informs
http://pubsonline.informs.org/journal/opre/

OPERATIONS RESEARCH
Vol. 67, No. 1, January–February 2019, pp. 267–280
ISSN 0030-364X (print), ISSN 1526-5463 (online)

Periodic Little's Law

Ward Whitt,[a] Xiaopei Zhang[a]

[a] Industrial Engineering and Operations Research, Columbia University, New York, New York 10027

Contact: ww2040@columbia.edu, http://orcid.org/0000-0003-4298-9964 (WW); xz2363@columbia.edu, http://orcid.org/0000-0002-9494-8112 (XZ)

Received: April 29, 2017
Revised: October 31, 2017; March 6, 2018
Accepted: April 20, 2018
Published Online in Articles in Advance: January 25, 2019

Subject Classifications: queues: nonstationary, limit theorems, applications
Area of Review: Stochastic Models

https://doi.org/10.1287/opre.2018.1766

Abstract. Motivated by our recent study of patient flow data from an Israeli emergency department (ED), we establish a sample path periodic Little's law (PLL), which extends the sample path Little's law (LL). The ED data analysis led us to propose a periodic stochastic process to represent the aggregate ED occupancy level, with the length of a periodic cycle being 1 week. Because we conducted the ED data analysis over successive hours, we construct our PLL in discrete time. The PLL helps explain the remarkable similarities between the simulation estimates of the average hourly ED occupancy level over a week using our proposed stochastic model fit to the data and direct estimates of the ED occupancy level from the data. We also establish a steady-state stochastic PLL, similar to the time-varying LL.

Funding: Financial support from the Division of Civil, Mechanical and Manufacturing Innovation of the National Science Foundation [Grants 1265070 and 1634133] is acknowledged.
Supplemental Material: The e-companion is available at https://doi.org/10.1287/opre.2018.1766.

Keywords: Little's law • $L = \lambda W$ • periodic queues • service systems • data analysis • emergency departments

1. Introduction

Many service systems with customer response times extending over hours or days can be modeled as periodic queues, with the length of a periodic cycle being 1 week. Examples are hospitals wards, order fulfillment systems, and loan-processing systems. In this paper, we establish a periodic version of Little's law (LL), which can provide insight into the performance of these periodic systems.

We formulate our periodic Little's law (PLL) in discrete time, assuming that there are d discrete time points within each periodic cycle. In discrete time, the PLL states that, under appropriate conditions,

$$L_k = \sum_{j=0}^{\infty} \lambda_{k-j} F^c_{k-j,j}, \quad k = 0, 1, \ldots, d-1, \tag{1}$$

where d is the number of time points within each periodic cycle, L_k is the long-run average number in the system at time k, λ_k is the long-run average number of arrivals at time k, and $F^c_{k,j}$, $j \geq 0$, is the long-run proportion of arrivals at time k that remain in the system for at least j time points, which can be viewed as the complementary cumulative distribution function (ccdf) of the length of stay (LoS) of an arbitrary arrival. The long-run averages are over all indices of the form $k + md$, $m \geq 0$. These quantities λ_k, $F^c_{k,}$ and L_k are periodic functions of the time index k, exploiting the extension of these periodic functions to all integers: negative as well as positive.

In many applications, time is naturally continuous, in which case the analog of (1) is

$$L(t) = \int_0^c \lambda(t) F^c(t-s, s)\, ds, \quad 0 \leq t < c, \tag{2}$$

where c is the length of each periodic cycle. When time is continuous, we can construct a discrete time version by letting there be d subintervals of equal length within each continuous time periodic cycle, which we can refer to as time periods. We then obtain discrete time processes by appropriately counting what happens in each time period. However, neither equal length time subintervals in continuous time nor a continuous time reference are needed to have a bona fide discrete time system.

However, if time is actually continuous, then we can use the discrete time sample path PLL to define what we mean by a corresponding continuous time sample path PLL: we say that a continuous time PLL holds with (2) if the discrete time PLL holds for all sequences of versions with d periods in each continuous time cycle with $d \to \infty$ and there is sufficient regularity in the limit functions so that the limits in (1) can serve successive Riemann sums converging to the integral (2) (Section 2.7 has additional discussion).

We were motivated to develop the PLL because of a remarkable similarity between two curves that we observed in our recent study of patient flow data from an Israeli emergency department (ED) in Whitt and Zhang (2017). As part of that study, we developed an aggregate stochastic model of an ED based on a statistical analysis of patient arrival and departure data from the ED of an Israeli hospital using 25 weeks of data from the data repository associated with the study by Armony et al. (2015). In section 6 of Whitt and Zhang (2017), we conducted simulation experiments to validate the aggregate model of ED patient flow. One of these comparisons compared direct estimates of the average ED occupancy level from data with estimations from simulations of the stochastic model, where the distributions of the daily number of arrivals, the arrival rate function, and the LoS distribution are estimated from the data. Figure 1 shows that the two curves are barely distinguishable. The PLL provides an explanation.

In Whitt and Zhang (2017), we suggested that this remarkable fit could be explained, at least in part, by the time-varying LL from Bertsimas and Mourtzinou (1997) and Fralix and Riano (2010). In this paper, we elaborate on that idea by providing the new sample path version of PLL, because we think that it may be important for constructing data-generated models of service systems more broadly. Although our primary focus here is on the PLL, in Section 3.4 and the e-companion, we provide evidence in support of the model that we proposed in Whitt and Zhang (2017).

The main contribution of this paper is the sample path PLL in discrete time, Theorem 1, extending the sample path LL ($L = \lambda W$) established by Stidham (1974) (also see Little 1961; Whitt 1991, 1992; El-Taha and Stidham 1999; Fiems and Bruneel 2002; Little 2011; and Wolfe and Yao 2014). This sample path PLL is different in detail from all previous sample path LL results (known to us). For example, in addition to the usual limits of averages of the arrival rates and LoS (waiting times), we need to assume a limit for the entire LoS distribution. The necessity of this condition is shown by Example 1 in Section 2.3.

We also establish steady-state stochastic versions of the PLL, which relate more directly to the time-varying LL in Bertsimas and Mourtzinou (1997) and Fralix and Riano (2010). This involves the usual two forms of stationarity associated with arrival times and arbitrary times that emerge from the Palm theory of stochastic point processes (e.g., Baccelli and Bremaud 1994, Sigman 1995), but now, both are in discrete time, such as in Miyazawa and Takahashi (1992) and section 1.7.4 of Baccelli and Bremaud (1994). Our steady-state stochastic versions of the PLL extend (and are consistent with) an early PLL for the $M_t/GI/1$ queue in proposition 2 of Rolski (1989).

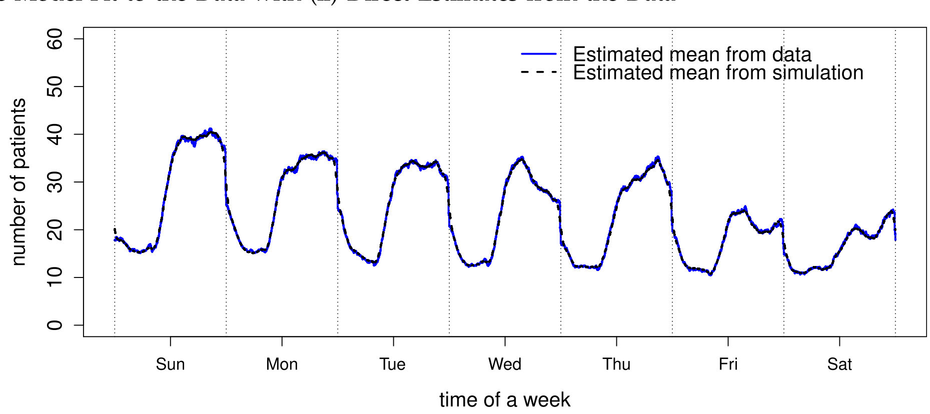
The rest of the paper is organized as follows: In Section 2, we state and discuss the sample path PLL. In Section 3, we establish the steady-state stochastic versions of the PLL. In Section 3.4 and the e-companion, we elaborate on the ED application, reviewing the model that we built in Whitt and Zhang (2017), illustrating how it relates to the PLL, and providing evidence that the conditions in the theorems are satisfied in our application. In Section 4, we provide the proofs of theorems in Section 2. Finally, in Section 5, we draw conclusions. Additional supporting material appears in the e-companion.

2. Sample-Path Version of the PLL

In this section, we develop the sample path PLL. This version is general in that (i) we do not directly make any stochastic assumptions and (ii) we do not directly impose any periodic structure. Instead, we assume that natural limits exist, which we take to be with probability 1 (w.p.1). It turns out that the periodicity of the limit emerges automatically from the assumed existence of the limits.

This section is organized as follows. In Section 2.1, we introduce our notation and definitions. In Section 2.2, we state our main limit theorem. In Section 2.3, we discuss our assumptions and give an example showing that the extra condition beyond what is needed for the LL is necessary. In Section 2.4, we establish a second limit theorem, showing that the natural indirect estimator for the average queue length based on the arrival rate and waiting time is consistent. In Section 2.5, we establish a limit for the departure process as a corollary to the main theorem. Finally, we conclude with some

Figure 1. (Color online) A Comparison of the Estimated Mean ED Occupancy Level from (i) Simulations of Multiple Replications of the Model Fit to the Data with (ii) Direct Estimates from the Data



additional discussion to add insight. In Section 2.6, we discuss the connection between our averages and associated cumulative processes. In Section 2.7, we discuss the different orderings of events at discrete time points and the relation between continuous time and discrete time.

2.1. Notation and Definitions

We consider discrete time points indexed by integers i, $i \geq 0$. Because multiple events can happen at these times, we need to carefully specify the order of events, just as in the large literature on discrete time queues (e.g., Bruneel and Kim 1993). We assume that all arrivals at one time occur before any departures. Moreover, we count the number of customers (patients in the ED in our intended application) in the system after the arrivals but before the departures. Thus, each arrival can spend time j in the system for any $j \geq 0$. Our convention yields a conservative upper bound on the occupancy. We discuss other possible orderings of events and the relation between continuous time and discrete time in Section 2.7.

With these conventions, we focus on a single sequence, $X \equiv \{X_{i,j} : i \geq 0; j \geq 0\}$, with $X_{i,j}$ denoting the number of arrivals at time i that have LoS j. We also could have customers at the beginning, but without loss of generality, we can view them as a part of the arrivals at time 0. We define other quantities of interest in terms of X. In particular, with $\equiv$ denoting equality by definition, the key quantities are:

- $Y_{i,j} \equiv \sum_{s=j}^{\infty} X_{i,s}$: the number of arrivals at time i with LoS greater than or equal to j, $j \geq 0$;
- $A_i \equiv Y_{i,0} = \sum_{s=0}^{\infty} X_{i,s}$: the total number of total arrivals at time i; and
- $Q_i \equiv \sum_{j=0}^{i} Y_{i-j,j} = \sum_{j=0}^{i} A_{i-j} \frac{Y_{i-j,j}}{A_{i-j}}$: the number in system at time i,

all for $i \geq 0$. In the last line and throughout the paper, we understand $0/0 \equiv 0$, so that we properly treat times with 0 arrivals.

We do not directly make any periodic assumptions, but with the periodicity in mind, we consider the following averages over n periods:

$$\begin{aligned}
\bar{\lambda}_k(n) &\equiv \frac{1}{n}\sum_{m=1}^{n} A_{k+(m-1)d},\\
\bar{Q}_k(n) &\equiv \frac{1}{n}\sum_{m=1}^{n} Q_{k+(m-1)d} = \frac{1}{n}\sum_{m=1}^{n}\left(\sum_{j=0}^{k+(m-1)d} Y_{k+(m-1)d-j,j}\right),\\
\bar{Y}_{k,j}(n) &\equiv \frac{1}{n}\sum_{m=1}^{n} Y_{k+(m-1)d,j}, \quad j \geq 0,\\
\bar{F}^c_{k,j}(n) &\equiv \frac{\bar{Y}_{k,j}(n)}{\bar{\lambda}_k(n)} = \frac{\sum_{m=1}^{n} Y_{k+(m-1)d,j}}{\sum_{m=1}^{n} A_{k+(m-1)d}}, \quad j \geq 0, \quad \text{and}\\
\bar{W}_k(n) &\equiv \sum_{j=0}^{\infty} \bar{F}^c_{k,j}(n), \quad 0 \leq k \leq d-1,
\end{aligned} \tag{3}$$

where d is a positive integer.

Clearly, $\bar{\lambda}_k(n)$ is the average number of arrivals at time k, $0 \leq k \leq d-1$, over the first n periods. Similarly, $\bar{Q}_k(n)$ is the average number of customers in the system at time k, whereas $\bar{Y}_{k,j}(n)$ is the average number of customers that arrive at time k that have an LoS greater than or equal to j. Thus, $\bar{F}^c_{k,j}(n)$ is the empirical ccdf, which is the natural estimator of the LoS ccdf of an arrival at time k. Finally, $\bar{W}_k(n)$ is the average LoS of customers that arrive at time k. We will let $n \to \infty$.

2.2. The Limit Theorem

With the framework introduced above, we can state our main theorem: the sample path version of the PLL. We first introduce our assumptions, which are just as in the sample path LL, with one exception. In particular, we assume that

$$\begin{aligned}
&\text{(A1)}\quad \bar{\lambda}_k(n) \to \lambda_k, \quad \text{w.p.1} \quad \text{as} \quad n \to \infty, \quad 0 \leq k \leq d-1,\\
&\text{(A2)}\quad \bar{F}^c_{k,j}(n) \to F^c_{k,j}, \quad \text{w.p.1} \quad \text{as} \quad n \to \infty,\\
&\qquad\qquad 0 \leq k \leq d-1,\ j \geq 0, \quad \text{and}\\
&\text{(A3)}\quad \bar{W}_k(n) \to W_k \equiv \textstyle\sum_{j=0}^{\infty} F^c_{k,j} \quad \text{w.p.1} \quad \text{as} \quad n \to \infty,\\
&\qquad\qquad 0 \leq k \leq d-1,
\end{aligned} \tag{4}$$

where the limits are deterministic and finite. For the sample path LL, $d = 1$, and we do not need (A2).

The assumptions above only assume the existence of limits within the first period, but the limits immediately extend to all $k \geq 0$, showing that the limit functions must be periodic functions. We then extend these periodic functions to the entire real line, including the negative time indices. We give a proof of the following in Section 4.1.

Lemma 1 (Periodicity of the Limits). *If the three assumptions in (4) hold, then the limits hold for all $k \geq 0$, with the limit functions being periodic with period d.*

We are now ready to state our main theorem; we give the proof in Section 4.2.

Theorem 1 (Sample Path PLL). *If the three assumptions (A1), (A2), and (A3) in (4) hold, then $\bar{Q}_k(n)$ defined in (3) converges* w.p.1 *as $n \to \infty$ to a limit that we call L_k. Moreover,*

$$L_k = \sum_{j=0}^{\infty} \lambda_{k-j} F^c_{k-j,j} < \infty, \quad 0 \leq k \leq d-1, \tag{5}$$

where λ_k and $F^c_{k,j}$ are the periodic limits in (A1) and (A2) extended to all integers, negative as well as positive.

Remark 1 (The Extension to Negative Indices). To have convenient notation, we have extended the periodic limit functions to the negative indices, but we do not consider the averages and their limits in assumptions (A1), (A2), and (A3) for negative indices.

2.3. The Assumptions in Theorem 1

When $d = 1$, the PLL reduces to the nontime-varying ordinary LL. In that case, $k = 0$ represents all time indices, because it is nontime varying. In Theorem 1, $L_0 \equiv \lim_{n\to\infty}\bar{Q}_0(n)$ is the limiting time average number of customers in the system, whereas $\lim_{n\to\infty}\bar{\lambda}_0(n) = \lambda_0$ is the limiting average number of arrivals at each time, and the right-hand side of (5) becomes

$$\sum_{j=0}^{\infty}\lambda_{k-j}F^c_{k-j,j} = \lambda_0\sum_{j=0}^{\infty}F^c_{0,j} = \lambda_0 W_0. \qquad (6)$$

Additionally, Theorem 1 claims that

$$L_0 = \lambda_0 W_0,$$

which is exactly the ordinary LL. Of course, the ordinary LL can be applied to the time-varying case as well, but then, we will lose the time structure and get overall averages.

There is a difference between our assumptions in (4) and the assumptions in the LL. For the LL, we let L be the limiting time average number in the system, λ be the limiting average arrival rate of customers, and W be the limiting customer average waiting time (time spent in the system or LoS). Then, if both λ and W exist and are finite, then L exists and is finite, and $L = \lambda W$. Our limit for $\bar{\lambda}_k(n)$ in (A1) is the natural extension; the only difference is that now we require that $\bar{\lambda}_k(n)$ converges w.p.1 for each k, $0 \le k \le d-1$. The third limit for $\bar{W}_k(n)$ in (A3) parallels the limit for the average waiting time, but again, we require that $\bar{W}_k(n)$ converges w.p.1 for each k, $0 \le k \le d-1$. However, these two limits alone are not sufficient to determine the number of customers for the periodic case. Now, we need to require that the LoS distribution converges for each k, $0 \le k \le d-1$, as stated in (A2). We show that this extra condition is needed in the following example.

Example 1 (The Need for Convergent ccdf's). We now show that we need to assume the limit $\bar{F}^c_{k,j}(n) \to F^c_{k,j}$ in (4). For simplicity, let $d = 2$, so that we have two time points in each periodic cycle. Suppose that we have two systems. In the first one, we deterministically have two arrivals at the first time of each periodic cycle (i.e., two arrivals at each even-indexed time), with one of them having LoS 0 and the other having LoS 2. In the other system, we also deterministically have two arrivals in the first time of each periodic cycle but with both of them having LoS 1. Suppose that there is no arrival at the odd indexed time for both of the systems. Now, the two systems have the same λ_k and W_k ($\lambda_0 = 2$, $\lambda_1 = 0$, $W_0 = 2$, and $W_1 = 0$). However, if we count the number of customers in the system, we have $\lim_{n\to\infty}\bar{Q}_0(n) = 3$ for the first system and $\lim_{n\to\infty}\bar{Q}_0(n) = 2$ for the second one.

2.4. Indirect Estimation of L_k via the PLL

The PLL in Theorem 1 provides an indirect way to estimate the long-run average occupancy level L_k through the right-hand side of (5) as discussed in Glynn and Whitt (1989b) for the ordinary LL. Here, we show that the indirect estimator for L_k is consistent with the direct estimator.

Because we only have data going forward in time from time 0, we start by rewriting (1) as

$$\sum_{j=0}^{\infty}\lambda_{k-j}F^c_{k-j,j} = \sum_{i=0}^{k}\lambda_i\sum_{l=0}^{\infty}F^c_{i,k-i+ld} + \sum_{i=k+1}^{d-1}\lambda_i\sum_{l=1}^{\infty}F^c_{i,k-i+ld}, \quad 0 \le k \le d-1. \qquad (7)$$

Guided by (7), we let our indirect estimator for L_k be

$$\bar{L}_k(n) \equiv \sum_{i=0}^{k}\bar{\lambda}_i(n)\sum_{l=0}^{\infty}\bar{F}^c_{i,k-i+ld}(n) + \sum_{i=k+1}^{d-1}\bar{\lambda}_i(n)\sum_{l=1}^{\infty}\bar{F}^c_{i,k-i+ld}(n), \quad 0 \le k \le d-1, \qquad (8)$$

where $\bar{\lambda}_i(n)$ and $\bar{F}^c_{i,j}(n)$ are defined in (3). With data, it is likely that the infinite sums in (8) would be truncated to finite sums but at a level growing with n; we do not address that truncation modification, which we regard as minor.

We now show that the estimator $\bar{L}_k(n)$ in (8) is asymptotically equivalent to the direct estimator $\bar{Q}_k(n)$ in (3); we will prove this result together with Theorem 1 in Section 4.2.

Theorem 2 (Indirect Estimation Through the PLL). *Under the conditions of Theorem 1,*

$$\lim_{n\to\infty}\bar{L}_k(n) = L_k \quad \text{w.p.1} \quad \textit{for} \quad 0 \le k \le d-1, \qquad (9)$$

where $\bar{L}_k(n)$ is defined in (8) and L_k is as in Theorem 1.

In applications, the LoS often can be considered to be bounded (i.e., for some $m > 1$, $X_{i,j} = 0$ when $j \ge md$). In that case, condition (A3) is directly implied by condition (A2), and it is possible to bound the error between the direct and indirect estimators for L_k, defined as

$$\bar{E}_k(n) \equiv |\bar{L}_k(n) - \bar{Q}_k(n)|, \qquad (10)$$

for $\bar{Q}_k(n)$ in (3) and $\bar{L}_k(n)$ in (8) as we show now.

Corollary 1 (The Bounded Case). *If, in addition to conditions (A1) and (A2) in Theorem 1, there exists some $m_u > 0$, such that $X_{i,j} = 0$ for $i \ge 0$, $j \ge dm_u$, then assumption (A3) is necessarily satisfied. If, in addition, there exists some $\lambda_u > 0$, such that $A_i \le \lambda_u$ for $i \ge 0$, then*

$$\bar{R}_n \equiv \max_{0\le k\le d-1}\{\bar{E}_k(n)\} \le \frac{\lambda_u d(m_u+2)^2}{2n}, \quad n \ge m_u, \qquad (11)$$

for $\bar{E}_k(n)$ in (10).

Proof. Here, we show the proof of the first part of the corollary (i.e., if the LoS is bounded, then assumption (A3) is implied from (A2)), and we postpone the second half of the proof to Section 4.3, because it depends on part of Proof of Theorem 1 and part of Proof of Theorem 2.

If $X_{i,j} = 0$ for $i \ge 0, j > dm_u$, then $\bar{F}_{k,j}(n) = 0$ for $0 \le k \le d-1$ and $j \ge dm_u$. Therefore,

$$\bar{W}_k(n) = \sum_{j=1}^{dm_u} \bar{F}^c_{k,j}(n), \quad 0 \le k \le d-1,$$

is a finite summation, and $F^c_{k,j} = 0$ for $0 \le k \le d-1$ and $j > dm_u$. Then,

$$\lim_{n\to\infty} \bar{W}_k(n) = \lim_{n\to\infty} \sum_{j=0}^{dm_u} \bar{F}^c_{k,j}(n) = \sum_{j=0}^{dm_u} \lim_{n\to\infty} \bar{F}^c_{k,j}(n) = \sum_{j=0}^{dm_u} F^c_{k,j} = W_k,$$

which is assumption (A3). □

Remark 2 (When (A2) Implies (A3)). In addition to the boundedness condition presented in Corollary 1, there are other mathematical conditions under which (A2) implies (A3) (i.e., under which we can interchange the order of the limits). Uniform integration is a standard condition for this purpose (cf. p. 185 of Billingsley 1995 and section 2.6 of El-Taha and Stidham 1999). We prefer (A3) plus (A2), because that makes our conditions easier to compare with the conditions in the ordinary LL.

2.5. Departure Processes

Other than relating the occupancy level with the arrival processes and the LoS as in the LL, we can also establish the relationship between the departure processes and the other quantities. This will also be helpful to understand the error between different ways of counting what happens at each time point as we will explain in Section 2.7.

Let $D_i \equiv \sum_{j=0}^{i} X_{i-j,j}, i \ge 0$, be the number of departures at time i. Given that the arrivals occur before departures at each time, it is easy to see that

$$D_i = Q_i - Q_{i+1} + A_{i+1} \quad \text{for} \quad i \ge 0. \tag{12}$$

Paralleling (3), we look at the averages

$$\bar{\delta}_k(n) \equiv \frac{1}{n}\sum_{m=1}^{n} D_{k+(m-1)d} \quad \text{for} \quad 0 \le k \le d-1. \tag{13}$$

Corollary 2 (Departure Averages). *Under the conditions of Theorem 1, $\bar{\delta}_k(n)$ defined in (13) converges w.p.1 as $n \to \infty$ to a periodic limit that we call δ_k. Moreover,*

$$\delta_k = \sum_{j=0}^{\infty} \lambda_{k-j} f_{k-j,j} \equiv \sum_{j=0}^{\infty} \lambda_{k-j}\,(F^c_{k-j,j} - F^c_{k-j,j+1}) \tag{14}$$

for $0 \le k \le d-1$, where λ_k and $F^c_{k,j}$ are the same periodic limits as in Theorem 1 and $f_{k,j} \equiv F^c_{k,j} - F^c_{k,j+1}$ is the discrete time probability mass function of the LoS.

Proof. The proof is easy given that we have Theorem 1 and Equation (12). By Equations (3) and (12),

$$\bar{\delta}_k(n) = \frac{1}{n}\sum_{m=1}^{n} D_{k+(m-1)d} = \frac{1}{n}\sum_{m=1}^{n} (Q_{k+(m-1)d} - Q_{k+1+(m-1)d} + A_{k+1+(m-1)d})$$
$$= \begin{cases} \bar{Q}_k(n) - \bar{Q}_{k+1}(n) + \bar{\lambda}_{k+1}(n), & 0 \le k < d-1, \\ \bar{Q}_{d-1}(n) - \bar{Q}_0(n+1) + \bar{\lambda}_0(n+1) + \frac{1}{n}Q_0 - \frac{1}{n}A_0, & k = d-1. \end{cases} \tag{15}$$

Because $\lim_{n\to\infty} \frac{1}{n} Q_0 = 0$ and $\lim_{n\to\infty}\frac{1}{n}A_0 = 0$, by Theorem 1 and (A1), we have

$$\begin{aligned}\lim_{n\to\infty} \bar{\delta}_k(n) &= L_k - L_{k+1} + \lambda_{k+1} \\ &= \sum_{j=0}^{\infty} \lambda_{k-j}F^c_{k-j,j} - \sum_{j=0}^{\infty}\lambda_{k+1-j}F^c_{k+1-j,j} + \lambda_{k+1} \\ &= \sum_{j=0}^{\infty}\lambda_{k-j}\,(F^c_{k-j,j} - f^c_{k-j,j+1}) - \lambda_{k+1} + \lambda_{k+1} \\ &= \sum_{j=0}^{\infty}\lambda_{k-j}f_{k-j,j}. \quad \square \end{aligned} \tag{16}$$

2.6. Connection to Cumulative Processes

The direct and indirect estimators of the occupancy level that we introduced in (3) and (8), respectively, can be related through the cumulative processes as depicted in Figure 2.

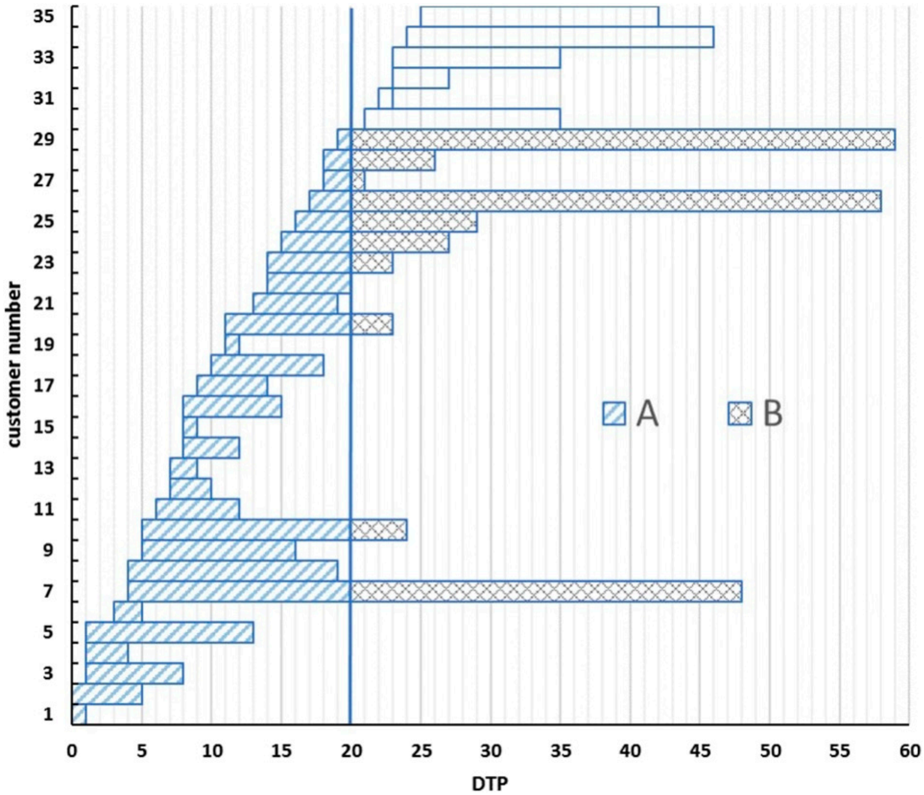
We focus on the two cumulative processes associated with the occupancy level and total LoS, respectively; that is,

$$\begin{aligned} C_Q(n) &\equiv \sum_{m=1}^{n}\sum_{k=0}^{d-1} Q_{k+(m-1)d} = \sum_{i=0}^{nd-1} Q_i, \\ C_L(n) &\equiv \sum_{m=1}^{n}\sum_{k=0}^{d-1}\sum_{j=0}^{\infty} (j+1)X_{k+(m-1)d,j} \\ &= \sum_{m=1}^{n}\sum_{k=0}^{d-1}\sum_{j=0}^{\infty} Y_{k+(m-1)d,j} \\ &= \sum_{i=0}^{nd-1}\sum_{j=0}^{\infty} Y_{i,j}. \end{aligned} \tag{17}$$

The first, $C_Q(n)$, is the cumulative occupancy level up to time n, whereas the second, $C_L(n)$, is the cumulative total LoS of customers that arrived up time n.

Figure 2 helps us understand the two cumulative quantities. In the figure, we plot the time intervals that each of the first 35 arrivals spends in the system as horizontal bars, each with height 1 placed in

Figure 2. (Color online) An Example of a Periodic Queueing System with $d = 5$ and $n = 4$



Notes. The vertical line is placed at discrete time period (DTP) $nd = 20$. Area A corresponds to $C_Q(4)$ in (18), whereas Area $A \cup B$ corresponds to $C_L(4)$ in (18).

order of the arrival times. The left end point is the arrival time, whereas the right end point is the departure time, which need not be in order of arrival. We can see that $C_Q(n)$ and $C_L(n)$ correspond to two areas.

We can further relate the two cumulative processes to the averages in (3) and (8) as stated in the following proposition.

Proposition 1. *The cumulative processes and the averages are related by*

$$Area(A) = C_Q(n) = n \sum_{k=0}^{d-1} \bar{Q}_k(n),$$
$$Area(A \cup B) = C_L(n) = n \sum_{k=0}^{d-1} \bar{\lambda}_k(n)\bar{W}_k(n) = n \sum_{k=0}^{d-1} \bar{L}_k(n). \tag{18}$$

Proof. The proof follows directly from the definitions, especially for $C_Q(n)$. For $C_L(n)$, observe that

$$C_L(n) = \sum_{m=1}^{n} \sum_{k=0}^{d-1} \sum_{j=0}^{\infty} Y_{k+(m-1)d,j} = n \sum_{k=0}^{d-1} \sum_{j=0}^{\infty} \bar{Y}_{k,j}(n)$$
$$= n \sum_{k=0}^{d-1} \sum_{j=0}^{\infty} \bar{\lambda}_k(n)\bar{F}^c_{k,j}(n) = n \sum_{k=0}^{d-1} \bar{\lambda}_k(n)\bar{W}_k(n). \tag{19}$$

By (8), if we sum over $0 \le k \le d-1$ and adjust the order of summation, we have

$$\sum_{k=0}^{d-1} \bar{L}_k(n) = \sum_{k=0}^{d-1} \sum_{j=0}^{\infty} \bar{\lambda}_k(n)\bar{F}^c_{k,j}(n) = \sum_{k=0}^{d-1} \bar{\lambda}_k(n)\bar{W}_k(n). \quad \square \tag{20}$$

In the context of Figure 2, Theorems 1 and 2 assert that

$$\bar{E}(n) \equiv \sum_{k=0}^{d-1} \bar{E}_k(n) = \text{Area}(B)/n \to 0 \quad \text{as} \quad n \to \infty, \tag{21}$$

where $\bar{E}_k(n)$ defined in (10).

We remark that Figure 2 is a variant of figure 1 in Whitt (1991) and figures 2 and 3 in Kim and Whitt (2013) as well as similar figures in earlier papers. The figures in Kim and Whitt (2013) are different because of the initial edge effect, which we avoid by treating arrivals before time 0 in the system as arrivals at time 0.

2.7. Different Orders of Events in Discrete Time

We have assumed that all arrivals occur before all departures at each time and that we count the number in the system after the arrivals but before the departures. That produces an upper bound on the system occupancy for all possible orderings. Suppose instead that all departures occur before all arrivals at each time and that we count the number in the system after the departures but before the arrivals. That obviously produces a lower bound. The consequence of any other ordering will fall in between these two.

For a discrete time system, the order may be given, but many applications start with a continuous time system. In that case, the modeler can choose which ordering to use in the discrete time version. We have chosen the conservative upper bound. In this section, we derive an expression for the alternative lower bound and the difference between the upper bound and the lower bound. From that difference, we can see that the difference between an initial continuous time system and a discrete time "approximation" will become asymptotically negligible as we refine the discrete time version by increasing the number of discrete time points within a fixed continuous time cycle.

For the alternative departure first (lower-bound) ordering, let Q_i^L be the number of customers in the system at time i, and let

$$\bar{Q}_k^L(n) \equiv \frac{1}{n}\sum_{m=1}^{n} Q_{k+(m-1)d}^L. \tag{22}$$

Assume that we keep the meaning of $X_{i,j}$ and $Y_{i,j}$.

Proposition 2 (The Lower-Bound Occupancy). *Under the conditions of Theorem 1, $\bar{Q}_k^L(n)$ converges* w.p.1 *as $n \to \infty$ to a limit that we call L_k^L, and we have*

$$L_k^L = L_k - \delta_k - 2\lambda_k + \lambda_k f_{k,0} \le L_k \quad \text{for} \quad 0 \le k \le d-1, \tag{23}$$

where L_k is in Theorem 1, λ_k is in (A1), and δ_k and $f_{k,0}$ are the same as in Corollary 2.

Proof. For the new departures first ordering, the number of customers at time i is

$$\begin{aligned} Q_i^L &= \sum_{j=1}^{i} Y_{i-j,j+1} - Y_{i,0} = \sum_{j=1}^{i} Y_{i-j,j} - \sum_{j=1}^{i} X_{i-j,j} - Y_{i,0} \\ &= \sum_{j=0}^{i} Y_{i-j,j} - \sum_{j=0}^{i} X_{i-j,j} - 2A_i + X_{i,0} \\ &= Q_i - D_i - 2A_i + X_{i,0}, \end{aligned} \tag{24}$$

where we regard the summation as 0 if the lower bound is larger than the upper bound. By Equations (22) and (24), we have

$$\begin{aligned} \bar{Q}_k^L(n) &= \frac{1}{n}\sum_{m=1}^{n} Q_{k+(m-1)d}^L \\ &= \frac{1}{n}\sum_{m=1}^{n}(Q_{k+(m-1)d} - D_{k+(m-1)d} - 2A_{k+(m-1)d} \\ &\quad + X_{k+(m-1)d,0}) \\ &= \bar{Q}_k(n) - \bar{\delta}_k(n) - 2\bar{\lambda}_k(n) + (\bar{Y}_{k,0}(n) - \bar{Y}_{k,1}(n)). \end{aligned} \tag{25}$$

Next, observe that (A1) and (A2) imply that $\lim_{n\to\infty} \bar{Y}_{k,j}(n) = \lambda_k F_{k,j}^c$, so that

$$\begin{aligned} \lim_{n\to\infty} \bar{Q}_k^L(n) &= L_k - \delta_k - 2\lambda_k + \lambda_k(F_{k,0}^c - F_{k,1}^c) \\ &= L_k - \delta_k - 2\lambda_k + \lambda_k f_{k,0}. \quad \square \end{aligned} \tag{26}$$

We can apply Proposition 2 to deduce that the error in an incorrect ordering of events is asymptotically negligible if we start with a continuous time system and choose sufficiently short time periods. For the arrival process $A(t)$, we assume that

$$t^{-1}A(t) \to \Lambda(t) \quad \text{as} \quad t \to \infty, \tag{27}$$

where

$$\Lambda(t) = \int_0^t \lambda(u)\,du \quad \text{and} \quad 0 < \lambda_U \le \lambda(u) \le \lambda_U < \infty \tag{28}$$

for all nonnegative u and t, with λ being a periodic function with period c. We make a similar assumption for the departure process $D(t)$ with the limiting rate being $\delta(t)$.

Corollary 3 (From Continuous to Discrete Time). *Suppose that we start with a continuous time periodic system with period c and construct a discrete time system by considering d evenly spaced time intervals within each periodic cycle. If the arrival and departure processes satisfy (27) and (28) and if the conditions of Theorem 1 hold for each d, then the error in the discrete time approximation owing to the order of events at each time point goes to zero as $d \to \infty$.*

Proof. From (23), we know that $0 \le L_k - L_k^L = \delta_k + 2\lambda_k - \lambda_k f_{k,0}$. Hence, if we consider a sequence of systems indexed by d, then the condition implies that $\lambda_{d,k} \to 0$ and $\delta_{d,k} \to 0$ as $d \to \infty$. $\square$

We can use Corollary 3 to define what we mean by a continuous time sample path PLL. We also exploit basic properties of Riemann integrals (section 6 of Rudin 1976).

Definition 1 (Continuous Time Sample Path PLL). Consider a continuous time system with arrival process having a periodic arrival rate and satisfying (27) and (28). A continuous time sample path PLL holds, yielding relation (2) if

(i) the discrete time PLL holds for each d when we form d time periods within each periodic cycle, and

(ii) the sequence of discrete time PLLs in (1) can be regarded as converging Riemann sum approximations for the continuous time relation in (2).

3. Steady-State Stochastic Versions of the PLL

We now discuss stochastic analogs of Theorem 1. In Section 3.1, we define a stationary framework in

Operations Research, 2019, vol. 67, no. 1, pp. 267–280, © 2019 INFORMS

continuous time. In Section 3.2, we derive a steady-state stochastic PLL by applying Theorem 1. In Section 3.3, we establish a version of PLL for the $G_t/GI_t/\infty$ time-varying infinite server model proposed in Whitt and Zhang (2017). In Section 3.4, we conduct additional analysis of the ED data to provide additional support for the stochastic model proposed in Whitt and Zhang (2017), despite the negligible support provided by Figure 1. Finally, in Section 3.5, we establish a continuous time stochastic PLL, which primarily follows from Rolski (1989) and Fralix and Riano (2010).

For this initial version, we aim for simplicity. Thus, we assume that the basic stochastic process is both stationary and ergodic, so that steady state means coincide with long-run averages. The main idea is that we now interpret the key quantities L_k and λ_k appearing in (1) as appropriate expected values of random variables associated with the system in periodic steady state. As we see it, there are two main issues:

(i) What is meant by periodic steady state?

(ii) What is $F^c_{k,j}$, or equivalently, what is the probability distribution of the LoS of an arbitrary arrival in period k?

3.1. Periodic Steady State

To construct periodic steady state, we assume that the basic stochastic process $\{\mathbf{Y}_n : n \in \mathbb{Z}\}$ with

$$\mathbf{Y}_n \equiv \{Y_{nd+k,j} : 0 \le k \le d-1; j \ge 0\} \tag{29}$$

introduced in Section 2.1 is a stationary sequence of nonnegative random elements indexed by the integer n. For each integer n, the random element $\mathbf{Y}_n$ takes values in the space $(Z^d)^\infty \equiv \mathbb{Z}^d \times \mathbb{Z}^d \times \cdots$ (chapter 6 of Breiman 1968, Baccelli and Bremaud 1994, and Sigman 1995 have background on stationary processes and their application to queues). Just as for the time-varying LL, as discussed in Fralix and Riano (2010), it is important to apply the Palm transformation, but we avoid that issue by exploiting the established limits for the averages.

Without loss of generality, we now regard our stochastic processes as stationary processes on the integers $\mathbb{Z}$, negative as well as positive (proposition 6.5 of Breiman 1968). As usual, we mean strictly stationary; that is, the finite-dimensional distributions are independent of time shifts, which in turn, means that, for each k and each k-tuple $(n_1, \ldots, n_k)$ of integers in $\mathbb{Z}$,

$$(\mathbf{Y}_{n_1}, \ldots, \mathbf{Y}_{n_k}) \stackrel{d}{=} (\mathbf{Y}_{n_1+m}, \ldots, \mathbf{Y}_{n_k+m}) \quad \text{for all} \quad m \in \mathbb{Z},$$

with $\stackrel{d}{=}$ denoting equality in distribution.

As a consequence of the stationarity assumed for $\{\mathbf{Y}_n : n \in \mathbb{Z}\}$, we also have stationarity for the associated stochastic process $\{(A_{nd+k}, Q_{nd+k}) : n \in \mathbb{Z}\}$, where A_{nd+k} is the number of arrivals at time $nd+k$ and Q_{nd+k} is the number of customers in the system at time $nd+k$, both of which are defined in Section 2.1, but now, we have stationarity on all integers, negative as well as positive. Thus, we have

$$A_{nd+k} \equiv Y_{nd+k,0} \quad \text{and} \quad Q_{nd+k} \equiv \sum_{j=0}^{\infty} Y_{nd+k-j,j}.$$

Hence, with some abuse of notation, we let $(\{Y_{k,j} : j \ge 0\}, A_k, Q_k)$ be a stationary random element. In this stochastic setting, we have

$$\lambda_k \equiv E[A_k] = E[Y_{k,0}] \quad \text{and}$$
$$L_k \equiv E[Q_k] = \sum_{j=0}^{\infty} E[Y_{k-j,j}]. \tag{30}$$

3.2. The Stochastic PLL

We now come to the second issue. In the stochastic setting, it remains to define $F^c_{k,j} \equiv P(W_k > j)$, where W_k is the time in the system for an arbitrary arrival in period k. It is natural to define $F^c_{k,j}$ by requiring that it agrees with the limit of the averages $\bar{F}^c_{k,j}(n)$ in (3). That limit is well defined if we assume that the basic sequence is ergodic as well as stationary, with $0 < E[Y_{k,0}] < \infty$ for all k.

With the stationary framework on all of the integers, positive and negative, the stochastic PLL becomes very elementary, because there are no edge effects.

Theorem 3 (Stochastic PLL). *Suppose that $\{\mathbf{Y}_n : n \in \mathbb{Z}\}$ in (29) is stationary and ergodic with $0 < \lambda_k \equiv E[Y_{k,0}] < \infty$, $0 \le k \le d-1$. Then, for each k, $0 \le k \le d-1$, $j \ge 0$,*

$$\bar{F}^c_{k,j}(n) \to F^c_{k,j} \equiv \frac{E[Y_{k,j}]}{E[Y_{k,0}]} \quad \text{w.p.1 as } n \to \infty, \tag{31}$$

and

$$L_k \equiv E[Q_k] = \sum_{j=0}^{\infty} \lambda_{k-j} F^c_{k-j,j}. \tag{32}$$

Proof. First, the stationary and ergodic condition with the specified moment assumptions implies that

$$\bar{Y}_{k,j}(n) \to E[Y_{k,j}] = \lambda_k F^c_{k,j} \quad \text{as} \quad n \to \infty \quad \text{w.p.1}$$

for all k and j in view of (3), (30), and the definition of $F^c_{k,j}$ in (31). Then, (31) follows immediately by continuity under the division operation by a quantity with a strictly positive limit. If we multiply and divide by λ_k within the representation for $E[Q_k]$ in (30), then we see that

$$E[Q_k] = \sum_{j=0}^{\infty} E[Y_{k-j,j}] = \sum_{j=0}^{\infty} \lambda_k \frac{E[Y_{k-j,j}]}{\lambda_k} = \sum_{j=0}^{\infty} \lambda_k F^c_{k-j,j}. \quad \square$$

In closing this section, we remark that the distribution $F_{k,j}$ can be seen to be correspond to an underlying Palm measure, but we do not develop that framework here (Sigman and Whitt 2018).

3.3. The Discrete-Time Periodic $G_t/GI_t/\infty$ Model

The candidate model for the ED proposed in Whitt and Zhang (2017) was a special case of the periodic $G_t/GI_t/\infty$ infinite server model, in which the LoS variables are mutually independent and independent of the arrival process, with a ccdf $F^c_{k,j} \equiv P(W_k \geq j)$ for a steady-state LoS in time period k that depends only on the time period k within the periodic cycle (a week). This strong local condition provides a sufficient condition for the steady-state stochastic PLL. That can be seen from the following proposition.

Proposition 3 (The $G_t/GI_t/\infty$ Special Case). *For the stationary $G_t/GI_t/\infty$ infinite server model specified above, where $\mathbf{Y} \equiv \{\mathbf{Y}_n : n \in \mathbb{Z}\}$ in (29) is strictly stationary with finite mean values, then*

$$E[Y_{k,j}] = F^c_{k,j} E[Y_{k,0}],$$

consistent with Equation (31).

Proof. Let $\{W_{k,i} : i \geq 1\}$ be a sequence of independent and identically distributed LoS variables associated with arrivals in time period k that is also independent of the arrival process and the other LoS variables. Under those conditions,

$$\begin{aligned} E[Y_{k,j}] &= \sum_{m=1}^{\infty} \sum_{i=1}^{m} P(W_{k,i} > j) P(A_k = m) \\ &= \sum_{m=1}^{\infty} m F^c_{k,j} P(A_k = m) = F^c_{k,j} E[A_k] \\ &= F^c_{k,j} E[Y_{k,0}]. \quad \square \end{aligned}$$

3.4. Additional Statistical Tests of the Infinite-Server Model

We now report results directly testing the GI_t assumption in the stochastic model proposed in Whitt and Zhang (2017). We briefly review the data analysis in the e-companion here.

Now, we first test whether the LoS distribution in period k can be regarded as being independent of the number of arrivals in period k. To be specific, let $A_k^{(m)}$ be the number of arrivals in hour k of week m, where in our ED case, $1 \leq k \leq 7 \times 24 = 168$ and m is from 1 to 25. Also, let $W_k^{(i)}$ be the average LoS of arrivals in hour k of week m. For each k, we compute the estimated (sample) Pearson correlation coefficients of $A_k^{(i)}$ and $W_k^{(m)}$ using samples where $A_k^{(m)} > 0$ (p. 169 of Casella and Berger 2002 has background). The plot at the top of Figure 3 shows the correlation coefficients (r_k) of all 168 hours in a week, where (if $A_k^{(m)} > 0$ for all m)

$$r_k = \frac{\sum_{m=1}^{25} (A_k^{(m)} - \bar{A}_k)(W_k^{(m)} - \bar{W}_k)}{\sqrt{\sum_{m=1}^{25} (A_k^{(m)} - \bar{A}_k)^2} \sqrt{\sum_{m=1}^{25} (W_k^{(m)} - \bar{W}_k)^2}},$$

where $\bar{A}_k \equiv (1/25)\sum_{i=1}^{25} A_k^{(m)}$ and $\bar{W}_k \equiv (1/25)\sum_{i=1}^{25} W_k^{(m)}$. For any m, such that $A_k^{(m)} = 0$, we remove those terms in the summation and the average correspondingly.

The middle left plot of Figure 3 compares the quantile of the coefficients with the normal distribution, and it indicates that the distribution of the 168 correlation coefficients is approximately normal, with sample mean 0.056 and sample standard error 0.21. It shows that there is no significant evidence indicating that the LoS distribution is related to the number of arrivals within each hour. The middle right plot of Figure 3 is a scatter plot of $(A_{36}^{(m)}, W_{36}^{(m)})$ (i.e., at noon on Monday, excluding samples with no arrivals), and the solid line is the mean of $W_{36}^{(m)}$ with the same number of arrivals. Finally, the bottom plot of Figure 3 shows the number of hours with no arrival for each hour in a week (i.e., the number of i, such that $A_k^{(m)} = 0$ for $k = 1, 2, \ldots, 7 \times 24 = 168$).

In summary, the statistical results in this section provide additional support for the stochastic model of the ED proposed in Whitt and Zhang (2017). In the e-companion to this paper, we review the data analysis in Whitt and Zhang (2017) and show the results of our studies that give evidence that the ED data are consistent with the conditions of Theorem 1 and Proposition 3.

3.5. Exploiting the Palm Theory in Continuous Time

Finally, we observe that the Palm theory in Rolski (1989) and Fralix and Riano (2010) also provides a general steady-state stochastic PLL in the conventional continuous time setting. For this steady-state stochastic result, we now assume that the arrival process is a simple point process (arrivals occur one a time) on the entire real line with a well-defined arrival rate $\lambda(t)$ at time t. This was satisfied in the ED, because the arrival times actually have very detailed time stamps.

Thus, letting $N(t) - N(s)$ be the number of arrivals in the interval $[s, t]$, we assume that

$$E[N(t)] - E[N(s)] = \int_s^t \lambda(s)\, ds, \tag{33}$$

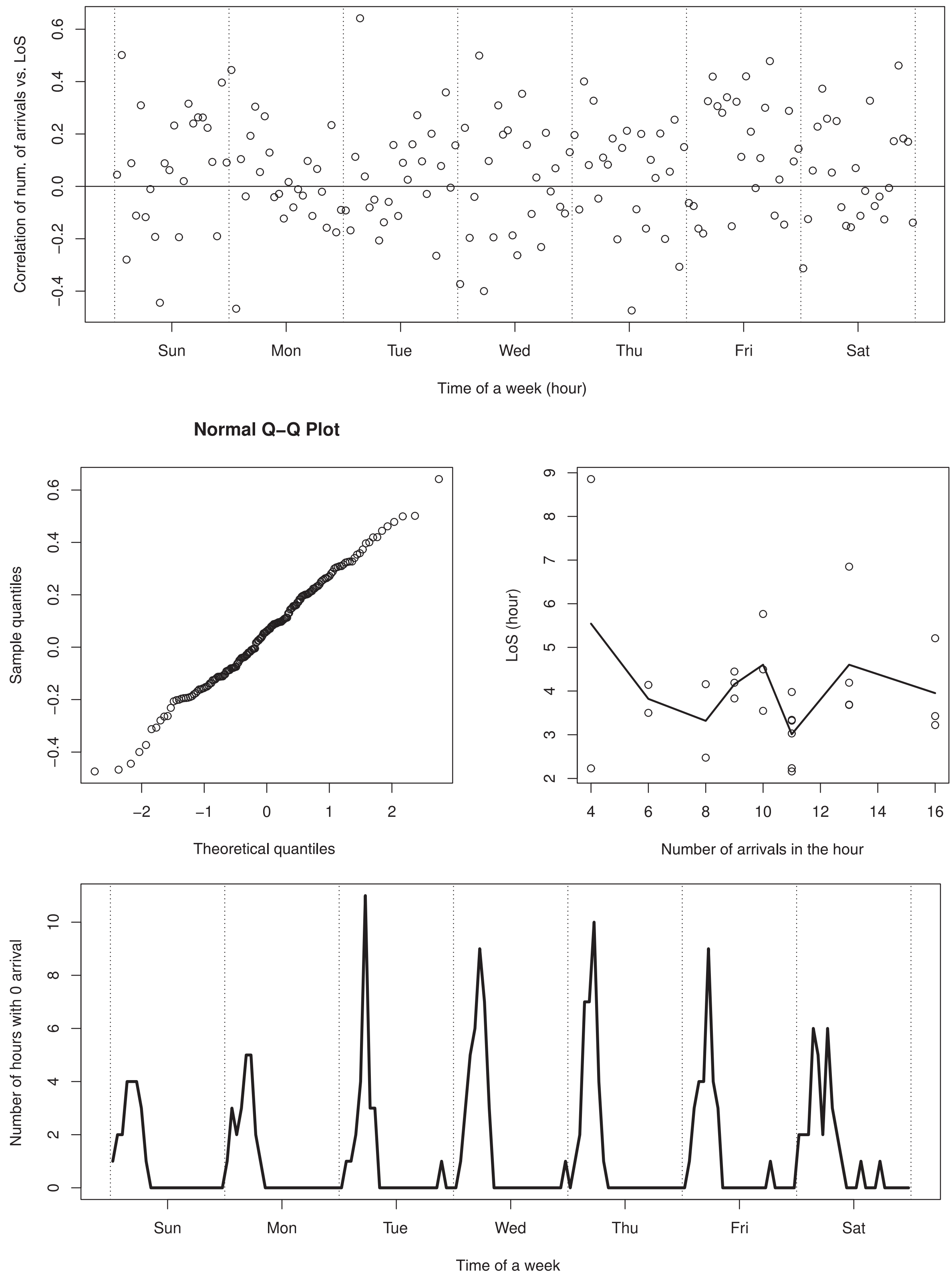
where the arrival rate function $\lambda(t)$ is a periodic function with periodic cycle length c, which is also right continuous with left limits.

As in section 2 of Fralix and Riano (2010), we let $W(t)$ be the waiting time of the last arrival before time t. That convention yields a well-defined waiting time process $\{W(t) : t \in \mathbb{R}\}$.

As a continuous time analog of the periodic stationarity assumed in Section 3.1, we assume that the queue length (number in system) process $Q \equiv \{Q(t) : t \in \mathbb{R}\}$ has a distribution that is invariant under time shifts by c. The arrival process is included by the upward jumps on Q. As a consequence of this c stationarity, the set of Palm measures $\{N_s : s \in \mathbb{R}\}$ associated with the arrival process N is periodic with period c. The mean queue length is expressed in terms of the tail probabilities

$$F^c_{t,x} \equiv P_t(W(t) > x) \tag{34}$$

Figure 3. Statistical Tests of the Infinite-Server Model



Notes. (Top) Estimated linear correlation between the number of arrivals and the mean LoS for each hour of a week. (Middle left) The *Q*–*Q* (quantile) plot of the correlation coefficients compared with a Gaussian distribution. (Middle right) An example of the relationship between the number of arrivals and the mean LoS at noon on Monday; the solid line is the average for each column of points. (Bottom) The number of hours with no arrival for each hour in a week.

under the Palm measures P_t, which are periodic with period c.

Then, paralleling the remark after theorem 3.1 of Fralix and Riano (2010), theorem 3.1 of Fralix and Riano (2010) implies the following continuous analog of (1), which was already given for the $M_t/GI/1$ special case in Rolski (1989).

Theorem 4 (Continuous Time PLL Following from Rolski 1989 and Fralix and Riano 2010). *Under the conditions above,*

$$E[Q(t)] = \int_0^\infty F^c_{t-s,s}\lambda(t-s)\,ds,$$

where $F^c_{t-s,s}$ *is defined in* (34).

4. Proofs

We now provide the postponed proofs of Lemma 1, Theorems 1 and 2, and Corollary 1. Here, all of the limits are in the sense of almost sure convergence. Hence, we can focus on one sample path where all of the limits exist.

4.1. Proof of Lemma 1

Proof. We will show that, under assumptions (A1), (A2), and (A3) in (4), we have $\lim_{n\to\infty}\bar{\lambda}_{k+ld}(n)$, $\lim_{n\to\infty}\bar{F}^c_{k+ld,j}(n)$, and $\lim_{n\to\infty}\bar{W}_{k+ld}(n)$ exist for all $0 \le k \le d-1$, $l \ge 0$, $j \ge 0$ and

$$\lim_{n\to\infty}\bar{\lambda}_{k+ld}(n) \equiv \lambda_{k+ld} = \lambda_k,$$
$$\lim_{n\to\infty}\bar{F}^c_{k+ld,j}(n) \equiv F^c_{k+ld,j} = F^c_{k,j}, \quad \text{and}$$
$$\lim_{n\to\infty}\bar{W}_{k+ld}(n) \equiv W_{k+ld} = W_k, \tag{35}$$

where λ_k, $F^c_{k,j}$, and W_k are the same constants as in (4).

By the definition of λ_k,

$$\begin{aligned}\lambda_k &= \lim_{n\to\infty}\bar{\lambda}_k(n) = \lim_{n\to\infty}\frac{1}{n}\sum_{m=1}^{n}A_{k+(m-1)d}\\ &= \lim_{n\to\infty}\frac{1}{n}\left(A_k + \sum_{m=1}^{n-1}A_{(k+d)+(m-1)d}\right)\\ &= \lim_{n\to\infty}\frac{n-1}{n}\,\frac{1}{n-1}\sum_{m=1}^{n-1}A_{(k+d)+(m-1)d}\\ &= \lambda_{k+d}.\end{aligned} \tag{36}$$

Next, by (3), we have $\bar{Y}_{k,j}(n) = \bar{\lambda}_k(n)\bar{F}^c_{k,j}(n)$. By assumptions (A1), (A2), and (A3) in (4), we know that $\lim_{n\to\infty}\bar{Y}_{k,j}(n) = \lambda_k F^c_{k,j}$ exists for all $0 \le k \le d-1$ and $j \ge 0$. Using the same argument as for λ_k, we know that $\lim_{n\to\infty}\bar{Y}_{k+d,j}(n) = \lim_{n\to\infty}\bar{Y}_{k,j}(n)$ for all $0 \le k \le d-1$, $j \ge 0$. Then,

$$F^c_{k+d,j} = \lim_{n\to\infty}\bar{F}^c_{k+d,j}(n) = \frac{\lim_{n\to\infty}\bar{Y}_{k+d,j}(n)}{\lim_{n\to\infty}\bar{\lambda}_{k+d}(n)} = \frac{\lambda_k F^c_{k,j}}{\lambda_k} = F^c_{k,j} \quad \text{for all } 0 \le k \le d-1,\ j \ge 0. \tag{37}$$

Similarly, for W_{k+d}, we have

$$W_{k+d} = \sum_{j=0}^{\infty}F^c_{k+d,j} = \sum_{j=0}^{\infty}F^c_{k,j} = W_k \quad \text{for all} \quad 0 \le k \le d-1. \tag{38}$$

By induction, we proved (35).

4.2. Proofs of Theorems 1 and 2

Proof. The proof is done in two steps. In Step 1, we show that $\lim_{n\to\infty}\bar{L}_k(n) = \sum_{j=0}^{\infty}\lambda_{k-j}F^c_{k-j,j}$ for $k = 0, 1, \ldots, d-1$, and then, in Step 2, we show that $L \equiv \lim_{n\to\infty}\bar{Q}_k(n) = \lim_{n\to\infty}\bar{L}_k(n)$, thus completing Proof of Theorem 1 and Proof of Theorem 2 together.

Step 1. Given a fixed $\epsilon > 0$, by assumption (A1), there exists N_1, such that, for any $n > N_1$, we have $\sup_{0\le k\le d-1}|\bar{\lambda}_k(n) - \lambda_k| < \epsilon$. Given assumption (A2), by the series form of Scheffé's lemma (p. 215 of Billingsley 1995), we know that assumption (A3) is equivalent to

$$\lim_{n\to\infty}\sum_{j=0}^{\infty}|\bar{F}^c_{k,j}(n) - F^c_{k,j}| = 0. \tag{39}$$

Therefore, there exists N_2, such that, for any $n > N_2$, we have $\sup_{0\le k\le d-1}\sum_{j=0}^{\infty}|\bar{F}^c_{k,j}(n) - \bar{F}^c_{k,j}| < \epsilon$. Let $N_3 = \max\{N_1, N_2\}$; then, when $n > N_3$,

$$\begin{aligned}&\left|\bar{L}_k(n) - \sum_{j=0}^{\infty}\lambda_{k-j}F^c_{k-j,j}\right|\\ &= \Bigg|\sum_{i=0}^{k}\bar{\lambda}_i(n)\sum_{l=0}^{\infty}\bar{F}^c_{i,k-i+ld}(n) + \sum_{i=k+1}^{d-1}\bar{\lambda}_i(n)\sum_{l=1}^{\infty}\bar{F}^c_{i,k-i+ld}(n)\\ &\quad -\sum_{i=0}^{k}\lambda_i\sum_{l=0}^{\infty}F^c_{i,k-i+ld} - \sum_{i=k+1}^{d-1}\lambda_i\sum_{l=1}^{\infty}F^c_{i,k-i+ld}\Bigg|\\ &= \Bigg|\sum_{j=0}^{k}\bar{\lambda}_{k-j}(n)\bar{F}^c_{k-j,j}(n) + \sum_{m=1}^{\infty}\sum_{j=1}^{d}\bar{\lambda}_{d-j}(n)\bar{F}^c_{d-j,(m-1)d+j+k}(n)\\ &\quad -\sum_{j=0}^{k}\lambda_{k-j}F^c_{k-j,j} - \sum_{m=1}^{\infty}\sum_{j=1}^{d}\lambda_{d-j}F^c_{d-j,(m-1)d+j+k}\Bigg|\\ &\le \Bigg|\sum_{j=0}^{k}\lambda_{k-j}\bar{F}^c_{k-j,j}(n) + \sum_{m=1}^{\infty}\sum_{j=1}^{d}\lambda_{d-j}\bar{F}^c_{d-j,(m-1)d+j+k}(n)\\ &\quad -\sum_{j=0}^{k}\lambda_{k-j}F^c_{k-j,j} - \sum_{m=1}^{\infty}\sum_{j=1}^{d}\lambda_{d-j}F^c_{d-j,(m-1)d+j+k}\Bigg|\\ &\quad +\epsilon\left(\sum_{j=0}^{k}\bar{F}^c_{k-j,j}(n) + \sum_{m=1}^{\infty}\sum_{j=1}^{d}\bar{F}^c_{d-j,(m-1)d+j+k}(n)\right)\end{aligned}$$

$$\leq \left(\max_{0\leq k\leq d-1}\lambda_k\right)\left(\sum_{j=0}^{k}|\bar{F}^c_{k-j,j}(n)-\bar{F}^c_{k-j,j}| + \sum_{m=1}^{\infty}\sum_{j=1}^{d}|\bar{F}^c_{d-j,(m-1)d+j+k}(n)-F^c_{d-j,(m-1)d+j+k}|\right) + \epsilon\left(\sum_{j=0}^{k}\bar{F}^c_{k-j,j}(n)+\sum_{m=1}^{\infty}\sum_{j=1}^{d}\bar{F}^c_{d-j,(m-1)d+j+k}(n)\right)$$

$$\leq \left(\max_{0\leq k\leq d-1}\lambda_k\right)d\epsilon + \epsilon d\left(\max_{0\leq k\leq d-1}W_k+\epsilon\right). \tag{40}$$

Hence, $\lim_{n\to\infty}|\bar{L}_k(n) - \sum_{j=0}^{\infty}\lambda_{k-j}F^c_{k-j,j}| = 0$ for all $0 \leq k \leq d-1$.

Step 2. Next, we show that $L_k \equiv \lim_{n\to\infty}\bar{Q}_k(n) = \lim_{n\to\infty}\bar{L}_k(n)$. To do so, actually, we will prove that

$$\bar{E}(n) \equiv \sum_{k=0}^{d-1}\bar{E}_k(n) \equiv \sum_{k=0}^{d-1}|\bar{L}_k(n)-\bar{Q}_k(n)| \to 0 \quad \text{as} \quad n\to\infty. \tag{41}$$

We further divide this step into two substeps. In the first substep, we compute the expression of $\bar{E}(n)$, and then, in the second substep, we show that it goes to zero as n goes to infinity.

Step 2.1. By some transformation, we know that

$$\begin{aligned}\bar{Q}_k(n) &\equiv \frac{1}{n}\sum_{m=1}^{n}Q_{k+(m-1)d} = \frac{1}{n}\sum_{m=1}^{n}\left(\sum_{j=0}^{k+(m-1)d}Y_{k+(m-1)d-j,j}\right)\\
&= \frac{1}{n}\sum_{m=1}^{n}\sum_{j=0}^{k}Y_{k-j+(m-1)d,j} + \frac{1}{n}\sum_{m=2}^{n}\sum_{j=k+1}^{k+(m-1)d}Y_{k-j+(m-1)d,j}\\
&= \frac{1}{n}\sum_{m=1}^{n}\sum_{j=0}^{k}Y_{k-j+(m-1)d,j} + \frac{1}{n}\sum_{m=1}^{n-1}\sum_{j=1}^{md}Y_{d-j+(m-1)d,j+k}\\
&= \frac{1}{n}\sum_{m=1}^{n}\sum_{j=0}^{k}Y_{k-j+(m-1)d,j} + \frac{1}{n}\sum_{m=1}^{n-1}\sum_{j=1}^{d}Y_{d-j+(m-1)d,j+k}\\
&\quad + \frac{1}{n}\sum_{m=1}^{n-2}\sum_{j=1}^{md}Y_{d-j+(m-1)d,j+k+d}\\
&= \frac{1}{n}\sum_{m=1}^{n}\sum_{j=0}^{k}Y_{k-j+(m-1)d,j}\\
&\quad + \frac{1}{n}\sum_{s=1}^{n-1}\sum_{m=1}^{n-s}\sum_{j=1}^{d}Y_{d-j+(m-1)d,j+k+(s-1)d}\end{aligned}$$

and

$$\begin{aligned}\bar{L}_k(n) &= \sum_{j=0}^{k}\bar{Y}_{k-j,j}(n) + \sum_{s=1}^{\infty}\sum_{j=1}^{d}\bar{Y}_{d-j,j+k+(s-1)d}(n)\\
&= \frac{1}{n}\sum_{m=1}^{n}\sum_{j=0}^{k}Y_{k-j+(m-1)d,j}\\
&\quad + \frac{1}{n}\sum_{s=1}^{\infty}\sum_{m=1}^{n}\sum_{j=1}^{d}Y_{d-j+(m-1)d,j+k+(s-1)d}.\end{aligned} \tag{42}$$

We may now study the absolute difference between $\bar{L}_k(n)$ and $\bar{Q}_k(n)$. Here,

$$\begin{aligned}\bar{E}_k(n) &\equiv |\bar{L}_k(n)-\bar{Q}_k(n)|\\
&= \frac{1}{n}\sum_{s=1}^{n-1}\sum_{m=n-s+1}^{n}\sum_{j=1}^{d}Y_{d-j+(m-1)d,j+k+(s-1)d}\\
&\quad + \frac{1}{n}\sum_{s=n}^{\infty}\sum_{m=1}^{n}\sum_{j=1}^{d}Y_{d-j+(m-1)d,j+k+(s-1)d}\\
&= \frac{1}{n}\sum_{m=1}^{n}\sum_{j=1}^{d}\sum_{s=n-m+1}^{\infty}Y_{d-j+(m-1)d,j+k+(s-1)d},\end{aligned} \tag{43}$$

and summing over $k = 0, 1, \ldots, d-1$ further gives

$$\begin{aligned}\bar{E}(n) &\equiv \sum_{k=0}^{d-1}\bar{E}_k(n) = \sum_{k=0}^{d-1}\frac{1}{n}\sum_{m=1}^{n}\sum_{j=1}^{d}\sum_{s=n-m+1}^{\infty}Y_{d-j+(m-1)d,j+k+(s-1)d}\\
&= \frac{1}{n}\sum_{m=1}^{n}\sum_{j=1}^{d}\sum_{s=(n-m)d}^{\infty}Y_{d-j+(m-1)d,j+s}.\end{aligned} \tag{44}$$

Step 2.2. Now it suffices to show that $\bar{E}(n)\to 0$ as $n\to\infty$. For that purpose, let N_1, N_2, and N_3 be the same as in the beginning of the proof, depending on given ϵ. Then, when $n > N_3$, we have

$$\begin{aligned}\left|\sum_{j=0}^{\infty}\bar{Y}_{k,j}(n) - \sum_{j=0}^{\infty}\lambda_k F^c_{k,j}\right| &= \left|\sum_{j=0}^{\infty}\bar{\lambda}_k(n)\bar{F}^c_{k,j}(n) - \sum_{j=0}^{\infty}\lambda_k F^c_{k,j}\right|\\
&\leq \left|\sum_{j=0}^{\infty}\lambda_k\bar{F}^c_{k,j}(n) - \sum_{j=0}^{\infty}\lambda_k F^c_{k,j}\right|\\
&\quad + \epsilon\sum_{j=0}^{\infty}\bar{F}^c_{k,j}(n)\\
&\leq \lambda_k\sum_{j=0}^{\infty}|\bar{F}^c_{k,j}(n) - F^c_{k,j}| + \epsilon(W_k+\epsilon)\\
&\leq \lambda_k\epsilon + \epsilon(W_k+\epsilon).\end{aligned} \tag{45}$$

Assumptions (A1) and (A2) indicate that $\lim_{n\to\infty}\bar{Y}_{k,j}(n) = \lambda_k F^c_{k,j}$. Again, by Scheffé's lemma, (45) is equivalent to

$$\lim_{n\to\infty}\sum_{j=0}^{\infty}|\bar{Y}_{k,j}(n) - \lambda_k F^c_{k,j}| = 0. \tag{46}$$

For any $\epsilon > 0$, because $\sum_{j=0}^{\infty}\lambda_k F^c_{k,j}$ converges for each $k = 0, 1, \ldots, d-1$, we know that there exists J, such that

$$\sum_{k=0}^{d-1}\sum_{j=J}^{\infty}\lambda_k F^c_{k,j} < \epsilon.$$

Let $N_4 \equiv \lceil J/d \rceil$, where $\lceil x \rceil$ means the smallest integer greater than x; then, when $n \geq N_4$, we have $nd > J$. By (46), there exists N_5, such that, when $n > N_5$,

$$\sum_{k=0}^{d-1} \sum_{j=0}^{\infty} |\bar{Y}_{k,j}(n) - \lambda_k F_{k,j}^c| < \epsilon.$$

Also, let $N_6 \equiv \max\{N_4, N_5\}$; then, when $n \geq N_6$,

$$\sum_{k=0}^{d-1} \sum_{j=N_6 d}^{\infty} \bar{Y}_{k,j}(n) \leq 2\epsilon. \tag{47}$$

Additionally, let $N_7 \equiv \lceil N_6/\epsilon \rceil$; then, when $n > N_7$, we have $(n - N_6)/n > 1 - \epsilon$. Finally, when $n > \max\{N_7, 2N_6\}$,

$$\begin{aligned}
\bar{E}(n) &= \frac{1}{n} \sum_{m=1}^{n} \sum_{j=1}^{d} \sum_{s=(n-m)d}^{\infty} Y_{d-j+(m-1)d,j+s} \\
&= \sum_{j=1}^{d} \frac{1}{n} \sum_{m=1}^{n-N_6} \sum_{s=(n-m)d}^{\infty} Y_{d-j+(m-1)d,j+s} \\
&\quad + \sum_{j=1}^{d} \frac{1}{n} \sum_{m=n-N_6+1}^{n} \sum_{s=(n-m)d}^{\infty} Y_{d-j+(m-1)d,j+s} \\
&\leq \sum_{j=1}^{d} \frac{1}{n} \sum_{m=1}^{n-N_6} \sum_{s=N_6 d}^{\infty} Y_{d-j+(m-1)d,j+s} \\
&\quad + \sum_{j=1}^{d} \frac{1}{n} \sum_{m=n-N_6+1}^{n} \sum_{s=0}^{\infty} Y_{d-j+(m-1)d,s} \\
&\leq \sum_{j=1}^{d} \frac{1}{n} \sum_{m=1}^{n} \sum_{s=N_6 d}^{\infty} Y_{d-j+(m-1)d,j+s} \\
&\quad + \sum_{j=1}^{d} \frac{1}{n} \sum_{m=n-N_6+1}^{n} \sum_{s=0}^{\infty} Y_{d-j+(m-1)d,s} \\
&= \sum_{j=1}^{d} \sum_{s=N_6 d}^{\infty} \bar{Y}_{d-j,j+s}(n) \\
&\quad + \sum_{j=1}^{d} \sum_{s=0}^{\infty} \left(\bar{Y}_{d-j,s}(n) - \frac{n-N_6}{n} \bar{Y}_{d-j,s}(n-N_6) \right) \\
&\leq 2\epsilon + 2\epsilon + \epsilon \left(\sum_{k=0}^{d-1} \sum_{j=0}^{\infty} \lambda_k F_{k,j}^c + \epsilon \right).
\end{aligned} \tag{48}$$

The inequality in the third line is just relaxing the index s, the next inequality is relaxing the index m for the first term, and the last inequality is given by the definition of N_6 and N_7, where the first term is bounded by 2ϵ by using (47) and the second term is bounded by

$$\begin{aligned}
&\sum_{j=1}^{d} \sum_{s=0}^{\infty} \left(\bar{Y}_{d-j,s}(n) - \frac{n-N_6}{n} \bar{Y}_{d-j,s}(n-N_6) \right) \\
&\leq \sum_{j=1}^{d} \sum_{s=0}^{\infty} (\bar{Y}_{d-j,s}(n) - (1-\epsilon)\bar{Y}_{d-j,s}(n-N_6)) \\
&= \sum_{j=1}^{d} \sum_{s=0}^{\infty} (\bar{Y}_{d-j,s}(n) - \bar{Y}_{d-j,s}(n-N_6)) + \epsilon \bar{Y}_{d-j,s}(n-N_6) \\
&\leq 2\epsilon + \epsilon \left(\sum_{k=0}^{d-1} \sum_{j=0}^{\infty} \lambda_k F_{k,j}^c + \epsilon \right).
\end{aligned} \tag{49}$$

Therefore, we have proved that $L_k \equiv \lim_{n\to\infty} \bar{Q}_k(n) = \lim_{n\to\infty} \bar{L}_k(n)$, which completes Proof of Theorem 1 and Proof of Theorem 2.

4.3. Proof of the Second Half of Corollary 1

Proof. Here, we give the second half of Proof of Corollary 1 (i.e., the explicit bound in (11)).

From Equation (43), we know that

$$\bar{E}_k(n) = \frac{1}{n} \sum_{m=1}^{n} \sum_{j=1}^{d} \sum_{s=n-m+1}^{\infty} Y_{d-j+(m-1)d,j+k+(s-1)d}.$$

By the condition, we know that $Y_{i,j} = 0$ for $i \geq 0$ and $j \geq dm_u$. Hence, when $n \geq m_u$,

$$\begin{aligned}
\bar{E}_k(n) &= \frac{1}{n} \sum_{j=1}^{d} \sum_{m=n-m_u}^{n} \sum_{s=n-m+1}^{m_u+1} Y_{d-j+(m-1)d,j+k+(s-1)d} \\
&\leq \frac{1}{n} \sum_{j=1}^{d} \sum_{m=n-m_u}^{n} \sum_{s=n-m+1}^{m_u+1} A_{d-j+(m-1)d} \\
&\leq \frac{1}{n} \sum_{j=1}^{d} \sum_{m=n-m_u}^{n} \sum_{s=n-m+1}^{m_u+1} \lambda_u \\
&= \frac{1}{n} \lambda_u d \frac{(m_u+1)(m_u+2)}{2} \leq \frac{\lambda_u d (m_u+2)^2}{2n}.
\end{aligned} \tag{50}$$

Therefore, we have proved Corollary 1.

5. Conclusions

In Sections 2 and 3, we have established sample path and stationary versions of a PLL; we think that they can add insight into the performance of periodic stochastic models, which are natural for many service systems. In particular, these new theorems explain the extraordinary model fit that we found in our data analysis of an ED in Whitt and Zhang (2017), which is shown in Figure 1. Nevertheless, in Section 3.4, we present additional evidence supporting the infinite server model proposed in Whitt and Zhang (2017).

There are many directions for future research. We ourselves have already established a central limit theorem (CLT) version of the PLL in Whitt and Zhang (2019), which parallels the CLT versions of LL in Glynn and Whitt (1986, 1987, 1988) and Whitt (2012); these have important statistical applications as in Glynn and Whitt (1989b) and Kim and Whitt (2013).

Related to Section 3, there should be more related good theory to develop associated with discrete time and periodic Palm measures and their application to queues, supplementing Whitt (1983), Miyazawa and Takahashi (1992), section 1.7.4 of Baccelli and Bremaud (1994), and section 1.4 and appendix D of Sigman (1995). Indeed, a contribution is in Sigman and Whitt (2018).

As noted in Glynn and Whitt (1989a), Whitt (1991), and El-Taha and Stidham (1999), there are many important generalizations of LL, such as the relation $H = \lambda G$. It remains to establish such results in a periodic setting (Sigman and Whitt 2018).

References

Armony M, Israelit S, Mandelbaum A, Marmor Y, Tseytlin Y, Yom-Tov G (2015) On patient flow in hospitals: A data-based queueing-science perspective. *Stochastic Systems* 5(1):146–194.

Baccelli F, Bremaud P (1994) *Elements of Queueing Theory* (Springer, Berlin).

Bertsimas D, Mourtzinou G (1997) Transient laws of non-stationary queueing systems and their applications. *Queueing Systems* 25(1–4):115–155.

Billingsley P (1995) *Probability and Measure*, 3rd ed. (Wiley, New York).

Breiman L (1968) *Probability* (Addison-Wesley, New York).

Bruneel H, Kim BG (1993) *Discrete-Time Models for Communication Systems Including ATM* (Kluwer Academic, Dordrecht, Netherlands).

Casella G, Berger RL (2002) *Statistical Inference*, 2nd ed. (Duxbury, Pacific Grove, CA).

El-Taha M, Stidham S (1999) *Sample-Path Analysis of Queueing Systems* (Kluwer, Boston).

Fiems D, Bruneel H (2002) A note on the discretization of Little's result. *Oper. Res. Lett.* 30(1):17–18.

Fralix BH, Riano G (2010) A new look at transient versions of Little's law. *J. Appl. Probab.* 47(2):459–473.

Glynn PW, Whitt W (1986) A central-limit-theorem version of L = λW. *Queueing Systems* 2(1):191–215.

Glynn PW, Whitt W (1987) Sufficient conditions for functional limit theorem: Versions of L = λW. *Queueing Systems* 1(3):279–287.

Glynn PW, Whitt W (1988) Ordinary CLT and WLLN versions of L = λW. *Math. Oper. Res.* 13(4):674–692.

Glynn PW, Whitt W (1989a) Extensions of the queueing relations L = λW and H = λG. *Oper. Res.* 37(4):634–644.

Glynn PW, Whitt W (1989b) Indirect estimation via L = λW. *Oper. Res.* 37(1) 82–103.

Kim S, Whitt W (2013) Statistical analysis with Little's law. *Oper. Res.* 61(4):1030–1045.

Kim S, Whitt W (2014a) Are call center and hospital arrivals well modeled by nonhomogeneous Poisson processes? *Manufacturing Service Oper. Management* 16(3):464–480.

Kim S, Whitt W (2014b) Choosing arrival process models for service systems: Tests of a nonhomogeneous Poisson process. *Naval Res. Logist.* 61(1):66–90.

Little JDC (1961) A proof of the queueing formula: L = λW. *Oper. Res.* 9(3):383–387.

Little JDC (2011) Little's law as viewed on its 50th anniversary. *Oper. Res.* 59(3):536–539.

Miyazawa M, Takahashi Y (1992) Rate conservation principle for discrete-time queues. *Queueing Systems* 12(3/4): 215–230.

Rolski T (1989) Relationships between characteristics in periodic Poisson queues. *Queueing Systems* 4(1):17–26.

Rudin W (1976) *Principles of Mathematical Analysis*, 3rd ed. (McGraw-Hill, New York).

Sigman K (1995) *Stationary Marked Point Processes: An Intuitive Approach* (Chapman and Hall/CRC, New York).

Sigman K, Whitt W (2018) Marked point processes in discrete time. Working paper, Columbia University, New York. Available at http://www.columbia.edu/~ww2040/allpapers.html.

Stidham S (1974) A last word on L = λW. *Oper. Res.* 22(2):417–421.

Whitt W (1983) Comparing batch delays and customer delays. *Bell System Tech. J.* 62(7):2001–2009.

Whitt W (1991) A review of L = λW. *Queueing Systems* 9(3):235–268.

Whitt W (1992) Correction note on L = λW. *Queueing Systems* 12(4): 431–432.

Whitt W (2012) Extending the FCLT version of L = λW. *Oper. Res. Lett.* 40(4):230–234.

Whitt W, Zhang X (2017) A data-driven model of an emergency department. *Oper. Res. Health Care* 12(1):1–15.

Whitt W, Zhang X (2019) A central-limit-theorem version of the periodic Little's law. *Queueing Systems* 91(1-2):15–47.

Wolfe RW, Yao Y-C (2014) Little's law when the average waiting time is infinite. *Queueing Systems* 76(3):267–281.

Ward Whitt is a professor in the Industrial Engineering and Operations Research Department at Columbia University. A major focus of his early work was the Queueing Network Analyzer performance analysis software tool, which is described in a 1983 paper in the *Bell Labs Technical Journal*. His new research explores ways to develop more effective approximations, drawing on new robust optimization methods as well as previous heavy traffic limits and indices of dispersion.

Xiaopei Zhang is a doctoral student in the Industrial Engineering and Operations Research Department at Columbia University. His primary research focus is on queueing theory, applied probability, and their applications to healthcare systems using stochastic modelling and simulation as well as statistics.