AALBORG UNIVERSITY

Nonparametric estimation of the stationary $M/G/1$ workload distribution function

by

Martin Bøgsted Hansen

R-2005-26 August 2005

NONPARAMETRIC ESTIMATION OF THE STATIONARY $M/G/1$ WORKLOAD DISTRIBUTION FUNCTION

Martin B. Hansen

Department of Mathematical Sciences
Aalborg University
Fredrik Bajers Vej 7G
DK-9220 Aalborg Ø, Denmark

ABSTRACT

In this paper it is demonstrated how a nonparametric estimator of the stationary workload distribution function of the $M/G/1$-queue can be obtained by systematic sampling the workload process. Weak convergence results and bootstrap methods for empirical distribution functions for stationary associated sequences are used to derive asymptotic results and bootstrap methods for inference about the workload distribution function. The potential of the method is illustrated by a simulation study of the $M/D/1$ model.

1 INTRODUCTION

Sampling (or probing) the workload in queuing systems is a standard tool for performance evaluation. This is e.g. used in call center evaluations, where the distribution of the waiting time to service is estimated by the empirical distribution function of the call answer time for repeated phone calls. Another application is when a call admission controller in an ATM network decides, whether there are sufficient resources to allow a new connection to be established, based on information obtained by sampling the workload at neighboring nodes. Reliable estimation of the cumulative distribution function is an important subject as a variety of characteristics can be estimated by functionals of the empirical cumulative distribution function (ecdf).

Throughout the paper we consider an $M/G/1$-queue, i.e. the inter-arrival times are independently and exponentially distributed with mean λ, the service times are independently and generally distributed with mean f_1, there is 1 server, and infinite waiting room.

For later purposes let $\{(T_n, S_n), n \geq 1\}$, denote the sequence of arrival and service times of the customers. Let S be a generic random variable with the same distribution as S_1.

Let the workload in the system at time t be denoted by V_t, i.e. V_t is the sum of the residual service times of the customer being presently served and the customers awaiting service. By convention, a workload process $\{V_t, t \in \mathbb{R}\}$ will be taken right-continuous with left-hand limits. For the $M/G/1$ queue, the evolution of V_t between two arrivals is described by Lindley's equation

$$V_t = (V_{T_n-} + S_n - (t - T_n)) \wedge 0, \qquad (1)$$

where $t \in [T_n, T_{n+1})$. In general a workload process V_t is defined only for $t \geq 0$. However, by Loynes' Theorem (Baccelli and Brémaud 2003, Theorem 2.1.1) it is possible to prove that under the stability condition $\rho = \lambda f_1 < 1$ there exists a unique stationary workload process $\{V_t, t \in \mathbb{R}\}$ satisfying (1). We will use, V_0 as a generic random variable with the stationary distribution.

In what follows a cumulative distribution function (cdf) is denoted by a capital letter, A, say. The k'th moment by a_k, the stationary excess distribution by $A_e(x) = a_1^{-1} \int_0^x (1 - A(y)) dy$, $x \geq 0$, the complementary cdf by $\bar{A}(x) = 1 - A(x)$, $x \geq 0$, and its Fourier transform by $\widehat{A}(t) = \int_{-\infty}^{\infty} A(x) \exp(-itx) dx$, $t \in \mathbb{R}$.

If we let F denote the cdf of the service time distribution function and G the cdf of the stationary workload distribution function, one can prove, under the stability condition $\rho < 1$ that Pollaczeck-Khintchine's formula holds (Asmussen 2003, Theorem X.5.2)

$$G(x) = (1 - \rho) \sum_{k=0}^{\infty} \rho^k F_e^{\star k}(x), \quad 0 \leq x < \infty, \quad (2)$$

where $\star k$ denotes k-fold convolution.

We now assume that it is possible to test the performance of the queue by sampling the workload, without loss of generality, at every positive integer time point, as other sampling intervals can be obtained by proper rescaling. This process is denoted by $\{V_i, i \geq 1\}$. The main objective of this paper is to infer G from the sampled workloads. We suggest the ecdf as an estimator for G

$$G_n(x) = n^{-1} \sum_{i=1}^{n} I(V_i \leq x), \quad 0 \leq x < \infty. \quad (3)$$

It's n'th *empirical process* counterpart is defined as

$$\beta_n(x) = n^{1/2}(G_n(x) - G(x)), \quad 0 \le x < \infty. \quad (4)$$

In the following we will provide sufficient conditions for the empirical process to converge weakly to a Gaussian process. This information will be used to make statistical inference about the workload distribution function.

The paper is organized as follows. In Section 2 the covariance structure of the workload process is studied. The weak convergence of the empirical process of subsampled $M/G/1$ workloads and the bootstrap method are given in Section 3. Section 4 discusses algorithmic aspects and gives an illustrative numerical example. Section 5 contains some remarks on possible extensions and other aspects of the results. All proofs are carried out in Section 6.

2 Dependency structure of sampled workload

In general, when one establish weak convergence of empirical processes, to appropriate Gaussian processes, a handle on the dependence structure of the process under study is needed.

One possibility is to utilize the regenerative structure of the workload process and proceed with an analysis along the lines of Datta and McCormick (1993). This will require a detailed study of the relation between the regenerative structure of the workload process and the sampled workload process along with the development of a theory for bootstrapping the empirical measure of regenerative processes. This track is currently under study and will be reported elsewhere, see Hansen and Pitts (2005a), Hansen and Pitts (2005b). More specifically, Hansen and Pitts (2005a) shows by utilizing the regenerative structure that weak convergence for the empirical process is ensured if $\mathrm{E}S^2 < \infty$. In Hansen and Pitts (2005b) the regenerative structure is used to show weak convergence for the empirical process and asymptotic results for a blocked bootstrap procedure. This indicates that the methods described in the present paper can be used on much more general queueing systems.

As it is straightforward to show that the sampled workload is associated a number of standard tools for associated sequences can be utilized. One of these tools is the moving block bootstrap procedure, and a clear advantage is that this is supported in most statistical software packages.

The theory of associated sequences was introduced by Esary, Proschan, and Walkup (1967) and has since found many applications in probability, statistics, and reliability, we refer to Newman (1984) for a survey and Vanichpun and Makowski (2002) for a recent review of applications within the analysis of queuing systems.

Definition 1 *A finite sequence of random variables, $X_1, \ldots, X_n$ is said to be associated if, the inequality*

$$\mathrm{Cov}(f(X_1, \ldots, X_n), g(X_1, \ldots, X_n)) \ge 0$$

holds for all coordinate-wise nondecreasing functions $f, g : \mathbb{R}^n \to \mathbb{R}$ for which the covariance is defined. An infinite sequence of random variables is said to be associated if every finite subsequence is associated.

We now turn to establishing that $\{V_i, \ i \ge 1\}$ is an associated sequence. Unfortunately, we are not able to show this directly, although it is sequentially increasing which implies associateness, see Definition 3 and Lemma 1 of Section 6 for details.

Proposition 1 *The sampled workload process $\{V_i, \ i \ge 1\}$ is an associated sequence of random variables.*

Under a regularity condition on G it is possible to bound $\mathrm{Cov}(G(V_1), G(V_n))$.

Assumption 1 *Let $\mathcal{B}(\mathbb{R}_+)$ denote the Borel σ-field. Define a σ-finite measure μ on the measurable space $(\mathbb{R}_+, \mathcal{B}(\mathbb{R}_+))$ by*

$$\mu(A) = 1_A(0) + \int_A dx, \quad A \in \mathcal{B}(\mathbb{R}_+). \quad (5)$$

Assume that G has a density g with respect to μ. In that case g can be decomposed as

$$g(x) = (1 - \rho)1_{\{0\}}(x) + g_{\mathrm{ac}}(x), \quad x \ge 0,$$

where g_{ac} is the absolutely continuous component of G. In addition assume that g_{ac} is essentially bounded with respect to Lebesgue measure.

Remark 1 *If the service time cdf is absolutely continuous we notice by Young's inequality*

$$\| f^{\star k} \|_\infty \le \| f \|_\infty \| f^{\star (k-1)} \|_1 = \| f \|_\infty,$$

where $\| \cdot \|_\infty$ is the essential supremum with respect to Lebesgue measure. Hence, the absolutely continuous part g_{ac} is bounded by

$$\| g_{\mathrm{ac}} \|_\infty \le \rho \| f \|_\infty.$$

Proposition 2 *Assume that G satisfies Assumption 1, then*

$$\mathrm{Cov}(G(V_1), G(V_n)) \le \| g_{\mathrm{ac}} \|_\infty^2 \, \mathrm{Cov}(V_1, V_n).$$

The covariance structure of the workload process for the $M/G/1$-queue has been under intensive investigation, see Beneš (1957), Ott (1977) and Abate and Whitt (1994). From these results it is possible to derive the following proposition (see Section 6 for details).

Proposition 3 *If* $\mathrm{E}S^{\nu+3} < \infty$, *for* $\nu \geq 1$, *then*

$$\mathrm{Cov}(V_1, V_n) = O(n^{-v-\epsilon}) \qquad (6)$$

for some $\epsilon > 0$.

The properties outlined in Proposition 1, 2 and 3 turn out to be sufficient information to apply several useful results from weak convergence and resampling methods of associated sequences.

3 Asymptotic normality and the bootstrap

The theory of empirical processes plays a central role in statistics and it has many applications ranging from parameter estimation to hypothesis testing. The literature of empirical processes is large and there are many profound results, see e.g. van der Vaart and Wellner (1996) for a comprehensive overview for independent random variables.

To deal with random variables such as time series that are dependent, one naturally asks whether results obtained under the independence assumption remain valid. Such asymptotic theory is evidently useful for statistical inference of stochastic processes.

Without the independence assumption, it is more demanding to develop weak convergence theory and bootstrap methods. There are three main directions: Via 1) mixing conditions, 2) martingale methods for causal processes, 3) structural properties of Markov chains (e.g. Harris recurrence), 4) conditions on the covariance structure of associated sequence or 5) regenerative processes by Hansen and Pitts (2005a), Hansen and Pitts (2005b). In the present paper we follow direction 4) based on Shao and Yu (1996), Peligrad (1998) and Louhichi (2000).

Let Q be the quantile function of G defined by

$$\begin{aligned} Q(t) &= \inf\{x: G(x) \geq t\}, \ 0 < t \leq 1 \qquad (7) \\ Q(0) &= Q(0+). \end{aligned}$$

This means, the quantile function Q is the left continuous generalized inverse (Billingsley 1968, Page 42) of the right continuous distribution function G.

In this paper we treat weak convergence $\to_{\mathcal{D}}$ in $D[0, T)$, where $[0, T)$ is a subset of the extended positive real line, and $D[0, T)$ the Skorohod space of all right-continuous functions on $[0, T)$ with left hand limits endowed with the metric induced by the supremum norm, see Jacod and Shiryaev (1987), Chapter IV.

Finally, let B be a tight random element in $D[0, \infty)$, satisfying $B(\infty) = 0$, whose marginal distributions are zero-mean normal and a covariance function specified

by

$$\begin{aligned} &\mathrm{E}(B(x)B(y)) \\ &= \sum_{k=1}^{\infty} \mathrm{Cov}(I(V_1 \leq x), I(V_k \leq y)). \qquad (8) \end{aligned}$$

We are now ready to present the weak approximation result for the n'th empirical process β_n defined in (4).

Theorem 1 *If* $\mathrm{E}S^{\nu+3} < \infty$ *for some* $\nu \geq 4$ *and* G *satisfies Assumption 1, then we have*

$$\beta_n \to_{\mathcal{D}} B \qquad in \ \ D[0, \infty). \qquad (9)$$

In order to assess the performance of the estimator G_n, some sort of confidence band is needed. Since we regard the estimate as an element of $D[0, \infty)$, it is natural to consider confidence regions in this function space. This leads to consideration of simultaneous confidence bands from the unknown distribution function.

Assume for the moment, that the distribution of $\| B \|_\infty$ is known and q is its quantile function (see (7)). Then, Theorem 1 implies that

$$\mathrm{P}(\| \beta_n \|_\infty \leq q(\alpha)) \to \mathrm{P}(\| B \|_\infty \leq q(\alpha)) = \alpha,$$

as $n \to \infty$. An $\alpha \cdot 100\%$-confidence band could then be calculated as

$$G_n \pm n^{-1/2}q(\alpha).$$

However, the quantile function is unknown. To deal with this problem we use the bootstrap. As the data is not iid, Efron's (Efron 1979) IID-bootstrap method is modified by Künsch's (Künsch 1989) moving block bootstrap (MBB) method. See, Lahiri (2003) for a recent and detailed account of bootstrap methods and their properties for dependent data.

Let k and l be two integers such that $n = kl$. Let $T_{n1}, \ldots, T_{nk}$ be iid random variables each having uniform distribution on $\{1, 2, \ldots, n\}$. Define the triangular array $\{V_{ni}, \ 1 \leq i \leq n\}$ by $V_{ni} = V_i$ for $1 \leq i \leq n$ and $V_{ni} = V_{i-n}$ for $n < i \leq n+l$. Then the bootstrapped estimator of the empirical distribution function is defined as

$$G_n^*(x) = \frac{1}{k} \sum_{i=1}^{k} \frac{1}{l} \sum_{j=T_{ni}}^{T_{ni}+l-1} I(V_{nj} \leq x), \quad 0 \leq x < \infty.$$

The n'th-bootstrapped empirical process is then defined as

$$\beta_n^*(x) = n^{1/2}(G_n^*(x) - G_n(x)), \quad 0 \leq x < \infty.$$

In the following P^* denotes the conditional probability given $(V_1, \ldots, V_n)$. The notation $\ll$ in the following Theorem is, for notational convenience, used to replace the O-notation.

Theorem 2 *If* $\mathrm{E}S^{\nu+3} < \infty$ *for some* $\nu \geq 3$, G *satisfies Assumption 1, and* l_n, k_n *are sequences of natural numbers satisfying*

$$n^h \ll l \ll n^{1/3-\alpha}$$

for some $0 < h < 1/3 - \alpha$, $0 < \alpha < 1/3$, $l_n = l(2^k)$ *for* $2^k \leq n < 2^{k+1}$, $l_n \to \infty$ *as* $n \to \infty$, *and* $n = k_n l_n$, *then the series in (8) is convergent and*

$$\beta_n^* \to_{\mathcal{D}} B, \quad in \ D[0, \infty)$$

holds P^*-*almost surely in the Skorohod Topology on* $D[0, \infty)$.

Remark 2 *As noted in Peligrad (1998), Remark 2.1, the central limit theorem of the bootstrapped empirical process is ensured for a larger class of processes, than the empirical process of the original data.*

From Theorem 2 the following result follows.

Corollary 1 *Under the conditions in Theorem 2, we get*

$$\mathrm{P}^* \left(\| \beta_n^* \|_\infty \leq q(\alpha) \right) \to \mathrm{P} \left(\| B \|_\infty \leq q(\alpha) \right) = \alpha,$$

as $n \to \infty$.

The confidence band is then constructed by simulating N independent replications $\beta_{n,i}^*$, $i = 1, \ldots, N$ of β_n^*, and $q(\alpha)$ is estimated by

$$\hat{q}_N(\alpha) = \inf \left\{ x : \ N^{-1} \sum_{i=1}^{N} I(\| \beta_{n,i}^* \|_\infty \leq x) \geq \alpha \right\}.$$

A $\alpha \cdot 100\%$ bootstrapped confidence band can then be constructed as

$$G_n \pm n^{-1/2} \hat{q}_N(\alpha). \tag{10}$$

4 Simulation results

The results above have interesting consequences for statistical inference about queuing systems. We shall in the present section see, how the presented weak convergence results enable statistical inference about the workload distribution function.

We will consider the widely used $M/D/1$ queue for our simulation study, which means the service times have a deterministic length. For the simulation study we will throughout this section assume $\rho = 0.5$ and $S \equiv 1$. (For simulation results on other service time distributions, see Hansen and Pitts (2005a), Hansen and Pitts (2005b).)

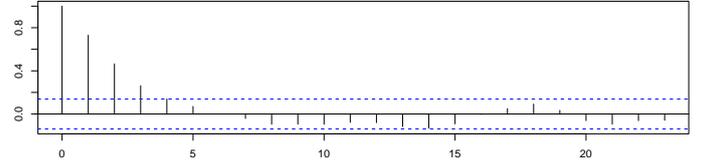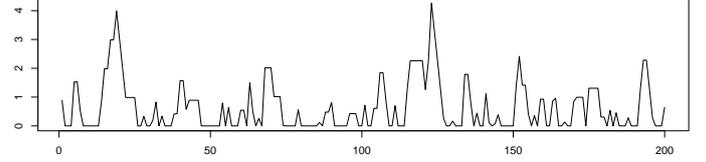Observe that all moments of S exists and that G is absolutely continuous except for the atom at zero. From

Figure 1: Upper panel: Simulated data set of size $n = 200$ with $\rho = 0.5$ and $S_n \equiv 1$, $n \geq 0$. Lower panel: Corresponding estimated acf.
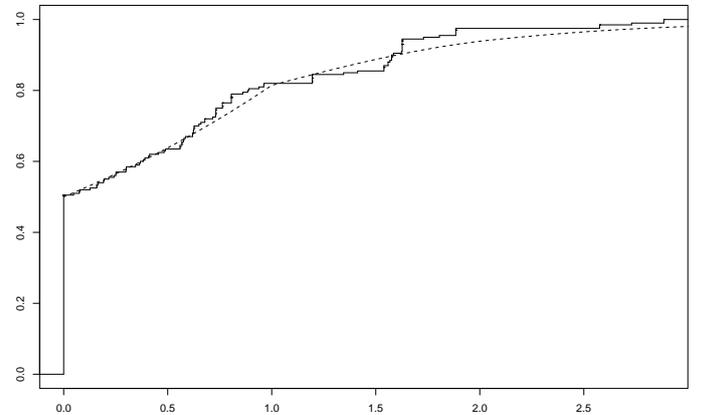

Figure 2: The theoretical distribution function G (broken line) and the estimate (solid line).

Remark 1 it follows that the absolute continuous part of G is bounded by ρ.

In order to make comparisons, we calculate the stationary cdf of the workload distribution. One approach to calculate the compound geometric distribution in (2), is 1. to notice that the stationary excess distribution of the constant random variable $S \equiv 1$, is the uniform distribution over $[0, 1]$, 2. derive the Fourier transform of F_e and 3. utilize that

$$G(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} (1 - \rho) \frac{1}{1 - \rho \widehat{F_e}(t)} \exp(-itx) dt.$$

However, inverting the Fourier transform is by no means trivial because of the discontinuity at 0. For numerically stable procedures we refer to Abate and Whitt (1992) and Grübel and Hermesmeier (1999).

Another and more simple possibility is to notice that, if F is one-sided and of lattice-type, i.e. concentrated
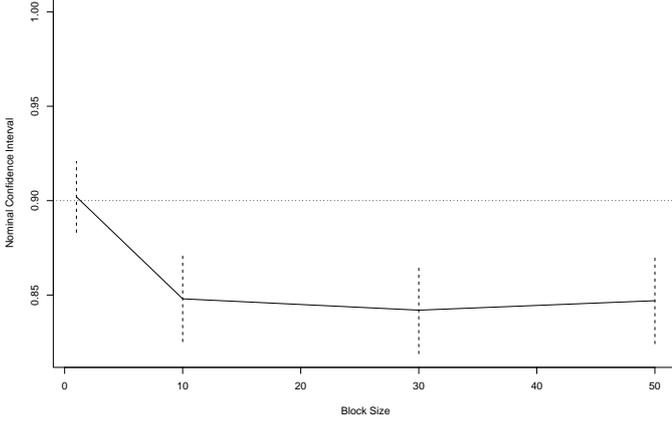
Figure 3: Sample mean of 1000 iid bootstrapped estimates of the nominal confidence interval, for block size $l = 1, 10, 30$, and 50. A ± 1.96 times the sample standard error, interval is centered around the sample mean, is shown by the vertical broken line segments.

on the non-negative integer multiples of some $h > 0$, then the compound geometric distribution (2) can be handled numerically by Panjer recursion (Panjer 1981). If F is not of lattice-type, then the use of Panjer recursion requires an initial discretization step which leads to a discretization error. Theoretical justifications for discretized Panjer recursion can be found in Grübel and Hermesmeier (1999), Grübel and Hermesmeier (2000). Using Panjer recursion instead of transform methods avoids problems that may arise if the Fourier transform winds about 0.

Algorithm 1 (Panjer Recursion)

1. *Choose a discretization level $h > 0$ and consider the corresponding lattice $\mathbb{P}_h = \{hz | z \in \mathbb{Z}\}$.*

2. *Discretize the density of the uniform distribution function F_e over $[0, 1]$ in the following way*

$$f(x) = \left\{ \begin{array}{ll} h & for\ x \in \mathbb{P}_h \cap (0, 1] \\ 0 & elsewhere \end{array} \right.$$

3. *Calculate an approximation to the density of G by the following recursion*

$$g(x)$$
$$= \left\{ \begin{array}{ll} 1 - \rho & x = 0 \\ \rho \sum_{j=1}^{x/h} f(jh)g(jh - x) & x \in \mathbb{P}_h \cap (0, \infty) \\ 0 & elsewhere \end{array} \right.$$

For the workload distribution function of the $M/D/1$ under study, approximated with $h = 1/1000$, see the broken line in Figure 2.
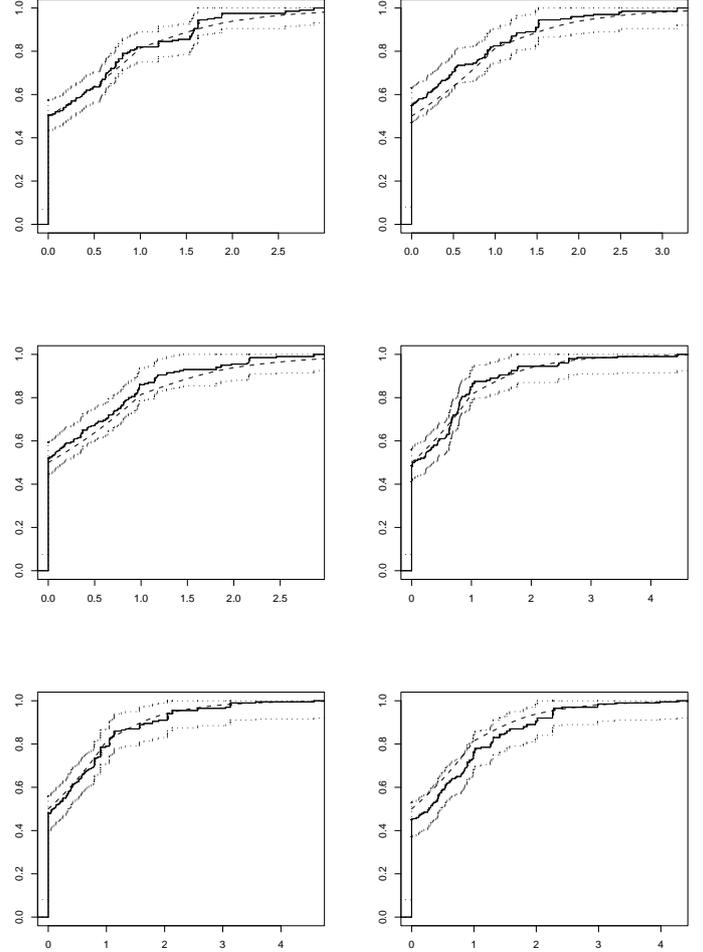

Figure 4: The distribution function G (broken line), the estimate (solid line), and the bootstrap confidence band (dotted lines) for $\rho = 0.5$, $n = 200$, $l = 1$, $k = 1000$ and $N = 1000$. The experiment was replicated on six independent simulations of the queue.

A stationary version of the $M/D/1$-workload process (1) can be started out at 0, by realizing, that the marginal distribution of V_0 equals the distribution of a geometrically-stopped uniformly-summed random variable. Recall that the stationary excess distribution F_e is the uniform distribution on $[0, 1]$ and the geometric distribution has mean $(1 - \rho)^{-1}$.

Figure 1 (upper panel) shows a simulated data set of size $n = 200$ from the stationary version of the $M/D/1$-model in (1) with $\rho = 0.5$ and $S \equiv 1$. The autocorrelation function, $\text{acf}(n) = \text{Cov}(V_1, V_n) / \text{Var}(V_0)$, is estimated and plotted in the lower panel (solid lines) together with the approximate 95% confidence limits for the hypothesis of no correlation. The plot shows positive and non-vanishing autocorrelations until lag 4, which indicates a quickly decaying positive autocorre-

lation function. This is expected as, from Proposition 3, we have $\mathrm{Cov}(V_1, V_n) = O(n^{-k-\epsilon})$ for all $k \geq 1$.

The ecdf (3) of the sample is illustrated by the solid line in Figure 2. We notice from this plot that the ecdf fits the distribution function quite well. To get an idea of the variability, see the six ecdf's based on independent replications of the experiment with the corresponding theoretical distribution function in Figure 4.

A practical problem of applying the asymptotic results of Section 3 lies in choosing the block size l. In the present case we use the 'calibration by adjusting the block size' method presented in Politis, Romano, and Wolf (1999), Algorithm 9.3.2. The rationale behind calibration methods is for various block sizes to calculate the nominal confidence interval. The word nominal is used to describe that we calculate the actual coverage probability for the chosen block size. However, as the distribution function is unknown, one uses the ecdf as an approximation to the theoretical cdf. To describe the calibration formally, see the following algorithm.

Algorithm 2 (Calibration)

1. *For each block size l*

2. *Generate I MBB distribution functions (using block size l)*
$$F_{n,i}^*, \ 1 \leq i \leq I,$$
   from the ecdf F_n.

3. *From each $F_{n,i}^*$ generate J MBB distribution functions (using block size l)*
$$F_{n,i}^{*j}, \ 1 \leq j \leq J.$$

4. *For each $F_{n,i}^*$ estimate the $\alpha \cdot 100\%$ confidence band by*

$$\hat{q}_{i,N}^*(\alpha)$$
$$= \inf \left\{ x : \ J^{-1} \sum_{j=1}^{J} I(\| F_{n,i}^{*j} - F_{n,i}^* \|_\infty \leq x) \geq \alpha \right\}$$
   and
$$F_{n,i}^* \pm \hat{q}_{i,N}^*(\alpha)$$

5. *Estimate the actual coverage probability by*
$$\frac{\sharp \left\{ i : \ \| F_{n,i}^* - F_n \|_\infty \leq q_{i,N}^*(\alpha) \right\}}{I}$$

Note, we have not claimed any asymptotic optimality of the described procedure. It should only be seen as a sensible way of choosing the block size in the small sample case. More formal analyses of calibration methods in the time series case, seem to be an open question.

In Figure 3, the 90% nominal coverage probability is estimated at block sizes $1, 10, 30$ and 50 and linearly interpolated in between them. In the simulation study the algorithm is based on $I = 1000$ and $J = 100$. A 95% confidence interval based on the 1000 independent samples of the confidence interval is indicated by the vertical broken lines. The conclusion seems to be that a small block size will do. This is not surprising as we noticed that the autocorrelations died out quickly. In our further analysis the block size was chosen to be 1.

As an estimate for the empirical distribution function we used the ecdf of the data and the MBB method to construct a 90% confidence band, by one iteration, i.e. $I = 1$ and $J = 1000$ in Algorithm 2.

In Figure 4 we started 6 independent queues in the stationary distribution and sampled $n = 200$ values. For each realization we calculated the ecdf and the 90% MBB confidence interval. We see, in each case, that the confidence region covers the true distribution function.

All programming and simulations have been carried out with standard routines in the freely available computational statistical software package R, see `<http://www.r-project.org>` for more details.

5 Discussion

Under strong requirements on the moment of the service time distribution we have proven weak convergence to a Gaussian process. Moreover, we have presented a recipe for analyzing data from the widely used $M/D/1$-model. However, by relaxing the moment conditions on S, the sum of the covariance function will not converge (Ott 1977, Theorem 1)

$$\sum_{n=1}^{\infty} \mathrm{Cov}(V_1, V_n) = \infty.$$

Processes of this type are described by various authors, as being long-range dependent. It is well-known that the scaling factor $n^{1/2}$ used in the empirical processes is substantially smaller for long-range dependent data. Furthermore, the limit distribution of the normalized ecdf can be nonnormal. Procedures for a rather special class of long-range dependent data are reviewed in Lahiri (2003), Chapter 10, but a systematic study of long-range dependent associated sequences seems to an open issue. As it is well documented that data in communication systems can be long-range dependent an interesting topic for further study is to see if similar results can be derived for sampling the workload of the $M/G/1$-queue with heavy tailed service times.

These and other open questions as well as applications are studied in the separate papers Hansen and Pitts (2005a) and Hansen and Pitts (2005b). In these

papers the moment condition on the service times are relaxed to $\mathrm{E}S^2 < \infty$. Quite interestingly this provides examples where the covariance function is not summable but we have indeed convergence to the normal distribution.

6 Proofs

In order to prove Proposition 1, we prove the slightly stronger property of being sequentially stochastically increasing.

Definition 2 *Let X, Y be $\mathbb{R}^n$-valued random vectors. We say that $X \leq Y$ in the sense of stochastically ordering (written $X \leq_{\mathrm{st}} Y$) if*

$$\mathrm{E}(f(X)) \leq \mathrm{E}(f(Y))$$

for all increasing measurable functions $f : \mathbb{R}^n \to \mathbb{R}$.

Definition 3 *The real-valued sequence $\{X_n, n \geq 1\}$ is said to be sequentially stochastically increasing (SSI) if for each $n = 1, 2, \ldots$ it holds that*

$$[X_n | (X_1, \ldots, X_{n-1}) = x] \leq_{\mathrm{st}} [X_n | (X_1, \ldots, X_{n-1}) = y]$$

for $x, y \in \mathbb{R}^{n-1}$ satisfying $x \leq y$ component-wise.

Lemma 1 *If a sequence of random variables $\{X_n, n \geq 1\}$ are sequentially stochastically increasing, then they are necessarily also associated.*

For a proof, see Theorem 4.7 of Barlow and Proschan (1975).

Proof of Proposition 1 First we notice that $\{V_i, i \geq 1\}$ is sequentially stochastically increasing. Let f be any real-valued measurable and increasing function and

$$(u_1, \ldots, u_{n-1}) \leq (v_1, \ldots, v_{n-1})$$

component-wise. Then by Lindley's equation (1)

$$
\begin{aligned}
&\mathrm{E}(f([V_n | V_1 = u_1, \ldots, V_{n-1} = u_{n-1}])) \\
&= \mathrm{E}(f([V_n | V_{n-1} = u_{n-1}])) \\
&\leq \mathrm{E}(f([V_n | V_{n-1} = v_{n-1}])) \\
&= \mathrm{E}(f([V_n | V_1 = v_1, \ldots, V_{n-1} = v_{n-1}])).
\end{aligned}
$$

Hence by Lemma 1 $\{V_i, i \geq 1\}$ is associated. $\qquad\square$

Proof of Proposition 2 Let $0 \leq x \leq y$, then

$$G(y) - G(x) \leq \| g_{\mathrm{ac}} \|_\infty (y - x).$$

From which we observe that the function f_1 defined by

$$f_1(x) = \| g_{\mathrm{ac}} \|_\infty x - G(x), \quad x \geq 0,$$

is increasing. Now, by successive applications of Definition 1 we get

$$
\begin{aligned}
\mathrm{Cov}(G(V_1), G(V_n)) &\leq \| g_{\mathrm{ac}} \|_\infty \, \mathrm{Cov}(V_1, G(V_n)) \\
&\leq \| g_{\mathrm{ac}} \|_\infty^2 \, \mathrm{Cov}(V_1, V_n).
\end{aligned}
$$

$\qquad\square$

Lemma 2 *If $\mathrm{E}S^{\nu+3} < \infty$, for some $\nu \geq 1$, then*

$$\mathrm{Cov}(V_0, V_t) = O(t^{-\nu-\epsilon}),$$

for some $\epsilon > 0$.

Proof From Abate and Whitt (1994), Proposition 1, it follows that V_t has one moment less than S. Furthermore, it also follows (Abate and Whitt 1994, Theorem 10) that

$$\mathrm{Cov}(V_0, V_t) = \mathrm{Var}(V_0)\bar{U}_{\mathrm{e}}(t).$$

The cdf U, has one moment less than V_t. Consequently, by the tail integration formula

$$
\begin{aligned}
u_{\nu+1} &= \int_0^\infty x^{\nu+1} U(dx) \\
&= \int_0^\infty (\nu + 1)x^\nu \bar{U}(x)dx \qquad (11) \\
&< \infty.
\end{aligned}
$$

Whereby, for some $\epsilon > 0$,

$$\bar{U}(x) = O(x^{-(\nu+1)-\epsilon}), \quad \text{for } x \to \infty.$$

From which

$$
\begin{aligned}
\bar{U}_{\mathrm{e}}(x) &= u_{\mathrm{e}1}^{-1} \int_x^\infty \bar{U}(y)dy \\
&= O(x^{-\nu-\epsilon}), \quad \text{for } x \to \infty,
\end{aligned}
$$

and some $\epsilon > 0$. $\qquad\square$

Proof of Proposition 3 Follows immediately from Lemma 2. $\qquad\square$

Lemma 3 *Assume $\mathrm{E}S^{\nu+3} < \infty$, for some $\nu \geq 1$ and G satisfies Assumption 1. Let $\{U'_n, n \geq 1\}$ be a sequence of independent random variables uniformly distributed on $[0, 1 - \rho]$. Then the sequence of random variables*

$$U_n = G(V_n)1(V_n > 0) + U'_n 1(V_n = 0)$$

is stationary, associated, uniformly distributed on $[0, 1]$ and

$$\mathrm{Cov}(U_1, U_n) = O(n^{-\nu-\epsilon})$$

for some $\epsilon > 0$.

Proof of Lemma 3 As $\{V_n,\ n \geq 1\}$ and $\{U_n',\ n \geq 1\}$ are stationary, stationarity of $\{U_n,\ n \geq 1\}$ is immediate. The uniform distribution of the U_n's, $n \geq 1$ on $[0, 1]$ follows directly by stationarity and their definition.

Let $f:\ \mathbb{R}^n \to \mathbb{R}$ be a measurable increasing function, then for $(u_1, \ldots, u_{n-1}) \leq (v_1, \ldots, v_{n-1})$ componentwise, in the same way as the proof of Proposition 1

$$\mathrm{E}f([U_n|U_1 = u_1, \ldots, U_{n-1} = u_{n-1}])$$
$$= \mathrm{E}f([U_n|U_{n-1} = u_{n-1}])$$
$$\leq \mathrm{E}f([U_n|U_{n-1} = v_{n-1}]).$$

Hence $\{U_n,\ n \geq 1\}$ is SSI and by Lemma 1 associated. For the covariance function consider

$$
\begin{aligned}
\mathrm{Cov}(U_1, U_n) &= \mathrm{Cov}(G(V_1)1(V_1 > 0), G(V_n)1(V_n > 0)) \\
&+ \mathrm{Cov}(G(V_1)1(V_1 > 0), U_n'1(V_n = 0)) \\
&+ \mathrm{Cov}(U_1'1(V_1 = 0), G(V_n)1(V_n > 0)) \\
&+ \mathrm{Cov}(U_1'1(V_1 = 0), U_n'1(V_n = 0)) \\
&= \mathrm{I} + \mathrm{II} + \mathrm{III} + \mathrm{IV}.
\end{aligned}
$$

By use of Cuadras' generalization (Cuadras 2002, Theorem 1) of the Hoeffding identity, Abate and Whitt (1994), the tail integration formula, Propositions 2 and 3 one obtains the following estimates

$$
\begin{aligned}
\mathrm{I} &= O(n^{-\nu-\epsilon}) \\
\mathrm{II} &= O(n^{-\nu-\epsilon}) \\
\mathrm{III} &= O(n^{-(\nu+1)-\epsilon}) \\
\mathrm{IV} &= O(n^{-\nu-\epsilon}).
\end{aligned}
$$

for some $\epsilon > 0$. Combining the estimates for I, II, III, and IV yields the result of the stated lemma. □

Proof of Theorem 1 Let $\{U_n,\ n \geq 1\}$ be defined as in Lemma 3. Let A be a tight element of $D[0, 1]$, satisfying $A(0) = A(1) = 0$, whose marginal distributions are zero-mean normal with a covariance function specified by

$$\mathrm{Cov}(A(t), A(s)) = \sum_{k=1}^{\infty} \mathrm{Cov}(I(U_1 \leq s), I(U_k \leq t)).$$

Define the ecdf of $\{U_i,\ 1 \leq i \leq n\}$ by

$$H_n(x) = n^{-1}\sum_{i=1}^{n} I(U_i \leq x), \quad x \in [0, 1],$$

and finally its n'th empirical process by

$$\alpha_n(x) = n^{1/2}(H_n(x) - x), \quad x \in [0, 1].$$

Then from Lemma 3 the conditions of Louhichi (2000), Theorem 1, are fulfilled for the sequence $\{U_n,\ n \geq 1\}$ and it follows that

$$\alpha_n \to_{\mathcal{D}} A \quad \text{in } D[0, 1].$$

If we define $h:\ D[0, 1] \to D[0, \infty)$ by $(hx)(t) = x(G(t))$ and let $G_n = h(H_n)$, the result follows from Corollary 1 to Theorem 5.1 in Billingsley (1968). □

Proof of Theorem 2 Use the same construction as in the proof of Theorem 1. Now, from Lemma 3 $\{U_n,\ n \geq 1\}$ satisfies the conditions of Peligrad (1998), Theorem 2.4). Again, as in the proof of Theorem 1 the result follows from Corollary 1 of Theorem 5.1 in Billingsley (1968). □

REFERENCES

Abate, J., and W. Whitt. 1992. The Fourier-series method for inverting transforms of probability distributions. *Queueing Systems* 10:5–88.

Abate, J., and W. Whitt. 1994. Transient behavior of the $M/G/1$ workload process. *Operations Research* 42:750–764.

Asmussen, S. 2003. *Applied probability and queues*. 2 ed. Springer Verlag, New York, NY.

Baccelli, F., and P. Brémaud. 2003. *Elements of queueing theory*. 2 ed. Springer Verlag, New York, NY.

Barlow, R., and F. Proschan. 1975. *Statistical theory of reliability and life testing*. Holt, Rinehart and Winston, New York, NY.

Beněs, V. 1957. On queues with Poisson arrivals. *The Annals of Mathematical Statistics* 28:670–677.

Billingsley, P. 1968. *Convergence of probability measures*. Wiley, New York, NY.

Cuadras, C. M. 2002. On the covariance between functions. *Journal of Multivariate Analysis* 81:19–27.

Datta, S., and W. McCormick. 1993. Regeneration-based bootstrap for Markov chains. *Canadian Journal of Statistics* 21:181–193.

Efron, B. 1979. Bootstrap methods: Another look at the jackknife. *The Annals of Statistics* 7:1–26.

Esary, J., F. Proschan, and D. W. Walkup. 1967. Association of random variables, with applications. *The Annals of Mathematical Statistics* 38:1466–1474.

Grübel, R., and R. Hermesmeier. 1999. Computation of compound distributions I: Aliasing errors and exponential tilting. *Astin Bulletin* 29:197–214.

Grübel, R., and R. Hermesmeier. 2000. Computation of compound distributions II: Discretization errors and Richardson extrapolation. *Astin Bulletin* 30:309–331.

Hansen, M. B., and S. M. Pitts. 2005a. Nonparametric inference from the $M/G/1$ workload. In preparation.

Hansen, M. B., and S. M. Pitts. 2005b. Regenerative block-bootstrap for empirical processes. In preparation.

Jacod, J., and A. N. Shiryaev. 1987. *Limit theorems for stochastic processes*. Grundlehren der mathematischen Wissenschaften. Springer Verlag, New York, NY.

Künsch, H. 1989. The jackknife and the bootstrap for general stationary observations. *The Annals of Statistics* 17:1217–1261.

Lahiri, S. N. 2003. *Resampling methods for dependent data*. Springer Verlag, New York, NY.

Louhichi, S. 2000. Weak convergence for empirical processes of associated sequences. *Annales de l'Institut Henri Poincare, Probabilités et Statistiques* 36:547–567.

Newman, C. M. 1984. Asymptotic independence and limit theorems for positively and negatively dependent random variables. In *Inequalities in Statistics and Probability*, ed. Y. L. Tong, 127–140. IMS, Hayward, CA.

Ott, T. J. 1977. The covariance function of the virtual waiting-time process in an $M/G/1$ queue. *Advances in Applied Probability* 9:158–168.

Panjer, H. H. 1981. Recursive evaluation of a family of compound distributions. *Astin Bulletin* 12:22–26.

Peligrad, M. 1998. On the blockwise bootstrap for empirical processes for stationary sequences. *The Annals of Probability* 26:877–901.

Politis, D. N., J. P. Romano, and M. Wolf. 1999. *Subsampling*. Springer Verlag, New York, NY.

Shao, Q.-M., and H. Yu. 1996. Weak convergence for weighted empirical processes of dependent sequences. *Annals of Statistics* 24:2098–2127.

van der Vaart, A. W., and J. A. Wellner. 1996. *Weak convergence and empirical processes*. Springer Verlag, New York, NY.

Vanichpun, S., and A. M. Makowski. 2002. Positive correlations and buffer occupancy: Lower bounds via supermodular ordering. In *Proceedings of IEEE INFOCOM*, Volume 16, 1298–1306.