a slight modification of (4.4) as it is based on the data

$$\mathcal{D}_T = \big\{(t_j : 0 \le t_j \le T), (\tau_j : -x_0 \le \tau_j \le T)\big\}. \tag{5.1}$$

We use this modification in order to compare our estimator with Brown's estimator (see the description below). The theoretical properties of $\hat{G}_{\mathrm{CL}}(x_0)$ coincide with those of $\hat{G}(x_0)$ defined in (4.4).

Brown's estimator

The estimator $\hat{G}^*_B(x_0)$ is defined as follows. For each output point $0 \le t_k \le T$ let z_k denote the distance from t_k to the closest input point τ_j to the left. Then

$$\hat{G}_B(x_0) = 1 - e^{\lambda x_0} \frac{\sum_{k\in\mathbb{Z}} \mathbf{1}_{(x_0,\infty)}(z_k)\mathbf{1}_{[0,T]}(t_k)}{\sum_{k\in\mathbb{Z}} \mathbf{1}_{[0,T]}(t_k)},$$

$$\hat{G}^*_B(x_0) = \Big\{\max_{u\le x_0} \hat{G}_B(u)\Big\}_{[0,1]}.$$

Local polynomial estimator

This estimator was proposed in [13], and it is based on observations of the queue-length process $\{X(t), t \in [0,T]\}$:

$$X(t) = \sum_{j\in\mathbb{Z}} \mathbf{1}\{\tau_j \le t, \sigma_j > t - \tau_j\}. \tag{5.2}$$

Assume that $\{X(t), t \in [0,T]\}$ is observed at the points of the regular grid on $[0,T]$: $i\delta$, $i = 0,\dots,n$, $n\delta = T$. We implemented the local polynomial estimator $\hat{G}_{\mathrm{LP}}(x_0)$ of the second order [for general definition see [13]].

The definition of $\hat{G}_{\mathrm{LP}}(x_0)$ is the following. Let $\hat{R}(k\delta)$, $k = 1,2,\dots$ be empirical covariances of the process $\{X(t)\}$,

$$\hat{R}(k\delta) = \frac{1}{n}\sum_{i=1}^{n-k}\big[X(i\delta)-\hat{\rho}\big]\big[X\big((i+k)\delta\big)-\hat{\rho}\big], \qquad \hat{\rho} = \frac{1}{n}\sum_{i=1}^{n} X(i\delta). \tag{5.3}$$

Fix $h > 2\delta$ and define $D_x = [x-h, x+h]$ if $h \le x \le T-h$, $D_x = [x, x+2h]$ if $0 < x < h$, and $D_x = [T-2h, T]$ if $T-h \le x \le T$. Let $M_{D_x} = \{k : k\delta \in D_x\}$, and $\{a_k(x)\}_{k\in M_{D_x}}$ be the weights solving the following optimization problem

$$\begin{aligned} &\min \sum_{k\in M_{D_x}} a_k^2(x) \\ &\text{subject to} \sum_{k\in M_{D_x}} a_k(x) = 0, \\ &\sum_{k\in M_{D_x}} a_k(x)(k\delta)^j = jx^{j-1}, \qquad j = 1,2. \end{aligned}$$

Then

$$\hat{G}_{\mathrm{LP}}(x_0) := 1 + \frac{1}{\lambda}\sum_{k\in M_{D_{x_0}}} a_k(x_0)\hat{R}(k\delta). \tag{5.4}$$

The final estimator $\hat{G}^*_{\mathrm{LP}}(x_0)$ is a monotonized and truncated version of $\hat{G}_{\mathrm{LP}}(x_0)$:

$$\hat{G}^*_{\mathrm{LP}}(x_0) = \Big\{\max_{u\le x_0} \hat{G}_{\mathrm{LP}}(u)\Big\}_{[0,1]}.$$

This estimator requires selection of the window width h; it will be specified later.

The goal of the experiments is to study influence of the arrival rate λ and the service time distribution tail on accuracy of the estimators. Thus we consider two different scenarios.

(a) For different values of the arrival rate $\lambda \in \{0.5, 1, 5, 15\}$, fixed exponential service time distribution $G(x) = 1 - e^{-x}$ and the observation horizon $T = 1000$ we estimate G by the three estimators at 100 equidistant points $\{x_i, i = 1, \dots, 100\}$ on the interval $[0, 4]$. At every simulation run, we compute the maximal error

$$\mathrm{Err}(\hat{G}) = \max_{x\in\{x_i\}} \big|\hat{G}(x) - G(x)\big|, \qquad \hat{G} \in \big\{\hat{G}^*_{\mathrm{CL}}, \hat{G}^*_{\mathrm{LP}}, \hat{G}^*_B\big\}. \tag{5.5}$$

(b) The arrival rate is fixed $\lambda = 1$, and we consider exponential service time distribution $G(x) = 1 - e^{-\mu x}$ with $\mu \in \{\frac{1}{2}, \frac{1}{5}, \frac{1}{10}, \frac{1}{15}\}$. The observation horizon is $T = 1000$ and the distribution G is estimated at 100 equidistant points $\{x_i\}$ on the interval $[0, 10]$. Similarly to (a), the accuracy is measured by the maximal error (5.5).

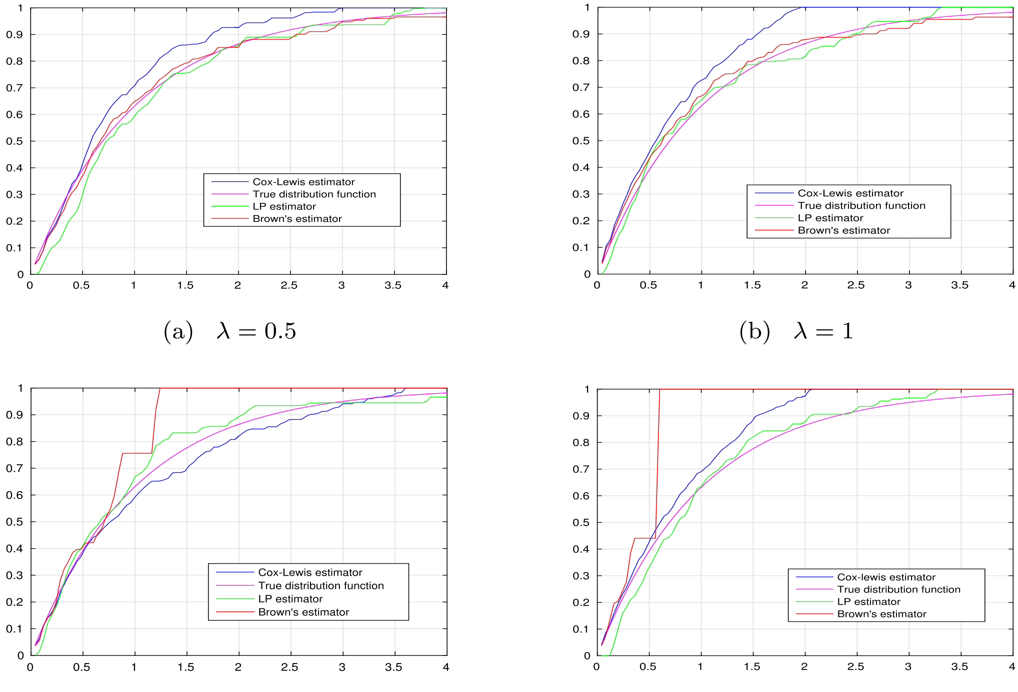
In all our experiments the local polynomial estimator $\hat{G}^*_{\mathrm{LP}}$ is computed on the basis of the queue-length process observations at equidistant points on $[0, T]$ with step size $\delta = 0.01$. It also requires specification of the bandwidth parameter h which was put to $h = 3\delta$. In fact, we do not tackle the question of "optimal" bandwidth selection, only little smoothing is applied. All the estimators are computed at the points of a regular grid on the corresponding intervals. The estimators are monotonized and truncated straightforwardly according to the definitions.

The results for scenario (a) are presented in Figures 1 and 2. Figure 1 shows typical realizations of the three estimators $\hat{G}^*_{\mathrm{CL}}$, $\hat{G}^*_B$ and $\hat{G}^*_{\mathrm{LP}}$ for different values of λ. In Figure 2, we present box-plots of the maximal errors $\mathrm{Err}(\hat{G})$, $\hat{G} \in \{\hat{G}^*_{\mathrm{CL}}, \hat{G}^*_{\mathrm{LP}}, \hat{G}^*_B\}$ of the estimators over 100 independent simulation runs for $\lambda \in \{0.5, 1, 5, 15\}$.

The numerical results show that both $\hat{G}^*_{\mathrm{CL}}$ and $\hat{G}^*_B$ behave poorly for large values of λ and x_0. This behavior of $\hat{G}^*_{\mathrm{CL}}$ is in full agreement with Theorem 2: the variance grows linearly in λ and quadratically in x_0. Figures 1 and 2 show that Brown's estimator performs well for $\lambda = 0.5$ and $\lambda = 1$, but its accuracy deteriorates very rapidly as λ increases. In our experiments for $\lambda = 5$, the median errors of $\hat{G}^*_{\mathrm{CL}}$ and $\hat{G}^*_B$ are close to each other, while in the case $\lambda = 15$ Brown's estimator is upset completely. In the light traffic regime (small λ), the matching of output points with the closest input points to the left are often reconstruct the "true" pair; perhaps this fact explains good performance of $\hat{G}^*_B$ in the light traffic regime. The local polynomial estimator $\hat{G}^*_{\mathrm{LP}}$ exhibits very stable behavior, and its accuracy is not affected by changes in the arrival rate.

Figures 3 and 4 show the corresponding results for the scenario (b). As it is seen, accuracy of all estimators is badly affected by heavy tails of the service time distribution. Brown's estimator is most sensitive in comparison with the other two, and the LP estimator is most stable.

In sum, although the local polynomial estimator requires specification of bandwidth, our experiments show that under considered scenarios the local polynomial estimator $\hat{G}^*_{\mathrm{LP}}$ with small



(a) $\lambda = 0.5$ (b) $\lambda = 1$

(c) $\lambda = 5$ (d) $\lambda = 15$

Figure 1. Typical realizations of $\hat{G}^*_{\mathrm{CL}}$, $\hat{G}^*_B$ and $\hat{G}^*_{\mathrm{LP}}$ for $G(x) = 1 - e^{-x}$ and different values of the arrival rate λ.

bandwidth parameter compares favorably with Brown's and the Cox–Lewis estimators. In general, the Cox–Lewis estimator exhibits larger variability than Brown's, although for large values of arrival rate λ its median behavior is better than that of Brown's.

6. Estimation of the expected service time

In this section, we discuss the problem of estimating the expected service time, $\alpha = \mathrm{E}_G[\sigma] = 1/\mu$ from the arrival–departure data.

It seems straightforward to base an estimator for α on the relationship (4.1). In particular, if we put $\varphi(\tau, t) = \mathbf{1}_{[0,T]}(\tau)\mathbf{1}_{[0,T]}(t)$ in (4.1) then

$$\mathrm{E}_G \iint \mathbf{1}_{[0,T]}(\tau)\mathbf{1}_{[0,T]}(t)\,\mathrm{d}M(\tau)\,\mathrm{d}N(t) = \lambda^2 T^2 + \lambda \int_0^T G(u)\,\mathrm{d}u,$$

and

$$\int_0^T \bigl[1 - G(u)\bigr]\,\mathrm{d}u = \lambda T^2 + T - \mathrm{E}_G \iint \mathbf{1}_{[0,T]}(\tau)\mathbf{1}_{[0,T]}(t)\,\mathrm{d}M(\tau)\,\mathrm{d}N(t).$$

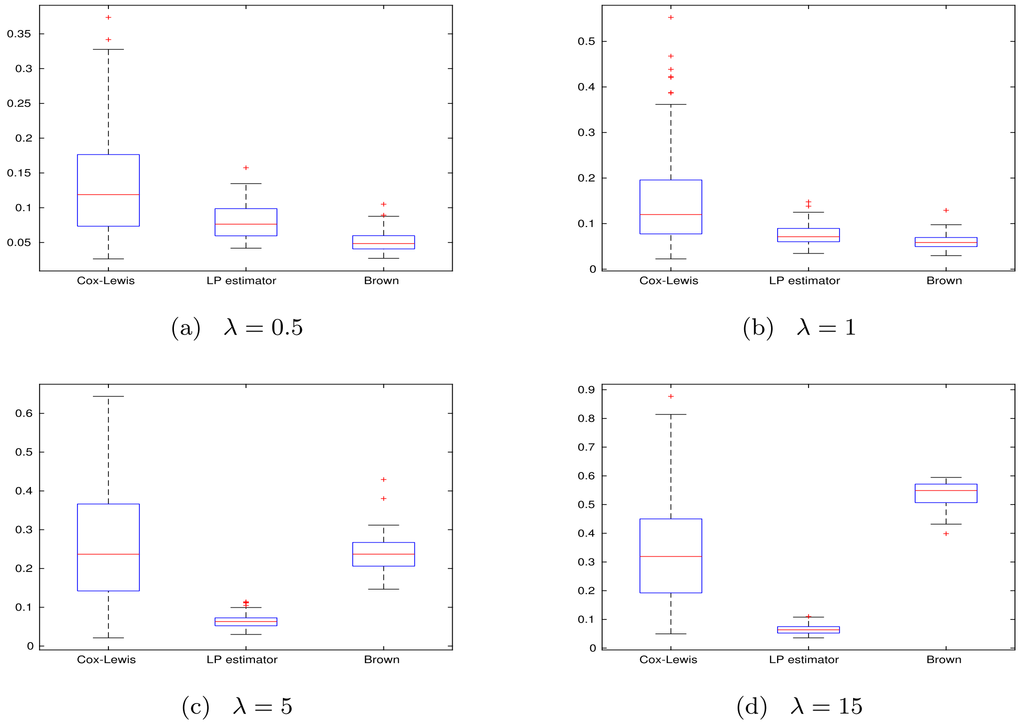


Figure 2. Boxplots of the maximal errors over $[0, 4]$ computed on the basis of 100 simulation runs for the estimators $\hat{G}^*_{\mathrm{CL}}$, $\hat{G}^*_B$ and $\hat{G}^*_{\mathrm{LP}}$ and different values of the arrival rate λ.

This shows that α can be estimated by

$$\tilde{\alpha} = \lambda T^2 + T - \sum_{j\in\mathbb{Z}}\sum_{k\in\mathbb{Z}} \mathbf{1}_{[0,T]}(\tau_j)\mathbf{1}_{[0,T]}(t_k).$$

However, even though under a moment condition the bias of $\tilde{\alpha}$, $\int_T^\infty [1 - G(u)]\,\mathrm{d}u$, is small, the estimator variance does not tend to zero as $T \to \infty$.

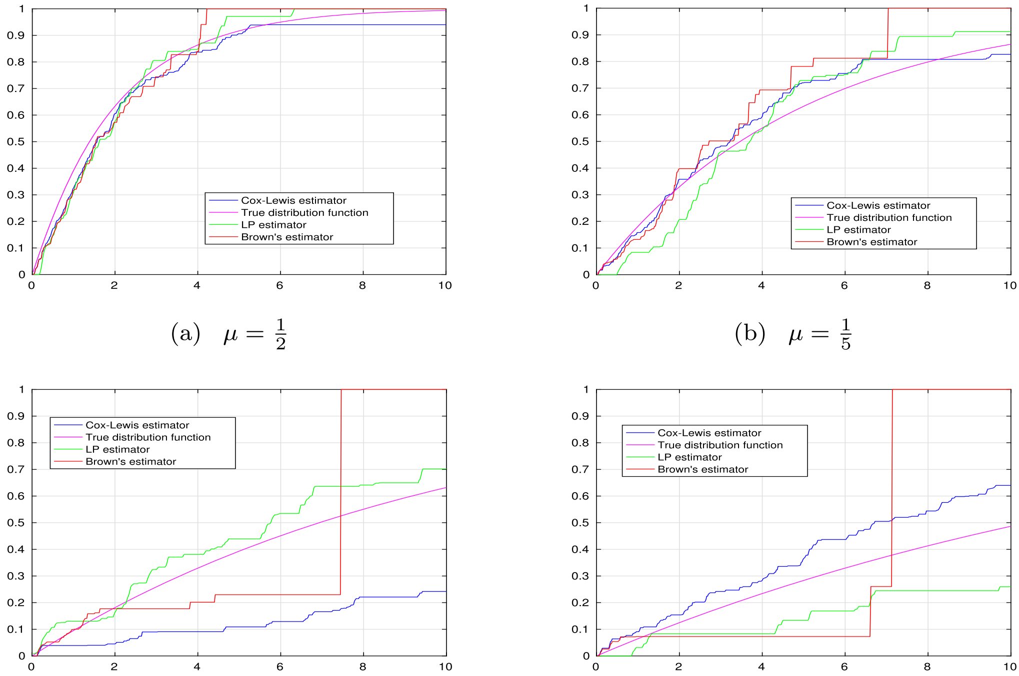
A natural estimator of α can be obtained by integration of the corresponding estimator of the service time distribution. In particular, for a real number $b > 0$ we set

$$\hat{\alpha} = \int_0^b \bigl[1 - \hat{G}(x)\bigr]\,\mathrm{d}x, \tag{6.1}$$

where $\hat{G}(\cdot)$ is defined in (4.4).

Let $\mathcal{M}_p(A)$ be the set of all distribution on $\mathbb{R}_+$ with pth moment bounded by $A < \infty$,

$$\mathcal{M}_p(A) := \biggl\{G : p\int_0^\infty x^{p-1}\bigl[1 - G(x)\bigr]\,\mathrm{d}x \le A < \infty\biggr\}, \qquad p > 1.$$



(a) $\mu = \frac{1}{2}$ (b) $\mu = \frac{1}{5}$

(c) $\mu = \frac{1}{10}$ (d) $\mu = \frac{1}{15}$

Figure 3. Typical realizations of $\hat{G}^*_{\mathrm{CL}}$, $\hat{G}^*_B$ and $\hat{G}^*_{\mathrm{LP}}$ for $G(x) = 1 - e^{-\mu x}$ with different values of the service rate μ.

We have the following theorem.

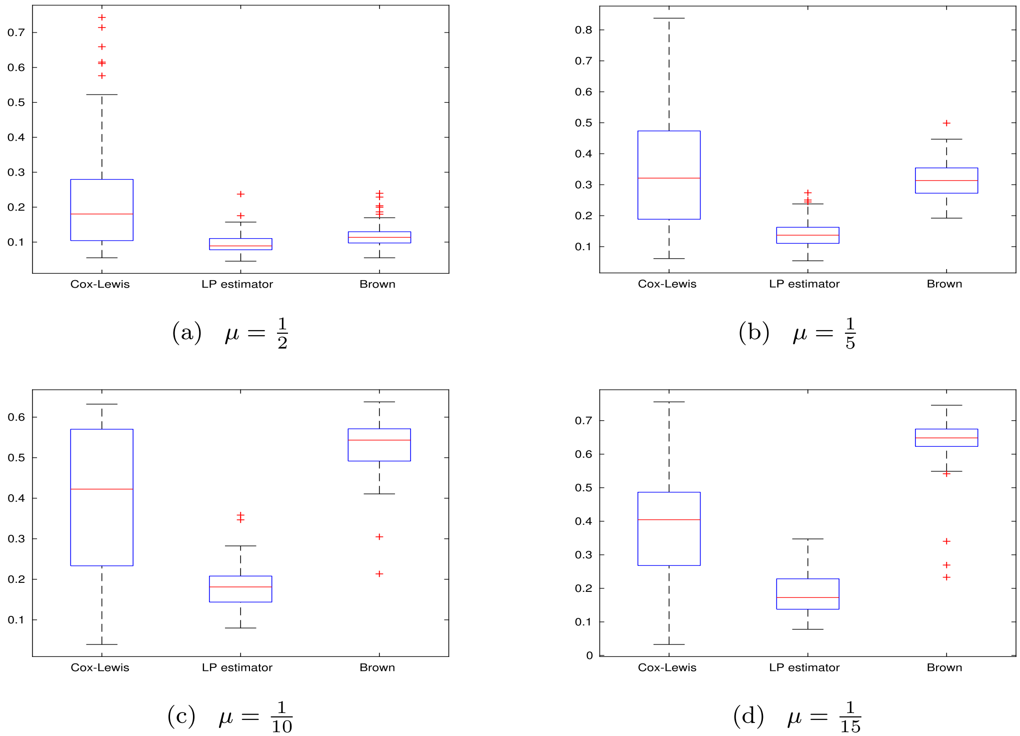
Theorem 4. *Let $\hat{\alpha}_*$ denote the estimator* (6.1) *associated with $b = b_* := (A/p)^{1/(p+1)}(T/\lambda)^{1/(2p+2)}$. Then for all $T \geq \lambda(1 \vee \lambda^{-2})^{2p+2}(A/p)^2$ one has*

$$\sup_{G \in \mathcal{M}_p(A)} \mathrm{E}_G |\hat{\alpha}_* - \alpha|^2 \leq C \left(\frac{A}{p}\right)^{4/(p+1)} \left(\frac{\lambda}{T}\right)^{(p-1)/(p+1)},$$

where C is an absolute constant.

Remark 4. (a) As discussed above, for a fixed x the estimator $\hat{G}(x)$ uses the data $\{(\tau_j : 0 \leq \tau_j \leq T), (t_k : 0 \leq t_k \leq T + x)\}$. Therefore $\hat{\alpha}$ is based on the observations $\{(\tau_j : 0 \leq \tau_j \leq T), (t_k : 0 \leq t_k \leq T + b_*)\}$. Because $b_* = o(T)$, the departure process is basically observed over an interval of the length $T(1 + o(1))$ as $T \to \infty$.

(b) The theorem demonstrates that the service time expectation is estimated with a nonparametric rate that depends on the tail behavior of G. This dependence stems from expression for



(a) $\mu = \frac{1}{2}$ (b) $\mu = \frac{1}{5}$ (c) $\mu = \frac{1}{10}$ (d) $\mu = \frac{1}{15}$

Figure 4. Boxplots of the maximal errors over [0, 10] computed on the basis of 100 simulation runs for the estimators $\hat{G}^*_{\mathrm{CL}}$, $\hat{G}^*_B$ and $\hat{G}^*_{\mathrm{LP}}$ and different values of the service rate μ.

the variance of the Cox–Lewis estimator for large values of x_0. Existence of the pth moment with $p > 1$ is sufficient for consistency of the estimator.

7. Concluding remarks

1. It is worth noting that the local polynomial estimator $\hat{G}_{\mathrm{LP}}(x_0)$ defined in (5.4) and based on the observations of the queue-length process (5.2) can be also computed on the basis of the arrival–departure data (5.1). Indeed, using the arrival–departure data we can define the following random process $\{V(t), t \geq 0\}$: let $V(0) = 0$ and

$$V(t) = \sum_{j\in\mathbb{Z}} \bigl[\mathbf{1}(0 \leq t_j \leq t) - \mathbf{1}(0 \leq \tau_j \leq t)\bigr], \qquad t > 0.$$

It is evident that $V(t) = X(t) - X(0)$; therefore setting $\tilde{\rho} = \frac{1}{n}\sum_{i=1}^n V(i\delta)$ we obtain

$$\frac{1}{n}\sum_{i=1}^{n-k} \bigl[V(i\delta) - \tilde{\rho}\bigr]\bigl[V\bigl((i+k)\delta\bigr) - \tilde{\rho}\bigr] = \hat{R}(k\delta),$$

where $\hat{R}(k\delta)$ is defined in (5.3). This shows that the local polynomial estimator $\hat{G}_{\text{LP}}(x_0)$ is applicable when the arrival–departure data is available.

On the other hand, there are striking differences in the problem of estimating the expected service time in the settings of arrival–departure and queue-length data. Although relative to the arrival–departure data, the only additional information in the queue-length data is the initial state of the system, it seems that this information is essential for improving accuracy of estimators of α. Specifically, suppose that the queue-length observations $\{X(k\delta), k=1,\dots,n\}$, $n\delta = T$ are given. Since the marginal distribution of $X(t)$ is Poisson with parameter $\rho = \alpha\lambda$, the natural estimator of α is $\check{\alpha} = \frac{1}{\lambda n}\sum_{k=1}^{n} X(k\delta)$. This estimator is unbiased, and if the second moment of G is finite then it can be shown that the variance of $\check{\alpha}$ converges to zero at the parametric rate T^{-1}. Thus, the information about initial state of the system leads to significant improvements in accuracy.

2. We considered the pointwise mean squared error as an estimation accuracy measure. In particular, Theorem 2 implies that the Cox–Lewis estimator (4.4) is pointwise consistent, that is, $\hat{G}(x)$ converges in P_G-probability to $G(x)$ as $T \to \infty$ for every x. We note however that for a properly modified version of the Cox–Lewis estimator the uniform convergence can be claimed. Indeed, let $\hat{G}(x)$ be the original Cox–Lewis estimator given in (8) and based on the data $\{(\tau_j : 0 \le \tau_j \le T), (t_k : 0 \le t_k \le T + x)\}$. Define

$$\hat{G}_T(x) = \begin{cases} \Big\{\max\limits_{u \le x} \hat{G}(u)\Big\}_{[0,1]}, & x < T^{1/4}, \\ 1, & x \ge T^{1/4}, \end{cases}$$

where $\{\cdot\}_{[0,1]}$ is the truncation operation defined in Section 5. Note that $\{\hat{G}_T\}$ is a sequence of probability distribution functions, and it follows from Theorem 2 that $\hat{G}_T(x)$ converges to $G(x)$ as $T \to \infty$ in P_G-probability for every x. Then invoking the proof of the Glivenko–Cantelli theorem we can prove that $\hat{G}_T$ converges to G uniformly in P_G-probability, i.e. $\lim_{T\to\infty} \mathrm{P}_G\{\sup_x |\hat{G}_T(x) - G(x)| > \varepsilon\} = 0$ for any $\varepsilon > 0$.

Appendix A: Proofs for Section 3

A.1. Proof of Proposition 1

Conditioning on $(\tau_j)_{j\in\mathbb{Z}}$ and using independence of $(\sigma_j)_{j\in\mathbb{Z}}$ on $(\tau_j)_{j\in\mathbb{Z}}$ and disjointness of $\{B_j\}$, we have

$$\begin{aligned} &\mathrm{E}_G\big[e^{\sum_{i=1}^m \eta_i M(A_i) + \sum_{l=1}^n \xi_l N(B_l)} | (\tau_j)_{j\in\mathbb{Z}}\big] \\ &\quad = e^{\sum_{i=1}^m \eta_i M(A_i)} \mathrm{E}_G\left[\exp\left\{\sum_{l=1}^n \xi_l \sum_{j\in\mathbb{Z}} \mathbf{1}_{B_l}(\tau_j + \sigma_j)\right\} \middle| (\tau_j)_{j\in\mathbb{Z}}\right] \\ &\quad = e^{\sum_{i=1}^m \eta_i M(A_i)} \prod_{j\in\mathbb{Z}} \mathrm{E}_G\left[\exp\left\{\sum_{l=1}^n \xi_l \mathbf{1}_{B_l - \tau_j}(\sigma_j)\right\} \middle| (\tau_j)_{j\in\mathbb{Z}}\right] \qquad \text{(A.1)} \end{aligned}$$

$$= e^{\sum_{i=1}^m \eta_i M(A_i)} \prod_{j\in\mathbb{Z}} \Biggl[\sum_{l=1}^n \bigl(e^{\xi_l}-1\bigr)\mathrm{P}_G(\sigma_j \in B_l - \tau_j) + 1\Biggr]$$

$$= e^{\sum_{i=1}^m \eta_i M(A_i)} \exp\Biggl\{\sum_{j\in\mathbb{Z}} \log\Biggl[\sum_{l=1}^n \bigl(e^{\xi_l}-1\bigr)\mathrm{P}_G(\sigma_j \in B_l - \tau_j) + 1\Biggr]\Biggr\}.$$

Now, if we put

$$f(x) := \sum_{i=1}^m \eta_i \mathbf{1}_{A_i}(x) + \log\Biggl[\sum_{l=1}^n \bigl(e^{\xi_l}-1\bigr)G(B_l - x) + 1\Biggr]$$

then the right-hand side of (A.1) is $\exp\{\sum_{j\in\mathbb{Z}} f(\tau_j)\}$, and applying Campbell's formula [see, e.g., [15]] we obtain

$$\log \mathrm{E}_G\bigl[e^{\sum_{i=1}^m \eta_i M(A_i) + \sum_{l=1}^n \xi_l N(B_l)}\bigr] = \lambda \int_{-\infty}^{\infty} \bigl[e^{f(x)} - 1\bigr]\,\mathrm{d}x.$$

It remains to compute the last integral. We have

$$\int_{-\infty}^{\infty} \bigl[e^{f(x)} - 1\bigr]\,\mathrm{d}x$$

$$= \int_{-\infty}^{\infty} \Biggl[e^{\sum_{i=1}^m \eta_i \mathbf{1}_{A_i}(x)} \Biggl(\sum_{l=1}^n \bigl(e^{\xi_l}-1\bigr)G(B_l - x) + 1\Biggr) - 1\Biggr]\,\mathrm{d}x$$

$$= \sum_{i=1}^m \int_{A_i} \Biggl[e^{\eta_i} \Biggl(\sum_{l=1}^n \bigl(e^{\xi_l}-1\bigr)G(B_l - x) + 1\Biggr) - 1\Biggr]\,\mathrm{d}x$$

$$+ \int_{\mathbb{R}\setminus\bigcup_{i=1}^m A_i} \sum_{l=1}^n \bigl(e^{\xi_l}-1\bigr)G(B_l - x)\,\mathrm{d}x$$

$$= \sum_{i=1}^m \bigl(e^{\eta_i}-1\bigr)|A_i| + \sum_{l=1}^n \bigl(e^{\xi_l}-1\bigr)|B_l| + \sum_{i=1}^m \sum_{l=1}^n \bigl(e^{\eta_i}-1\bigr)\bigl(e^{\xi_l}-1\bigr) \int_{A_i} G(B_l - x)\,\mathrm{d}x,$$

where we have used that $\int_{-\infty}^{\infty} G(B_l - x)\,\mathrm{d}x = |B_l|$. This completes the proof.

A.2. Covariance measures of (M, N)

In the next lemma, we present expressions for the covariance measures of the process (M, N). Corollary 1 is restated as the part (i) of Lemma 1.

Lemma 1. *Let $Q(\cdot,\cdot)$ be given by* (3.2); *then the following statements hold.*

(i) *For any two intervals A and B, one has*

$$\mathrm{E}_G\big[M(A)N(B)\big] = \lambda^2|A||B| + \lambda Q(A, B). \tag{A.2}$$

In particular, for $\mathrm{d}M(\tau) = M((\tau, \tau + \mathrm{d}\tau])$ *and* $\mathrm{d}N(t) = N((t, t + \mathrm{d}t])$ *we have*

$$\mathrm{E}_G\big[\mathrm{d}M(\tau)\,\mathrm{d}N(t)\big] = \lambda^2\,\mathrm{d}\tau\,\mathrm{d}t + \lambda\,\mathrm{d}G(t - \tau)\,\mathrm{d}\tau.$$

(ii) *If A_1, A_2 and B are intervals such that $A_1 \cap A_2 = \varnothing$, then*

$$\mathrm{E}_G\big[M(A_1)M(A_2)N(B)\big] = \lambda^3|A_1||A_2||B| + \lambda^2 Q(A_1, B)|A_2| + \lambda^2 Q(A_2, B)|A_1|, \tag{A.3}$$

and for $\tau_1 \neq \tau_2$

$$\begin{aligned}&\mathrm{E}_G\big[\mathrm{d}M(\tau_1)\,\mathrm{d}M(\tau_2)\,\mathrm{d}N(t)\big]\\ &\quad = \lambda^3\,\mathrm{d}\tau_1\,\mathrm{d}\tau_2\,\mathrm{d}t + \lambda^2\,\mathrm{d}G(t - \tau_1)\,\mathrm{d}\tau_1\,\mathrm{d}\tau_2 + \lambda^2\,\mathrm{d}G(t - \tau_2)\,\mathrm{d}\tau_2\,\mathrm{d}\tau_1.\end{aligned} \tag{A.4}$$

Similarly, for intervals B_1, B_2 and A such that $B_1 \cap B_2 = \varnothing$

$$\mathrm{E}_G\big[M(A)N(B_1)N(B_2)\big] = \lambda^3|A||B_1||B_2| + \lambda^2 Q(A, B_1)|B_2| + \lambda^2 Q(A, B_2)|B_1|, \tag{A.5}$$

and for $t_1 \neq t_2$

$$\begin{aligned}&\mathrm{E}_G\big[\mathrm{d}M(\tau)\,\mathrm{d}N(t_1)\,\mathrm{d}N(t_2)\big]\\ &\quad = \lambda^3\,\mathrm{d}\tau\,\mathrm{d}t_1\,\mathrm{d}t_2 + \lambda^2\,\mathrm{d}G(t_1 - \tau)\,\mathrm{d}\tau\,\mathrm{d}t_2 + \lambda^2\,\mathrm{d}G(t_2 - \tau)\,\mathrm{d}\tau\,\mathrm{d}t_1.\end{aligned} \tag{A.6}$$

(iii) *If $A_1 \cap A_2 = \varnothing$ and $B_1 \cap B_2 = \varnothing$, then*

$$\begin{aligned}&\mathrm{E}_G\big[M(A_1)M(A_2)N(B_1)N(B_2)\big] - \mathrm{E}_G\big[M(A_1)N(B_1)\big]\mathrm{E}_G\big[M(A_2)N(B_2)\big]\\ &\quad = \lambda^3\big[Q(A_1, B_2)|A_2||B_1| + Q(A_2, B_1)|A_1||B_2|\big] + \lambda^2 Q(A_1, B_2)Q(A_2, B_1).\end{aligned}$$

In particular, for $\tau_1 \neq \tau_2$ and $t_1 \neq t_2$

$$\begin{aligned}&\mathrm{E}_G\big[\mathrm{d}M(\tau_1)\,\mathrm{d}M(\tau_2)\,\mathrm{d}N(t_1)\,\mathrm{d}N(t_2)\big] - \mathrm{E}_G\big[\mathrm{d}M(\tau_1)\,\mathrm{d}N(t_1)\big]\mathrm{E}_G\big[\mathrm{d}M(\tau_2)\,\mathrm{d}N(t_2)\big]\\ &\quad = \lambda^3\big[\mathrm{d}G(t_2 - \tau_1)\,\mathrm{d}\tau_1\,\mathrm{d}\tau_2\,\mathrm{d}t_1 + \mathrm{d}G(t_1 - \tau_2)\,\mathrm{d}\tau_2\,\mathrm{d}\tau_1\,\mathrm{d}t_2\big]\\ &\qquad + \lambda^2\,\mathrm{d}G(t_2 - \tau_1)\,\mathrm{d}\tau_1\,\mathrm{d}G(t_1 - \tau_2)\,\mathrm{d}\tau_2.\end{aligned} \tag{A.7}$$

Proof. The proof follows by straightforward though tedious differentiation of (3.1).

(i) Write for brevity $\psi := \mathrm{E}_G \exp\{\eta M(A) + \xi N(B)\}$, $Q := Q(A, B)$, and

$$\begin{aligned}U = U(\eta, \xi) &:= |A|\big(e^\eta - 1\big) + |B|\big(e^\xi - 1\big) + \big(e^\eta - 1\big)\big(e^\xi - 1\big)Q(A, B),\\ \Sigma_A(\xi) &:= |A| + \big(e^\xi - 1\big)Q, \qquad \Sigma_B(\eta) := |B| + \big(e^\eta - 1\big)Q.\end{aligned}$$

It follows from (3.1) that $\psi = e^{\lambda U}$. Note that $U(0,0) = 0$, $\Sigma_A(0) = |A|$ and $\Sigma_B(0) = |B|$.

With the introduced notation

$$\frac{\partial U}{\partial \eta} = e^{\eta}\Sigma_A(\xi), \qquad \frac{\partial U}{\partial \xi} = e^{\xi}\Sigma_B(\eta),$$
$$\frac{\partial \Sigma_A(\xi)}{\partial \xi} = e^{\xi} Q, \qquad \frac{\partial \Sigma_B(\eta)}{\partial \eta} = e^{\eta} Q.$$

Using these relations we obtain

$$\begin{aligned} \frac{\partial \psi}{\partial \eta} &= \lambda e^{\lambda U} e^{\eta}\Sigma_A(\xi), \qquad \frac{\partial^2 \psi}{\partial \eta^2} = \lambda^2 e^{\lambda U} e^{2\eta}\Sigma_A^2(\xi) + \lambda e^{\lambda U} e^{\eta}\Sigma_A(\xi), \\ \frac{\partial \psi}{\partial \xi} &= \lambda e^{\lambda U} e^{\xi}\Sigma_B(\eta), \qquad \frac{\partial^2 \psi}{\partial \xi^2} = \lambda^2 e^{\lambda U} e^{2\xi}\Sigma_B^2(\eta) + \lambda e^{\lambda U} e^{\xi}\Sigma_B(\eta), \\ \frac{\partial^2 \psi}{\partial \eta \partial \xi} &= \lambda^2 e^{\lambda U} e^{\eta+\xi}\Sigma_A(\xi)\Sigma_B(\eta) + \lambda Q e^{\lambda U} e^{\eta+\xi}. \end{aligned} \tag{A.8}$$

Then (A.2) follows immediately from (A.8) by substituting $\eta = \xi = 0$.

(ii) From (3.1), we have

$$\begin{aligned} &\frac{1}{\lambda}\log \psi(\eta_1, \eta_2, \xi_1, \xi_2) \\ &\quad := \frac{1}{\lambda}\log \mathrm{E}_G \exp\{\eta_1 M(A_1) + \eta_2 M(A_2) + \xi_1 N(B_1) + \xi_2 N(B_2)\} \\ &\quad = |A_1|(e^{\eta_1} - 1) + |A_2|(e^{\eta_2} - 1) + |B_1|(e^{\xi_1} - 1) + |B_2|(e^{\xi_2} - 1) \\ &\qquad + \sum_{i=1}^2 \sum_{j=1}^2 (e^{\eta_i} - 1)(e^{\xi_j} - 1) Q(A_i, B_j). \end{aligned} \tag{A.9}$$

Let $U = U(\eta_1, \eta_2, \xi_1, \xi_2)$ stand for the right-hand side of (A.9) and put for brevity

$$\Sigma(A_i) := \sum_{j=1}^2 (e^{\xi_j} - 1) Q(A_i, B_j), \qquad S(B_j) := \sum_{i=1}^2 (e^{\eta_i} - 1) Q(A_i, B_j).$$

Then

$$\begin{aligned} \frac{\partial \psi}{\partial \eta_1} &= e^{\lambda U}\lambda e^{\eta_1}\bigl[|A_1| + \Sigma(A_1)\bigr], \qquad \frac{\partial \psi}{\partial \xi_1} = e^{\lambda U}\lambda e^{\xi_1}\bigl[|B_1| + \Sigma(B_1)\bigr], \\ \frac{\partial^2 \psi}{\partial \eta_1 \partial \eta_2} &= e^{\lambda U}\lambda^2 e^{\eta_1+\eta_2}\bigl[|A_1| + \Sigma(A_1)\bigr]\bigl[|A_2| + \Sigma(A_2)\bigr], \\ \frac{\partial^2 \psi}{\partial \xi_1 \partial \xi_2} &= e^{\lambda U}\lambda^2 e^{\xi_1+\xi_2}\bigl[|B_1| + \Sigma(B_1)\bigr]\bigl[|B_2| + \Sigma(B_2)\bigr], \end{aligned}$$

$$\begin{aligned}\frac{\partial^3\psi}{\partial\eta_1\partial\eta_2\partial\xi_1} &= e^{\lambda U}\lambda^3 e^{\eta_1+\eta_2+\xi_1}\big[|A_1|+\Sigma(A_1)\big]\big[|A_2|+\Sigma(A_2)\big]\big[|B_1|+S(B_1)\big]\\ &\quad + e^{\lambda U}\lambda^2 e^{\eta_1+\eta_2+\xi_1} Q(A_1,B_1)\big[|A_2|+\Sigma(A_2)\big]\\ &\quad + e^{\lambda U}\lambda^2 e^{\eta_1+\eta_2+\xi_1} Q(A_2,B_1)\big[|A_1|+\Sigma(A_1)\big],\end{aligned} \tag{A.10}$$

$$\begin{aligned}\frac{\partial^3\psi}{\partial\eta_1\partial\xi_1\partial\xi_2} &= e^{\lambda U}\lambda^3 e^{\xi_1+\xi_2+\eta_1}\big[|B_1|+\Sigma(B_1)\big]\big[|B_2|+\Sigma(B_2)\big]\big[|A_1|+S(A_1)\big]\\ &\quad + e^{\lambda U}\lambda^2 e^{\xi_1+\xi_2+\eta_1} Q(A_1,B_1)\big[|B_2|+\Sigma(B_2)\big]\\ &\quad + e^{\lambda U}\lambda^2 e^{\xi_1+\xi_2+\eta_1} Q(A_1,B_2)\big[|B_1|+\Sigma(B_1)\big].\end{aligned} \tag{A.11}$$

Now substituting $\eta_1=\eta_2=\xi_1=\xi_2=0$ in (A.10) and (A.11) we come to (A.3) and (A.5).

(iii) Further differentiation yields

$$\begin{aligned}&\frac{\partial^4\psi}{\partial\eta_1\partial\eta_2\partial\xi_1\partial\xi_2}\\ &\quad= e^{\lambda U}\lambda^4 e^{\eta_1+\eta_2+\xi_1+\xi_2}\big[|A_1|+\Sigma(A_1)\big]\big[|A_2|+\Sigma(A_2)\big]\big[|B_1|+S(B_1)\big]\big[|B_2|+S(B_2)\big]\\ &\qquad + e^{\lambda U}\lambda^3 e^{\eta_1+\eta_2+\xi_1+\xi_2} Q(A_1,B_2)\big[|A_2|+\Sigma(A_2)\big]\big[|B_1|+S(B_1)\big]\\ &\qquad + e^{\lambda U}\lambda^3 e^{\eta_1+\eta_2+\xi_1+\xi_2}\big[|A_1|+\Sigma(A_1)\big]Q(A_2,B_2)\big[|B_1|+S(B_1)\big]\\ &\qquad + e^{\lambda U}\lambda^3 e^{\eta_1+\eta_2+\xi_1+\xi_2} Q(A_1,B_1)\big[|A_2|+\Sigma(A_2)\big]\big[|B_2|+S(B_2)\big]\\ &\qquad + e^{\lambda U}\lambda^3 e^{\eta_1+\eta_2+\xi_1+\xi_2} Q(A_2,B_1)\big[|A_1|+\Sigma(A_1)\big]\big[|A_2|+S(B_2)\big]\\ &\qquad + e^{\lambda U}\lambda^2 e^{\eta_1+\eta_2+\xi_1+\xi_2} Q(A_1,B_1)Q(A_2,B_2)\\ &\qquad + e^{\lambda U}\lambda^2 e^{\eta_1+\eta_2+\xi_1+\xi_2} Q(A_2,B_1)Q(A_1,B_2).\end{aligned}$$

Computing the fourth derivative at $\eta_1=\eta_2=\xi_1=\xi_2=0$ we obtain

$$\begin{aligned}&\mathrm{E}_G\big[M(A_1)M(A_2)N(B_1)N(B_2)\big]\\ &\quad= \lambda^4|A_1||A_2||B_1||B_2|\\ &\qquad + \lambda^3\big[Q(A_1,B_2)|A_2||B_1| + Q(A_2,B_2)|A_1||B_1|\\ &\qquad + Q(A_1,B_1)|A_2||B_2| + Q(A_2,B_1)|A_1||B_2|\big]\\ &\qquad + \lambda^2\big[Q(A_1,B_1)Q(A_2,B_2)+Q(A_1,B_2)Q(A_2,B_1)\big].\end{aligned}$$

By (A.2)

$$\begin{aligned}&\mathrm{E}_G\big[M(A_1)N(B_1)\big]\mathrm{E}_G\big[M(A_2)N(B_2)\big]\\ &\quad= \lambda^4|A_1||B_1||A_2||B_2| + \lambda^2 Q(A_1,B_1)Q(A_2,B_2)\\ &\qquad + \lambda^3\big[|A_1||B_1|Q(A_2,B_2)+|A_2||B_2|Q(A_1,B_1)\big];\end{aligned}$$

therefore

$$\mathrm{E}_G\big[M(A_1)M(A_2)N(B_1)N(B_2)\big]-\mathrm{E}_G\big[M(A_1)N(B_1)\big]\mathrm{E}_G\big[M(A_2)N(B_2)\big]$$
$$=\lambda^3\big(Q(A_1,B_2)|A_2||B_1|+Q(A_2,B_1)|A_1||B_2|\big)+\lambda^2 Q(A_1,B_2)Q(A_2,B_1),$$

as claimed. □

A.3. Covariance measures of S

The next result establishes expression for the covariance measures of the superposed process S. Note that the part (i) of the lemma is the restatement of Corollary 2.

Let $Q(\cdot,\cdot)$ be given by (3.2), and for any pair of disjoint intervals A_i and A_j define

$$\tilde{Q}_{i,j}=\tilde{Q}_{j,i}:=\frac{1}{2}\big[Q(A_i,A_j)+Q(A_j,A_i)\big].$$

Lemma 2. *The following statements hold.*

(i) *Let A_1 and A_2 be disjoint intervals*; *then*

$$\mathrm{E}_G\big[S(A_1)S(A_2)\big]=4\lambda^2|A_1||A_2|+\lambda\big[Q(A_1,A_2)+Q(A_2,A_1)\big]. \tag{A.12}$$

In particular, for $A_1=(\tau,\tau+\mathrm{d}\tau]$ and $A_2=(t,t+\mathrm{d}t]$ with $\tau\neq t$ one has

$$\mathrm{E}_G\big[\mathrm{d}S(\tau)\,\mathrm{d}S(t)\big]=4\lambda^2\,\mathrm{d}\tau\,\mathrm{d}t+\lambda\big[\mathrm{d}G(t-\tau)\,\mathrm{d}\tau+\mathrm{d}G(\tau-t)\,\mathrm{d}t\big].$$

(ii) *If A_1, A_2 and A_3 are disjoint intervals*, *then*

$$\mathrm{E}_G\big[S(A_1)S(A_2)S(A_3)\big]=8\lambda^3|A_1||A_2||A_3|+4\lambda^2\big[\tilde{Q}_{1,2}|A_3|+\tilde{Q}_{1,3}|A_2|+\tilde{Q}_{2,3}|A_1|\big],$$

and for all distinct x, y and z

$$\begin{aligned}\mathrm{E}_G\big[\mathrm{d}S(x)\,\mathrm{d}S(y)\,\mathrm{d}S(z)\big]=8\lambda^3\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z&+2\lambda^2\,\mathrm{d}y\big[\mathrm{d}G(x-z)\,\mathrm{d}z+\mathrm{d}G(z-x)\,\mathrm{d}x\big]\\&+2\lambda^2\,\mathrm{d}x\big[\mathrm{d}G(z-y)\,\mathrm{d}y+\mathrm{d}G(y-z)\,\mathrm{d}z\big]\\&+2\lambda^2\,\mathrm{d}z\big[\mathrm{d}G(x-y)\,\mathrm{d}y+\mathrm{d}G(y-x)\,\mathrm{d}x\big].\end{aligned} \tag{A.13}$$

(iii) *If A_1, A_2, A_3, A_4 are disjoint intervals*, *then*

$$\begin{aligned}&\mathrm{E}_G\big[S(A_1)S(A_2)S(A_3)S(A_4)\big]-\mathrm{E}_G\big[S(A_1)S(A_2)\big]\mathrm{E}_G\big[S(A_3)S(A_4)\big]\\&\quad=8\lambda^3\big[\tilde{Q}_{1,4}|A_2||A_3|+\tilde{Q}_{2,4}|A_1||A_3|+\tilde{Q}_{1,3}|A_2||A_4|+\tilde{Q}_{2,3}|A_1||A_4|\big]\\&\qquad+4\lambda^2[\tilde{Q}_{1,3}\tilde{Q}_{2,4}+\tilde{Q}_{2,3}\tilde{Q}_{1,4}].\end{aligned}$$

In particular, if $A_1 = (\tau_1, \tau_1 + d\tau_1]$, $A_2 = (t_1, t_1 + \mathrm{d}t_1]$, $A_3 = (\tau_2, \tau_2 + \mathrm{d}\tau_2]$ *and* $A_4 = (t_2, t_2 + \mathrm{d}t_2]$ *with all distinct* τ_1, t_1, τ_2 *and* t_2 *then*

$$\begin{aligned}
&\mathrm{E}_G\big[\mathrm{d}S(\tau_1)\,\mathrm{d}S(t_1)\,\mathrm{d}S(\tau_2)\,\mathrm{d}S(t_2)\big] - \mathrm{E}_G\big[\mathrm{d}S(\tau_1)\,\mathrm{d}S(t_1)\big]\mathrm{E}_G\big[\mathrm{d}S(\tau_2)\,\mathrm{d}S(t_2)\big] \\
&\quad = 8\lambda^3\big[\tilde{Q}(\mathrm{d}\tau_1, \mathrm{d}t_2)\,\mathrm{d}t_1\,\mathrm{d}\tau_2 + \tilde{Q}(\mathrm{d}\tau_1, \mathrm{d}\tau_2)\,\mathrm{d}t_1\,\mathrm{d}t_2 \\
&\qquad + \tilde{Q}(\mathrm{d}t_1, \mathrm{d}\tau_2)\,\mathrm{d}\tau_1\,\mathrm{d}t_2 + \tilde{Q}(\mathrm{d}t_1, \mathrm{d}t_2)\,\mathrm{d}\tau_1\,\mathrm{d}\tau_2\big] \\
&\qquad + 4\lambda^2\big[\tilde{Q}(\mathrm{d}\tau_1, \mathrm{d}\tau_2)\tilde{Q}(\mathrm{d}t_1, \mathrm{d}t_2) + \tilde{Q}(\mathrm{d}t_1, \mathrm{d}\tau_2)\tilde{Q}(\mathrm{d}\tau_1, \mathrm{d}t_2)\big],
\end{aligned}$$

where $\tilde{Q}(\mathrm{d}x, \mathrm{d}y) := \frac{1}{2}[\mathrm{d}G(x-y)\,\mathrm{d}y + \mathrm{d}G(y-x)\,\mathrm{d}x]$.

Proof. (i) It follows from Proposition 2 that

$$\begin{aligned}
&\psi = \psi(\eta_1, \eta_2) := \mathrm{E}_G \exp\big\{\eta_1 S(A_1) + \eta_2 S(A_2)\big\} =: \exp\{\lambda U\}, \\
&U = U(\eta_1, \eta_2) \\
&\quad = 2\big(e^{\eta_1} - 1\big)|A_1| + 2\big(e^{\eta_2} - 1\big)|A_2| + \big(e^{\eta_1} - 1\big)^2 Q_{1,1} + \big(e^{\eta_2} - 1\big)^2 Q_{2,2} \\
&\qquad + \big(e^{\eta_1} - 1\big)\big(e^{\eta_2} - 1\big)[Q_{1,2} + Q_{2,1}],
\end{aligned}$$

where, for the sake of brevity, $Q_{i,j}$ stands for $Q(A_i, A_j)$. Denote also

$$\begin{aligned}
\Sigma_1(\eta_1, \eta_2) &:= \big(e^{\eta_1} - 1\big)Q_{1,1} + \big(e^{\eta_2} - 1\big)\frac{1}{2}(Q_{1,2} + Q_{2,1}), \\
\Sigma_2(\eta_1, \eta_2) &:= \big(e^{\eta_2} - 1\big)Q_{2,2} + \big(e^{\eta_1} - 1\big)\frac{1}{2}(Q_{1,2} + Q_{2,1}).
\end{aligned}$$

With this notation, we have

$$\begin{aligned}
\frac{\partial \psi}{\partial \eta_1} &= \lambda e^{\lambda U}\frac{\partial U}{\partial \eta_1} = \lambda e^{\lambda U} 2e^{\eta_1}\big[|A_1| + \Sigma_1(\eta_1, \eta_2)\big], \\
\frac{\partial^2 \psi}{\partial \eta_1 \partial \eta_2} &= \lambda^2 e^{\lambda U}\frac{\partial U}{\partial \eta_1}\frac{\partial U}{\partial \eta_2} + \lambda e^{\lambda U}\frac{\partial^2 U}{\partial \eta_1 \partial \eta_2} \\
&= \lambda^2 e^{\lambda U} 4e^{\eta_1 + \eta_2}\big[|A_1| + \Sigma_1(\eta_1, \eta_2)\big]\big[|A_2| + \Sigma_2(\eta_1, \eta_2)\big] + \lambda e^{\lambda U} e^{\eta_1 + \eta_2}(Q_{1,2} + Q_{2,1}).
\end{aligned}$$

The last formula yields

$$\mathrm{E}_G\big[S(A_1)S(A_2)\big] = \frac{\partial^2 \psi}{\partial \eta_1 \partial \eta_2}\bigg|_{\eta_1 = \eta_2 = 0} = 4\lambda^2|A_1||A_2| + \lambda(Q_{1,2} + Q_{2,1}),$$

as claimed in (A.12).

(ii) Now we set $\psi := \mathrm{E}_G \exp\{\sum_{i=1}^4 \eta_i S(A_i)\}$,

$$U := \sum_{i=1}^{4} \bigl(e^{\eta_i} - 1\bigr)|A_i| + \frac{1}{2}\sum_{\substack{i=1 \\ i\neq j}}^{4}\sum_{j=1}^{4}\bigl(e^{\eta_i}-1\bigr)\bigl(e^{\eta_j}-1\bigr)Q_{i,j} + \frac{1}{2}\sum_{i=1}^{4}\bigl(e^{\eta_i}-1\bigr)Q_{i,i},$$

$$\tilde{Q}_{i,j} := \tilde{Q}_{j,i} = \frac{1}{2}(Q_{i,j} + Q_{j,i}), \qquad i, j = 1, 2, 3, 4,$$

$$\Sigma_k := |A_k| + \sum_{j=1}^{4}\bigl(e^{\eta_j}-1\bigr)\tilde{Q}_{j,k}, \qquad k = 1, 2, 3, 4.$$

According to Proposition 2, with this notation $\psi = e^{2\lambda U}$, $\frac{\partial U}{\partial \eta_k} = e^{\eta_k}\Sigma_k$, $k = 1, \dots, 4$, and therefore

$$\begin{aligned}
\frac{\partial \psi}{\partial \eta_1} &= 2\lambda e^{2\lambda U} e^{\eta_1}\Sigma_1, \\
\frac{\partial^2 \psi}{\partial \eta_1 \partial \eta_2} &= e^{2\lambda U} e^{\eta_1+\eta_2}\bigl[4\lambda^2 \Sigma_1\Sigma_2 + 2\lambda \tilde{Q}_{1,2}\bigr], \\
\frac{\partial^3 \psi}{\partial \eta_1 \partial \eta_2 \partial \eta_3} &= e^{2\lambda U} e^{\eta_1+\eta_2+\eta_3}\bigl[8\lambda^3 \Sigma_1\Sigma_2\Sigma_3 + 4\lambda^2\tilde{Q}_{1,3}\Sigma_2 + 4\lambda^2\tilde{Q}_{2,3}\Sigma_1 + 4\lambda^2\tilde{Q}_{1,2}\Sigma_3\bigr].
\end{aligned} \tag{A.14}$$

Then

$$\begin{aligned}
\mathrm{E}_G\bigl[S(A_1)S(A_2)S(A_3)\bigr] &= \frac{\partial^3 \psi}{\partial \eta_1 \partial \eta_2 \partial \eta_3}\bigg|_{\eta_1=\eta_2=\eta_3=0} \\
&= 8\lambda^3|A_1||A_2||A_3| + 4\lambda^2\bigl[\tilde{Q}_{1,2}|A_3| + \tilde{Q}_{1,3}|A_2| + \tilde{Q}_{2,3}|A_1|\bigr],
\end{aligned}$$

as claimed.

(iii) Finally, differentiating (A.14) we obtain

$$\begin{aligned}
&\frac{\partial^4 \psi}{\partial \eta_1 \partial \eta_2 \partial \eta_3 \partial \eta_4} \\
&\quad = e^{2\lambda U} e^{\eta_1+\eta_2+\eta_3+\eta_4}\bigl[16\lambda^4\Sigma_1\Sigma_2\Sigma_3\Sigma_4 + 8\lambda^3\tilde{Q}_{1,4}\Sigma_2\Sigma_3 + 8\lambda^3\tilde{Q}_{2,4}\Sigma_1\Sigma_3 + 8\lambda^3\tilde{Q}_{3,4}\Sigma_1\Sigma_2 \\
&\qquad + 8\lambda^3\tilde{Q}_{1,3}\Sigma_2\Sigma_4 + 8\lambda^3\tilde{Q}_{2,3}\Sigma_1\Sigma_4 + 8\lambda^3\tilde{Q}_{1,2}\Sigma_3\Sigma_4 + 4\lambda^2\tilde{Q}_{1,3}\tilde{Q}_{2,4} \\
&\qquad + 4\lambda^2\tilde{Q}_{2,3}\tilde{Q}_{1,4} + 4\lambda^2\tilde{Q}_{1,2}\tilde{Q}_{2,4}\bigr].
\end{aligned}$$

The last formula yields

$$\begin{aligned}
&\mathrm{E}_G\bigl[S(A_1)S(A_2)S(A_3)S(A_4)\bigr] \\
&\quad = \frac{\partial^4 \phi}{\partial \eta_1 \partial \eta_2 \partial \eta_3 \partial \eta_4}\bigg|_{\eta_1=\eta_2=\eta_3=\eta_4=0}
\end{aligned}$$

$$
\begin{aligned}
&= 16\lambda^4 \prod_{i=1}^{4} |A_i| + 4\lambda^2[\tilde{Q}_{1,3}\tilde{Q}_{2,4} + \tilde{Q}_{2,3}\tilde{Q}_{1,4} + \tilde{Q}_{1,2}\tilde{Q}_{3,4}] \\
&\quad + 8\lambda^3\big[\tilde{Q}_{1,4}|A_2||A_3| + \tilde{Q}_{2,4}|A_1||A_3| + \tilde{Q}_{3,4}|A_1||A_2| \\
&\quad + \tilde{Q}_{1,3}|A_2||A_4| + \tilde{Q}_{2,3}|A_1||A_4| + \tilde{Q}_{1,2}|A_3||A_4|\big].
\end{aligned}
$$

Taking into account (A.12) we obtain

$$
\begin{aligned}
&\mathrm{E}_G\big[S(A_1)S(A_2)S(A_3)S(A_4)\big] - \mathrm{E}_G\big[S(A_1)S(A_2)\big]\mathrm{E}_G\big[S(A_3)S(A_4)\big] \\
&\quad = 8\lambda^3\big[\tilde{Q}_{1,4}|A_2||A_3| + \tilde{Q}_{2,4}|A_1||A_3| + \tilde{Q}_{1,3}|A_2||A_4| + \tilde{Q}_{2,3}|A_1||A_4|\big] \\
&\qquad + 4\lambda^2[\tilde{Q}_{1,3}\tilde{Q}_{2,4} + \tilde{Q}_{2,3}\tilde{Q}_{1,4}],
\end{aligned}
$$

as claimed. □

Appendix B: Proofs for Sections 4 and 6

B.1. Proof of Theorem 1

The fact that $\hat{\theta}_I$ is unbiased follows immediately from (4.1). We compute the variance of $\hat{\theta}_I$:

$$
\mathrm{var}_G\{\hat{\theta}_I\} = \frac{1}{\lambda^2 T^2}\left[\mathrm{E}_G\left\{\sum_{j\in\mathbb{Z}}\sum_{k\in\mathbb{Z}}\varphi_*(\tau_j, t_k)\right\}^2 - \left(\iint \varphi_*(\tau, t)\mathrm{E}_G\big[\mathrm{d}M(\tau)\,\mathrm{d}N(t)\big]\right)^2\right]. \tag{B.1}
$$

The expression in the square brackets on the right-hand side of (B.1) is decomposed as

$$
\begin{aligned}
&\mathrm{E}_G \sum_{j\in\mathbb{Z}}\sum_{k\in\mathbb{Z}} \varphi_*^2(\tau_j, t_k) + \mathrm{E}_G \sum_{j_1\in\mathbb{Z}}\sum_{\substack{j_2\in\mathbb{Z}\\ j_1\neq j_2}}\sum_{k\in\mathbb{Z}} \varphi_*(\tau_{j_1}, t_k)\varphi_*(\tau_{j_2}, t_k) \\
&\quad + \mathrm{E}_G \sum_{j\in\mathbb{Z}}\sum_{\substack{k_1\in\mathbb{Z}\\ k_1\neq k_2}}\sum_{k_2\in\mathbb{Z}} \varphi_*(\tau_j, t_{k_1})\varphi_*(\tau_j, t_{k_2}) \\
&\quad + \left\{\mathrm{E}_G \sum_{j_1\in\mathbb{Z}}\sum_{k_1\in\mathbb{Z}}\sum_{\substack{j_2\in\mathbb{Z}\\ j_1\neq j_2, k_1\neq k_2}}\sum_{k_2\in\mathbb{Z}} \varphi_*(\tau_{j_1}, t_{k_1})\varphi_*(\tau_{j_2}, t_{k_2}) - \left(\iint \varphi_*(\tau, t)\mathrm{E}_G\big[\mathrm{d}M(\tau)\,\mathrm{d}N(t)\big]\right)^2\right\} \\
&=: J_1 + J_2 + J_3 + J_4.
\end{aligned} \tag{B.2}
$$

Now we compute the terms J_i, $i = 1, \dots, 4$ separately.

Let

$$
\chi(\tau, t) := \begin{cases} 1, & \tau \neq t, \\ 0, & \text{otherwise}, \end{cases} \qquad \tau, t \in \mathbb{R}.
$$

By (4.1) we have

$$J_1 = \mathrm{E}_G \iint \mathbf{1}_{\mathcal{T}_M}(\tau)\mathbf{1}_I(t-\tau)\,\mathrm{d}M(\tau)\,\mathrm{d}N(t) = \lambda^2|I|T + \lambda G(I)T. \tag{B.3}$$

Next, using (A.4) we obtain

$$\begin{aligned} J_2 &= \mathrm{E}_G \iiint \mathbf{1}_{\mathcal{T}_M}(\tau_1)\mathbf{1}_{\mathcal{T}_M}(\tau_2)\mathbf{1}_I(t-\tau_1)\mathbf{1}_I(t-\tau_2)\chi(\tau_1,\tau_2)\,\mathrm{d}M(\tau_1)\,\mathrm{d}M(\tau_2)\,\mathrm{d}N(t) \\ &= \lambda^3 \iiint \mathbf{1}_{\mathcal{T}_M}(\tau_1)\mathbf{1}_{\mathcal{T}_M}(\tau_2)\mathbf{1}_I(t-\tau_1)\mathbf{1}_I(t-\tau_2)\,\mathrm{d}\tau_1\,\mathrm{d}\tau_2\,\mathrm{d}t \\ &\quad + \lambda^2 \iiint \mathbf{1}_{\mathcal{T}_M}(\tau_1)\mathbf{1}_{\mathcal{T}_M}(\tau_2)\mathbf{1}_I(t-\tau_1)\mathbf{1}_I(t-\tau_2)\,\mathrm{d}G(t-\tau_1)\,\mathrm{d}\tau_1\,\mathrm{d}\tau_2 \\ &\quad + \lambda^2 \iiint \mathbf{1}_{\mathcal{T}_M}(\tau_1)\mathbf{1}_{\mathcal{T}_M}(\tau_2)\mathbf{1}_I(t-\tau_1)\mathbf{1}_I(t-\tau_2)\,\mathrm{d}G(t-\tau_2)\,\mathrm{d}\tau_1\,\mathrm{d}\tau_2 \\ &=: J_2^{(1)} + 2J_2^{(2)}, \end{aligned}$$

where

$$J_2^{(2)} = \lambda^2 \iiint \mathbf{1}_{\mathcal{T}_M}(\tau_1)\mathbf{1}_{\mathcal{T}_M}(\tau_2)\mathbf{1}_I(u)\mathbf{1}_I(u+\tau_1-\tau_2)\,\mathrm{d}G(u)\,\mathrm{d}\tau_1\,\mathrm{d}\tau_2.$$

Here we have taken into account that the last two integrals in the expression for J_2 coincide. The integrals $J_2^{(1)}$ and $J_2^{(2)}$ are evaluated as follows. We have

$$\begin{aligned} &\int \mathbf{1}_I(t-\tau_1)\mathbf{1}_I(t-\tau_2)\,\mathrm{d}t \\ &\quad = \int \mathbf{1}_I(u)\mathbf{1}_I(u+\tau_1-\tau_2)\,\mathrm{d}u \\ &\quad = \big[|I|-(\tau_1-\tau_2)\big]\mathbf{1}\big(0\le\tau_1-\tau_2\le|I|\big) + \big[|I|+(\tau_1-\tau_2)\big]\mathbf{1}\big(-|I|\le\tau_1-\tau_2\le 0\big). \end{aligned}$$

Substituting this expression we obtain

$$J_2^{(1)} = \lambda^3\int_0^{|I|}\big(|I|-u\big)(T-u)\,\mathrm{d}u + \lambda^3\int_{-|I|}^0\big(|I|+u\big)(T+u)\,\mathrm{d}u = \lambda^3|I|^2\bigg(T-\frac{1}{3}|I|\bigg).$$

Similarly, for $I=(a,b]$

$$\begin{aligned} \int \mathbf{1}_I(u)\mathbf{1}_I(u+\tau_2-\tau_1)\,\mathrm{d}G(u) &= \big[G(b)-G(a+\tau_2-\tau_1)\big]\mathbf{1}\big\{0\le\tau_2-\tau_1\le|I|\big\} \\ &\quad + \big[G(b+\tau_2-\tau_1)-G(a)\big]\mathbf{1}\big\{-|I|\le\tau_2-\tau_1<0\big\}. \end{aligned}$$

Hence the integral $J_2^{(2)}$ takes the form

$$\begin{aligned}
&\lambda^2 \iiint \mathbf{1}_{\mathcal{T}_M}(\tau_1)\mathbf{1}_{\mathcal{T}_M}(\tau_2)\mathbf{1}_I(u)\mathbf{1}_I(u+\tau_1-\tau_2)\,\mathrm{d}G(u)\,\mathrm{d}\tau_1\,\mathrm{d}\tau_2 \\
&\quad = \lambda^2 \int_0^{|I|} (T-u)\big[G(b) - G(a+u)\big]\,\mathrm{d}u + \lambda^2 \int_{-|I|}^{0} (T+u)\big[G(b+u) - G(a)\big]\,\mathrm{d}u \\
&\quad = \lambda^2 T \int_0^{|I|} \left(1 - \frac{u}{T}\right)\big[G(I) + G(b-u) - G(a+u)\big]\,\mathrm{d}u,
\end{aligned}$$

so that

$$J_2 = \lambda^3 |I|^2 \left(T - \frac{1}{3}|I|\right) + 2\lambda^2 T \int_0^{|I|} \left(1 - \frac{u}{T}\right)\big[G(I) + G(b-u) - G(a+u)\big]\,\mathrm{d}u. \tag{B.4}$$

The similar calculation that uses (A.6) yields

$$\begin{aligned}
J_3 &= \mathrm{E}_G \iiint \mathbf{1}_{\mathcal{T}_M}(\tau)\mathbf{1}_I(t_1-\tau)\mathbf{1}_I(t_2-\tau)\chi(t_1,t_2)\,\mathrm{d}M(\tau)\,\mathrm{d}N(t_1)\,\mathrm{d}N(t_2) \\
&= \lambda^3 \iiint \mathbf{1}_{\mathcal{T}_M}(\tau)\mathbf{1}_I(t_1-\tau)\mathbf{1}_I(t_2-\tau)\,\mathrm{d}\tau\,\mathrm{d}t_1\,\mathrm{d}t_2 \\
&\quad + \lambda^2 \iiint \mathbf{1}_{\mathcal{T}_M}(\tau)\mathbf{1}_I(t_1-\tau)\mathbf{1}_I(t_2-\tau)\,\mathrm{d}G(t_1-\tau)\,\mathrm{d}\tau\,\mathrm{d}t_2 \qquad \text{(B.5)} \\
&\quad + \lambda^2 \iiint \mathbf{1}_{\mathcal{T}_M}(\tau)\mathbf{1}_I(t_1-\tau)\mathbf{1}_I(t_2-\tau)\,\mathrm{d}G(t_2-\tau)\,\mathrm{d}\tau\,\mathrm{d}t_1 \\
&= \lambda^3 |I|^2 T + 2\lambda^2 |I| G(I) T.
\end{aligned}$$

Finally, letting $\Omega = \{(\tau_1,\tau_2,t_1,t_2) \in \mathbb{R}^4 : \tau_1 \neq \tau_2, t_1 \neq t_2\}$, and using (A.7) we have

$$\begin{aligned}
J_4 &:= \iiiint_\Omega \varphi_*(\tau_1,t_1)\varphi_*(\tau_2,t_2) \\
&\qquad \times \big\{\mathrm{E}_G\big[\mathrm{d}M(\tau_1)\,\mathrm{d}M(\tau_2)\,\mathrm{d}N(t_1)\,\mathrm{d}N(t_2)\big] - \mathrm{E}_G\big[\mathrm{d}M(\tau_1)\,\mathrm{d}N(t_1)\big]\mathrm{E}_G\big[\mathrm{d}M(\tau_2)\,\mathrm{d}N(t_2)\big]\big\} \\
&= \lambda^3 \iiiint_\Omega \varphi_*(\tau_1,t_1)\varphi_*(\tau_2,t_2)\,\mathrm{d}G(t_2-\tau_1)\,\mathrm{d}\tau_1\,\mathrm{d}\tau_2\,\mathrm{d}t_1 \\
&\quad + \lambda^3 \iiiint_\Omega \varphi_*(\tau_1,t_1)\varphi_*(\tau_2,t_2)\,\mathrm{d}G(t_1-\tau_2)\,\mathrm{d}\tau_2\,\mathrm{d}\tau_1\,\mathrm{d}t_2 \\
&\quad + \lambda^2 \iiiint_\Omega \varphi_*(\tau_1,t_1)\varphi_*(\tau_2,t_2)\,\mathrm{d}G(t_2-\tau_1)\,\mathrm{d}\tau_1\,\mathrm{d}G(t_1-\tau_2)\,\mathrm{d}\tau_2 \\
&=: J_4^{(1)} + J_4^{(2)} + J_4^{(3)}.
\end{aligned}$$

Furthermore,

$$
\begin{aligned}
J_4^{(1)} &= \lambda^3 \iiiint_\Omega \mathbf{1}_{\mathcal{T}_M}(\tau_1)\mathbf{1}_{\mathcal{T}_M}(\tau_2)\mathbf{1}_I(t_1-\tau_1)\mathbf{1}_I(t_2-\tau_2)\,\mathrm{d}G(t_2-\tau_1)\,\mathrm{d}\tau_1\,\mathrm{d}\tau_2\,\mathrm{d}t_1 \\
&= \lambda^3|I| \iiint \mathbf{1}_{\mathcal{T}_M}(\tau_1)\mathbf{1}_{\mathcal{T}_M}(\tau_2)\mathbf{1}_I(u+\tau_1-\tau_2)\,\mathrm{d}G(u)\,\mathrm{d}\tau_1\,\mathrm{d}\tau_2 \\
&= \lambda^3|I| \int_0^T \int_0^T G(I+\tau_2-\tau_1)\,\mathrm{d}\tau_1\,\mathrm{d}\tau_2 \\
&= \lambda^3|I|T \int_{-T}^T G(I+u)\left(1-\frac{|u|}{T}\right)\mathrm{d}u.
\end{aligned}
\tag{B.6}
$$

The same expression holds for $J_4^{(2)}$. As for $J_4^{(3)}$, we have

$$
\begin{aligned}
J_4^{(3)} &= \lambda^2 \iiiint \mathbf{1}_{\mathcal{T}_M}(\tau_1)\mathbf{1}_{\mathcal{T}_M}(\tau_2)\mathbf{1}_I(t_1-\tau_1)\mathbf{1}_I(t_2-\tau_2)\,\mathrm{d}G(t_2-\tau_1)\,\mathrm{d}\tau_1\,\mathrm{d}G(t_1-\tau_2)\,\mathrm{d}\tau_2 \\
&= \lambda^2 \iiiint \mathbf{1}_{\mathcal{T}_M}(u_1)\mathbf{1}_{\mathcal{T}_M}(u_2)\mathbf{1}_I(w_1+u_2-u_1) \\
&\quad\times \mathbf{1}_I(w_2+u_1-u_2)\,\mathrm{d}G(w_2)\,\mathrm{d}u_1\,\mathrm{d}G(w_1)\,\mathrm{d}u_2 \\
&= \lambda^2 \int_0^T \int_0^T G(I+u_1-u_2)G(I+u_2-u_1)\,\mathrm{d}u_1\,\mathrm{d}u_2 \\
&= \lambda^2 T \int_{-T}^T G(I+u)G(I-u)\left(1-\frac{|u|}{T}\right)\mathrm{d}u.
\end{aligned}
\tag{B.7}
$$

Combining (B.1)–(B.7) we complete the proof.

B.2. Proof of Theorem 3

1^0. By definition,

$$
\tilde{G}(x_0) = \frac{1}{\lambda T}\iint \mathbf{1}_{[0,T]}(\tau)\mathbf{1}_{[0,x_0]}(t-\tau)\chi(\tau,t)\,\mathrm{d}S(\tau)\,\mathrm{d}S(t) - 4\lambda x_0,
$$

and it easily follows from (4.6) that $\tilde{G}(x_0)$ is an unbiased estimator of $G(x_0)$

The variance of $\tilde{G}(x_0)$ is

$$
\mathrm{var}_G\{\tilde{G}(x_0)\} = \frac{1}{\lambda^2 T^2}\left\{\left[\sum_{j\in\mathbb{Z}}\sum_{k\in\mathbb{Z}}\varphi_*(s_j,s_k)\right]^2 - \left[\mathrm{E}_G \iint \varphi_*(\tau,t)\,\mathrm{d}S(\tau)\,\mathrm{d}S(t)\right]^2\right\}.
$$

The expression in the curly brackets on the right-hand side of the previous display formula is decomposed as

$$\begin{aligned}
&\mathrm{E}_G \sum_{j\in\mathbb{Z}}\sum_{k\in\mathbb{Z}} \varphi_*^2(s_j, s_k) + \mathrm{E}_G \sum_{j\in\mathbb{Z}}\sum_{k\in\mathbb{Z}} \varphi_*(s_j, s_k)\varphi_*(s_k, s_j)\\
&\quad + \mathrm{E}_G \sum_{j\in\mathbb{Z}}\sum_{\substack{k_1\in\mathbb{Z}\\ k_1\neq k_2}}\sum_{k_2\in\mathbb{Z}} \varphi_*(s_j, s_{k_1})\varphi_*(s_j, s_{k_2}) + \mathrm{E}_G \sum_{j\in\mathbb{Z}}\sum_{\substack{k_1\in\mathbb{Z}\\ j\neq k_2}}\sum_{k_2\in\mathbb{Z}} \varphi_*(s_j, s_{k_1})\varphi_*(s_{k_1}, s_{k_2})\\
&\quad + \mathrm{E}_G \sum_{j_1\in\mathbb{Z}}\sum_{\substack{k\in\mathbb{Z}\\ k\neq j_2}}\sum_{j_2\in\mathbb{Z}} \varphi_*(s_{j_1}, s_k)\varphi_*(s_{j_2}, s_{j_1}) + \mathrm{E}_G \sum_{j_1\in\mathbb{Z}}\sum_{\substack{k\in\mathbb{Z}\\ j_1\neq j_2}}\sum_{j_2\in\mathbb{Z}} \varphi_*(s_{j_1}, s_k)\varphi_*(s_{j_2}, s_k)\\
&\quad + \mathrm{E}_G \sum_{j_1\in\mathbb{Z}}\sum_{\substack{k_1\in\mathbb{Z}\\ j_1\neq j_2, j_1\neq k_2\\ k_1\neq j_2, k_1\neq k_2}}\sum_{j_2\in\mathbb{Z}}\sum_{k_2\in\mathbb{Z}} \varphi_*(s_{j_1}, s_{k_1})\varphi_*(s_{j_2}, s_{k_2}) - \left[\mathrm{E}_G \iint \varphi_*(\tau, t)\,\mathrm{d}S(\tau)\,\mathrm{d}S(t)\right]^2 =: \sum_{k=1}^{7} J_k,
\end{aligned}$$

where

$$\begin{aligned}
J_1 &= \mathrm{E}_G \iint \mathbf{1}_{[0,T]}\mathbf{1}_{[0,x_0]}(t-\tau)\chi(\tau, t)\,\mathrm{d}S(\tau)\,\mathrm{d}S(t),\\
J_2 &= \mathrm{E}_G \iint \mathbf{1}_{[0,T]}(\tau)\mathbf{1}_{[0,x_0]}(t-\tau)\mathbf{1}_{[0,T]}(t)\mathbf{1}_{[0,x_0]}(\tau-t)\chi(\tau, t)\,\mathrm{d}S(\tau)\,\mathrm{d}S(t),\\
J_3 &= \mathrm{E}_G \iiint \mathbf{1}_{[0,T]}(\tau)\mathbf{1}_{[0,x_0]}(t_1-\tau)\mathbf{1}_{[0,x_0]}(t_2-\tau)\chi(\tau, t_1)\chi(\tau, t_2)\,\mathrm{d}S(\tau)\,\mathrm{d}S(t_1)\,\mathrm{d}S(t_2),\\
J_4 &= \mathrm{E}_G \iiint \mathbf{1}_{[0,T]}(\tau)\mathbf{1}_{[0,x_0]}(t_1-\tau)\mathbf{1}_{[0,T]}(t_1)\\
&\qquad \times \mathbf{1}_{[0,x_0]}(t_2-t_1)\chi(\tau, t_1)\chi(t_1, t_2)\,\mathrm{d}S(\tau)\,\mathrm{d}S(t_1)\,\mathrm{d}S(t_2),\\
J_5 &= \mathrm{E}_G \iiint \mathbf{1}_{[0,T]}(\tau_1)\mathbf{1}_{[0,x_0]}(t-\tau_1)\mathbf{1}_{[0,T]}(\tau_2)\\
&\qquad \times \mathbf{1}_{[0,x_0]}(\tau_2-\tau_1)\chi(\tau_1, t)\chi(\tau_1, \tau_2)\,\mathrm{d}S(\tau_1)\,\mathrm{d}S(\tau_2)\,\mathrm{d}S(t),\\
J_6 &= \mathrm{E}_G \iiint \mathbf{1}_{[0,T]}(\tau_1)\mathbf{1}_{[0,x_0]}(t-\tau_1)\mathbf{1}_{[0,T]}(\tau_2)\\
&\qquad \times \mathbf{1}_{[0,x_0]}(t-\tau_2)\chi(\tau_1, t)\chi(\tau_2, t)\,\mathrm{d}S(\tau_1)\,\mathrm{d}S(t)\,\mathrm{d}S(\tau_2),\\
J_7 &= \mathrm{E}_G \iiiint_{\Omega} \mathbf{1}_{[0,T]}(\tau_1)\mathbf{1}_{[0,x_0]}(t_1-\tau_1)\mathbf{1}_{[0,T]}(\tau_2)\mathbf{1}_{[0,x_0]}(t_2-\tau_2)\,\mathrm{d}S(\tau_1)\,\mathrm{d}S(t_1)\,\mathrm{d}S(\tau_2)\,\mathrm{d}S(t_2)\\
&\qquad - \left[\mathrm{E}_G \iint \mathbf{1}_{[0,T]}(\tau)\mathbf{1}_{[0,x_0]}(t-\tau)\chi(\tau, t)\,\mathrm{d}S(\tau)\,\mathrm{d}S(t)\right]^2,
\end{aligned}$$

and Ω is the set of all vectors $(\tau_1, t_1, \tau_2, t_2)$ with all distinct elements.

We compute all integrals J_k, $k=1,\dots,7$ separately. Although the computations are routine and straightforward, they are often tedious.

2^0. In view of (4.6), we have

$$J_1 = \mathrm{E}_G \iint \mathbf{1}_{[0,T]}\mathbf{1}_{[0,x_0]}(t-\tau)\chi(\tau,t)\,\mathrm{d}S(\tau)\,\mathrm{d}S(t) = 4\lambda^2 T x_0 + \lambda T G(x_0).$$

Note that $\mathbf{1}_{[0,x_0]}(t-\tau)\mathbf{1}_{[0,x_0]}(\tau-t)\chi(\tau,t)=0$; hence $J_2=0$.

Using (A.13) after straightforward calculation, we obtain

$$J_3 = 8\lambda^3 T x_0^2 + 4\lambda^2 x_0 G(x_0) T + 4\lambda^2 T \int_0^{x_0} G(x_0-u)\,\mathrm{d}u.$$

In order to compute J_4 we again use (A.13). In particular,

$$\begin{aligned}
&8\lambda^3 \iiint \mathbf{1}_{[0,T]}(\tau)\mathbf{1}_{[0,x_0]}(t_1-\tau)\mathbf{1}_{[0,T]}(t_1)\mathbf{1}_{[0,x_0]}(t_2-t_1)\,\mathrm{d}\tau\,\mathrm{d}t_1\,\mathrm{d}t_2\\
&\quad= 8\lambda^3 x_0 \iint \mathbf{1}_{[0,T]}(\tau)\mathbf{1}_{[0,x_0]}(t_1-\tau)\mathbf{1}_{[0,T]}(t_1)\,\mathrm{d}\tau\,\mathrm{d}t_1 = 8\lambda^3 x_0^2\biggl[T-\frac{1}{2}x_0\biggr],
\end{aligned}$$

and

$$\begin{aligned}
&2\lambda^2 \iiint \mathbf{1}_{[0,T]}(\tau)\mathbf{1}_{[0,x_0]}(t_1-\tau)\mathbf{1}_{[0,T]}(t_1)\mathbf{1}_{[0,x_0]}(t_2-t_1)\,\mathrm{d}t_1\,\mathrm{d}G(\tau-t_2)\,\mathrm{d}t_2\\
&\quad= 2\lambda^2 \iint \bigl[G(t_1-t_2)-G(t_1-t_2-x_0)\bigr]\mathbf{1}_{[x_0,T]}(t_1)\mathbf{1}_{[0,x_0]}(t_2-t_1)\,\mathrm{d}t_1\,\mathrm{d}t_2\\
&\qquad+ 2\lambda^2 \iint \bigl[G(t_1-t_2)-G(-t_2)\bigr]\mathbf{1}_{[0,x_0]}(t_1)\mathbf{1}_{[0,x_0]}(t_2-t_1)\,\mathrm{d}t_1\,\mathrm{d}t_2 = 0
\end{aligned}$$

because $G(0)=0$ and $x_0>0$. Furthermore,

$$\begin{aligned}
&2\lambda^2 \iiint \mathbf{1}_{[0,T]}(\tau)\mathbf{1}_{[0,x_0]}(t_1-\tau)\mathbf{1}_{[0,T]}(t_1)\mathbf{1}_{[0,x_0]}(t_2-t_1)\,\mathrm{d}t_1\,\mathrm{d}G(t_2-\tau)\,\mathrm{d}\tau\\
&\quad= 2\lambda^2 \iint \mathbf{1}_{[0,T]}(\tau)\mathbf{1}_{[0,x_0]}(t_1-\tau)\mathbf{1}_{[0,T]}(t_1)\bigl[G(t_1-\tau+x_0)-G(t_1-\tau)\bigr]\,\mathrm{d}t_1\,\mathrm{d}\tau\\
&\quad= 2\lambda^2 T \int_0^{x_0} \bigl[G(u+x_0)-G(u)\bigr]\biggl(1-\frac{u}{T}\biggr)\,\mathrm{d}u,\\
&2\lambda^2 \iiint \mathbf{1}_{[0,T]}(\tau)\mathbf{1}_{[0,x_0]}(t_1-\tau)\mathbf{1}_{[0,T]}(t_1)\mathbf{1}_{[0,x_0]}(t_2-t_1)\,\mathrm{d}\tau\,\mathrm{d}G(t_2-t_1)\,\mathrm{d}t_1\\
&\quad= 2\lambda^2 G(x_0) x_0\biggl[T-\frac{1}{2}x_0\biggr],\\
&2\lambda^2 \iiint \mathbf{1}_{[0,T]}(\tau)\mathbf{1}_{[0,x_0]}(t_1-\tau)\mathbf{1}_{[0,T]}(t_1)\mathbf{1}_{[0,x_0]}(t_2-t_1)\,\mathrm{d}\tau\,\mathrm{d}G(t_1-t_2)\,\mathrm{d}t_2 = 0,
\end{aligned}$$

$$2\lambda^2 \iiint \mathbf{1}_{[0,T]}(\tau)\mathbf{1}_{[0,x_0]}(t_1-\tau)\mathbf{1}_{[0,T]}(t_1)\mathbf{1}_{[0,x_0]}(t_2-t_1)\,\mathrm{d}t_2\,\mathrm{d}G(\tau-t_1)\,\mathrm{d}t_1 = 0,$$

$$2\lambda^2 \iiint \mathbf{1}_{[0,T]}(\tau)\mathbf{1}_{[0,x_0]}(t_1-\tau)\mathbf{1}_{[0,T]}(t_1)\mathbf{1}_{[0,x_0]}(t_2-t_1)\,\mathrm{d}t_2\,\mathrm{d}G(t_1-\tau)\,\mathrm{d}\tau$$
$$= 2\lambda^2 x_0 T \int_0^{x_0} \left(1-\frac{u}{T}\right)\mathrm{d}G(u).$$

Combining these expressions we obtain

$$J_4 = 8\lambda^3 x_0^2\left[T-\frac{1}{2}x_0\right] + 2\lambda^2 G(x_0)x_0\left[T-\frac{1}{2}x_0\right]$$
$$+ 2\lambda^2 x_0 T\int_0^{x_0}\left(1-\frac{u}{T}\right)\mathrm{d}G(u) + 2\lambda^2 T\int_0^{x_0}\big[G(u+x_0)-G(u)\big]\left(1-\frac{u}{T}\right)\mathrm{d}u.$$

Now we note that the expression for J_5 up to a change in notation of integration variables coincides with that for J_4; hence $J_5 = J_4$.

We proceed with computation of J_6:

$$8\lambda^3 \iiint \mathbf{1}_{[0,T]}(\tau_1)\mathbf{1}_{[0,x_0]}(t-\tau_1)\mathbf{1}_{[0,T]}(\tau_2)\mathbf{1}_{[0,x_0]}(t-\tau_2)\,\mathrm{d}\tau_1\,\mathrm{d}t\,\mathrm{d}\tau_2$$
$$= 8\lambda^3 \iiint \mathbf{1}_{[0,T]}(\tau_1)\mathbf{1}_{[\tau_1,\tau_1+x_0]}(t)\mathbf{1}_{[\tau_2,\tau_2+x_0]}(t)\mathbf{1}_{[0,T]}(\tau_2)\mathbf{1}_{[0,x_0]}(\tau_2-\tau_1)\,\mathrm{d}\tau_1\,\mathrm{d}t\,\mathrm{d}\tau_2$$
$$+ 8\lambda^3 \iiint \mathbf{1}_{[0,T]}(\tau_1)\mathbf{1}_{[\tau_1,\tau_1+x_0]}(t)\mathbf{1}_{[\tau_2,\tau_2+x_0]}(t)\mathbf{1}_{[0,T]}(\tau_2)\mathbf{1}_{[0,x_0]}(\tau_1-\tau_2)\,\mathrm{d}\tau_1\,\mathrm{d}t\,\mathrm{d}\tau_2$$
$$= 8\lambda^3 \iint \mathbf{1}_{[0,T]}(\tau_1)[x_0+\tau_1-\tau_2]\mathbf{1}_{[0,T]}(\tau_2)\mathbf{1}_{[0,x_0]}(\tau_2-\tau_1)\,\mathrm{d}\tau_1\,\mathrm{d}\tau_2$$
$$+ 8\lambda^3 \iint \mathbf{1}_{[0,T]}(\tau_1)[x_0+\tau_2-\tau_1]\mathbf{1}_{[0,T]}(\tau_2)\mathbf{1}_{[0,x_0]}(\tau_1-\tau_2)\,\mathrm{d}\tau_1\,\mathrm{d}\tau_2$$
$$= 16\lambda^3 \iint [x_0-u]\mathbf{1}_{[0,T]}(\tau)\mathbf{1}_{[-u,T-u]}(\tau)\mathbf{1}_{[0,x_0]}(u)\,\mathrm{d}\tau\,\mathrm{d}u = 8\lambda^3 x_0^2\left[T-\frac{1}{3}x_0\right].$$

Furthermore,

$$2\lambda^2 \iiint \mathbf{1}_{[0,T]}(\tau_1)\mathbf{1}_{[0,x_0]}(t-\tau_1)\mathbf{1}_{[0,T]}(\tau_2)\mathbf{1}_{[0,x_0]}(t-\tau_2)\,\mathrm{d}\tau_1\big[\mathrm{d}G(t-\tau_2)\,\mathrm{d}\tau_2 + \mathrm{d}G(\tau_2-t)\,\mathrm{d}t\big]$$
$$= 2\lambda^2 T\int_0^{x_0}\big[G(x_0-u)+G(x_0)-G(u)\big]\left(1-\frac{u}{T}\right)\mathrm{d}u,$$

$$2\lambda^2 \iiint \mathbf{1}_{[0,T]}(\tau_1)\mathbf{1}_{[0,x_0]}(t-\tau_1)\mathbf{1}_{[0,T]}(\tau_2)\mathbf{1}_{[0,x_0]}(t-\tau_2)\,\mathrm{d}\tau_2\big[\mathrm{d}G(t-\tau_1)\,\mathrm{d}\tau_1 + \mathrm{d}G(\tau_1-t)\,\mathrm{d}t\big]$$
$$= 2\lambda^2 T\int_0^{x_0}\big[G(x_0-u)+G(x_0)-G(u)\big]\left(1-\frac{u}{T}\right)\mathrm{d}u,$$

and

$$
\begin{aligned}
&2\lambda^2 \iiint \mathbf{1}_{[0,T]}(\tau_1)\mathbf{1}_{[0,x_0]}(t-\tau_1)\mathbf{1}_{[0,T]}(\tau_2) \\
&\quad\times \mathbf{1}_{[0,x_0]}(t-\tau_2)\,\mathrm{d}t\big[\mathrm{d}G(\tau_2-\tau_1)\,\mathrm{d}\tau_1+\mathrm{d}G(\tau_1-\tau_2)\,\mathrm{d}\tau_2\big] \\
&= 2\lambda^2 \iint [x_0+\tau_1-\tau_2]\mathbf{1}_{[0,T]}(\tau_1)\mathbf{1}_{[0,T]}(\tau_2) \\
&\quad\times \mathbf{1}_{[0,x_0]}(\tau_2-\tau_1)\big[\mathrm{d}G(\tau_2-\tau_1)\,\mathrm{d}\tau_1+\mathrm{d}G(\tau_1-\tau_2)\,\mathrm{d}\tau_2\big] \\
&\quad+ 2\lambda^2 \iint [x_0+\tau_2-\tau_1]\mathbf{1}_{[0,T]}(\tau_1)\mathbf{1}_{[0,T]}(\tau_2) \\
&\quad\times \mathbf{1}_{[0,x_0]}(\tau_1-\tau_2)\big[\mathrm{d}G(\tau_2-\tau_1)\,\mathrm{d}\tau_1+\mathrm{d}G(\tau_1-\tau_2)\,\mathrm{d}\tau_2\big] \\
&= 4\lambda^2 T\int_0^{x_0}\left(1-\frac{u}{T}\right)(x_0-u)\,\mathrm{d}G(u).
\end{aligned}
$$

Therefore,

$$
\begin{aligned}
J_6 &= 8\lambda^3 x_0^2\left[T-\frac{1}{3}x_0\right]+4\lambda^2 T\int_0^{x_0}\big[G(x_0-u)+G(x_0)-G(u)\big]\left(1-\frac{u}{T}\right)\mathrm{d}u \\
&\quad+4\lambda^2 T\int_0^{x_0}\left(1-\frac{u}{T}\right)(x_0-u)\,\mathrm{d}G(u).
\end{aligned}
$$

Then, combining the above expressions for J_i, $i=1,\dots,6$, after straightforward bounding we obtain

$$
\frac{1}{\lambda^2 T}\sum_{i=1}^{6} J_i \le 32\lambda x_0^2+36x_0+\frac{1}{\lambda}G(x_0). \tag{B.8}
$$

3^0. It remains to compute J_7. If $\tilde{Q}(\mathrm{d}x,\mathrm{d}y) := \frac{1}{2}[\mathrm{d}G(x-y)\,\mathrm{d}y+\mathrm{d}G(y-x)\,\mathrm{d}x]$ then by Lemma 2(iii)

$$
\begin{aligned}
J_7 = \iiiint \varphi_*(\tau_1,t_1)\varphi_*(\tau_2,t_2)\big\{&\mathrm{E}_G\big[\mathrm{d}S(\tau_1)\,\mathrm{d}S(\tau_2)\,\mathrm{d}S(t_1)\,\mathrm{d}S(t_2)\big] \\
&-\mathrm{E}_G\big[\mathrm{d}S(\tau_1)\,\mathrm{d}S(t_1)\big]\mathrm{E}_G\big[\mathrm{d}S(\tau_2)\,\mathrm{d}S(t_2)\big]\big\} =: \sum_{i=1}^{6} K_i,
\end{aligned}
$$

where

$$
\begin{aligned}
K_1 &:= 8\lambda^3 \iiiint \mathbf{1}_{[0,T]}(\tau_1)\mathbf{1}_{[0,x_0]}(t_1-\tau_1)\mathbf{1}_{[0,T]}(\tau_2)\mathbf{1}_{[0,x_0]}(t_2-\tau_2)\,\tilde{Q}(\mathrm{d}\tau_1,\mathrm{d}t_2)\,\mathrm{d}\tau_2\,\mathrm{d}t_1, \\
K_2 &:= 8\lambda^3 \iiiint \mathbf{1}_{[0,T]}(\tau_1)\mathbf{1}_{[0,x_0]}(t_1-\tau_1)\mathbf{1}_{[0,T]}(\tau_2)\mathbf{1}_{[0,x_0]}(t_2-\tau_2)\,\tilde{Q}(\mathrm{d}\tau_1,\mathrm{d}\tau_2)\,\mathrm{d}t_2\,\mathrm{d}t_1,
\end{aligned}
$$

$$K_3 := 8\lambda^3 \iiiint \mathbf{1}_{[0,T]}(\tau_1)\mathbf{1}_{[0,x_0]}(t_1-\tau_1)\mathbf{1}_{[0,T]}(\tau_2)\mathbf{1}_{[0,x_0]}(t_2-\tau_2)\tilde{Q}(\mathrm{d}\tau_2,\mathrm{d}t_1)\,\mathrm{d}\tau_1\,\mathrm{d}t_2,$$

$$K_4 := 8\lambda^3 \iiiint \mathbf{1}_{[0,T]}(\tau_1)\mathbf{1}_{[0,x_0]}(t_1-\tau_1)\mathbf{1}_{[0,T]}(\tau_2)\mathbf{1}_{[0,x_0]}(t_2-\tau_2)\tilde{Q}(\mathrm{d}t_1,\mathrm{d}t_2)\,\mathrm{d}\tau_1\,\mathrm{d}\tau_2,$$

$$K_5 := 4\lambda^2 \iiiint \mathbf{1}_{[0,T]}(\tau_1)\mathbf{1}_{[0,x_0]}(t_1-\tau_1)\mathbf{1}_{[0,T]}(\tau_2)\mathbf{1}_{[0,x_0]}(t_2-\tau_2)\tilde{Q}(\mathrm{d}\tau_1,\mathrm{d}\tau_2)\tilde{Q}(\mathrm{d}t_1,\mathrm{d}t_2),$$

$$K_6 := 4\lambda^2 \iiiint \mathbf{1}_{[0,T]}(\tau_1)\mathbf{1}_{[0,x_0]}(t_1-\tau_1)\mathbf{1}_{[0,T]}(\tau_2)\mathbf{1}_{[0,x_0]}(t_2-\tau_2)\tilde{Q}(\mathrm{d}\tau_1,\mathrm{d}t_2)\tilde{Q}(\mathrm{d}\tau_2,\mathrm{d}t_1).$$

First, we note that $K_1 = K_3$; then standard calculations yield

$$\begin{aligned}
K_1 = K_3 &= 4\lambda^3 x_0 T \int_{-T}^{T}\big[G(x_0+u)-G(u)\big]\left(1-\frac{|u|}{T}\right)\mathrm{d}u + 4\lambda^3 x_0 \int_0^T \int_0^{x_0} G(v-u)\,\mathrm{d}u\,\mathrm{d}v \\
&\le 12\lambda^3 x_0^2 T, \\
K_2 &= 8\lambda^3 x_0^2 T \int_0^T \left(1-\frac{u}{T}\right)\mathrm{d}G(u) \le 8\lambda^3 x_0^2 T.
\end{aligned}$$

The computation of K_4 is more involved. Although exact expression for K_4 can be derived, we will compute exactly only the main term which grows with T; for other terms we provide upper bounds. The computation is based on the following formula

$$\begin{aligned}
&\iint \mathbf{1}_{[0,T]}(\tau_1)\mathbf{1}_{[0,x_0]}(t_1-\tau_1)\mathbf{1}_{[0,T]}(\tau_2)\mathbf{1}_{[0,x_0]}(t_2-\tau_2)\,\mathrm{d}\tau_1\,\mathrm{d}\tau_2 \\
&\quad = \big[t_1\mathbf{1}_{[0,x_0]}(t_1) + x_0\mathbf{1}_{[x_0,T]}(t_1) + (T-t_1+x_0)\mathbf{1}_{[T,T+x_0]}(t_1)\big] \\
&\qquad \times \big[t_2\mathbf{1}_{[0,x_0]}(t_2) + x_0\mathbf{1}_{[x_0,T]}(t_2) + (T-t_2+x_0)\mathbf{1}_{[T,T+x_0]}(t_2)\big] \\
&\quad =: \big[L_1(t_1)+L_2(t_1)+L_3(t_1)\big]\big[L_4(t_2)+L_5(t_2)+L_6(t_2)\big].
\end{aligned}$$

Then

$$\begin{aligned}
&8\lambda^3 \iint L_1(t_1)L_4(t_2)\tilde{Q}(\mathrm{d}t_1,\mathrm{d}t_2) \\
&\quad = 4\lambda^3 \int_0^{x_0}\int_0^{x_0} t_1 t_2\big[\mathrm{d}G(t_1-t_2)\,\mathrm{d}t_2 + \mathrm{d}G(t_2-t_1)\,\mathrm{d}t_1\big] \\
&\quad = 8\lambda^3 \int_0^{x_0} \frac{1}{3}(x_0-u)^2\left(x_0+\frac{1}{2}u\right)\mathrm{d}G(u) \le 4\lambda^3 x_0^3 G(x_0), \\
&8\lambda^3 \iint L_1(t_1)L_5(t_2)\tilde{Q}(\mathrm{d}t_1,\mathrm{d}t_2) \\
&\quad = 4\lambda^3 x_0 \int_0^{x_0}\int_{x_0}^{T} t_1\,\mathrm{d}G(t_2-t_1)\,\mathrm{d}t_1
\end{aligned}$$

$$= 4\lambda^3 x_0 \left\{ \int_0^{x_0} \left(x_0 - \frac{1}{2} u \right) u \, \mathrm{d}G(u) + \frac{1}{2} x_0^2 \big[G(T - x_0) - G(x_0) \big] + \frac{1}{2} \int_{T-x_0}^{T} (T - u)^2 \, \mathrm{d}G(u) \right\}$$
$$\le 8\lambda^3 x_0^3 G(T),$$
$$8\lambda^3 \iint L_1(t_1) L_6(t_2) \tilde{Q}(\mathrm{d}t_1, \mathrm{d}t_2)$$
$$= -4\lambda^3 \int_0^{x_0} \int_0^{x_0} t_1 u \, \mathrm{d}G(T + x_0 - u - t_1) \, \mathrm{d}t_1 \le 4\lambda^3 x_0^3 G(T + x_0).$$

The following integrals are evaluated similarly:

$$8\lambda^3 \iint L_3(t_1) L_4(t_2) \tilde{Q}(\mathrm{d}t_1, \mathrm{d}t_2) = 8\lambda^3 \iint L_1(t_1) L_6(t_2) \tilde{Q}(\mathrm{d}t_1, \mathrm{d}t_2) \le 4\lambda^3 x_0^3 G(T + x_0),$$
$$8\lambda^3 \iint L_3(t_1) L_6(t_2) \tilde{Q}(\mathrm{d}t_1, \mathrm{d}t_2) = 8\lambda^3 \iint L_1(t_1) L_4(t_2) \tilde{Q}(\mathrm{d}t_1, \mathrm{d}t_2) \le 4\lambda^3 x_0^3 G(x_0),$$
$$8\lambda^3 \iint L_2(t_1) L_4(t_2) \tilde{Q}(\mathrm{d}t_1, \mathrm{d}t_2) = 8\lambda^3 \iint L_1(t_1) L_5(t_2) \tilde{Q}(\mathrm{d}t_1, \mathrm{d}t_2) \le 8\lambda^3 x_0^3 G(T),$$
$$8\lambda^3 \iint L_2(t_1) L_6(t_2) \tilde{Q}(\mathrm{d}t_1, \mathrm{d}t_2) = 8\lambda^3 \iint L_3(t_1) L_5(t_2) \tilde{Q}(\mathrm{d}t_1, \mathrm{d}t_2)$$
$$= -4\lambda^3 x_0 \int_{x_0}^{T} \int_0^{x_0} u \, \mathrm{d}G(T + x_0 - u - t_1) \, \mathrm{d}t_1$$
$$\le 2\lambda^3 x_0^3 G(T),$$

and finally,

$$8\lambda^3 \iint L_2(t_2) L_5(t_2) \tilde{Q}(\mathrm{d}t_1, \mathrm{d}t_2) = 4\lambda^3 x_0^2 \int_{x_0}^{T} \int_{x_0}^{T} \big[\mathrm{d}G(t_1 - t_2) \, \mathrm{d}t_2 + \mathrm{d}G(t_2 - t_1) \, \mathrm{d}t_2 \big]$$
$$= 8\lambda^3 x_0^2 \int_0^{T-x_0} G(u) \, \mathrm{d}u.$$

Combining these expressions, we see that

$$K_4 \le 8\lambda^3 x_0^2 \int_0^{T-x_0} G(u) \, \mathrm{d}u + 36\lambda^3 x_0^3 \le 8\lambda^3 x_0^2 T + 36\lambda^3 x_0^3.$$

Next, we proceed with computation of K_5.

$$K_5 = \lambda^2 \iiiint \mathbf{1}_{[0,T]}(\tau_1) \mathbf{1}_{[0,x_0]}(t_1 - \tau_1) \mathbf{1}_{[0,T]}(\tau_2)$$
$$\times \mathbf{1}_{[0,x_0]}(t_2 - \tau_2) \, \mathrm{d}G(\tau_1 - \tau_2) \, \mathrm{d}\tau_2 \, \mathrm{d}G(t_1 - t_2) \, \mathrm{d}t_2$$
$$+ \lambda^2 \iiiint \mathbf{1}_{[0,T]}(\tau_1) \mathbf{1}_{[0,x_0]}(t_1 - \tau_1) \mathbf{1}_{[0,T]}(\tau_2)$$

$$\times \mathbf{1}_{[0,x_0]}(t_2-\tau_2)\,\mathrm{d}G(\tau_1-\tau_2)\,\mathrm{d}\tau_2\,\mathrm{d}G(t_2-t_1)\,\mathrm{d}t_1$$
$$+\lambda^2\iiiint \mathbf{1}_{[0,T]}(\tau_1)\mathbf{1}_{[0,x_0]}(t_1-\tau_1)\mathbf{1}_{[0,T]}(\tau_2)$$
$$\times \mathbf{1}_{[0,x_0]}(t_2-\tau_2)\,\mathrm{d}G(\tau_2-\tau_1)\,\mathrm{d}\tau_1\,\mathrm{d}G(t_1-t_2)\,\mathrm{d}t_2$$
$$+\lambda^2\iiiint \mathbf{1}_{[0,T]}(\tau_1)\mathbf{1}_{[0,x_0]}(t_1-\tau_1)\mathbf{1}_{[0,T]}(\tau_2)$$
$$\times \mathbf{1}_{[0,x_0]}(t_2-\tau_2)\,\mathrm{d}G(\tau_2-\tau_1)\,\mathrm{d}\tau_1\,\mathrm{d}G(t_2-t_1)\,\mathrm{d}t_1$$
$$=: \lambda^2\big(K_5^{(1)}+K_5^{(2)}+K_5^{(3)}+K_5^{(4)}\big).$$

Integrating first with respect to t_1, we have

$$K_5^{(1)}=\iiint \mathbf{1}_{[0,T]}(\tau_1)\big[G(x_0+\tau_1-t_2)-G(\tau_1-t_2)\big]$$
$$\times \mathbf{1}_{[0,T]}(\tau_2)\mathbf{1}_{[0,x_0]}(t_2-\tau_2)\,\mathrm{d}G(\tau_1-\tau_2)\,\mathrm{d}\tau_2\,\mathrm{d}t_2.$$

Define $F(t):=\int_0^{x_0}[G(x_0+t-u)-G(t-u)]\,\mathrm{d}u$; then

$$K_5^{(1)}=\iint \mathbf{1}_{[0,T]}(\tau_1)\mathbf{1}_{[0,T]}(\tau_2)F(\tau_1-\tau_2)\,\mathrm{d}G(\tau_1-\tau_2)\,\mathrm{d}\tau_2=\int_0^T(T-v)F(v)\,\mathrm{d}G(v).$$

The same argument yields

$$K_5^{(2)}=\iint \mathbf{1}_{[0,T]}(\tau_1)\mathbf{1}_{[0,T]}(\tau_2)F(\tau_2-\tau_1)\,\mathrm{d}G(\tau_1-\tau_2)\,\mathrm{d}\tau_2=\int_0^T(T-v)F(-v)\,\mathrm{d}G(v).$$

Moreover, by symmetry $K_5^{(3)}=K_5^{(2)}$ and $K_5^{(4)}=K_5^{(1)}$. Therefore,

$$K_5=2\lambda^2T\int_0^T\left(1-\frac{v}{T}\right)\big[F(v)+F(-v)\big]\,\mathrm{d}G(v)\le 4\lambda^2x_0T.$$

It remains to compute K_6:

$$K_6=\lambda^2\iiiint \mathbf{1}_{[0,T]}(\tau_1)\mathbf{1}_{[0,x_0]}(t_1-\tau_1)\mathbf{1}_{[0,T]}(\tau_2)$$
$$\times \mathbf{1}_{[0,x_0]}(t_2-\tau_2)\,\mathrm{d}G(\tau_1-t_2)\,\mathrm{d}t_2\,\mathrm{d}G(\tau_2-t_1)\,\mathrm{d}t_1$$
$$+\lambda^2\iiiint \mathbf{1}_{[0,T]}(\tau_1)\mathbf{1}_{[0,x_0]}(t_1-\tau_1)\mathbf{1}_{[0,T]}(\tau_2)$$
$$\times \mathbf{1}_{[0,x_0]}(t_2-\tau_2)\,\mathrm{d}G(\tau_1-t_2)\,\mathrm{d}t_2\,\mathrm{d}G(t_1-\tau_2)\,\mathrm{d}\tau_2$$
$$+\lambda^2\iiiint \mathbf{1}_{[0,T]}(\tau_1)\mathbf{1}_{[0,x_0]}(t_1-\tau_1)\mathbf{1}_{[0,T]}(\tau_2)$$
$$\times \mathbf{1}_{[0,x_0]}(t_2-\tau_2)\,\mathrm{d}G(t_2-\tau_1)\,\mathrm{d}\tau_1\,\mathrm{d}G(\tau_2-t_1)\,\mathrm{d}t_1$$

$$
\begin{aligned}
&+\lambda^2 \iiiint \mathbf{1}_{[0,T]}(\tau_1)\mathbf{1}_{[0,x_0]}(t_1-\tau_1)\mathbf{1}_{[0,T]}(\tau_2) \\
&\quad\times \mathbf{1}_{[0,x_0]}(t_2-\tau_2)\,\mathrm{d}G(t_2-\tau_1)\,\mathrm{d}\tau_1\,\mathrm{d}G(t_1-\tau_2)\,\mathrm{d}\tau_2 \\
=: {}&\lambda^2\big(K_6^{(1)}+K_6^{(2)}+K_6^{(3)}+K_6^{(4)}\big).
\end{aligned}
$$

First, we note that $K_6^{(1)}=0$. Indeed, the integrand does not vanish only if the following inequalities hold simultaneously $0\le t_1-\tau_1\le x_0$, $\tau_1\ge t_2$, $0\le t_2-\tau_2\le x_0$, $\tau_2\ge t_1$, which is possible only if all four variables coincide. Furthermore,

$$
\begin{aligned}
K_6^{(2)} &= \iiint \big[G(x_0+\tau_1-\tau_2)-G(\tau_1-\tau_2)\big]\mathbf{1}_{[0,T]}(\tau_1)\mathbf{1}_{[0,T]}(\tau_2) \\
&\quad\times \mathbf{1}_{[0,x_0]}(t_2-\tau_2)\,\mathrm{d}G(\tau_1-t_2)\,\mathrm{d}t_2\,\mathrm{d}\tau_2 \\
&= T\int_0^{x_0}\bigg[\int_0^T\Big(1-\frac{u}{T}\Big)\big[G(x_0+u)-G(u)\big]\,\mathrm{d}G(u-v) \\
&\quad+\int_T^{T+x_0}\big[G(x_0+u)-G(u)\big]\,\mathrm{d}G(u-v)\bigg]\,\mathrm{d}v.
\end{aligned}
$$

By symmetry $K_6^{(3)}=K_6^{(2)}$, and

$$
\begin{aligned}
K_6^{(4)} &= \iint \mathbf{1}_{[0,T]}(\tau_1)\big[G(\tau_1+x_0-\tau_2)-G(\tau_1-\tau_2)\big] \\
&\quad\times \mathbf{1}_{[0,T]}(\tau_2)\big[G(\tau_2+x_0-\tau_1)-G(\tau_2-\tau_1)\big]\,\mathrm{d}\tau_1\,\mathrm{d}\tau_2 \\
&= T\int_{-T}^{T}\Big(1-\frac{|u|}{T}\Big)\big[G(x_0+u)-G(u)\big]\big[G(x_0-u)-G(-u)\big]\,\mathrm{d}u.
\end{aligned}
$$

Thus, we obtain

$$
\begin{aligned}
K_6 &\le \lambda^2 T\int_{-T}^{T}\Big(1-\frac{|u|}{T}\Big)\big[G(x_0+u)-G(u)\big]\big[G(x_0-u)-G(-u)\big]\,\mathrm{d}u \\
&\quad+2\lambda^2 T\int_0^{T+x_0}\bigg\{\int_0^{x_0}\big[G(x_0+u+v)-G(u+v)\big]\,\mathrm{d}v\bigg\}\,\mathrm{d}G(u)\le 4\lambda^2 x_0 T.
\end{aligned}
$$

Combining expressions for K_i, $i=1,\dots,6$ we come to

$$
\frac{J_7}{\lambda^2 T}=\frac{1}{\lambda^2 T}\sum_{i=1}^{6}K_i\le 40\lambda x_0^2+8x_0.
$$

This relationship together with (B.8) lead to the announced result.

B.3. Proof of Theorem 4

We have

$$\mathrm{E}_G|\hat{\alpha}-\alpha|^2 \le 2\mathrm{E}_G \int_0^b \int_0^b \big[\hat{G}(x)-G(x)\big]\big[\hat{G}(y)-G(y)\big]\,\mathrm{d}x\,\mathrm{d}y + 2\left\{\int_b^\infty \big[1-G(x)\big]\,\mathrm{d}x\right\}^2.$$

By the Cauchy–Schwarz inequality and Theorem 2

$$\mathrm{E}_G\big[\hat{G}(x)-G(x)\big]\big[\hat{G}(y)-G(y)\big] \le \frac{c_1}{T}\left[\lambda xy + (\sqrt{x}+\sqrt{y})\left(\sqrt{xy}+\frac{1}{\sqrt{\lambda}}\right)+\frac{1}{\lambda}\right],$$

where c_1 is an absolute constant. Therefore,

$$\sup_{G\in\mathcal{M}_p} \mathrm{E}_G|\hat{\alpha}-\alpha|^2 \le \frac{c_2}{T}\big[\lambda b^4 + b^{7/2} + \lambda^{-1/2}b^{3/2} + \lambda^{-1}b^2\big] + c_3(A/p)^2 b^{-2p+2}.$$

The lower bound on T in the premise of the theorem implies that with $b=b_*$ the first term on the right-hand side of the previous display formula is dominating. The announced result follows by substitution of $b=b_*$.

Acknowledgements

The research was supported by the grants BSF 2010466 and ISF 361/15. The author is grateful to Rui Castro for useful discussions and valuable remarks.

References

[1] Beneš, V.E. (1957). Fluctuations of telephone traffic. *Bell Syst. Tech. J.* **36** 965–973.
[2] Bingham, N.H. and Dunham, B. (1997). Estimating diffusion coefficients from count data: Einstein–Smoluchowski theory revisited. *Ann. Inst. Statist. Math.* **49** 667–679. MR1621845
[3] Bingham, N.H. and Pitts, S.M. (1999). Non-parametric estimation for the $M/G/\infty$ queue. *Ann. Inst. Statist. Math.* **51** 71–97. MR1704647
[4] Blanghaps, N., Nov, Y. and Weiss, G. (2013). Sojourn time estimation in an $M/G/\infty$ queue with partial information. *J. Appl. Probab.* **50** 1044–1056. MR3161372
[5] Brillinger, D.R. (1974). Cross-spectral analysis of processes with stationary increments including the stationary $G/G/\infty$ queue. *Ann. Probab.* **2** 815–827. MR0359221
[6] Brillinger, D.R. (1975). *Statistical Inference for Stationary Point Processes* 55–99. New York: Academic Press. MR0381201
[7] Brillinger, D.R. (1975). The identification of point process systems. *Ann. Probab.* **3** 909–929. MR0394865
[8] Brillinger, D.R. (1976). Estimation of the second-order intensities of a bivariate stationary point process. *J. Roy. Statist. Soc. Ser. B* **38** 60–66. MR0400595
[9] Brown, M. (1970). An $M/G/\infty$ estimation problem. *Ann. Math. Stat.* **41** 651–654.

[10] Cox, D.R. and Lewis, P.A.W. (1972). Multivariate point processes. In *Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability*, *Vol*. *III*: *Probability Theory* 401–448. Berkeley: Univ. California Press. MR0413254
[11] Daley, D.J. and Vere-Jones, D. (2003). *An Introduction to the Theory of Point Processes*. *Vol I*: *Elementary Theory and Methods*, 2nd ed. *Probability and Its Applications* (*New York*). New York: Springer. MR1950431
[12] Doob, J.L. (1953). *Stochastic Processes*. New York: Wiley.
[13] Goldenshluger, A. (2016). Nonparametric estimation of the service time distribution in the $M/G/\infty$ queue. *Adv. in Appl*. *Probab*. **48** 1117–1138.
[14] Hall, P. and Park, J. (2004). Nonparametric inference about service time distribution from indirect observations. *J*. *R*. *Statist*. *Soc*. *B* **66** 861–875.
[15] Kingman, J.F.C. (1993). *Poisson Processes*. Oxford: Clarendon Press.
[16] Lindley, D.V. (1956). The estimation of velocity distributions from counts. In *Proceedings of the International Congress of Mathematicians*, 1954, *Amsterdam*, *Vol*. *III* 427–444. North-Holland: Erven P. Noordhoff N.V. MR0087309
[17] Mandjes, M. and Zuraniewski, P. (2011). $M/G/\infty$ transience, and its applications to overload detection. *Perform*. *Eval*. **68** 507–527.
[18] Milne, R.K. (1970). Identifiability for random translations of Poisson processes. *Z*. *Wahrsch*. *Verw*. *Gebiete* **15** 195–201.
[19] Milne, R.K. and Westcott, M. (1972). Further results for Gauss–Poisson processes. *Adv. in Appl*. *Probab*. **4** 151–176. MR0314111
[20] Mori, T. (1975). Ergodicity and identifiability for random translations of stationary point processes. *J*. *Appl*. *Probab*. **12** 734–743. MR0402916
[21] Newman, D.S. (1970). A new family of point processes which are characterized by their second moment properties. *J*. *Appl*. *Probab*. **7** 338–358.
[22] Petty, K.F., Bickel, P., Ostland, M., Rice, J., Schoenberg, F. and Ritov, Y. (1998). Accurate estimation of travel times from single–loop detectors. *Transp*. *Res*., *Part A Policy Pract*. **32** 1–17.
[23] Pickands, J. III and Stine, R.A. (1997). Estimation for an $M/G/\infty$ queue with incomplete information. *Biometrika* **84** 295–308. MR1467048
[24] Rothschild, L. (1953). A new method for measuring the activity of spermatozoa. *J*. *Exp*. *Biol*. **30** 178–199.
[25] Schweer, S. and Wichelhaus, C. (2015). Nonparametric estimation of the service time distribution in the discrete-time $GI/G/\infty$ queue with partial information. *Stochastic Process*. *Appl*. **125** 233–253. MR3274698

Received June 2016 and revised December 2016