Queueing System manuscript No.
(will be inserted by the editor)

A Central-Limit-Theorem Version of the Periodic Little's Law

Ward Whitt · Xiaopei Zhang

Received: June 27, 2017; revision: May 10, 2018

Abstract We establish a central-limit-theorem (CLT) version of the periodic Little's law (PLL) in discrete time, which complements the sample-path and stationary versions of the PLL we recently established, motivated by data analysis of a hospital emergency department. Our new CLT version of the PLL extends previous CLT versions of LL. As with the LL, the CLT version of the PLL is useful for statistical applications.

Keywords Little's law · $L = \lambda W$ · periodic queues · central limit theorem · emergency departments · weak convergence in $(\ell_1)^d$

1 Introduction

Little's law ($L = \lambda W$) states that, under weak conditions, the long-run average number of customers in a system (L) is equal to the long-run average arrival rate (λ) multiplied by the long-run customer-average sojourn time in the system (W). Little's law (LL) provides an important consistency check, like double-entry bookkeeping. Such consistency checks are often regarded as trivial, because they are quite intuitive, but it has been suggested that the 1494 book by Luca Pacioli [24], which contains the first codified account of double-entry bookkeeping, might be the most influential work in the history of capitalism; see [14] and [17]. It evidently took Philip M. Morse to realize that it would be good to have a proof of LL; see the endnote by John Little in [21].

Ward Whitt

Industrial Engineering and Operations Research Department, Columbia University

E-mail: ww2040@columbia.edu

Xiaopei Zhang

Industrial Engineering and Operations Research Department, Columbia University

E-mail: xz2363@columbia.edu

After the seminal paper by Little in [19], LL has been further studied, notably by Stidham [25], and is frequently used as a fundamental tool in queuing theory; see [7,20,26,27] for general review. In [9,10,29], Glynn and Whitt established a functional central limit theorem (FCLT) version of LL by exploiting the continuous mapping theorem, and as a corollary, a CLT version of LL, by applying a projection. In [11], they derived a CLT version directly, without exploiting FCLT's. The CLT version of LL provides the convergence rate in the sample-path version of LL and has important statistical applications, as discussed in [12,16].

In [31], we established a sample-path periodic Little's law (PLL) in discrete time, extending the sample-path version of LL, and a stationary PLL, extending the time-varying Little's law (TVLL) in [3,8], which refines the LL in another direction. In doing so, we were motivated by our statistical analysis of patient arrival and departure data from the emergency department (ED) of an Israeli hospital, using 25 weeks of data from the data repository associated with the study by Armony et al. [2]. Based on our data analysis in [30], we concluded that stochastic models of that ED should be periodic with the week serving as the relevant period.

In the present paper we establish a CLT version of the PLL in discrete time. With the periodic structure, it is natural to think about the relation between the direct estimator of the occupancy level, obtained by directly averaging over periods, and the indirect estimator based on the arrival process and length of stay (LoS) via the PLL. The main story of this paper is that, given a joint CLT for the CLT-scaled arrival process and LoS, the CLT-scaled indirect estimator also converges, but we need more conditions to ensure the CLT-scaled direct estimator having the same limit. We give both a simple practical version, assuming bounded arrivals and LoS distributions, and a more general version without the boundedness restriction, but involving more complex mathematics. We also provide reasonable sufficient conditions such that the two estimators are asymptotically equivalent.

For the PLL in [31], just as for the TVLL in [3,8], the relation requires considering the entire LoS distribution function instead of just the mean LoS. The proof of the main theorem here is still by the continuous mapping theorem, but the analysis here for the general case with unbounded distributions is nonstandard. In particular, in order to directly exploit the continuous mapping theorem, we use the Banach space ℓ_1, which includes all the absolutely summable sequences. While the Banach space ℓ_1 is standard within functional analysis, it is not standard within probability theory. Neverthelesss, there is substantial literature for us to draw upon; e.g., [1,18,22]. We specialize the general weak convergence theory to this specific space ℓ_1 and give a sufficient condition for weak convergence in this space. We also specialize the CLT for i.i.d. random elements in general Banach spaces to ℓ_1, which may be useful to find new conditions for the CLT version of PLL as well as in other contexts.

This paper as organized as follows: In §2 we review the sample-path PLL from [31]. In §3 we state the new CLT versions of the PLL and then we discuss the statistical applications. In §4 we establish sufficient conditions for the general CLT version of the PLL. In §5, we discuss the weak convergence theory for random elements of $(\ell_1)^d$. Finally, in §6 we provide the longer proofs.

2 Review of the Periodic Little's Law

In this section we review the sample-path PLL from [31]. We use the notation introduced there. We consider discrete time points indexed by the nonnegative integers k. Since multiple events can happen at each time, we need to carefully specify the order of events. We assume that all arrivals at each time occur before any departures. Moreover, we count the number of customers in the system (patients in the ED) at each time after the arrivals but before the departures. Thus, each arrival can spend time j in the system for any $j \geq 0$. We discuss other orders of events in §2.7 of [31].

We start with 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 leave the system at time $i + j$. 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 derive all the other quantities in terms of this sequence. In particular, with $\equiv$ denoting equality by definition, let:

$$Y_{i,j} \equiv \sum_{s=j}^{\infty} X_{i,s}, \quad \text{the number of arrivals at time } i \text{ with LoS greater or equal to } j, j \geq 0,$$

$$A_i \equiv Y_{i,0} = \sum_{s=0}^{\infty} X_{i,s}, \quad \text{the total number of arrivals at time } i,$$

$$Q_i \equiv \sum_{j=0}^{i} Y_{i-j,j} = \sum_{j=0}^{i} A_{i-j} \frac{Y_{i-j,j}}{A_{i-j}}, \, j \geq 0, \quad \text{the number in system at time } i, \quad \text{and}$$

$$D_i \equiv \sum_{j=0}^{i} X_{i-j,j} = Q_i - Q_{i+1} + A_{i+1}, \quad \text{the number of departures at time } i, \quad i \geq 0.$$

In the third line 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:

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

where d is a positive integer. (Note that our conventions about the order of events implies that an arrival can depart in the same period, in which case the time spent in the system is counted as 1. That explains why the final sum above starts at 0 instead of at 1.)

Clearly, $\bar{\lambda}_k(n)$ and $\bar{\delta}_k(n)$ are the average arrival and departure rates, respectively, at time k within a period, averaged over n periods; we think of it applying to all the time periods $(m-1)d+k$ for $0 \le k \le d-1$ and $m \ge 1$. Similarly, $\bar{Q}_k(n)$ is the average number of customers in the system at time k, within a period, averaged over n periods; while $\bar{Y}_{k,j}(n)$ is the average number of customers that arrive at time k that have a LoS greater or equal to j. Thus, $\bar{F}^c_{k,j}(n)$ is the empirical complementary cumulative distribution function (ccdf), which is the natural estimator of the LoS ccdf of an arrival in time period k. Finally, $\bar{W}_k(n)$ is the sample mean LoS of customers that arrive at time k within a period, averaged over n periods. We write n as a parameter to indicate that the estimator is computed by averaging over n periodic cycles. We will let $n \to \infty$.

We make the following three assumptions, which parallel or extend the assumptions used in the ordinary Little's law. We assume that

$$\begin{aligned}
&(A1) \quad \bar{\lambda}_k(n) \to \lambda_k, \quad \text{w.p.1} \quad \text{as} \quad n \to \infty, \quad 0 \le k \le d-1,\\
&(A2) \quad \bar{F}^c_{k,j}(n) \to F^c_{k,j}, \quad \text{w.p.1} \quad \text{as} \quad n \to \infty, \quad 0 \le k \le d-1, \quad j \ge 0, \quad \text{and}\\
&(A3) \quad \bar{W}_k(n) \to W_k \equiv \sum_{j=0}^{\infty} F^c_{k,j} \quad \text{w.p.1} \quad \text{as} \quad n \to \infty, \quad 0 \le k \le d-1,
\end{aligned} \tag{2}$$

where the limits are deterministic and finite.

Paralleling the assumptions in the LL [25], assumptions ($A1$) and ($A3$) state that the average arrival rates and LoS converge, but for each k because of the extension to the periodic case. Assumption ($A2$) has no counterpart in the LL; it requires that the empirical ccdfs converge. Lemma 1 in [31] shows that if the three assumptions in (2) hold, then the limits hold for all $k \geq 0$, with the limit functions being periodic with period d. We extend these periodic functions to the entire real line, including the negative time indices.

To focus on the indirect estimator, we also add another estimator. It has a more complex form to account for the fact that in practice we only have data going forward in time. In particular, let

$$\bar{L}_k(n) \equiv \sum_{i=0}^{k} \bar{\lambda}_i(n) \sum_{l=0}^{\infty} \bar{F}^c_{i,k-i+ld}(n) + \sum_{i=k+1}^{d-1} \bar{\lambda}_i(n) \sum_{l=1}^{\infty} \bar{F}^c_{i,k-i+ld}(n), \quad 0 \leq k \leq d-1, \tag{3}$$

where $\bar{\lambda}_i(n)$ and $\bar{F}^c_{i,j}(n)$ are defined in (1).

The following combines Theorems 1 and 2 and Corollary 2 of [31].

Theorem 2.1 (*sample-path PLL from [31]*) *If the three assumptions* ($A1$), ($A2$) *and* ($A3$) *in* (2) *hold, then* $(\bar{Q}_k(n), \bar{\delta}_k(n), \bar{L}_k(n))$ *defined in* (1) *and* (3) *converges w.p.1 in* $\mathbb{R}^3$ *as* $n \to \infty$ *to a limit that we denote by* (L_k, δ_k, L_k). *Moreover,*

$$\begin{aligned} L_k &= \sum_{j=0}^{\infty} \lambda_{k-j} F^c_{k-j,j} < \infty \quad \text{and} \\ \delta_k &= \sum_{j=0}^{\infty} \lambda_{k-j} f_{k-j,j} \equiv \sum_{j=0}^{\infty} \lambda_{k-j} (F^c_{k-j,j} - F^c_{k-j,j+1}) \end{aligned} \tag{4}$$

for $0 \leq k \leq d-1$, *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, while* $f_{k,j} \equiv F^c_{k,j} - F^c_{k,j+1}$ *is the LoS probability mass function.*

The fact that the first and third terms of the limit (L_k, δ_k, L_k) coincide in Theorem 2.1 implies that, without extra conditions, the direct and indirect estimators are consistent. In this paper, we develop the CLT version, showing that the stochastic limit of the CLT-scaled versions are the same random variable under stronger conditions. As in [9], this can be understood by recognizing that there is an important link between associated cumulative processes, which lies behind the relation between the averages in LL; see §2.6 of [31].

3 Central-Limit-Theorem Version of the PLL

We now establish the CLT versions of the PLL, paralleling the CLT versions in [9,10,11,29]. We look for the relationship between the CLT-scaled arrival rates, LoS distributions and occupancy level in the periodic setting, so we now require stronger assumptions than for the sample-path PLL in Theorem 2.1, but we obtain a rate of convergence for the sample-path PLL. We will show that linking the indirect estimator of occupancy level with the arrival rates and LoS distributions is straightforward by the continuous mapping theorem, however, stronger conditions are needed such that the CLT-scaled direct estimator converges to the same limit.

One simple and practical case, where we assume that both the number of arrivals at each time and the LoS of each arrival are bounded and the limits are Gaussian, is provided first in §3.2. It applies to the ED in [30] and evidently to most practical cases. Then two more complicated sufficient conditions are stated in §4. The CLT versions of PLL also have statistical applications, as in [12,16], which we discuss in §3.4.

3.1 More Notation and Definitions

In this section we assume the LoSs are bounded by J, i.e. $X_{i,j} = 0$ for $j \geq J$ and all i. All the vectors are understood to be column vectors.

For $0 \leq k \leq d-1$, let

$$\begin{aligned}\bar{F}^c_k(n) &\equiv (\bar{F}^c_{k,0}(n), \bar{F}^c_{k,1}(n), \cdots, \bar{F}^c_{k,J-1}(n)) \in \mathbb{R}^J,\\ F^c_k &\equiv (F^c_{k,0}, F^c_{k,1}, \cdots, F^c_{k,J-1}) \in \mathbb{R}^J.\end{aligned} \tag{5}$$

Let the law-of-large-numbers-scaled (LLN-scaled) averages be

$$\begin{aligned}\bar{\boldsymbol{\lambda}}(n) &\equiv (\bar{\lambda}_0(n), \bar{\lambda}_1(n), \cdots, \bar{\lambda}_{d-1}(n)) \in \mathbb{R}^d,\\ \bar{\boldsymbol{\delta}}(n) &\equiv (\bar{\delta}_0(n), \bar{\delta}_1(n), \cdots, \bar{\delta}_{d-1}(n)) \in \mathbb{R}^d,\\ \bar{\boldsymbol{F}}^{\boldsymbol{c}}(n) &\equiv (\bar{F}^c_0(n)^T, \bar{F}^c_1(n)^T, \cdots, \bar{F}^c_{d-1}(n)^T) \in \mathbb{R}^{d\times J},\\ \bar{\boldsymbol{W}}(n) &\equiv (\bar{W}_0(n), \bar{W}_1(n), \cdots, \bar{W}_{d-1}(n)) \in \mathbb{R}^d,\\ \bar{\boldsymbol{Q}}(n) &\equiv (\bar{Q}_0(n), \bar{Q}_1(n), \cdots, \bar{Q}_{d-1}(n)) \in \mathbb{R}^d,\\ \bar{\boldsymbol{L}}(n) &\equiv (\bar{L}_0(n), \bar{L}_1(n), \cdots, \bar{L}_{d-1}(n)) \in \mathbb{R}^d,\end{aligned} \tag{6}$$

and the CLT-scaled averages be

$$\begin{aligned}
\hat{\boldsymbol{\lambda}}(n) &\equiv \sqrt{n}(\bar{\boldsymbol{\lambda}}(n) - \boldsymbol{\lambda}) \in \mathbb{R}^d, \\
\hat{\boldsymbol{\delta}}(n) &\equiv \sqrt{n}(\bar{\boldsymbol{\delta}}(n) - \boldsymbol{\delta}) \in \mathbb{R}^d, \\
\hat{\boldsymbol{F}}^{\boldsymbol{c}}(n) &\equiv \sqrt{n}(\bar{\boldsymbol{F}}^{\boldsymbol{c}}(n) - \boldsymbol{F}^{\boldsymbol{c}}) \in \mathbb{R}^{d\times J}, \\
\hat{\boldsymbol{W}}(n) &\equiv \sqrt{n}(\bar{\boldsymbol{W}}(n) - \boldsymbol{W}) \in \mathbb{R}^d, \\
\hat{\boldsymbol{Q}}(n) &\equiv \sqrt{n}(\bar{\boldsymbol{Q}}(n) - \boldsymbol{L}) \in \mathbb{R}^d, \\
\hat{\boldsymbol{L}}(n) &\equiv \sqrt{n}(\bar{\boldsymbol{L}}(n) - \boldsymbol{L}) \in \mathbb{R}^d,
\end{aligned} \tag{7}$$

where the deterministic centering constants are

$$\begin{aligned}
\boldsymbol{\lambda} &\equiv (\lambda_0, \lambda_1, \cdots, \lambda_{d-1}) \in \mathbb{R}^d, \\
\boldsymbol{\delta} &\equiv (\delta_0, \delta_1, \cdots, \delta_{d-1}) \in \mathbb{R}^d, \\
\boldsymbol{F}^{\boldsymbol{c}} &\equiv ((F_0^c)^T, (F_1^c)^T, \cdots, (F_{d-1}^c)^T) \in \mathbb{R}^{d\times J}, \\
\boldsymbol{W} &\equiv (W_0, W_1, \cdots, W_{d-1}) \in \mathbb{R}^d, \\
\boldsymbol{L} &\equiv (L_0, L_1, \cdots, L_{d-1}) \in \mathbb{R}^d.
\end{aligned} \tag{8}$$

Again, all the above constant vectors and matrices in (8) can be extended as periodic functions with period d, but for convenience, throughout this paper, we use the modulo function, $[x] = x \mod d$, to treat k beyond $0 \le k \le d-1$.

3.2 A Practical Version for Applications

We now state our first CLT version of the PLL. Because it is a special case of the more general one later, the proof is not given separately. The statement is proved after we introduce Theorem 3.2, Proposition 3.1 and the two corollaries right after them. Let $\Rightarrow$ denote convergence in distribution.

Theorem 3.1 (*practical CLT version of the PLL*) *If the following conditions hold:*

$(E1)$ $\quad (\hat{\boldsymbol{\lambda}}(n), \hat{\boldsymbol{F}}^{\boldsymbol{c}}(n)) \Rightarrow (\boldsymbol{\Lambda}, \boldsymbol{\Gamma})$ *in* $\mathbb{R}^d \times \mathbb{R}^{d\times J}$ *as* $n \to \infty$,

$(E2)$ *the number of arrivals in a single discrete time period is bounded, and*

$$(E3) \quad W_k = \sum_{j=0}^{J-1} F_{k,j}^c \quad \textit{and} \quad L_k = \sum_{j=0}^{J-1} \lambda_{[k-j]} F_{[k-j],j}^c \quad \textit{for} \quad 0 \le k \le d-1, \tag{9}$$

for $(\hat{\boldsymbol{\lambda}}(n), \hat{\boldsymbol{F}}^{\boldsymbol{c}}(n))$ *in* (7) *and* (W_k, L_k) *in* (8)*, where the limit* $(\boldsymbol{\Lambda}, \boldsymbol{\Gamma})$ *in* $(E1)$ *is a zero-mean Gaussian random vector, then*

$$\left(\hat{\boldsymbol{\lambda}}(n), \hat{\boldsymbol{\delta}}(n), \hat{\boldsymbol{F}}^{\boldsymbol{c}}(n), \hat{\boldsymbol{Q}}(n), \hat{\boldsymbol{L}}(n), \hat{\boldsymbol{W}}(n)\right) \Rightarrow (\boldsymbol{\Lambda}, \boldsymbol{\Delta}, \boldsymbol{\Gamma}, \boldsymbol{\Upsilon}, \boldsymbol{\Upsilon}, \boldsymbol{\Omega}) \quad in \quad \mathbb{R}^{2d} \times \mathbb{R}^{d\times J} \times \mathbb{R}^{3d}, \tag{10}$$

where $\boldsymbol{\Omega} = (\Omega_0, \Omega_1, \cdots, \Omega_{d-1})$*,* $\boldsymbol{\Upsilon} = (\Upsilon_0, \Upsilon_1, \cdots, \Upsilon_{d-1})$ *and* $\boldsymbol{\Delta} = (\Delta_0, \Delta_1, \cdots, \Delta_{d-1})$ *are given by*

$$\begin{aligned} &\Omega_k = \sum_{j=0}^{J-1} \Gamma_{k,j}, \quad \Upsilon_k = \sum_{j=0}^{J-1} \lambda_{[k-j]} \Gamma_{[k-j],j} + \sum_{j=0}^{J-1} \Lambda_{[k-j]} F^c_{[k-j],j} \quad and \\ &\Delta_k = \Upsilon_{[k]} - \Upsilon_{[k+1]} + \Lambda_{[k+1]} \quad for \quad 0 \le k \le d-1, \end{aligned} \tag{11}$$

and $(\boldsymbol{\Lambda}, \boldsymbol{\Delta}, \boldsymbol{\Gamma}, \boldsymbol{\Upsilon}, \boldsymbol{\Upsilon}, \boldsymbol{\Omega})$ *is also jointly zero-mean Gaussian distributed.*

Theorem 3.1 implies that, given the joint convergence of the CLT-scaled arrival and LoS processes $(\hat{\boldsymbol{\lambda}}(n), \hat{\boldsymbol{F}}^{\boldsymbol{c}}(n))$ in (E1), with the associated regularity conditions, we get the associated convergence for the CLT-scaled direct estimate of occupancy $\hat{\boldsymbol{Q}}(n)$ as well as the indirect estimate $\hat{\boldsymbol{L}}(n)$, jointly with the other processes. We get the consistency requirement in $(E3)$ from the PLL in Theorem 2.1.

3.3 The General Version

In this section we introduce the more general CLT version of the PLL, where we allow both the number of arrivals and the LoS distributions to be unbounded. Thus it involves countably-infinite-dimensional spaces. We show that the connection among the CLT-scaled indirect estimator of occupancy level $\hat{\boldsymbol{L}}(n)$, the arrival rates $\hat{\boldsymbol{\lambda}}(n)$ and the LoS distributions $\hat{\boldsymbol{F}}^{\boldsymbol{c}}(n)$ can be established using the continuous mapping theorem. More conditions are needed to ensure that the CLT-scaled direct estimator $\hat{\boldsymbol{Q}}(n)$ has the same limit as $\hat{\boldsymbol{L}}(n)$.

Let $\mathbb{R}^d$ be the d-dimensional real space with usual topology; let $\mathbb{R}^\infty$ be the space of sequences $x = (x_0, x_1, \cdots)$ of real numbers with the topology determined by the convergence of all finite-dimensional projections (which is induced by the metric in Example 1.2 of [5]); let $\ell_1 \subseteq \mathbb{R}^\infty$ be the subspace of $\mathbb{R}^\infty$ which contains sequences with finite absolute sums; and let $\ell_\infty \subseteq \mathbb{R}^\infty$ be the subspace of $\mathbb{R}^\infty$ which contains sequences with bounded values, i.e.

$$\begin{aligned} \ell_1 &\equiv \{\boldsymbol{x} = (x_0, x_1, \cdots) \in \mathbb{R}^\infty : ||\boldsymbol{x}||_1 \equiv \sum_{i=0}^{\infty} |x_i| < \infty\}, \\ \ell_\infty &\equiv \{\boldsymbol{x} = (x_0, x_1, \cdots) \in \mathbb{R}^\infty : ||\boldsymbol{x}||_\infty \equiv \sup_i |x_i| < \infty\} \end{aligned} \tag{12}$$

We equip ℓ_1 and ℓ_∞ with the norms $||\cdot||_1$ and $||\cdot||_\infty$ as above, under which both are Banach spaces. (We remark that $\ell_1 \subseteq \ell_\infty$ and the dual space of ℓ_1 is ℓ_∞.) Then we can define $(\ell_1)^d$ as the d-fold product space of ℓ_1 with the norm equal to the maximum of the ℓ_1-norms of each component, i.e. if $\boldsymbol{y} \equiv (\boldsymbol{y}_0, \boldsymbol{y}_1, \cdots, \boldsymbol{y}_{d-1}) \in (\ell_1)^d$, where $\boldsymbol{y}_i \equiv (y_{i,0}, y_{i,1}, \cdots) \in \ell_1$, then $||\boldsymbol{y}||_{1,d} = \max_{i=0,\cdots,d-1}\{||\boldsymbol{y}_i||_1\}$, and we use the topology induced by this norm; see §5 for more discussion of this space.

All the quantities related to LoS distributions are now in those infinite dimensional spaces. To be specific,

$$\begin{aligned}
\bar{F}^c_k(n) &\equiv (\bar{F}^c_{k,0}(n), \bar{F}^c_{k,1}(n), \bar{F}^c_{k,2}(n), \cdots) \in \ell_1, \quad 0 \le k \le d-1, \\
F^c_k &\equiv (F^c_{k,0}, F^c_{k,1}, F^c_{k,2}, \cdots) \in \mathbb{R}^\infty, \quad 0 \le k \le d-1, \\
\bar{\boldsymbol{F}}^{\boldsymbol{c}}(n) &\equiv (\bar{F}^c_0(n), \bar{F}^c_1(n), \cdots, \bar{F}^c_{d-1}(n)) \in (\ell_1)^d, \\
\hat{\boldsymbol{F}}^{\boldsymbol{c}}(n) &\equiv \sqrt{n}(\bar{\boldsymbol{F}}^{\boldsymbol{c}}(n) - \boldsymbol{F}^{\boldsymbol{c}}) \in (\mathbb{R}^\infty)^d.
\end{aligned} \tag{13}$$

Other quantities are still the same as we defined in (7) and (8).

We also define the LLN-scaled and CLT-scaled difference between the direct and indirect occupancy estimators as

$$\begin{aligned}
\bar{\boldsymbol{R}}(n) &\equiv \bar{\boldsymbol{L}}(n) - \bar{\boldsymbol{Q}}(n) \in \mathbb{R}^d, \\
\hat{\boldsymbol{R}}(n) &\equiv \hat{\boldsymbol{L}}(n) - \hat{\boldsymbol{Q}}(n) \in \mathbb{R}^d.
\end{aligned} \tag{14}$$

Just as in [9], the continuous mapping theorem plays a key role in the proof of the theorem. Hence, we start by introducing the key mappings and show that they are continuous.

For $0 \le k \le d-1$, let $\boldsymbol{x} \in \mathbb{R}^d$ and $\boldsymbol{y} \in (\ell_1)^d$ and define $h_k : \mathbb{R}^d \times (\ell_1)^d \to \mathbb{R}$ as

$$h_k(\boldsymbol{x}, \boldsymbol{y}) = \sum_{j=0}^{\infty} x_{[k-j]} y_{[k-j],j}, \tag{15}$$

where, again, $[k] = k \mod d$ is the modulo function. As a critical condition, in the following lemma we show two functions that will be used as mapping functions are continuous. The proof can be found in §6.1.

Lemma 3.1 *For a given constant $\boldsymbol{z} \in (\ell_1)^d$, let $f_{\boldsymbol{z}} : \mathbb{R}^d \times \mathbb{R}^d \times (\ell_1)^d \to \mathbb{R}^d$ and $g : (\ell_1)^d \to \mathbb{R}^d$ be defined by*

$$\begin{aligned}
f_{\boldsymbol{z}}(\boldsymbol{x}^{(1)}, \boldsymbol{x}^{(2)}, \boldsymbol{y}) &\equiv (f_{\boldsymbol{z},0}(\boldsymbol{x}^{(1)}, \boldsymbol{x}^{(2)}, \boldsymbol{y}), f_{\boldsymbol{z},1}(\boldsymbol{x}^{(1)}, \boldsymbol{x}^{(2)}, \boldsymbol{y}), \cdots, f_{\boldsymbol{z},d-1}(\boldsymbol{x}^{(1)}, \boldsymbol{x}^{(2)}, \boldsymbol{y})), \\
g(\boldsymbol{y}) &\equiv (\sum_{j=0}^{\infty} y_{0,j}, \sum_{j=0}^{\infty} y_{1,j}, \cdots, \sum_{j=0}^{\infty} y_{d-1,j}),
\end{aligned} \tag{16}$$

where $f_{\boldsymbol{z},k} : \mathbb{R}^d \times \mathbb{R}^d \times (\ell_1)^d \to \mathbb{R}$ *is defined as*

$$f_{\boldsymbol{z},k}(\boldsymbol{x}^{(1)}, \boldsymbol{x}^{(2)}, \boldsymbol{y}) \equiv h_k(\boldsymbol{x}^{(1)}, \boldsymbol{y}) + h_k(\boldsymbol{x}^{(2)}, \boldsymbol{z}), \quad 0 \le k \le d-1,$$

with h_k *defined in* (15)*. The functions* $f_{\boldsymbol{z}}(\boldsymbol{x}^{(1)}, \boldsymbol{x}^{(2)}, \boldsymbol{y})$ *and* $g(\boldsymbol{y})$ *in* (16) *are continuous.*

The following is a counterexample to show the two functions above are not continuous if we replace ℓ_1 by $\mathbb{R}^\infty$. Let $\mathbf{0} = (0, 0, \cdots, 0)$ be the zero vector in proper spaces depending on the context.

Example 3.1 (discontinuity of f and g when $\boldsymbol{y} \in (\mathbb{R}^\infty)^d$) It suffices to see that $g_0(\boldsymbol{y}_0) \equiv \sum_{j=0}^{\infty} y_{0,j}$ from $\mathbb{R}^\infty \to \mathbb{R}$ is not continuous in general. Let $y_{0,j}^{(i)} \equiv I_{\{i=j\}}$ for all $i, j \ge 0$, so that $\boldsymbol{y}_0^{(i)} \equiv (y_{0,0}^{(i)}, y_{0,1}^{(i)}, \cdots)$ are all 0 except the i^{th} component. Under the metric of $\mathbb{R}^\infty$ as in Example 1.2 of [5], i.e. $\rho(\boldsymbol{x}^{(1)}, \boldsymbol{x}^{(2)}) = \sum_{i=1}^{\infty} \min(1, |x_i^{(1)} - x_i^{(2)}|)/2^i$, where $\boldsymbol{x}^{(j)} = (x_1, x_2, \cdots)$, $j = 1, 2$, $\boldsymbol{y}_0^{(i)} \to \mathbf{0}$ as $i \to \infty$, however $\lim_{i\to\infty} g_0(\boldsymbol{y}_0^{(i)}) = 1 \ne 0 = g_0(\mathbf{0})$.

We now state our general CLT version of the PLL with indirect estimator of the occupancy level. The proof appears in §6.1.

Theorem 3.2 (*CLT version of the PLL with indirect estimator*) *If the following conditions hold:*

$$\begin{aligned} &(C1) \qquad (\hat{\boldsymbol{\lambda}}(n), \hat{\boldsymbol{F}}^{\boldsymbol{c}}(n)) \Rightarrow (\boldsymbol{\Lambda}, \boldsymbol{\Gamma}) \quad \textit{in} \quad \mathbb{R}^d \times (\ell_1)^d \quad \textit{as} \quad n \to \infty, \\ &(C2) \qquad \boldsymbol{W} = g(\boldsymbol{F}^{\boldsymbol{c}}) \quad \textit{and} \quad \boldsymbol{L} = f_{\boldsymbol{F}^{\boldsymbol{c}}}(\mathbf{0}, \boldsymbol{\lambda}, \mathbf{0}), \end{aligned} \tag{17}$$

for $(\hat{\boldsymbol{\lambda}}(n), \hat{\boldsymbol{F}}^{\boldsymbol{c}}(n))$ *in* (7) *and* $(\boldsymbol{\lambda}, \boldsymbol{F}^{\boldsymbol{c}}, \boldsymbol{W}, \boldsymbol{L})$ *in* (8)*, using Lemma* 3.1*, then*

$$\begin{aligned} &\Big(\big(\bar{\boldsymbol{\lambda}}(n), \bar{\boldsymbol{F}}^{\boldsymbol{c}}(n), \bar{\boldsymbol{L}}(n), \bar{\boldsymbol{W}}(n)\big), \big(\hat{\boldsymbol{\lambda}}(n), \hat{\boldsymbol{F}}^{\boldsymbol{c}}(n), \hat{\boldsymbol{L}}(n), \hat{\boldsymbol{W}}(n)\big) \Big) \\ &\quad \Rightarrow \big((\boldsymbol{\lambda}, \boldsymbol{F}^{\boldsymbol{c}}, \boldsymbol{L}, \boldsymbol{W}), (\boldsymbol{\Lambda}, \boldsymbol{\Gamma}, \boldsymbol{\Upsilon}, \boldsymbol{\Omega}) \big) \quad \textit{in} \quad (\mathbb{R}^d \times (\ell_1)^d \times \mathbb{R}^{2d}) \times (\mathbb{R}^d \times (\ell_1)^d \times \mathbb{R}^{2d}), \end{aligned} \tag{18}$$

where $\boldsymbol{\Omega} = g(\boldsymbol{\Gamma})$ *and* $\boldsymbol{\Upsilon} = f_{\boldsymbol{F}^{\boldsymbol{c}}}(\boldsymbol{\lambda}, \boldsymbol{\Lambda}, \boldsymbol{\Gamma})$*; in particular,*

$$\Omega_k = \sum_{j=0}^{\infty} \Gamma_{k,j}, \quad \Upsilon_k = \sum_{j=0}^{\infty} \lambda_{[k-j]} \Gamma_{[k-j],j} + \sum_{j=0}^{\infty} \Lambda_{[k-j]} F^c_{[k-j],j}. \tag{19}$$

Note that condition (C1) requires that the random elements actually belong to the specified space. Also the limit for the first five elements in (18) yield a weak LLN, consistent with the PLL.

Just as in the ordinary Little's law and the PLL, it is interesting to consider when can we add the direct estimator of occupancy level $\hat{\boldsymbol{Q}}(n)$ into the the joint convergence. Clearly, if we can show that $\hat{\boldsymbol{R}}(n)$ defined in (14) goes to $\mathbf{0}$ in distribution, then $\hat{\boldsymbol{Q}}(n)$ converges to the same limit as $\hat{\boldsymbol{L}}(n)$ in distribution as $n \to \infty$ and can be added into the joint convergence in (18) by applying Theorem 3.1 of [5]. For clarity, we summarize that observation in the following proposition.

Proposition 3.1 *If, in addition to condition* $(C1)$ *and* $(C2)$*, we have* $\hat{\boldsymbol{R}}(n) \Rightarrow \boldsymbol{0}$*, then*

$$\Big(\big(\bar{\boldsymbol{\lambda}}(n), \bar{\boldsymbol{\delta}}(n), \bar{\boldsymbol{F}}^{\boldsymbol{c}}(n), \bar{\boldsymbol{Q}}(n), \bar{\boldsymbol{L}}(n), \bar{\boldsymbol{W}}(n) \big), \big(\hat{\boldsymbol{\lambda}}(n), \hat{\boldsymbol{\delta}}(n), \hat{\boldsymbol{F}}^{\boldsymbol{c}}(n), \hat{\boldsymbol{Q}}(n), \hat{\boldsymbol{L}}(n), \hat{\boldsymbol{W}}(n), \hat{\boldsymbol{R}}(n) \big) \Big)$$
$$\Rightarrow \big((\boldsymbol{\lambda}, \boldsymbol{\delta}, \boldsymbol{F}^{\boldsymbol{c}}, \boldsymbol{L}, \boldsymbol{L}, \boldsymbol{W}), (\boldsymbol{\varLambda}, \boldsymbol{\varDelta}, \boldsymbol{\varGamma}, \boldsymbol{\varUpsilon}, \boldsymbol{\varUpsilon}, \boldsymbol{\varOmega}, \boldsymbol{0}) \big) \quad \text{in} \quad (\mathbb{R}^{2d} \times (\ell_1)^d \times \mathbb{R}^{3d}) \times (\mathbb{R}^{2d} \times (\ell_1)^d \times \mathbb{R}^{4d}), \tag{20}$$

where $\boldsymbol{\delta}$ *is in (8),* $\boldsymbol{\varDelta} = (\varDelta_0, \varDelta_1, \cdots, \varDelta_{d-1})$ *is given by*

$$\varDelta_k = \varUpsilon_{[k]} - \varUpsilon_{[k+1]} + \varLambda_{[k+1]} \quad \text{for} \quad 0 \le k \le d-1, \tag{21}$$

and all the other variables have the same meaning as in Theorem 3.2.

The proof is given together with Lemma 3.1 and Theorem 3.2 in §6.1.

The following two corollaries show that boundedness is a simple yet practical condition such that $\hat{\boldsymbol{R}}(n) \Rightarrow \boldsymbol{0}$, so that Theorem 3.1 is covered by Theorem 3.2 and Proposition 3.1 as a special case. The proofs are in §6.2 and §6.3 respectively.

Corollary 3.1 (*Gaussian limits*) *If, in addition to the conditions of Theorem 3.2,* $(\boldsymbol{\varLambda}, \boldsymbol{\varGamma})$ *has a zero-mean Gaussian distribution with covariance and cross-covariance* $Cov(\boldsymbol{\varLambda}, \boldsymbol{\varLambda}) = \varSigma^{\varLambda}$, $Cov(\varGamma_k, \varGamma_l) = \varSigma^{\varGamma:k,l}$ *and* $Cov(\boldsymbol{\varLambda}, \varGamma_k) = \varSigma^{\varLambda,\varGamma:k}$, $0 \le k, l \le d-1$*, then,* $(\boldsymbol{\varOmega}, \boldsymbol{\varUpsilon})$ *also has a zero-mean Gaussian distribution with*

$$\begin{aligned}
Cov(\boldsymbol{\varOmega}, \boldsymbol{\varOmega})_{k,l} &= \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} \varSigma^{\varGamma:k,l}_{i,j}, \\
Cov(\boldsymbol{\varUpsilon}, \boldsymbol{\varUpsilon})_{k,l} &= \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \lambda_{[k-i]} \lambda_{[l-j]} \varSigma^{\varGamma:[k-i],[l-j]}_{i+1,j+1} + \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \lambda_{[k-i]} F^c_{[l-j],j} \varSigma^{\varLambda,\varGamma:[k-j]}_{[l-j]+1,i+1} \\
&\quad + \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} F^c_{[k-i],i} \lambda_{[l-j]} \varSigma^{\varLambda,\varGamma:[l-j]}_{[k-i]+1,j+1} + \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} F^c_{[k-i],i} F^c_{[l-j],j} \varSigma^{\varLambda}_{[k-i]+1,[l-j]+1}, \\
Cov(\boldsymbol{\varOmega}, \boldsymbol{\varUpsilon})_{k,l} &= \sum_{i=1}^{\infty} \sum_{j=0}^{\infty} \lambda_{[l-j]} \varSigma^{\varGamma:k,[l-j]}_{i,j+1} + \sum_{i=1}^{\infty} \sum_{j=0}^{\infty} F^c_{[l-j],j} \varSigma^{\varLambda,\varGamma:k}_{[l-j]+1,i}.
\end{aligned} \tag{22}$$

Corollary 3.2 (*bounded LoS*) *Suppose that the number of arrivals in a discrete time period and the LoS are bounded (by* J*). Then condition* $(C1)$ *reduces to convergence in* $\mathbb{R}^d \times (\mathbb{R}^J)^d$ *for some finite* J*. If, in addition, conditions* $(C1)$ *and* $(C2)$ *hold, then* $\hat{\boldsymbol{R}}(n) \Rightarrow \boldsymbol{0}$*, so that the joint convergence (20) in Proposition 3.1 holds.*

If we don't want to strengthen the conditions with boundedness, two other reasonable sufficient conditions are given in §4 such that $\hat{\boldsymbol{R}}(n) \Rightarrow \boldsymbol{0}$ and a full version of CLT-PLL can be achieved with both CLS-scaled direct and indirect estimator of occupancy level in the jointly convergence as in Proposition 3.1. But before we go to that, we make a remark that connect the CLT-PLL to the ordinary CLT version of LL which is studied in [11]. We also give potential statistical applications of our theorems.

Remark 3.1 (connection to the ordinary CLT version of LL in [11]) When $d = 1$, Theorem 3.2 reduces to an ordinary CLT version of LL, which can be compared to the earlier one in [11]. Specifically, when $d = 1$, we have $\boldsymbol{\Upsilon} = f_{\boldsymbol{F}^c}(\boldsymbol{\lambda}, \boldsymbol{\Lambda}, \boldsymbol{\Gamma}) = \boldsymbol{\lambda\Omega} + \boldsymbol{W\Lambda}$. All the discussion in [11] is in continuous time, but it is not difficult to translate everything into discrete time. To avoid notation conflicts and make things clear, we add tildes on all the variables in [11]. Λ is the limit of the time-averaged arrival rate, so it corresponds to the second term of (1.2) in [11], i.e. $\boldsymbol{\Lambda} = -\tilde{\lambda}^{3/2}\tilde{U}$. $\boldsymbol{\Omega}$ is the limit of customer-averaged LoS, so in the notation of [11], assuming Theorem 1 of [11] holds, we should write

$$\begin{aligned}
\tilde{t}^{1/2}\Big(\frac{\sum_{k=1}^{\tilde{N}(\tilde{t})}\tilde{W}_k}{\tilde{N}(\tilde{t})} - \tilde{w}\Big) &= \frac{\tilde{t}}{\tilde{N}(\tilde{t})}\tilde{t}^{-1/2}\Big(\sum_{k=1}^{\tilde{N}(\tilde{t})}\tilde{W}_k - \tilde{N}(\tilde{t})\tilde{w}\Big) \\
&= \frac{\tilde{t}}{\tilde{N}(\tilde{t})}\tilde{t}^{-1/2}\Big(\sum_{k=1}^{\tilde{N}(\tilde{t})}\tilde{W}_k - \tilde{\lambda}\tilde{t}\tilde{w} + \tilde{\lambda}\tilde{t}\tilde{w} - \tilde{N}(\tilde{t})\tilde{w}\Big) \\
&= \frac{\tilde{t}}{\tilde{N}(\tilde{t})}\big(\tilde{t}^{-1/2}\big(\sum_{k=1}^{\tilde{N}(\tilde{t})}\tilde{W}_k - \tilde{\lambda}\tilde{t}\tilde{w}\big) - \tilde{t}^{-1/2}\tilde{w}(\tilde{N}(\tilde{t}) - \tilde{\lambda}\tilde{t})\big) \\
&\Rightarrow \tilde{\lambda}^{-1}(\tilde{\lambda}^{1/2}(\tilde{W} - \tilde{w}\tilde{U}) + \tilde{w}\tilde{\lambda}^{3/2}\tilde{U}) = \boldsymbol{\Omega},
\end{aligned}$$

so that $\tilde{\lambda}^{1/2}\tilde{W} = \tilde{\lambda}\boldsymbol{\Omega} + \tilde{\lambda}^{1/2}\tilde{w}\tilde{U} - \tilde{w}\tilde{\lambda}^{3/2}\tilde{U}$. Finally, $\boldsymbol{\Upsilon}$ as the limit of $\hat{\boldsymbol{L}}(n)$ corresponds to the sixth term in (1.2) of [11], and if we further have $\hat{\boldsymbol{R}}(n) \Rightarrow \boldsymbol{0}$, then $\boldsymbol{\Upsilon}$ as the limit of $\hat{\boldsymbol{Q}}(n)$ is the time-averaged occupancy level of the system which corresponds to the eighth term in (1.2) of [11]. Both the sixth and the eighth term have the common limit

$$\begin{aligned}
\tilde{\lambda}^{1/2}(\tilde{W} - \tilde{w}\tilde{U}) &= \tilde{\lambda}\boldsymbol{\Omega} + \tilde{\lambda}^{1/2}\tilde{w}\tilde{U} - \tilde{w}\tilde{\lambda}^{3/2}\tilde{U} - \tilde{\lambda}^{1/2}\tilde{w}\tilde{U} \\
&= \tilde{\lambda}\boldsymbol{\Omega} - \tilde{w}\tilde{\lambda}^{3/2}\tilde{U} = \tilde{\lambda}\boldsymbol{\Omega} + \tilde{w}\boldsymbol{\Lambda}.
\end{aligned}$$

Note that $\tilde{\lambda}$ and $\tilde{w}$ being the arrival rate and mean LoS correspond to $\boldsymbol{\lambda}$ and $\boldsymbol{W}$, respectively, in our notation, so the terms in the two theorems match perfectly.

In the notation of [11], our Theorem 3.2 in the case $d = 1$ actually states that if

$$\tilde{t}^{-1/2}\big(\tilde{N}(\tilde{t}) - \tilde{\lambda}\tilde{t}, \sum_{k=1}^{\tilde{N}(\tilde{t})}\tilde{W}_k - \tilde{N}(\tilde{t})\tilde{w}\big) \Rightarrow (\boldsymbol{\Lambda}, \boldsymbol{\Omega}),$$

and Theorem 2(f) of [11] (which is exactly $(C1)$) holds, then we have the joint convergence of the second, sixth and eighth terms in (1.2) of [11].

Unlike Theorem 1 of [11], Theorem 3.2 and Proposition 3.1 do not require stationarity. That is explained by the extra convergence $\hat{\boldsymbol{R}}(n) \Rightarrow \mathbf{0}$, which corresponds to Theorem 2(f) of [11]. That extra convergence in [11] is implied by the stationarity. Inspired by that observation, we will obtain a similar (stronger) sufficient condition for $\hat{\boldsymbol{R}}(n) \Rightarrow \mathbf{0}$ involving stationarity in Theorem 4.2.

3.4 Statistical Applications

The CLT versions of the PLL have statistical applications. First, Theorem 3.2 supports using the indirect estimator of the occupancy level via the PLL. Moreover, confidence areas for the occupancy-level estimators can be constructed. In particular, if Theorem 3.1 or Corollary 3.1 holds, then we know that we can have confidence ellipsoids for $\boldsymbol{L}$. To be specific, $\hat{\boldsymbol{L}}(n) \Rightarrow \boldsymbol{\Upsilon} \sim N(\mathbf{0}, \Sigma^{\Upsilon})$ in $\mathbb{R}^d$ where Σ^{Υ}, the covariance matrix of $\boldsymbol{\Upsilon}$, is determined by (22). Then when n is large $n(\boldsymbol{L}-\bar{\boldsymbol{L}}(n))^T(\Sigma^{\Upsilon})^{-1}(\boldsymbol{L}-\bar{\boldsymbol{L}}(n))$ has approximately a standard normal distribution. Let q_α be such that $P(|N(0,1)| \le q_\alpha) = 1-\alpha$, where $0 \le \alpha < 1$, then

$$\{\boldsymbol{x} \in \mathbb{R}^d : n(\boldsymbol{x} - \bar{\boldsymbol{L}}(n))^T(\Sigma^{\Upsilon})^{-1}(\boldsymbol{x} - \bar{\boldsymbol{L}}(n)) \le q_\alpha\}, \tag{23}$$

which is an ellipsoid centered at $\bar{\boldsymbol{L}}(n)$, is a confidence ellipsoid for $\boldsymbol{L}$ with approximate confidence level $1-\alpha$. By (19) and (22), the more negatively related $\boldsymbol{\Lambda}$ and $\boldsymbol{\Gamma}$ are, the more asymptotically efficient the indirect estimator of the number of customer in the system becomes. However, unlike the case for the ordinary LL discussed in [12], there are many covariance terms in (22) even if we only consider the variance of $\boldsymbol{\Upsilon}_k$ for a given k, so it is not straightforward to compare the asymptotic efficiency of the estimator when we change the other two elements in the PLL.

Socondly, Theorem 3.1 as well as the two more sufficient conditions we will introduce soon in §4 tell us when will the indirect estimator $\bar{\boldsymbol{L}}(n)$ and the natural direct estimator $\bar{\boldsymbol{Q}}(n)$ for the number of customers in the system have the same asymptotic efficiency, i.e. $\hat{\boldsymbol{L}}(n)$ and $\hat{\boldsymbol{Q}}(n)$ converge to the same random variable. Then the confidence area analysis above in (23) also holds for $\bar{\boldsymbol{Q}}(n)$.

To apply (23), we need to estimate Σ^{Υ}. The expression is complicated, but we will soon establish a useful sufficient condition. In particular, under the conditions of Theorem 4.1, we can estimate Σ^{Υ} by (25) using $\bar{\lambda}_k(n)$ to estimate λ_k, $(n-1)^{-1}\sum_{m=1}^{n}(\boldsymbol{A}_m - \bar{\boldsymbol{\lambda}}(n))(\boldsymbol{A}_m - \bar{\boldsymbol{\lambda}}(n))^T$ to estimate Σ^{Λ} and the empirical distributions to estimate $F_{k,j}$ and $F^c_{k,j}$.

4 Sufficient Conditions for the CLT Version of the PLL

In this section we provide convenient sufficient conditions for the two conditions in 3.2 and $\hat{\mathbf{R}}(n) \Rightarrow \mathbf{0}$ to be satisfied, so that we have Proposition 3.1 and the joint convergence in (20).

The first sufficient condition is independence. To state it, for $n = 0, 1, 2, \cdots$, let $\boldsymbol{A}_n$ be the vector of arrivals at the d times within period n, i.e.,

$$\boldsymbol{A}_n = (A_{0+nd}, A_{1+nd}, \cdots, A_{d-1+nd}) \in \mathbb{R}^d. \tag{24}$$

For vectors, let $>$ mean strict order for all components.

Theorem 4.1 (*independent case*) *If the following conditions hold:*

(*I*1) $\{\boldsymbol{A}_n\}$, $n = 0, 1, 2, \cdots$, *are i.i.d. with* $\mathbb{E}\boldsymbol{A}_0 = \boldsymbol{\lambda} > 0$ *and* $Var(\boldsymbol{A}_0) = \Sigma^{\Lambda}$,

(*I*2) *the LoS are mutually independent and independent of the arrival process, having a cdf that depends only on the discrete time period* k, *and*

(*I*3) *the LoS distribution satisfies* $F^c_{k,j} \sim O(j^{-(3+\delta)})$ *for all* k *and some* $\delta > 0$,

then conditions $(C1)$, $(C2)$ *and* $\hat{\mathbf{R}}(n) \Rightarrow \mathbf{0}$ *are satisfied, so that the joint convergence in (20) holds. Further,* $(\boldsymbol{\Lambda}, \boldsymbol{\Gamma})$ *has a zero-mean Gaussian distribution in* $\mathbb{R}^d \times (\ell_1)^d$ *with* $\boldsymbol{\Lambda} \sim N(\mathbf{0}, \Sigma^{\Lambda})$, $Cov(\Gamma_{k,j}, \Gamma_{k,s}) = \lambda_k F_{k,j} F^c_{k,s}$ *for* $0 \le k \le d-1$ *and* $0 \le j \le s$, *and* $\boldsymbol{\Lambda}, \boldsymbol{\Gamma}_0, \boldsymbol{\Gamma}_1, \cdots, \boldsymbol{\Gamma}_{d-1}$ *are independent. As a special case of Corollary 3.1,* $(\boldsymbol{\Omega}, \boldsymbol{\Upsilon})$ *also has a zero-mean Gaussian distribution with, for* $1 \le k \le l \le d-1$,

$$\begin{aligned}
Cov(\boldsymbol{\Omega}, \boldsymbol{\Omega})_{k,l} &= \begin{cases} \lambda_k \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} F_{k,\min\{i,j\}} F^c_{k,\max\{i,j\}}, & \text{for} \quad k = l, \\ 0, & \text{for} \quad k \neq l, \end{cases} \\
Cov(\boldsymbol{\Upsilon}, \boldsymbol{\Upsilon})_{k,l} &= \sum_{s=0}^{k} \lambda_s^3 \Big(\sum_{m=0}^{\infty} \sum_{n=0}^{\infty} F_{s,\min\{k-s+md, l-s+nd\}} F^c_{s,\max\{k-s+md, l-s+nd\}}\Big) \\
&\quad + \sum_{s=k+1}^{l} \lambda_s^3 \Big(\sum_{m=1}^{\infty} \sum_{n=0}^{\infty} F_{s,\min\{k-s+md, l-s+nd\}} F^c_{s,\max\{k-s+md, l-s+nd\}}\Big) \\
&\quad + \sum_{s=l+1}^{d-1} \lambda_s^3 \Big(\sum_{m=1}^{\infty} \sum_{n=1}^{\infty} F_{s,\min\{k-s+md, l-s+nd\}} F^c_{s,\max\{k-s+md, l-s+nd\}}\Big) \\
&\quad + \sum_{i=0}^{d-1} \sum_{j=0}^{d-1} c_{k,i} c_{k,j} \Sigma^{\Lambda}_{i,j}, \\
Cov(\boldsymbol{\Omega}, \boldsymbol{\Upsilon})_{k,l} &= \sum_{i=1}^{\infty} \sum_{m=0}^{\infty} \lambda_k^2 F_{k,\min\{i, l-k+md\}} F^c_{k,\max\{i, l-k+md\}}.
\end{aligned} \tag{25}$$

where

$$c_{k,j} = \begin{cases} \sum_{m=0}^{\infty} F^c_{j,k-j+md} & \textit{for} \quad 0 \le j \le k, \\ \sum_{m=1}^{\infty} F^c_{j,k-j+md} & \textit{for} \quad k+1 \le j \le d-1. \end{cases}$$

The proofs can be found in §6.4.

Remark 4.1 (necessity of condition $(I3)$) In condition $(I3)$ we assume a $3+\delta$ rate of decay of the tail distribution of the LoS, which is stronger than an ordinary CLT requires. However, this is needed in the proof as can be seen in (63). More generally, it remains to determine if condition $(I3)$ is necessary, i.e., if it can be replaced by a $2+\delta$ rate of decay.

It is significant that Theorem 4.1 can be applied to the $M_T/GI_t/\infty$ queueing model we built for the ED patient flow analysis in [30], which had mutually independent Gaussian daily totals for the arrivals, but dependence among the hourly arrivals within each day. However, we conjecture that more data would show dependence among the daily totals as well. Moreover, we want to treat queueing models that are not infinite-server models. For infinite-server models, each LoS (sojourn time) coincides with the service time, but that is not the case for other service systems. Hence, we next show that Theorem 3.2 and Proposition 3.1 can be applied to more general queueing models by replacing the iid conditions with stationarity plus appropriate mixing, as in Chapter 4 of [5].

We start with the framework used in §3 of [31], where we consider the composite arrival+service input process stochastic process $\boldsymbol{Y} \equiv \{\boldsymbol{Y}_n : n \in \mathbb{N}\}$ with

$$\boldsymbol{Y}_n \equiv \{Y_{nd+k,j} : 0 \le k \le d-1;\ j \ge 0\}. \tag{26}$$

For simplicity, we assume the LoS distributions are bounded, i.e. $Y_{k,j} = 0$, $j \ge J$, for some constant $J > 0$, then each $\boldsymbol{Y}_n$ becomes a finite dimensional vector, and we can regard $\boldsymbol{Y}_n$ as a $d \times J$ random matrix

$$\boldsymbol{Y_n} = \begin{pmatrix} Y_{nd+0,0} & Y_{nd+0,1} & \cdots & Y_{nd+0,J-1} \\ Y_{nd+1,0} & Y_{nd+1,1} & \cdots & Y_{nd+1,J-1} \\ \vdots & \vdots & \vdots & \vdots \\ Y_{nd+d-1,0} & Y_{nd+d-1,1} & \cdots & Y_{nd+d-1,J-1} \end{pmatrix} \in \mathbb{R}^{d \times J}. \tag{27}$$

We will start by assuming $\{\boldsymbol{Y}_n : n \in \mathbb{N}\}$ is a stationary process. By adding a suitable mixing condition, we can show that it satisfies a multivariate CLT, then we exploit the relationship between $\{\boldsymbol{Y}_n : n \in \mathbb{N}\}$ and $(\hat{\boldsymbol{\lambda}}(n), \hat{\boldsymbol{F}}^{\boldsymbol{c}}(n))$ and show that $(C1)$ and $(C2)$ hold. Finally, we show that $\hat{\boldsymbol{R}}(n) \Rightarrow \boldsymbol{0}$ so that Theorem 3.2 and Proposition 3.1 hold.

Before we state the theorem, we make a few definitions. For convenience, we directly use $\{\boldsymbol{Y}_n, n \in \mathbb{N}\}$ in (26) as the example. We say the process $\{\boldsymbol{Y}_n, n \in \mathbb{N}\}$ is strictly stationary if

$$(\boldsymbol{Y}_0, \boldsymbol{Y}_1, \cdots, \boldsymbol{Y}_m) \stackrel{d}{=} (\boldsymbol{Y}_{0+n}, \boldsymbol{Y}_{1+n}, \cdots, \boldsymbol{Y}_{m+n}), \quad \text{for all} \quad m \geq 0 \quad \text{and} \quad n \geq 0, \tag{28}$$

where $\stackrel{d}{=}$ means equal in distribution. For $m \leq n$, let $\mathcal{F}_m^n \equiv \sigma(\boldsymbol{Y}_m, \boldsymbol{Y}_{m+1}, \cdots, \boldsymbol{Y}_n)$ be the sigma-algebra generated by the family of random variables, where we allow $n = \infty$. And we say $\{\boldsymbol{Y}_n, n \in \mathbb{N}\}$ is strongly mixing (or α-mixing) if $\alpha(n) \to 0$, where

$$\alpha(n) \equiv \sup_{m \in \mathbb{N}, A \in \mathcal{F}_0^m, B \in \mathcal{F}_{m+n}^\infty} |P(A \cap B) - P(A)P(B)|, \tag{29}$$

is called the strong mixing coefficient, following Theorem 18.5.3 of [15] and Theorem 0 of [6].

Now we have the following theorem, whose proof is provided in §6.5.

Theorem 4.2 (*stationary+bounded case*) *If the following conditions hold:*

($S1$) *the composite process* $\{\boldsymbol{Y}_n, n \in \mathbb{N}\}$ *in (26) is strictly stationary and strongly mixing,*

($S2$) $\mathbb{E}\boldsymbol{A}_0 = \boldsymbol{\lambda} > 0$ *and there exists some* $\delta > 0$, *such that* $\mathbb{E}||\boldsymbol{A}_0||_\infty^{2+\delta} < \infty$ *and* $\sum_{n=1}^\infty \alpha(n)^{\delta/(2+\delta)} < \infty$, *and*

($S3$) *The LoS distributions are bounded (by* J*),*

for $\boldsymbol{Y}_n$ *in* (26) *and* $\boldsymbol{A}_0$ *in* (24), *then condtions* ($C1$), ($C2$) *and* $\hat{\boldsymbol{R}}(n) \Rightarrow \boldsymbol{0}$ *are satisfied, so that we have the joint convergence in* (20). *Furthermore,* $(\boldsymbol{\Lambda}, \boldsymbol{\Gamma})$ *has a zero-mean Gaussian distribution in* $\mathbb{R}^d \times (\mathbb{R}^J)^d$, *so that Corollary 3.1 is also satisfied.*

We remark that there is a large literature on the CLT under weak dependence so that many generalizations of Theorem 4.2 are possible. For example, in Chapter 4 of [5] CLTs under mixing conditions in §19 are reduced to CLTs for martingale-difference sequences in §18.

Finally, other than the two conditions we discussed in this section, there could be many other sufficient conditions such that Proposition 3.1 holds.

5 Weak Convergence of Random Elements of $(\boldsymbol{\ell_1})^{\boldsymbol{d}}$

In this section we provide background on the weak convergence of random elements of $(\boldsymbol{\ell_1})^{\boldsymbol{d}}$. We used this space so that we could apply the continuous mapping theorem to prove Theorem 3.2. In particular, condition ($C1$) in Theorem 3.2 requires the convergence $\hat{\boldsymbol{F}}^{\boldsymbol{c}}(n) \Rightarrow \boldsymbol{\Gamma}$ in the space $(\ell_1)^d$.

Within functional analysis, $(\ell_1)^d$ is quite standard, as it is just a finite product space associated with the well known Banach space ℓ_1. However, neither ℓ_1 nor $(\ell_1)^d$ are standard in weak convergence theory, especially in applications on queueing systems. In fact, we know of no previous use in queueing theory. Of course, we have the well-established general and powerful weak convergence theory in general metric spaces, such as [5], but we would like to specialize the general theory to the space $(\ell_1)^d$. In addition, there is a substantial literature on Banach spaces and weak convergence in Banach spaces, [1,18,22]. We will draw on this developed theory.

We will give a practical criterion for checking weak convergence in the space $(\ell_1)^d$. The criterion is not used directly in our main theorem, but we hope it helps readers better understand the meaning of weak convergence in $(\ell_1)^d$ and could be useful in future research. We will also discuss the established CLT for Banach-space random variables and specialize it to the space ℓ_1, which, by giving this special case, helps to understand when we can have such convergence as in $(C1)$ and Gaussian limits as in Corollary 3.1.

5.1 A Criterion for Weak Convergence in $(\ell_1)^d$

The ℓ_1 space is a well-defined separable Banach space. Throughout this paper, we always use B^* to represent the dual space of the Banach space B. As a special case of a general metric space, the weak convergence of probability distributions on it is well defined, just as in general metric spaces; see [5]. For more on random variables in Banach spaces and convergence of those random variables, see [1] and [18]. Here we briefly state some definitions and results about Banach spaces and probability measures defined on them, which can be found in [22,18]. Then we exploit the general results in the special case of $(\ell_1)^d$ space.

The $(\ell_1)^d$ space is the product (or in some literature called direct sum and denoted as $\ell_1 \oplus \ell_1 \oplus \cdots \oplus \ell_1$) of ℓ_1 spaces. We refer to [22] for general Banach spaces. Since ℓ_1 is a separable Banach space, $(\ell_1)^d$ is also a separable Banach space with the norm $||\cdot||_{1,d}$ we defined earlier in §3.1. The dual of a product of Banach spaces is the product of the corresponding duals. Because $(\ell_1)^* = \ell_\infty$, thus the dual of $(\ell_1)^d$ is $(\ell_\infty)^d$.

We want to have a good sufficient condition for weak convergence in $(\ell_1)^d$. For that purpose, let $\boldsymbol{U}^{(n)} \equiv (\boldsymbol{U}_0^{(n)}, \ldots, \boldsymbol{U}_{d-1}^{(n)}) \in (\ell_1)^d$, where $\boldsymbol{U}_k^{(n)} \equiv (\boldsymbol{U}_{k,j}^{(n)} : j \geq 0) \in \ell_1$, and similarly for a prospective limit $\boldsymbol{U}$. A result from [18] states that $\{\boldsymbol{U}^{(n)}\}$ converges weakly to $\boldsymbol{U}$ as soon as $h(\boldsymbol{U}^{(n)})$ converges weakly (as a sequence of real valued random variables) to $h(\boldsymbol{U})$ for every h in a weakly dense subset of $((\ell_1)^d)^* = (\ell_\infty)^d$, and $\{\boldsymbol{U}^{(n)}\}$ is tight, i.e. for each $\epsilon > 0$, there exists a compact set $K \subseteq (\ell_1)^d$ such that $P(\boldsymbol{U}^{(n)} \in K) \geq 1 - \epsilon$ for all n. Now we transform the above conditions to more explicit ones in the case of $(\ell_1)^d$ space.

For $n \geq 1$, denote

$$O_n \equiv \{(\boldsymbol{y} \in (\mathbb{R}^\infty)^d : y_{k,j} = 0, 0 \leq k \leq d-1, j \geq n-1\},$$

then $O = \cup_{n=1}^\infty O_n$ is the subset of $(\ell_\infty)^d$ with only finite many non-zero components.

Lemma 5.1 *The subset O of $(\ell_\infty)^d$ containing elements with only finitely many non-zero components is a weakly dense subset of $(\ell_\infty)^d$.*

Proof It suffices to show that for any $\boldsymbol{y} \in (\ell_\infty)^d$, there exists a sequence $\{\boldsymbol{y}^{(i)}\} \subseteq O$ which weakly converges to $\boldsymbol{y}$. Let

$$y^{(i)}_{k,j} = \begin{cases} y_{k,j}, & \text{if} \quad j < i, \\ 0, & \text{otherwise}, \end{cases} \tag{30}$$

for all $0 \le k \le d-1$ and $i \ge 1$. Then $\boldsymbol{y}^{(i)} \in O$. Now we show that it weakly converges to $\boldsymbol{y}$.

Note that $((\ell_\infty)^d)^* = (\ell_{\infty,s})^d$, where $\ell_{\infty,s} \equiv \{\boldsymbol{x} = (x_0, x_1, \cdots) \in \mathbb{R}^\infty : \sum_{i=0}^\infty x_i < \infty\}$ is the set of all summable (but not necessarily absolutely summable) sequences. For any $\boldsymbol{x} \in ((\ell_\infty)^d)^* = (\ell_{\infty,s})^d$,

$$\boldsymbol{x}(\boldsymbol{y}^{(i)}) = \sum_{k=0}^{d-1}\sum_{j=0}^{\infty} x_{k,j} y^{(i)}_{k,j} = \sum_{k=0}^{d-1}\sum_{j=0}^{i-1} x_{k,j} y_{k,j} \to \sum_{k=0}^{d-1}\sum_{j=0}^{\infty} x_{k,j} y_{k,j} = \boldsymbol{x}(\boldsymbol{y}) \quad \text{as} \quad i \to \infty, \tag{31}$$

which indicates $\boldsymbol{y}^{(i)}$ weakly converges to $\boldsymbol{y}$, hence O is weakly dense in $(\ell_\infty)^d$. □

Next, we describe the compact sets in $(\ell_1)^d$ space.

Lemma 5.2 (*compact sets in $(\ell_1)^d$ space.*) *A set $K \subseteq (\ell_1)^d$ is compact if and only if K is bounded and closed and for each $\epsilon > 0$, there exists J_ϵ such that for all $\boldsymbol{y} \in K$, $\sum_{k=0}^{d-1}\sum_{j=J_\epsilon}^{\infty} |y_{k,j}| < \epsilon$.*

Proof If $K \subseteq (\ell_1)^d$ is compact, then K must be closed and bounded. Further, it is well known that a subset in metric space is compact if and only if it is complete and totally bounded (see for example Theorem 45.1 of [23]). Totally bounded means that, for any $\epsilon > 0$ given, we can find a finite set of $\{\boldsymbol{y}^{(i)}\}_{i=1}^N \subseteq K$ such that K is covered by the N ϵ-balls centered at those points. Because the set is finite, we can find J_ϵ large enough such that $\sum_{k=0}^{d-1}\sum_{j=J_\epsilon}^{\infty} |y^{(i)}_{k,j}| < \epsilon$ for all $\boldsymbol{y}^{(i)} \in K$. Then for any $\boldsymbol{y} \in K$, we can find $\boldsymbol{y}^{(i)}$ such that $||\boldsymbol{y} - \boldsymbol{y}^{(i)}||_{1,d} < \epsilon$, so

$$\begin{aligned} \sum_{k=0}^{d-1}\sum_{j=J_\epsilon}^{\infty} |y_{k,j}| &\le \sum_{k=0}^{d-1}\sum_{j=J_\epsilon}^{\infty} |y_{k,j} - y^{(i)}_{k,j}| + \sum_{k=0}^{d-1}\sum_{j=J_\epsilon}^{\infty} |y^{(i)}_{k,j}| \le \sum_{k=0}^{d-1}\sum_{j=0}^{\infty} |y_{k,j} - y^{(i)}_{k,j}| + \sum_{k=0}^{d-1}\sum_{j=J_\epsilon}^{\infty} |y^{(i)}_{k,j}| \\ &\le d||\boldsymbol{y} - \boldsymbol{y}^{(i)}||_{1,d} + \sum_{k=0}^{d-1}\sum_{j=J_\epsilon}^{\infty} |y^{(i)}_{k,j}| \le d\epsilon + \epsilon. \end{aligned} \tag{32}$$

Conversely, assume $K \subseteq (\ell_1)^d$ is bounded and closed and for each $\epsilon > 0$, there exists J_ϵ such that for all $\boldsymbol{y} \in K$, $\sum_{k=0}^{d-1}\sum_{j=J_\epsilon}^{\infty} |y_{k,j}| < \epsilon$. Since $(\ell_1)^d$ is a Banach space and K is closed, K is a complete subset. Assume K is bounded by C/d, i.e. $||\boldsymbol{y}||_{1,d} \le C/d$ for all $\boldsymbol{y} \in K$. For any $\epsilon > 0$ fixed, let J_ϵ be as above. Let $K_0 \equiv \{\boldsymbol{x} \in (\mathbb{R}^{J_\epsilon})^d : ||\boldsymbol{x}||_1 \le C\}$, which is bounded, thus totally bounded since it has finite dimension. So there exists a finite set $\{\boldsymbol{x}^{(i)}\}_{i=1}^{N} \subseteq (\mathbb{R}^{J_\epsilon})^d$ such that the ϵ-balls centered at those points cover K_0. Assume $\{\boldsymbol{y}^{(i)}\}_{i=1}^{N} \subseteq (\ell_1)^d$ are the corresponding points of $\{\boldsymbol{x}^{(i)}\}_{i=1}^{N}$ when we naturally embed $(\mathbb{R}^{J_\epsilon})^d$ to $(\ell_1)^d$, i.e. $y_{k,j}^{(i)} = x_{k,j}^{(i)}$ for $j \leq J_\epsilon$ and 0 otherwise. For any $\boldsymbol{y} \in K$, let $\boldsymbol{x}$ be the natural projection of $\boldsymbol{y}$ on $(\mathbb{R}^{J_\epsilon})^d$. Notice that $||\boldsymbol{x}||_1 \leq d||\boldsymbol{y}||_{1,d} \leq C$, which means $\boldsymbol{x} \in K_0$, so there is $\boldsymbol{x}^{(i)}$ such that $||\boldsymbol{x} - \boldsymbol{x}^{(i)}||_1 \leq \epsilon$. Then,

$$\begin{aligned} ||\boldsymbol{y} - \boldsymbol{y}^{(i)}||_{1,d} &\leq \sum_{k=0}^{d-1}\sum_{j=0}^{J_\epsilon-1} |y_{k,j} - y_{k,j}^{(i)}| + \sum_{k=0}^{d-1}\sum_{j=J_\epsilon}^{\infty} |y_{k,j} - y_{k,j}^{(i)}| \leq ||\boldsymbol{x} - \boldsymbol{x}^{(i)}||_1 + \sum_{k=0}^{d-1}\sum_{j=J_\epsilon}^{\infty} |y_{k,j}| \\ &\leq \epsilon + \epsilon, \end{aligned} \tag{33}$$

which implies that K is covered by the finite number of ϵ-balls in $(\ell_1)^d$ centered at $\{\boldsymbol{y}^{(i)}\}_{i=1}^{N}$. So K is totally bounded, thus compact. □

Now we can give an easy-to-check sufficient condition for weak convergence of random elements in $(\ell_1)^d$, which can be used to establish $(C1)$ when applying Theorem 3.2. For any $m, C > 0$, let

$$K_{m,C} \equiv \{\boldsymbol{x} \in (\ell_1)^d : ||x||_{1,d} \leq C, \sum_{k=0}^{d-1}\sum_{j=m^n}^{\infty} |x_{k,j}| < n^{-1} \quad \text{for all} \quad n \geq 1\} \tag{34}$$

which is obviously a compact set by Lemma 5.2.

Theorem 5.1 (*criterion for convergence of random elements of* $(\ell_1)^d$) *Convergence in distribution* $\boldsymbol{U}^{(n)} \Rightarrow \boldsymbol{U}$ *in* $(\ell_1)^d$ *as* $n \to \infty$ *holds if*

(i) *For all* $\epsilon > 0$, *there exists* m, C *and corresponding* $K_{m,C}$, *such that*

$$P(\boldsymbol{U}^{(n)} \in K_{m,C}) > 1 - \epsilon \tag{35}$$

and

(ii) *for all* J, $0 \leq J < \infty$,

$$(\boldsymbol{U}_{k,j}^{(n)} : 0 \leq k \leq d-1; 0 \leq j \leq J) \Rightarrow (\boldsymbol{U}_{k,j} : 0 \leq k \leq d-1; 0 \leq j \leq J) \quad \textit{in} \quad \mathbb{R}^{dJ}. \tag{36}$$

Condition (ii) *holds if and only if* (iii) *for all* J, $0 \leq J < \infty$, *and for all sets of real number* $\{a_{k,j} : 1 \leq k \leq d-1, 0 \leq j \leq J\}$

$$\sum_{k=0}^{d-1}\sum_{J=0}^{J} a_{k,j}\boldsymbol{U}_{k,j}^{(n)} \Rightarrow \sum_{k=0}^{d-1}\sum_{j=0}^{J} a_{k,j}\boldsymbol{U}_{k,j}. \tag{37}$$

Proof Condition (i) ensures that $\{\boldsymbol{U}^{(n)}\}$ is tight. Condition (ii) is equivalent to condition (iii) by the familiar Cramer-Wold device; see p. 382 of [4]. And condition (iii) mean $\boldsymbol{y}(\boldsymbol{U}^{(n)}) \to \boldsymbol{y}(\boldsymbol{U})$ in distribution for all $\boldsymbol{y} \in O$, which by Lemma 5.1 is a weakly dense set in $((\ell_1)^d)^*$. So that condition (i) with condition (iii) ensures $\boldsymbol{U}^{(n)} \Rightarrow \boldsymbol{U}$ in $(\ell_1)^d$. □

5.2 Central Limit Theorem for i.i.d. ℓ_1-Valued Random Variables

In this section we specialize the CLT in general Banach spaces for i.i.d. random elements to ℓ_1. First, we emphasize that, even for i.i.d. random elements, the classical CLT does not hold in all Banach spaces, but only for some "good" Banach spaces. Fortunately, ℓ_1 is such a "good" Banach space. To make this clear, we do a quick review; more related theory can be found in [1] and [18].

We first introduce two concepts: cotype of a Banach space and pre-Gaussian random variables. A separable Banach space B with norm $||\cdot||$ is said to be of cotype q if there is a constant C_q such that for all finite sequences $x_i \in B$,

$$(\sum_i ||x_i||^q)^{1/q} \leq C_q || \sum_i \epsilon_i x_i ||, \tag{38}$$

where $\{\epsilon_i\}$ is a Rademacher sequance, i.e., i.i.d. random variables with $P(\epsilon_i = 1) = P(\epsilon_i = -1) = 1/2$. It is known that ℓ_1 is of cotype 2; see page 274 in Section 9.2 of [18].

We say that a random variable $\boldsymbol{U}$ in a Banach space B has a Gaussian distribution if, for every h in the dual space B^*, $h(\boldsymbol{U})$ has a one-dimensional Gaussian distribution. A random variable $\boldsymbol{U}$ in B, with $\mathbb{E}h(\boldsymbol{U}) = 0$ and $\mathbb{E}h^2(\boldsymbol{U}) < \infty$ for every h in B^* (i.e. weakly centered and square integrable), is pre-Gaussian if its covariance is also the covariance of a Gaussian Borel probability measure on B. A weakly centered and square integrable random variable $\boldsymbol{U} = (U_0, U_1, \cdots)$ in ℓ_1 is pre-Gaussian if and only if

$$\sum_{k=0}^{\infty} (\mathbb{E}|U_k|^2)^{1/2} < \infty; \tag{39}$$

see page 261 of [18].

Theorem 5.2 (*CLT for Banach-space random variables from [18] Theorem 10.7*) *If $\boldsymbol{U}$ is pre-Gaussian with values in a separable cotype-2 Banach space, then $\boldsymbol{U}$ satisfies the CLT, i.e., $n^{-1/2}\sum_{i=1}^n \boldsymbol{U}^{(i)}$ where $\boldsymbol{U}^{(i)}$ are i.i.d. copies of $\boldsymbol{U}$, converges in distribution, where the limit is Gaussian.*

We remark that the limit distribution must have a Gaussian distribution because if $\boldsymbol{U}$ satisfies the CLT in B, then $h(\boldsymbol{U})$ satisfies the ordinary CLT with a Gaussian limit for $h \in B^*$. If we include (39), then we get the following corollary.

Corollary 5.1 (CLT for ℓ_1-value random variables) *If $\boldsymbol{U} = (U_0, U_1, \cdots) \in \ell_1$ is pre-Gaussian, then $\boldsymbol{U}$ satisfies the CLT. The limit has a Gaussian distribution with the same covariance structure as $\boldsymbol{U}$.*

We now discuss how this background theory is relevant here. We will be considering a discrete random variable Y taking values in the non-negative integers. Let $F_j^c \equiv P(Y \geq j)$ for $k = 0, 1, \cdots$ be the ccdf of Y and $F_j \equiv 1 - F_j^c$. Let $Y^{(i)}$ be i.i.d. random variables each distributed as Y and let

$$U_j^{(i)} \equiv I_{\{Y^{(i)} \geq j\}} - F_j^c.$$

Then $\boldsymbol{U}^{(i)}$ are i.i.d. random variables in ℓ_1 distributed as $\boldsymbol{U}$ with $Eh(\boldsymbol{U}) = 0$ for all $h \in \ell_\infty$ (i.e., for all $h \in (\ell_1)^*$). We want $\boldsymbol{U}$ to be pre-Gaussian, so we require that (39) holds, i.e.,

$$\sum_{k=0}^{\infty} (\mathbb{E}(I_{\{Y^{(1)} \geq j\}} - F_j^c)^2)^{1/2} = \sum_{k=0}^{\infty} (F_j F_j^c)^{1/2} < \infty. \tag{40}$$

Hence, a sufficient condition for (40) is

$$F_j^c \sim O(j^{-(2+\epsilon)}) \quad \text{for some} \quad \epsilon > 0, \tag{41}$$

which is actually implied by condition $(I3)$ in Theorem 4.1. Condition (41) is also sufficient for $\mathbb{E}h(\boldsymbol{U}) = 0$ and $\mathbb{E}h^2(\boldsymbol{U}) < \infty$ for every h in ℓ_∞. Hence, $\boldsymbol{U}$ is pre-Gaussian.

6 Proofs

We now provide the postponed proofs of Lemma 3.1, Theorem 3.2, Proposition 3.1, Corollary 3.1, Corollary 3.2, Theorem 4.1 and Theorem 4.2.

6.1 Proof of Lemma 3.1, Theorem 3.2 and Proposition 3.1

Proof of Lemma 3.1. To prove $f_{\boldsymbol{z}}(\boldsymbol{x}^{(1)}, \boldsymbol{x}^{(2)}, \boldsymbol{y})$ is continuous from $\mathbb{R}^d \times \mathbb{R}^d \times (\ell_1)^d$ to $\mathbb{R}^d$, it suffices to show that $f_{\boldsymbol{z},k}(\boldsymbol{x}^{(1)}, \boldsymbol{x}^{(2)}, \boldsymbol{y})$ is continuous from $\mathbb{R}^d \times \mathbb{R}^d \times (\ell_1)^d$ to $\mathbb{R}$, where $\boldsymbol{x}^{(i)} = (x_0^{(i)}, x_1^{(i)}, \cdots, x_{d-1}^{(i)}) \in \mathbb{R}^d$ for $i = 1, 2$, and $\boldsymbol{y} = (\boldsymbol{y}_0, \boldsymbol{y}_1, \cdots, \boldsymbol{y}_{d-1}) \in (\ell_1)^d$ are variables and $\boldsymbol{z} = (\boldsymbol{z}_0, \boldsymbol{z}_1, \cdots, \boldsymbol{z}_{d-1}) \in (\ell_1)^d$ is a constant. Further, it suffices to show that $h_k(\boldsymbol{x}, \boldsymbol{y})$ is continuous from $\mathbb{R}^d \times (\ell_1)^d$ to $\mathbb{R}$, where $\boldsymbol{x} = (x_0, x_1, \cdots, x_{d-1}) \in \mathbb{R}^d$. For convenience, we use the maximum norm in $\mathbb{R}^d$, i.e. $||\boldsymbol{x}||_\infty \equiv \max_{0 \leq k \leq d-1} |x_k|$, which is an equivalent norm to the usual Euclidean distance, and for the space $\mathbb{R}^d \times \mathbb{R}^d \times (\ell_1)^d$ and $\mathbb{R}^d \times (\ell_1)^d$, we use the metric induced by the norm $||(\boldsymbol{x}^{(1)}, \boldsymbol{x}^{(2)}, \boldsymbol{y})||_{\mathbb{R}^d \times \mathbb{R}^d \times (\ell_1)^d} \equiv ||\boldsymbol{x}^{(1)}||_\infty + ||\boldsymbol{x}^{(2)}||_\infty + ||\boldsymbol{y}||_{1,d}$ and $||(\boldsymbol{x}, \boldsymbol{y})||_{\mathbb{R}^d \times (\ell_1)^d} \equiv ||\boldsymbol{x}||_\infty + ||\boldsymbol{y}||_{1,d}$ respectively. Because the notation of infinity norm is the same in $\mathbb{R}^d$ and ℓ_∞, it should not cause confusion.

First we observe that h_k can be written as a inner product of two projection functions:

$$h_k(\boldsymbol{x}, \boldsymbol{y}) = \boldsymbol{\Pi}_k^1(\boldsymbol{x}) \cdot \boldsymbol{\Pi}_k^2(\boldsymbol{y}), \tag{42}$$

where

$$\begin{aligned} &\boldsymbol{\Pi}_k^1(\boldsymbol{x}) \equiv (x_k, x_{k-1}, \cdots, x_0, x_{d-1}, x_{d-2}, \cdots, x_0, x_{d-1}, x_{d-2}, \cdots) \quad \text{and} \\ &\boldsymbol{\Pi}_k^2(\boldsymbol{y}) \equiv (y_{k,0}, y_{k-1,1}, \cdots, y_{0,k}, y_{d-1,k+1}, y_{d-2,k+1}, \cdots, y_{0,k+d}, y_{d-1,k+d+1}, \cdots). \end{aligned} \tag{43}$$

Note that $||\boldsymbol{\Pi}_k^1(\boldsymbol{x})||_\infty = ||\boldsymbol{x}||_\infty < \infty$, and $||\boldsymbol{\Pi}_k^2(\boldsymbol{y})||_1 \le \sum_{k=0}^{d-1} ||\boldsymbol{y}_k||_1 < \infty$. So obviously $\boldsymbol{\Pi}_k^1(\boldsymbol{x})$ is continuous from $(\mathbb{R}^d, ||\cdot||_\infty)$ to $(\ell_\infty, ||\cdot||_\infty)$ and $\boldsymbol{\Pi}_k^2(\boldsymbol{y})$ is continuous from $((\ell_1)^d, ||\cdot||_{1,d})$ to $(\ell_1, ||\cdot||_1)$. We also have that $\boldsymbol{\Pi}_k^1(\boldsymbol{x}) - \boldsymbol{\Pi}_k^1(\tilde{\boldsymbol{x}}) = \boldsymbol{\Pi}_k^1(\boldsymbol{x} - \tilde{\boldsymbol{x}})$ and $\boldsymbol{\Pi}_{\boldsymbol{k}}^{\boldsymbol{2}}(\boldsymbol{y}) - \boldsymbol{\Pi}_k^2(\tilde{\boldsymbol{y}}) = \boldsymbol{\Pi}_k^2(\boldsymbol{y} - \tilde{\boldsymbol{y}})$ for $\boldsymbol{x}, \tilde{\boldsymbol{x}} \in \mathbb{R}^d$ and $\boldsymbol{y}, \tilde{\boldsymbol{y}} \in (\ell_1)^d$. For any $\delta > 0$ and $(\boldsymbol{x}, \boldsymbol{y})$ fixed, choose $(\tilde{\boldsymbol{x}}, \tilde{\boldsymbol{y}})$ such that

$$||(\boldsymbol{x}, \boldsymbol{y}) - (\tilde{\boldsymbol{x}}, \tilde{\boldsymbol{y}})||_{\mathbb{R}^d \times (\ell_1)^d} = ||\boldsymbol{x} - \tilde{\boldsymbol{x}}||_\infty + \sum_{k=0}^{d-1} ||\boldsymbol{y}_k - \tilde{\boldsymbol{y}}_k||_1 < \delta,$$

then

$$\begin{aligned} &|h_k(\boldsymbol{x}, \boldsymbol{y}) - h_k(\tilde{\boldsymbol{x}}, \tilde{\boldsymbol{y}})| = |\boldsymbol{\Pi}_k^1(\boldsymbol{x}) \cdot \boldsymbol{\Pi}_k^2(\boldsymbol{y}) - \boldsymbol{\Pi}_k^1(\tilde{\boldsymbol{x}}) \cdot \boldsymbol{\Pi}_k^2(\tilde{\boldsymbol{y}})| \\ &\le |\boldsymbol{\Pi}_k^1(\boldsymbol{x}) \cdot \boldsymbol{\Pi}_k^2(\boldsymbol{y}) - \boldsymbol{\Pi}_k^1(\boldsymbol{x}) \cdot \boldsymbol{\Pi}_k^2(\tilde{\boldsymbol{y}})| + |\boldsymbol{\Pi}_k^1(\boldsymbol{x}) \cdot \boldsymbol{\Pi}_k^2(\tilde{\boldsymbol{y}}) - \boldsymbol{\Pi}_k^1(\tilde{\boldsymbol{x}}) \cdot \boldsymbol{\Pi}_k^2(\tilde{\boldsymbol{y}})| \\ &\le ||\boldsymbol{\Pi}_k^1(\boldsymbol{x})||_\infty ||\boldsymbol{\Pi}_k^2(\boldsymbol{y} - \tilde{\boldsymbol{y}})||_1 + ||\boldsymbol{\Pi}_k^1(\boldsymbol{x} - \tilde{\boldsymbol{x}})||_\infty ||\boldsymbol{\Pi}_k^2(\tilde{\boldsymbol{y}})||_1 \\ &\le ||\boldsymbol{x}||_\infty \delta + \delta \sum_{k=0}^{d-1} ||\tilde{\boldsymbol{y}}_k||_1 \le ||\boldsymbol{x}||_\infty \delta + \delta \sum_{k=0}^{d-1} (||\boldsymbol{y}_k||_1 + ||\boldsymbol{y}_k - \tilde{\boldsymbol{y}}_k||_1) \\ &\le ||\boldsymbol{x}||_\infty \delta + \delta \sum_{k=0}^{d-1} ||\boldsymbol{y}_k||_1 + \delta^2 \to 0 \quad \text{as} \quad \delta \to 0. \end{aligned} \tag{44}$$

Given that each function $f_{\boldsymbol{z},k}$ is continuous from $(\mathbb{R}^d \times \mathbb{R}^d \times (\ell_1)^d, ||\cdot||_{\mathbb{R}^d \times \mathbb{R}^d \times (\ell_1)^d})$ to $(\mathbb{R}, |\cdot|)$, we can conclude that $f_{\boldsymbol{z}}$ is continuous from $(\mathbb{R}^d \times \mathbb{R}^d \times (\ell_1)^d, ||\cdot||_{\mathbb{R}^d \times \mathbb{R}^d \times (\ell_1)^d})$ to $(\mathbb{R}^d, ||\cdot||_\infty)$. Finally, note that we can write $g(\boldsymbol{y}) = f_{\boldsymbol{0}}(\boldsymbol{e}, \boldsymbol{0}, \boldsymbol{y})$, where $\boldsymbol{e} = (1, 1, \cdots, 1) \in \mathbb{R}^d$, so that g is also continuous from $((\ell_1)^d, ||\cdot||_{1,d})$ to $(\mathbb{R}^d, ||\cdot||_\infty)$ and the lemma is proved. ■

Proof of Theorem 3.2. Condition (C1) implies that $\hat{\boldsymbol{F}}^{\boldsymbol{c}}(n) \in (\ell_1)^d$ as well as $\boldsymbol{F}^{\boldsymbol{c}} \in (\ell_1)^d$, which indicates that the limiting distributions all have finite means. Moreover, (C1) implies that

$$(\bar{\boldsymbol{\lambda}}(n), \bar{\boldsymbol{F}}^{\boldsymbol{c}}(n)) \Rightarrow (\boldsymbol{\lambda}, \boldsymbol{F}^{\boldsymbol{c}}) \quad \text{in} \quad \mathbb{R}^d \times (\ell_1)^d. \tag{45}$$

Note that $\bar{\boldsymbol{W}}(n) = g(\bar{\boldsymbol{F}}^{\boldsymbol{c}}(n))$ and $\boldsymbol{W} = g(\boldsymbol{F}^{\boldsymbol{c}})$, so by continuous mapping theorem,

$$(\bar{\boldsymbol{\lambda}}(n), \bar{\boldsymbol{W}}(n), \bar{\boldsymbol{F}}^{\boldsymbol{c}}(n)) \Rightarrow (\boldsymbol{\lambda}, \boldsymbol{W}, \boldsymbol{F}^{\boldsymbol{c}}) \quad \text{in} \quad \mathbb{R}^{2d} \times (\ell_1)^d. \tag{46}$$

By the definition of $\bar{\boldsymbol{L}}(n)$ in (6) and (3), $\bar{\boldsymbol{L}}(n)$ is a function of $(\bar{\boldsymbol{\lambda}}(n), \bar{\boldsymbol{F}}^{\boldsymbol{c}}(n))$, i.e. $\bar{\boldsymbol{L}}(n) = f_{\boldsymbol{0}}(\bar{\boldsymbol{\lambda}}(n), \boldsymbol{0}, \bar{\boldsymbol{F}}^{\boldsymbol{c}}(n))$ where f is defined in (16). By Lemma 3.1, f is a continuous function, hence

$$(\bar{\boldsymbol{\lambda}}(n), \bar{\boldsymbol{W}}(n), \bar{\boldsymbol{L}}(n), \bar{\boldsymbol{F}}^{\boldsymbol{c}}(n)) \Rightarrow (\boldsymbol{\lambda}, \boldsymbol{W}, \boldsymbol{L}, \boldsymbol{F}^{\boldsymbol{c}}) \quad \text{in} \quad \mathbb{R}^{3d} \times (\ell_1)^d. \tag{47}$$

For the CLT-scaled terms, notice that $\hat{\boldsymbol{W}}(n) = g(\hat{\boldsymbol{F}}^{\boldsymbol{c}}(n))$, so by the continuous mapping theorem,

$$(\hat{\boldsymbol{\lambda}}(n), \hat{\boldsymbol{W}}(n), \hat{\boldsymbol{F}}^{\boldsymbol{c}}(n)) \Rightarrow (\boldsymbol{\varLambda}, \boldsymbol{\varOmega}, \boldsymbol{\varGamma}) \quad \text{in} \quad \mathbb{R}^{2d} \times (\ell_1)^d, \tag{48}$$

where $\boldsymbol{\varOmega} = g(\boldsymbol{\varGamma})$.

Combining (47) and (48), by Theorem 3.9 of [5], we have

$$(\bar{\boldsymbol{\lambda}}(n), \bar{\boldsymbol{W}}(n), \bar{\boldsymbol{L}}(n), \hat{\boldsymbol{\lambda}}(n), \hat{\boldsymbol{W}}(n), \bar{\boldsymbol{F}}^{\boldsymbol{c}}(n), \hat{\boldsymbol{F}}^{\boldsymbol{c}}(n)) \Rightarrow (\boldsymbol{\lambda}, \boldsymbol{W}, \boldsymbol{L}, \boldsymbol{\varLambda}, \boldsymbol{\varOmega}, \boldsymbol{F}^{\boldsymbol{c}}, \boldsymbol{\varGamma}) \quad \text{in} \quad \mathbb{R}^{5d} \times (\ell_1)^{2d}. \tag{49}$$

Now we turn to $\hat{\boldsymbol{L}}(n)$. Note that for $k = 0, 1, \cdots, d-1$, we can write the k^{th} component of $\hat{\boldsymbol{L}}(n)$ as

$$\begin{aligned}
\sqrt{n}(\bar{L}_k(n) - L_k) &= \sqrt{n}\Big(\sum_{j=0}^{\infty} \bar{\lambda}_{[k-j]}(n)\bar{F}^c_{[k-j],j}(n) - \sum_{j=0}^{\infty} \lambda_{[k-j]}F^c_{[k-j],j}\Big) \\
&= \sqrt{n}\sum_{j=0}^{\infty}\big(\bar{\lambda}_{[k-j]}(n)\bar{F}^c_{[k-j],j}(n) - \bar{\lambda}_{[k-j]}(n)F^c_{[k-j],j} + \bar{\lambda}_{[k-j]}(n)F^c_{[k-j],j} - \lambda_{[k-j]}F^c_{[k-j],j}\big) \\
&= h_k(\bar{\boldsymbol{\lambda}}(n), \hat{\boldsymbol{F}}^{\boldsymbol{c}}(n)) + h_k(\hat{\boldsymbol{\lambda}}(n), \hat{\boldsymbol{F}}^{\boldsymbol{c}}) \\
&= f_{\boldsymbol{F}^{\boldsymbol{c}},k}(\bar{\boldsymbol{\lambda}}(n), \hat{\boldsymbol{\lambda}}(n), \hat{\boldsymbol{F}}^{\boldsymbol{c}}(n)),
\end{aligned} \tag{50}$$

so

$$\hat{\boldsymbol{L}}(n) = f_{\boldsymbol{F}^{\boldsymbol{c}}}\big(\bar{\boldsymbol{\lambda}}(n), \hat{\boldsymbol{\lambda}}(n), \hat{\boldsymbol{F}}^{\boldsymbol{c}}(n)\big). \tag{51}$$

By Lemma 3.1, $f_{\boldsymbol{F}^{\boldsymbol{c}}}$ is a continuous function. In addition, $\boldsymbol{\varGamma} \in (\ell_1)^d$ w.p.1, so we can apply the continuous mapping theorem again to get

$$\begin{aligned}
(\bar{\boldsymbol{\lambda}}(n), \bar{\boldsymbol{W}}(n), \bar{\boldsymbol{L}}(n), \hat{\boldsymbol{\lambda}}(n), \hat{\boldsymbol{W}}(n), \hat{\boldsymbol{L}}(n), \bar{\boldsymbol{F}}^{\boldsymbol{c}}(n), \hat{\boldsymbol{F}}^{\boldsymbol{c}}(n)) &\Rightarrow (\boldsymbol{\lambda}, \boldsymbol{W}, \boldsymbol{L}, \boldsymbol{\varLambda}, \boldsymbol{\varOmega}, \boldsymbol{\varUpsilon}, \boldsymbol{F}^{\boldsymbol{c}}, \boldsymbol{\varGamma}) \\
&\quad \text{in} \quad \mathbb{R}^{6d} \times (\ell_1)^{2d},
\end{aligned} \tag{52}$$

where $\boldsymbol{\Upsilon} = f_{\boldsymbol{F^c}}(\boldsymbol{\lambda}, \boldsymbol{\Lambda}, \boldsymbol{\Gamma})$, which is what we want as in (18). ∎

Proof of Proposition 3.1. If $\hat{\boldsymbol{R}}(n) \Rightarrow 0$, then both $\bar{\boldsymbol{R}}(n) = \bar{\boldsymbol{L}}(n) - \bar{\boldsymbol{Q}}(n) \to \boldsymbol{0}$ and $\hat{\boldsymbol{L}}(n) - \hat{\boldsymbol{Q}}(n) \to 0$ in probability. By applying Theorem 3.1 of [5] based on (52), we have

$$(\bar{\boldsymbol{\lambda}}(n), \bar{\boldsymbol{W}}(n), \bar{\boldsymbol{Q}}(n), \bar{\boldsymbol{L}}(n), \hat{\boldsymbol{\lambda}}(n), \hat{\boldsymbol{W}}(n), \hat{\boldsymbol{Q}}(n), \hat{\boldsymbol{L}}(n), \bar{\boldsymbol{F}}^{\boldsymbol{c}}(n), \hat{\boldsymbol{F}}^{\boldsymbol{c}}(n), \hat{\boldsymbol{R}}(n)) \Rightarrow (\boldsymbol{\lambda}, \boldsymbol{W}, \boldsymbol{L}, \boldsymbol{L}, \boldsymbol{\Lambda}, \boldsymbol{\Omega}, \boldsymbol{\Upsilon}, \boldsymbol{\Upsilon}, \boldsymbol{F^c}, \boldsymbol{\Gamma}, \boldsymbol{0}) \quad \text{in} \quad \mathbb{R}^{9d} \times (\ell_1)^{2d}. \tag{53}$$

Now consider the departure processes $\bar{\boldsymbol{\delta}}(n)$ and $\hat{\boldsymbol{\delta}}(n)$. Note that

$$\begin{aligned}
\bar{\delta}_k(n) &= \frac{1}{n}\sum_{m=1}^{n} D_{k+(m-1)d} = \frac{1}{n}\sum_{m=1}^{n}(Q_{k+(m-1)d} - Q_{k+1+(m-1)d} + A_{k+1+(m-1)d}) \\
&= \begin{cases} \bar{Q}_k(n) - \bar{Q}_{k+1}(n) + \bar{\lambda}_{k+1}(n), & 0 \le k < d-1, \\ \bar{Q}_{d-1}(n) - \bar{Q}_0(n+1) + \bar{\lambda}_0(n+1) + \dfrac{1}{n}Q_0 - \dfrac{1}{n}A_0, & k = d-1, \end{cases}
\end{aligned} \tag{54}$$

and

$$\begin{aligned}
\hat{\delta}_k(n) &= \begin{cases} \sqrt{n}(\bar{Q}_k(n) - \bar{Q}_{k+1}(n) + \bar{\lambda}_{k+1}(n) - L_k + L_{k+1} - \lambda_{k+1}) \\ \sqrt{n}(\bar{Q}_{d-1}(n) - \bar{Q}_0(n+1) + \bar{\lambda}_0(n+1) - L_k + L_{k+1} - \lambda_{k+1} + \dfrac{1}{n}Q_0 - \dfrac{1}{n}A_0) \end{cases} \\
&= \begin{cases} \hat{Q}_k(n) - \hat{Q}_{k+1} + \hat{\lambda}_{k+1}(n), & 0 \le k < d-1, \\ \hat{Q}_{d-1}(n) - \hat{Q}_0 + \hat{\lambda}_0(n) + \dfrac{1}{\sqrt{n}}Q_0 - \dfrac{1}{\sqrt{n}}A_0, & k = d-1. \end{cases}
\end{aligned} \tag{55}$$

Once again by continuous mapping theorem, we get the joint convergence (20) in Proposition 3.1, where $\boldsymbol{\delta}$ is given by (8) and $\Delta_k = \Upsilon_{[k]} - \Upsilon_{[k+1]} + \Lambda_{[k+1]}$, $0 \le k \le d-1$. ∎

6.2 Proof of Corollary 3.1

To make the proof clear, we first establish two lemmas.

Lemma 6.1 *Assume $X_n \sim N(\mu_n, \sigma_n^2)$, $n = 1, 2, \cdots$, and $X_n \to X$ almost surely as $n \to \infty$. Then $X \sim N(\mu, \sigma^2)$ where $\mu = \lim_{n\to\infty} \mu_n$, $\sigma^2 = \lim_{n\to\infty} \sigma_n^2$ and $X_n \to X$ in L^2 as $n \to \infty$.*

Proof Let $\phi_{X_n}(t) = \mathbb{E}e^{itX_n} = \exp(i\mu_n t - 2^{-1}\sigma_n^2 t^2)$ be the characteristic function of X_n. Since $X_n \to X$ almost surely, by dominated convergence theorem, we know that $\lim_{n\to\infty} \mathbb{E}e^{itX_n} = \mathbb{E}e^{itX}$ for each t. So we must have $\lim_{n\to\infty} \mu_i = \mu$, $\lim_{n\to\infty} \sigma_i^2 = \sigma^2$ for some μ and $\sigma \ge 0$. (Note that we cannot have $\lim_{n\to\infty} \mu_n = \infty$ or $\lim_{n\to\infty} \sigma_n^2 = \infty$, which would contradict Lévy's continuity theorem.) Then we know that $\mathbb{E}e^{itX} = \exp(i\mu t - 2^{-1}\sigma^2 t^2)$, which shows that $X \sim N(\mu, \sigma^2)$.

Note that $\lim_{n\to\infty} \mu_n = \mu$ and $\lim_{n\to\infty} \sigma_n^2 = \sigma^2$ also implies that $\sup_n \mathbb{E}X_n^4 = \sup_n(\mu_n^4 + 6\mu_n^2\sigma_n^2 + 3\sigma_n^4) < \infty$. So $\{X_n^2\}$ is uniformly integrable. Hence $X_n \to X$ in L^2 as $n \to \infty$. □

Lemma 6.2 *Assume $\boldsymbol{N} \in \ell_1$ is a zero-mean normally distributed random variable with $Cov(\boldsymbol{N}, \boldsymbol{N}) = \Sigma^N$ and $\boldsymbol{a}, \boldsymbol{b} \in \ell_0$ are constants. Let $X \equiv \sum_{i=1}^{\infty} a_i N_i = \boldsymbol{a}^T \cdot \boldsymbol{N}$ and $Y \equiv \sum_{i=1}^{\infty} b_i N_i = \boldsymbol{b}^T \cdot \boldsymbol{N}$, then (X, Y) is jointly zero-mean normal distributed with*

$$\begin{aligned}
Cov(X,X) = Var(X) &= \sum_{i=1}^{\infty}\sum_{j=1}^{\infty} a_i a_j \Sigma_{i,j}^N < \infty, \\
Cov(Y,Y) = Var(Y) &= \sum_{i=1}^{\infty}\sum_{j=1}^{\infty} b_i b_j \Sigma_{i,j}^N < \infty \quad \textit{and} \\
Cov(X,Y) = \mathbb{E}(XY) &= \sum_{i=1}^{\infty}\sum_{j=1}^{\infty} a_i b_j \Sigma_{i,j}^N < \infty.
\end{aligned}$$

Proof By the definition of normal distribution in ℓ_1 space, X and Y are normally distributed random variables on $\mathbb{R}$. We only need to show that they are also jointly normally distributed, i.e. for any given α, $\beta \in \mathbb{R}$, $\alpha X + \beta Y$ is a zero-mean normal distributed random variable.

Denote $X_n = \sum_{i=1}^{n} a_i N_i$ and $Y_n = \sum_{i=1}^{n} b_i N_i$. Since X and Y are well defined almost surely, we know that $\alpha X_n + \beta Y_n \to \alpha X + \beta Y$ almost surely. Note that $\alpha X_n + \beta Y_n$ is normal distributed with mean zero and variance $\sum_{i=1}^{n}\sum_{j=1}^{n}(\alpha a_i + \beta b_i)(\alpha a_j + \beta b_j)\Sigma_{i,j}^N$. And we can apply Lemma 6.1 and know that $\alpha X + \beta Y$ has normal distribution with mean zero and variance $\sum_{i=1}^{\infty}\sum_{j=1}^{\infty}(\alpha a_i + \beta b_i)(\alpha a_j + \beta b_j)\Sigma_{i,j}^N$. So (X, Y) is jointly normally distributed.

Now we derive $\mathbb{E}(XY)$. By Lemma 6.1 we know that $X_n \to X$ and $Y_n \to Y$ in L^2 as well, so $X_n Y_n \to XY$ in L^1 and we have $\mathbb{E}(XY) = \lim_{n\to\infty} \mathbb{E}(X_n Y_n) = \lim_{n\to\infty} \sum_{i=1}^{n}\sum_{j=1}^{n} a_i b_j \Sigma_{i,j}^N = \sum_{i=1}^{\infty}\sum_{j=1}^{\infty} a_i b_j \Sigma_{i,j}^N < \infty$. Because X and Y are zero-mean normal distributed, we have $\mathrm{Cov}(X, Y) = \mathbb{E}(XY)$, where $\mathrm{Cov}(X, X)$ and $\mathrm{Cov}(Y, Y)$ are special cases with $X = Y$. □

Proof of Corollary 3.1. Lemma 6.2 can be easily generalized to $\boldsymbol{N} \in (\ell_1)^d$ only with some tedious steps. Since $\boldsymbol{\Gamma} \in (\ell_1)^d$ almost surely, $\boldsymbol{\Omega} = g(\boldsymbol{\Gamma})$ and $\boldsymbol{\Upsilon} = f_{\boldsymbol{F^c}}(\boldsymbol{\lambda}, \boldsymbol{\Lambda}, \boldsymbol{\Gamma})$ are well defined a.s. In addition, with the Gaussian assumption, we may apply the generalized version of Lemma 6.2 and know that $(\boldsymbol{\Omega}, \boldsymbol{\Upsilon})$ has a zero-mean jointly normal distribution with each element being well defined a.s. and in L^2, where

$$
\begin{aligned}
(\mathrm{Cov}(\boldsymbol{\Omega},\boldsymbol{\Omega}))_{k,l} &= \mathrm{Cov}(\sum_{j=0}^{\infty}\Gamma_{k,j},\sum_{j=0}^{\infty}\Gamma_{l,j}) = \sum_{i=1}^{\infty}\sum_{j=1}^{\infty}\Sigma_{i,j}^{\Gamma:k,l},\\
(\mathrm{Cov}(\boldsymbol{\Upsilon},\boldsymbol{\Upsilon}))_{k,l} &= \mathrm{Cov}(f_{\boldsymbol{F^c},k}(\boldsymbol{\lambda},\boldsymbol{\Lambda},\boldsymbol{\Gamma}), f_{\boldsymbol{F^c},l}(\boldsymbol{\lambda},\boldsymbol{\Lambda},\boldsymbol{\Gamma}))\\
&= \mathrm{Cov}(\sum_{j=0}^{\infty}\lambda_{[k-j]}\Gamma_{[k-j],j} + \sum_{j=0}^{\infty}F^c_{[k-j],j}\Lambda_{[k-j]}, \sum_{j=0}^{\infty}\lambda_{[l-j]}\Gamma_{[l-j],j} + \sum_{j=0}^{\infty}F^c_{[l-j],j}\Lambda_{[l-j]})\\
&= \sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\lambda_{[k-i]}\lambda_{[l-j]}\Sigma_{i+1,j+1}^{\Gamma:[k-i],[l-j]} + \sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\lambda_{[k-i]}F^c_{[l-j],j}\Sigma_{[l-j]+1,i+1}^{\Lambda,\Gamma:[k-j]}\\
&\quad + \sum_{i=0}^{\infty}\sum_{j=0}^{\infty}F^c_{[k-i],i}\lambda_{[l-j]}\Sigma_{[k-i]+1,j+1}^{\Lambda,\Gamma:[l-j]} + \sum_{i=0}^{\infty}\sum_{j=0}F^c_{[k-i],i}F^c_{[l-j],j}\Sigma_{[k-i]+1,[l-j]+1}^{\Lambda},\\
(\mathrm{Cov}(\boldsymbol{\Omega},\boldsymbol{\Upsilon}))_{k,l} &= \mathrm{Cov}(\sum_{j=0}^{\infty}\Gamma_{k,j},\sum_{j=0}^{\infty}\lambda_{[l-j]}\Gamma_{[l-j],j} + \sum_{j=0}^{\infty}F^c_{[l-j,j]}\Lambda_{[l-j]})\\
&= \sum_{i=1}^{\infty}\sum_{j=0}^{\infty}\lambda_{[l-j]}\Sigma_{i+1,j+1}^{\Gamma:k,[l-j]} + \sum_{i=1}^{\infty}\sum_{j=0}^{\infty}F^c_{[l-j],j}\Sigma_{[l-j]+1,i+1}^{\Lambda,\Gamma:k}. \quad \blacksquare
\end{aligned}
$$

6.3 Proof of Corollary 3.2

If the LoS is bounded by J, then $\bar{F}^c_{k,j}(n) = 0$ for all $0 \le k \le d-1$ and n when $j > J$. So if we have $(C1)$, then $F^c_{k,j} = 0$ must holds for all k and $j > J$ and $\hat{\boldsymbol{F}}^{\boldsymbol{c}}(n) \in (\ell_1)^d$ w.p. 1. Hence $\Gamma_{k,j} = 0$ for all k and $j > J$, and $\boldsymbol{\Gamma} \in (\ell_1)^d$ w.p. 1.

To see that $\hat{\boldsymbol{R}}(n) \Rightarrow 0$ holds, by (7), $\hat{\boldsymbol{R}}(n) = \hat{\boldsymbol{L}}(n) - \hat{\boldsymbol{Q}}(n) = \sqrt{n}(\bar{\boldsymbol{L}}(n) - \bar{\boldsymbol{Q}}(n))$, so it suffices to show that $\sqrt{n}E(n) \to 0$, where $E(n) \equiv ||\bar{\boldsymbol{L}}(n) - \bar{\boldsymbol{Q}}(n)||_1$. We take M such that $(M-1)d < J \le Md$. When $n > M$, because $Y_{k,j} = 0$ for $j > Md \ge J$, we have

$$
\begin{aligned}
\sqrt{n}E(n) &= \sqrt{n}\frac{1}{n}\sum_{m=1}^{n}\sum_{j=1}^{d}\sum_{s=(n-m)d}^{\infty}Y_{d-j+(m-1)d,j+s} = n^{-1/2}\sum_{m=n-M+1}^{n}\sum_{j=1}^{d}\sum_{s=(n-m)d}^{Md}Y_{d-j+(m-1)d,j+s}\\
&\le \sum_{j=1}^{d}n^{-1/2}\sum_{m=n-M+1}^{n}MdA_{d-j+(m-1)d} \le n^{-1/2}d^2M^2C \to 0 \quad \text{as} \quad n \to \infty,
\end{aligned}
\tag{56}
$$

where C is the upper bound for the number of arrivals within a discrete time period. This establishes $\hat{\boldsymbol{R}}(n) \Rightarrow 0$. $\blacksquare$

6.4 Proof of Theorem 4.1

To prove Theorem 4.1, we need a lemma to establish $(C1)$.

Lemma 6.3 *Suppose W_k, $k = 0, 1, \cdots, d-1$, are non-negative integer-valued random variables with $F^c_{k,j} \equiv 1 - F_{k,j} \equiv P(W_k \geq j) \sim O(j^{-(3+\delta)})$, $j \geq 0$, for some $\delta > 0$. $W_k^{(i)}$ are i.i.d. samples of W_k, denote $\boldsymbol{I}_k^{(i)} \equiv (I_{W_k^{(i)} \geq 0}, I_{W_k^{(i)} \geq 1}, \cdots)$, $\boldsymbol{F}_k^c \equiv (F^c_{k,0}, F^c_{k,1}, \cdots)$, and $\boldsymbol{X}_k^{(i)} \equiv \boldsymbol{I}_k^{(i)} - \boldsymbol{F}_k^c$. Assume $\boldsymbol{Y}^{(i)}$ are i.i.d. non-negative integer-valued random variables in $\mathbb{R}^d$ with $\mathbb{E}\boldsymbol{Y}^{(i)} = \boldsymbol{\mu_Y} \equiv (\mu_{Y,0}, \mu_{Y,1}, \cdots, \mu_{Y,d-1}) > \mathbf{0}$ and $Var(\boldsymbol{Y}^{(i)}) = \Sigma^Y$, where all the $W_k^{(i)}$ and $\boldsymbol{Y}^{(j)}$ are independent for $k = 0, 1, \cdots, d-1$, and $i, j \geq 1$. Let*

$$\boldsymbol{S}(n) \equiv (S_0(n), S_1(n), \cdots, S_{d-1}(n)) \equiv \sum_{i=1}^{n} \boldsymbol{Y}^{(i)}$$

and

$$\boldsymbol{G}(n) \equiv (\boldsymbol{G}_0(n), \boldsymbol{G}_1(n), \cdots, \boldsymbol{G}_{d-1}(n)) \equiv (\sum_{i=1}^{S_0(n)} \boldsymbol{X}_0^{(i)}, \sum_{i=1}^{S_1(n)} \boldsymbol{X}_1^{(i)}, \cdots, \sum_{i=1}^{S_{d-1}(n)} \boldsymbol{X}_{d-1}^{(i)}).$$

We claim that

$$n^{-1/2}(\boldsymbol{S}(n) - n\boldsymbol{\mu_Y}, \boldsymbol{G}(n)) \Rightarrow (\boldsymbol{\Lambda}, \boldsymbol{\Gamma}) \quad in \quad \mathbb{R}^d \times (\ell_1)^d, \tag{57}$$

where $\boldsymbol{\Gamma} = (\boldsymbol{\Gamma}_0, \boldsymbol{\Gamma}_1, \cdots, \boldsymbol{\Gamma}_{d-1})$ and $(\boldsymbol{\Lambda}, \boldsymbol{\Gamma})$ has a zero-mean Gaussian distribution in $\mathbb{R}^d \times (\ell_1)^d$ with $\boldsymbol{\Lambda} \sim N(\mathbf{0}, \Sigma^Y)$, $Cov(\Gamma_{k,j}, \Gamma_{k,s}) = \mu_{Y,k} F_{k,j} F^c_{k,s}$ for $0 \leq k \leq d-1$ and $0 \leq j \leq s$, and $\boldsymbol{\Lambda}, \boldsymbol{\Gamma}_0, \boldsymbol{\Gamma}_1, \cdots, \boldsymbol{\Gamma}_{d-1}$ are independent.

Proof The classical multivariate CLT implies that

$$n^{-1/2}(\boldsymbol{S}(n) - n\boldsymbol{\mu_Y}) \Rightarrow \boldsymbol{\Lambda} \sim N(\mathbf{0}, \Sigma^Y). \tag{58}$$

Let $\boldsymbol{Z}_k \in \ell_1$ for $0 \leq k \leq d-1$ be zero-mean Gaussian distributed random variables with $\mathrm{Cov}(Z_{k,j}, Z_{k,l}) = F_{k,j} F^c_{k,l}$. By Theorem 1.1 of [13],

$$(S_k(n))^{-1/2} \sum_{i=1}^{S_k(n)} \boldsymbol{X}_k^{(i)} \Rightarrow \boldsymbol{Z}_k \quad \text{in} \quad \ell_1 \quad \text{for} \quad 0 \leq k \leq d-1, \tag{59}$$

then after applying Slutsky's theorem, we have

$$n^{-1/2} \sum_{i=1}^{S_k(n)} \boldsymbol{X}_k^{(i)} \Rightarrow \boldsymbol{\Gamma}_k = \mu_{Y,k}^{1/2} \boldsymbol{Z}_k \quad \text{in} \quad \ell_1 \quad \text{for} \quad 0 \leq k \leq d-1, \tag{60}$$

so that $\boldsymbol{\Gamma}_k$ has a zero-mean Gaussian distribution with $\mathrm{Cov}(\Gamma_{k,j}, \Gamma_{k,s}) = \mu_{Y,k} F_{k,j} F^c_{k,s}$ for $j \leq s$. Unfortunately we cannot get the joint convergence directly since they are not independent of each other, however, we can use independent copies of $\boldsymbol{Y}^{(i)}$ as a bridge to prove it.

Assume that $\{\tilde{\boldsymbol{Y}}^{(i)}\}$ are i.i.d. copies of $\boldsymbol{Y}^{(i)}$ and are also independent of other variables. Let

$$\tilde{\boldsymbol{S}}(n) \equiv (\tilde{S}_0(n), \tilde{S}_1(n), \cdots, \tilde{S}_{d-1}(n)) \equiv \sum_{i=1}^{n} \tilde{\boldsymbol{Y}}^{(i)}.$$

Because $n^{-1/2}(\boldsymbol{S}(n) - n\boldsymbol{\mu}_{\boldsymbol{Y}})$ is independent of $n^{-1/2}\tilde{\boldsymbol{G}}_0(n) \equiv n^{-1/2}\sum_{i=1}^{\tilde{S}_0(n)} \boldsymbol{X}_0^{(i)}$, we can apply Theorem 11.4.4 of [28] to obtain

$$n^{-1/2}(\boldsymbol{S}(n) - n\boldsymbol{\mu}_{\boldsymbol{Y}}, \tilde{\boldsymbol{G}}_0(n)) \Rightarrow (\boldsymbol{\Lambda}, \boldsymbol{\Gamma}_0) \quad \text{in} \quad \mathbb{R}^d \times \ell_1, \tag{61}$$

where we can make $\boldsymbol{\Lambda}$ be independent of $\boldsymbol{\Gamma}_0$. For what we want, we need to show that $n^{-1/2}||\boldsymbol{G}_0(n) - \tilde{\boldsymbol{G}}_0(n)||_1 \to 0$ in probability as $n \to \infty$ and then apply Theorem 3.1 from [5]. For any $\epsilon > 0$,

$$\begin{aligned}
&P(n^{-1/2}||\boldsymbol{G}_0(n) - \tilde{\boldsymbol{G}}_0(n)||_1 > \epsilon) = P(||\sum_{i=1}^{S_0(n)} \boldsymbol{X}_0^{(i)} - \sum_{i=1}^{\tilde{S}_0(n)} \boldsymbol{X}_0^{(i)}||_1 > n^{1/2}\epsilon)\\
\le& P(||\sum_{i=1}^{S_0(n)} \boldsymbol{X}_0^{(i)} - \sum_{i=1}^{\tilde{S}_0(n)} \boldsymbol{X}_0^{(i)}||_1 > n^{1/2}\epsilon, |S_0(n) - \tilde{S}_0(n)| \le n^{3/4}) + P(|S_0(n) - \tilde{S}_0(n)| > n^{3/4})\\
\le& 2P(\max_{I=1,2,\cdots,\lfloor n^{3/4}\rfloor} ||\sum_{i=1}^{I} \boldsymbol{X}_0^{(i)}||_1 > n^{1/2}\epsilon) + P(|S_0(n) - \tilde{S}_0(n)| > n^{3/4})\\
=& 2P(\max_{I=1,2,\cdots,\lfloor n^{3/4}\rfloor} \sum_{j=0}^{\infty} |\sum_{i=1}^{I} X_{0,j}^{(i)}| > n^{1/2}\epsilon) + P(|S_0(n) - \tilde{S}_0(n)| > n^{3/4})\\
\le& 2P(\sum_{j=0}^{\infty} \max_{I=1,2,\cdots,\lfloor n^{3/4}\rfloor} |\sum_{i=1}^{I} X_{0,j}^{(i)}| > n^{1/2}\epsilon) + P(|S_0(n) - \tilde{S}_0(n)| > n^{3/4}).
\end{aligned} \tag{62}$$

For the first part, let $\delta_1 \le \delta/2$ and $C \equiv (\sum_{j=0}^{\infty} j^{-(1+\delta_1)})^{-1}$ be constants, note that $\text{Var}(X_{0,j}^{(i)}) = F_{0,j}F_{0,j}^c$, so

$$\begin{aligned}
&P(\sum_{j=0}^{\infty} \max_{I=1,2,\cdots,\lfloor n^{3/4}\rfloor} |\sum_{i=1}^{I} X_{0,j}^{(i)}| > n^{1/2}\epsilon) \le \sum_{j=0}^{\infty} P(\max_{I=1,2,\cdots,\lfloor n^{3/4}\rfloor} |\sum_{i=1}^{I} X_{0,j}^{(i)}| > Cn^{1/2}\epsilon j^{-(1+\delta_1)})\\
\le& 3\sum_{j=0}^{\infty} \max_{I=1,2,\cdots,\lfloor n^{3/4}\rfloor} P(|\sum_{i=1}^{I} X_{0,j}^{(i)}| > Cn^{1/2}\epsilon j^{-(1+\delta_1)}/3) \le 27\sum_{j=0}^{\infty} \max_{I=1,2,\cdots,\lfloor n^{3/4}\rfloor} \frac{\text{Var}(\sum_{i=1}^{I} X_{0,j}^{(i)})}{C^2 n\epsilon^2 j^{-(2+2\delta_1)}}\\
\le& 27\sum_{j=0}^{\infty} C^{-2}n^{-1}\epsilon^{-2}j^{2+2\delta_1}\lfloor n^{3/4}\rfloor F_{0,j}F_{0,j}^c\\
\le& 27C^{-2}\epsilon^{-2}n^{-1}\lfloor n^{3/4}\rfloor \sum_{j=0}^{\infty} j^{2+2\delta_1} F_{0,j}^c \to 0 \quad \text{as} \quad n \to \infty,
\end{aligned} \tag{63}$$

because as a consequence of $F_{k,j}^c \sim O(j^{-(3+\delta)})$, $\sum_{j=0}^{\infty} j^{2+2\delta_1} F_{0,j}^c \le \infty$, and the second and third inequalities follow from Etemadi's inequality (see page 256 of [5]) and Chebyshev's inequality respectively.

As for the second part, again using Chebyshev's inequality, we have

$$P(|S_0(n) - \tilde{S}_0(n)| > n^{3/4}) \leq \frac{\mathrm{Var}(S_0(n) - \tilde{S}_0(n))}{n^{3/2}} \leq \frac{2n\Sigma^Y_{1,1}}{n^{3/2}} \to 0 \quad \text{as} \quad n \to \infty. \tag{64}$$

So (62) goes to 0 as $n \to \infty$, hence

$$n^{-1/2}(\boldsymbol{S}(n) - n\boldsymbol{\mu_Y}, \boldsymbol{G}_0(n)) \Rightarrow (\boldsymbol{\Lambda}, \boldsymbol{\Gamma}_0) \quad \text{in} \quad \mathbb{R}^d \times \ell_1. \tag{65}$$

Using the same argument (making new i.i.d. copies of $\boldsymbol{Y}^{(i)}$), we can add $\boldsymbol{G}_1(n), \boldsymbol{G}_2(n), \cdots, \boldsymbol{G}_{d-1}(n)$ one by one and in the end get (57). □

Proof of Theorem 4.1. Take $\boldsymbol{\lambda} = \mathbb{E}(A_{0+(m-1)d}, A_{1+(m-1)d}, \cdots, A_{d-1+(m-1)d})$ and $\boldsymbol{F^c}$ to be the complementary distribution functions of the LoS. Let $\boldsymbol{W}$ and $\boldsymbol{L}$ be as in (C2), in which case $\boldsymbol{W}$ is the vector of mean LoS for a period.

We can use Lemma 6.3 to establish $(C1)$ by letting $\boldsymbol{Y}^{(i)}$ be the $\boldsymbol{A}_i$ and $W^{(i)}_k$ be the LoS of i^{th} customer that arrived at discrete time period $k+(m-1)d$ for all $m = 1, 2, \cdots$. (So $\mathbb{E}W^{(i)}_k = W_k$ and $\sigma^2_{W,k} \equiv \mathrm{Var}(W^{(i)}_k) < \infty$.) The conclusion of Lemma 6.3 is exactly $(C1)$.

Then we need to establish $\hat{\boldsymbol{R}}(n) \Rightarrow 0$. Note that $(I1)$ and $(I2)$ imply that the SLLN holds, i.e. equation (2) hold, so that we have Theorem 2.1. Since $\hat{\boldsymbol{R}}(n) = \hat{\boldsymbol{L}}(n) - \hat{\boldsymbol{Q}}(n) = n^{1/2}(\bar{\boldsymbol{L}}(n) - \bar{\boldsymbol{Q}}(n))$, it suffices to show that for each $0 \leq k \leq d-1$, $n^{1/2}(\bar{L}_k(n) - \bar{Q}_k(n)) \to 0$ in probability as $n \to \infty$.

From the proof of Theorem 2.1 in [31], we know that

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

So for any $\epsilon > 0$,

$$\begin{aligned}
&P(n^{1/2}(\bar{L}_k(n) - \bar{Q}_k(n)) > \epsilon) \\
=&P(n^{-1/2} \sum_{m=1}^{n} \sum_{j=1}^{d} \sum_{s=n-m+1}^{\infty} Y_{d-j+(m-1)d, j+k+(s-1)d} > \epsilon) \\
\leq&\sum_{j=1}^{d} P(\sum_{m=1}^{n} \sum_{s=n-m+1}^{\infty} Y_{d-j+(m-1)d, j+k+(s-1)d} > n^{1/2}d^{-1}\epsilon) \\
\leq&\sum_{j=1}^{d} (P(\sum_{m=1}^{n-\lfloor n^{1/4} \rfloor} \sum_{s=n-m+1}^{\infty} Y_{d-j+(m-1)d, j+k+(s-1)d} > n^{1/2}d^{-1}\epsilon) \\
&+ P(\sum_{m=n-\lfloor n^{1/4} \rfloor+1}^{n} \sum_{s=n-m+1}^{\infty} Y_{d-j+(m-1)d, j+k+(s-1)d} > n^{1/2}d^{-1}\epsilon)).
\end{aligned} \tag{66}$$

Then we only need to prove that each of the two probabilities go to zero as $n \to \infty$.

For the first part, we have

$$
\begin{aligned}
&P(\sum_{m=1}^{n-\lfloor n^{1/4}\rfloor} \sum_{s=n-m+1}^{\infty} Y_{d-j+(m-1)d,j+k+(s-1)d} > n^{1/2}d^{-1}\epsilon) \\
\leq &P(\sum_{m=1}^{n-\lfloor n^{1/4}\rfloor} \sum_{s=\lfloor n^{1/4}\rfloor+1}^{\infty} Y_{d-j+(m-1)d,j+k+(s-1)d} > n^{1/2}d^{-1}\epsilon) \\
\leq &P(\sum_{m=1}^{n} \sum_{s=\lfloor n^{1/4}\rfloor+1}^{\infty} Y_{d-j+(m-1)d,j+k+(s-1)d} > n^{1/2}d^{-1}\epsilon) \\
\leq &\frac{\sum_{m=1}^{n} \sum_{s=\lfloor n^{1/4}\rfloor+1}^{\infty} \mathbb{E}Y_{d-j+(m-1)d,j+k+(s-1)d}}{n^{1/2}d^{-1}\epsilon} \leq n^{1/2}d\epsilon^{-1} \sum_{s=\lfloor n^{1/4}\rfloor+1}^{\infty} \lambda_{d-j}F^c_{d-j,j+k+(s-1)d} \\
\leq &n^{1/2}d\lambda_{d-j}\epsilon^{-1} \sum_{s=\lfloor n^{1/4}\rfloor+1}^{\infty} F^c_{d-j,s} \leq n^{1/2}Cd\lambda_{d-j}\epsilon^{-1}\int_{\lfloor n^{1/4}\rfloor}^{\infty} s^{-(3+\delta)}ds \\
= &n^{1/2}Cd\lambda_{d-j}(2+\delta)^{-1}\epsilon^{-1}\lfloor n^{1/4}\rfloor^{-(2+\delta)} \to 0 \quad \text{as} \quad n \to \infty. \qquad (67)
\end{aligned}
$$

For the second part, let $S_k(n) = \sum_{m=1}^{n} A_{k+(m-1)d}$. Then

$$
\begin{aligned}
&P(\sum_{m=n-\lfloor n^{1/4}\rfloor+1}^{n} \sum_{s=n-m+1}^{\infty} Y_{d-j+(m-1)d,j+k+(s-1)d} > n^{1/2}d^{-1}\epsilon) \\
\leq &P(\sum_{m=n-\lfloor n^{1/4}\rfloor+1}^{n} \sum_{s=0}^{\infty} Y_{d-j+(m-1)d,s} > n^{1/2}d^{-1}\epsilon) = P(\sum_{i=1}^{S_{d-j}(\lfloor n^{1/4}\rfloor)} W^{(i)}_{d-j} > n^{1/2}d^{-1}\epsilon) \\
\leq &\frac{\mathrm{Var}(\sum_{i=1}^{S_{d-j}(\lfloor n^{1/4}\rfloor)} W^{(i)}_{d-j}) + (\mathbb{E}(\sum_{i=1}^{S_{d-j}(\lfloor n^{1/4}\rfloor)} W^{(i)}_{d-j}))^2}{nd^{-2}\epsilon^2} \\
= &n^{-1}d^2\epsilon^{-2}(\lfloor n^{1/4}\rfloor \Sigma^{\Lambda}_{d-j+1,d-j+1}W_{d-j} + \lfloor n^{1/4}\rfloor\lambda_{d-j}\sigma^2_{W,d-j} + \lfloor n^{1/4}\rfloor^2\lambda^2_{d-j}W^2_{d-j}) \\
&\to 0 \quad \text{as} \quad n \to \infty. \qquad (68)
\end{aligned}
$$

Hence, we have proved that $n^{1/2}(\bar{L}_k(n) - \bar{Q}_k(n)) \to 0$ in probability as $n \to \infty$, i.e., $\hat{\boldsymbol{R}}(n) \Rightarrow 0$. Since we have established all three conditions in Theorem 3.2 and Proposition 3.1, so their conclusions follow. As a special case of Corollary 3.1, the covariance matrix of $(\boldsymbol{\Omega}, \boldsymbol{\Upsilon})$ has the same form as in (22) with $\Sigma^{\Lambda:k,l} = 0$ for $k \neq l$, $\Sigma^{\Lambda,\Gamma:k} = 0$ for all k and $\Sigma^{\Gamma:k,k}_{i+1,j+1} = \lambda_k F_{k,i}F^c_{k,j}$ for $0 \leq i \leq j$. ∎

6.5 Proof of Theorem 4.2

As we discussed before Theorem 4.2, we will first use a CLT for stationary processes with mixing conditions to $\{\boldsymbol{Y}_n : n \in \mathbb{N}\}$ (Step 1). Such type of CLT was established by Ibragimov; see Theorem 18.5.3 in [15] or

Theorem 0 in [6]. We apply it in $\mathbb{R}^{dJ}$ by utilizing Cramér-Wold device; see Theorem 29.4 in [4]. Then, in Step 2, we show that $(C1)$ holds. Finally, in Step 3, we show $\hat{\boldsymbol{R}}(n) \Rightarrow \boldsymbol{0}$ to complete the proof.

Step 1: Firstly, by the definition of $Y_{i,j}$ and A_i, we observe that $\mathbb{E}||\boldsymbol{A}_0||_\infty^{2+\delta} < \infty$ in $(S2)$ implies that $\mathbb{E}|Y_{k+nd,j}|^{2+\delta} < \infty$ for all $n \geq 0$, $0 \leq k \leq d-1$ and $0 \leq j \leq J-1$. Let $F_{k,j}^c \equiv \dfrac{\mathbb{E}Y_{k,j}}{\mathbb{E}Y_{k,0}} = \dfrac{\mathbb{E}Y_{k,j}}{\lambda_k}$ for $0 \leq k \leq d-1$ and $0 \leq j \leq J-1$, where $\boldsymbol{\lambda} = (\lambda_0, \lambda_1, \cdots, \lambda_{d-1})$ is in $(S2)$. Now the $\boldsymbol{F}$ we defined in (8) can also be reduced to finite dimensional space $\mathbb{R}^{d\times J}$, i.e.

$$\boldsymbol{F^c} = \begin{pmatrix} F_{0,0} & F_{0,1} & \cdots & F_{0,J-1} \\ F_{1,0} & F_{1,1} & \cdots & F_{1,J-1} \\ \vdots & \vdots & \vdots & \vdots \\ F_{d-1,0} & F_{d-1,1} & \cdots & F_{d-1,J-1} \end{pmatrix}. \tag{69}$$

We want to show that $\{\boldsymbol{Y}_n : n \in \mathbb{N}\}$ satisfies a CLT. By Cramér-Wold device, we only need to show that $\sum_{k=0}^{d-1}\sum_{j=0}^{J-1}\theta_{k,j}Y_{k+nd,j}$ satisfies the corresponding CLT for each $\boldsymbol{\theta} = (\theta_{k,j})_{d\times J} \in \mathbb{R}^{d\times J}$. Let

$$\{Z_n(\boldsymbol{\theta}) \equiv \sum_{k=0}^{d-1}\sum_{j=0}^{J-1}\theta_{k,j}(Y_{k+nd,j} - \lambda_k F_{k,j}^c), n \in \mathbb{N}\}, \tag{70}$$

be the centralized strictly stationary process, and we need to show that it satisfies Theorem 0 in [6], i.e. for some $\delta > 0$, $\mathbb{E}|Z_n(\boldsymbol{\theta})|^{2+\delta} < \infty$ and $\sum_{n=1}^{\infty}\alpha_{Z_{\boldsymbol{\theta}}}(n)^{\delta/(2+\delta)}$, where $\alpha_{Z_{\boldsymbol{\theta}}}(n)$ is the strong mixing coefficient for $\{Z_n\}$ like we introduced in (29). We take the δ as in $(S2)$. Note that

$$\begin{aligned}\mathbb{E}|Z_n(\boldsymbol{\theta})|^{2+\delta} &= \mathbb{E}|\sum_{k=0}^{d-1}\sum_{j=0}^{J-1}\theta_{k,j}(Y_{k+nd,j} - \lambda_k F_{k,j}^c)|^{2+\delta} \leq \sum_{k=0}^{d-1}\sum_{j=0}^{J-1}|\theta_{k,j}|^{2+\delta}\mathbb{E}|Y_{k+nd,j} - \lambda_k F_{k,j}^c|^{2+\delta} \\ &\leq \sum_{k=0}^{d-1}\sum_{j=0}^{J-1}|\theta_{k,j}|^{2+\delta}((\mathbb{E}|Y_{k+nd,j}|^{2+\delta})^{1/(2+\delta)} + \lambda_k F_{k,j}^c)^{2+\delta} < \infty, \end{aligned} \tag{71}$$

where the first inequality uses the linearity of expectation and the second is by Minkowski's inequality. Since $Z_n(\boldsymbol{\theta})$ is a linear combination of the elements of $\boldsymbol{Y}_n$, the corresponding sigma-algebra generated by it is smaller than the one generated by $\boldsymbol{Y}_n$, so that by the definition of strong mixing coefficient we know that $\alpha_{Z_{\boldsymbol{\theta}}}(n) \leq \alpha(n)$. Hence, given $\sum_{n=1}^{\infty}\alpha(n)^{\delta/(2+\delta)} < \infty$ as in $(S2)$, we have $\sum_{n=1}^{\infty}\alpha_{Z_{\boldsymbol{\theta}}}(n)^{\delta/(2+\delta)} < \infty$. By Theorem 0 in [6], we know that $\sigma_{\boldsymbol{\theta}}^2 = \mathbb{E}Z_0(\boldsymbol{\theta})^2 + 2\sum_{i=1}^{\infty}\mathbb{E}(Z_0(\boldsymbol{\theta})Z_i(\boldsymbol{\theta}))$ exists with the sum being absolutely convergent and $n^{-1/2}\sum_{i=0}^{n-1}Z_i(\boldsymbol{\theta}) \Rightarrow N(0, \sigma_{\boldsymbol{\theta}}^2)$. Let

$$\hat{\boldsymbol{Y}}(n) \equiv \sqrt{n}(\bar{\boldsymbol{Y}}(n) - diag(\boldsymbol{\lambda})\boldsymbol{F^c}) \equiv \sqrt{n}(\frac{1}{n}\sum_{i=0}^{n-1}\boldsymbol{Y}_i - diag(\boldsymbol{\lambda})\boldsymbol{F^c}), \tag{72}$$

where $diag(\boldsymbol{\lambda}) \in \mathbb{R}^{d\times d}$ is the diagonal matrix with $\boldsymbol{\lambda}$ on the diagonal. By Cramér-Wold device, we know that

$$\hat{\boldsymbol{Y}}(n) \Rightarrow \boldsymbol{\psi} \equiv \begin{pmatrix} \psi_{0,0} & \psi_{0,1} & \cdots & \psi_{0,J-1} \\ \psi_{1,0} & \psi_{1,1} & \cdots & \psi_{1,J-1} \\ \vdots & \vdots & \vdots & \vdots \\ \psi_{d-1,0} & \psi_{d-1,1} & \cdots & \psi_{d-1,J-1} \end{pmatrix} \quad \text{as} \quad n \to \infty, \tag{73}$$

where $\boldsymbol{\psi}$ is normal distributed with mean 0 and covariance

$$\mathrm{Cov}(\psi_{k,j}, \psi_{l,s}) = \mathbb{E}((Y_{k,j} - \lambda_k, F^c_{k,j})(Y_{l,s} - \lambda_l F^c_{l,s})) + 2\sum_{i=1}^{\infty} \mathbb{E}((y_{k,j} - \lambda_k F^c_{k,j})(Y_{l+id,j} - \lambda_l F^c_{l,s})). \tag{74}$$

Step 2: Now we show that (73) implies conditon $(C1)$ holds. Actually it suffices to show that $(\hat{\boldsymbol{\lambda}}(n), \hat{\boldsymbol{F}}^{\boldsymbol{c}}(n))$ is actually a continuous function of $\hat{\boldsymbol{Y}}(n)$. It is trivial to see that $\hat{\boldsymbol{\lambda}}(n) = \hat{\boldsymbol{Y}}(n)\boldsymbol{e}_1$, where $\boldsymbol{e}_1 = (1, 0, 0, \cdots, 0)^T \in \mathbb{R}^J$ is a continuous map from $\mathbb{R}^{d\times J}$ to $\mathbb{R}^d$. For $\hat{\boldsymbol{F}}^{\boldsymbol{c}}(n) = \sqrt{n}(\bar{\boldsymbol{F}}^{\boldsymbol{c}}(n) - \boldsymbol{F}^{\boldsymbol{c}})$, note that it is, by $(S3)$, a finite dimensional matrix, so it suffices to show that each element of it is a continuous function of $\hat{\boldsymbol{Y}}(n)$. To see it, observe that

$$\begin{aligned} \hat{F}^c_{k,j}(n) &= \sqrt{n}(\bar{F}^c_{k,j}(n) - F^c_{k,j}) = \sqrt{n}\Big(\frac{\bar{Y}_{k,j}(n)}{\bar{Y}_{k,0}(n)} - F^c_{k,j}\Big) = \sqrt{n}\frac{(\bar{Y}_{k,j}(n) - \lambda_k F^c_{k,j}) - F^c_{k,j}(\bar{Y}_{k,0}(n) - \lambda_k)}{\bar{Y}_{k,0}(n)} \\ &= \frac{\hat{Y}_{k,j}(n) - F^c_{k,j}\hat{Y}_{k,0}(n)}{\bar{Y}_{k,0}(n)}. \end{aligned} \tag{75}$$

In addition, (73) implies that $\bar{Y}_{k,0}(n) \to \lambda_k$ in probability, so by continuous mapping theorem, Slutskys theorem and Cramér-Wold device, we know that

$$(\hat{\boldsymbol{\lambda}}(n), \hat{\boldsymbol{F}}^{\boldsymbol{c}}(n)) \Rightarrow (\boldsymbol{\Lambda}, \boldsymbol{\Gamma}) \quad \text{in} \quad \mathbb{R}^d \times \mathbb{R}^{d\times J}, \tag{76}$$

where $\boldsymbol{\Lambda} = (\psi_{0,0}, \psi_{1,0}, \cdots, \psi_{d-1,0})$ and $\Gamma_{k,j} = \dfrac{\psi_{k,j} - F^c_{k,j}\psi_{k,0}}{\lambda_k}$, $0 \le k \le d-1$, $0 \le j \le J-1$.

Step 3: Finally, since the LoSs are bounded, we only need to show $\hat{\boldsymbol{R}}(n) \Rightarrow \boldsymbol{0}$ to establish Theorem 3.2 and Proposition 3.1. Moreover, if $(\boldsymbol{\Lambda}, \boldsymbol{\Gamma})$ has a Gaussian distribution, Theorem 3.1 holds as well.

Analogously to the proof of Corollary 3.2, it suffices to show that $\sqrt{n}E(n) \to 0$ in probability. To see that, for any $\epsilon > 0$, take the same M as in the proof of Corollary 3.2, and based on (56), we have

$$\begin{aligned} P(\sqrt{n}E(n) > \epsilon) &\le P(n^{-1/2}Md\sum_{j=1}^{d}\sum_{m=n-M+1}^{n} A_{d-j+(m-1)d} > \epsilon) \\ &\le P(n^{-1/2}Md^2\sum_{m=n-M+1}^{n} ||\boldsymbol{A}_{(m-1)}||_\infty > \epsilon) \le \frac{M^2d^2\mathbb{E}||\boldsymbol{A}_0||_\infty}{\epsilon n^{1/2}} \to 0 \quad \text{as} \quad n \to \infty, \end{aligned} \tag{77}$$

where in the last inequality we apply Markov inequality and exploit the stationarity of $\{\boldsymbol{A}_n\}$. So we know $\hat{\boldsymbol{R}}(n) \Rightarrow \mathbf{0}$ and together with Step 2, we have proved the theorem. ■

Acknowledgment. Support was received from NSF grants CMMI 1634133.

References

1. Araujo, A., Giné, E.: The Central Limit Theorem for Real and Banach Valued Random Variables. John Wiley & Sons (1980)
2. Armony, M., Israelit, S., Mandelbaum, A., Marmor, Y., Tseytlin, Y., Yom-Tov, G.: On patient flow in hospitals: A data-based queueing-science perspective. Stochastic Systems **5**(1), 146–194 (2015)
3. Bertsimas, D., Mourtzinou, G.: Transient laws of non-stationary queueing systems and their applications. Queueing Systems **25**, 115–155 (1997)
4. Billingsley, P.: Probability and Measure, third edn. Wiley, New York (1995)
5. Billingsley, P.: Convergence of Probability Measures, second edn. Wiley, New York (1999)
6. Bradley, R.C.: On the central limit question under absolute regularity. The Annals of Probability pp. 1314–1325 (1985)
7. El-Taha, M., Stidham, S.: Sample-Path Analysis of Queueing Systems. Kluwer, Boston (1999)
8. Fralix, B.H., Riano, G.: A new look at transient versions of Little's law. Journal of Applied Probability **47**, 459–473 (2010)
9. Glynn, P.W., Whitt, W.: A central-limit-theorem version of $L = \lambda W$. Queueing Systems **2**, 191–215 (1986). (See Correction Note on $L = \lambda W$, Queueing Systems, 12 (4), 1992, 431-432. The results are correct; minor but important change needed in proofs.)
10. Glynn, P.W., Whitt, W.: Sufficient Conditions for Functional Limit Theorem Versions of $L = \lambda W$. Queueing Systems **1**, 279–287 (1987)
11. Glynn, P.W., Whitt, W.: Ordinary CLT and WLLN versions of $L = \lambda W$. Mathematics of Operations Research **13**, 674–692 (1988)
12. Glynn, P.W., Whitt, W.: Indirect estimation via $L = \lambda W$. Operations Research **37**(1), 82–103 (1989)
13. Gut, A.: Stopped Random Walks: Limit Theorems and Applications, 2 edn. Springer Series in Operations Research and Financial Engineering. Springer (2009). DOI 10.1007/978-0-387-87835-5
14. Harford, T.: Is this the most influential work in the history of capitalism? (2017). 50 Things that Made the Modern Economy, BBC World Service, 23 October 2017
15. Ibragimov, I.A., Linnik, Y.V.: Independent and stationary sequences of random variables. Groningen: Wolters-Noordhoff (1971)
16. Kim, S., Whitt, W.: Statistical analysis with Little's law. Operations Research **61**(4), 1030–1045 (2013)
17. Lauwers, L., Willekens, M.: Five hundred years of bookkeeping: a portrait of Luca Pacioli. Tijdschrift voor Economie en Management **39**(3), 289–304 (1994)
18. Ledoux, M., Talagrand, M.: Probability in Banach Spaces: Isoperimetry and Processes. Springer-Verlag (1991)
19. Little, J.D.C.: A proof of the queueing formula: $L = \lambda W$. Operations Research **9**, 383–387 (1961)
20. Little, J.D.C.: Little's law as viewed on its 50^{th} anniversary. Operations Research **59**, 536–539 (2011)
21. Little, J.D.C., Graves, S.C.: Little's law. In: D. Chhajed, T.J. Lowe (eds.) Building Intuition: Insights from Basic Opertations Management Models and Principles, chap. 5, pp. 81–100. Springer, New York (2008)
22. Megginson, R.E.: An Introduction to Banach Space Theory, vol. 183. Springer Science & Business Media (2012)

23. Munkres, J.R.: Topology, 2 edn. Prentice Hall (2000)

24. Pacioli, L.: Summa de Arithmetica, Geometria, Proportioni et Proportionalita. Venice (1494)

25. Stidham, S.: A last word on $L = \lambda W$. Operations Research **22**, 417–421 (1974)

26. Whitt, W.: A review of $L = \lambda W$. Queueing Systems **9**, 235–268 (1991)

27. Whitt, W.: Correction Note on $L = \lambda W$. Queueing Systems **12**, 431–432 (1992). (The results in the previous papers are correct, but minor important changes are needed in some proofs.)

28. Whitt, W.: Stochastic-Process Limits. Springer, New York (2002)

29. Whitt, W.: Extending the FCLT version of $L = \lambda W$. Operations Research Letters **40**, 230–234 (2012)

30. Whitt, W., Zhang, X.: A Data-Driven Model of an Emergency Department. Operations Research for Health Care **12**(1), 1–15 (2017)

31. Whitt, W., Zhang, X.: Periodic Little's Law (2017). Submitted to *Operations Research*, available at Columbia University, http://www.columbia.edu/∼ww2040/allpapers.html