Asymptotic Variability Analysis in Tandem Queues

December 2, 2024

Yongjiang Guo[1], Yunan Liu[2], Kaiming Yang[3]

Abstract

We study a multi-stage tandem queueing system and develop the law of the iterated logarithm (LIL) for performance measures including the queue length, workload, busy time, idle time and departure processes. These LIL results can help quantify the level of stochastic variability of these performance functions. Using a strong approximation method, which transforms the renewal-process-based performance functions to their continuous Brownian motion approximations, we establish all the LIL limits and express them as simple functions of model parameters (e.g, means and variances) of the interarrival and service times. Our LIL results reveal clear-cut insights on how the stochastic variability received from upstream stages can be propagated to the downsteam echelons in the tandem queue model; we show that stages that are underloaded, overloaded and critically loaded play distinct roles. An underloaded stage simply *transfers* all received upstream variability to the downstream stages; its own service-time variability makes no impact on any succeeding echelons. An overloaded stage overrides the variability received from upstream stages; it resets the propagation process by feeding its successive stages with its own service-time variability alone. A critically loaded *inherits* the variability received from upstream stages, which it *modifies* using its own service-time variability.

Keywords: tandem queue; the law of the iterated logarithm (LIL); the strong approximation method; Brownian motion

1 Introduction

In this paper, we study a K-stage tandem queue in which customers arrive according to a renewal process at station 1, and depart from the last station after completing their services at all K stations. Each stage is operated by a single server and the service times at the K stages are mutually independent and modeled by K sequences of *independent and identically distributed* (iid) random variables. We develop a *law of the iterated logarithm*

[1]School of Science, Beijing University of Posts and Telecommunications, Beijing, 100876, China; Key Laboratory of Mathematics and Information Networks (Beijing University of Posts and Telecommunications), Ministry of Education, China Email: yongerguo@bupt.edu.cn.

[2]Department of Industrial and Systems Engineering, North Carolina State University, Raleigh, NC 27695-7906, USA. Email: yliu48@ncsu.edu.

[3]School of Science, Beijing University of Posts and Telecommunications, Beijing 100876, P.R. China. Email: kaiming.yang0823@outlook.com.

(LIL) limit to characterize the asymptotic variability for this tandem queueing system; we focus on studying the common performance measures including the queue length, workload, busy time, idle time and departure processes. Understanding and quantifying the stochastic variability in a queueing system is important: First, it plays a critical role in the design of staffing functions subject to service-level constraints [17, 24, 26, 37]. For example, in a two-term square-root staffing formula [24, 26], the first (nominal) staffing term is determined by the means of the arrival and service times, while the second (safety) staffing term is set by the stochastic variability of the model. Second, the variability can be implemented in effective scheduling rules [25]. Last, the system's variability (e.g., variance of the waiting time) has also been used to measure fairness of the service process [5].

1.1 Background and Related Literature

Tandem queueing systems have been widely studied by researchers for many years due to their practical relevance to manufacture systems, wireless communication networks, and service systems [9, 20]. Results for tandem queues have been developed in several aspects: See [33] for transient behavior analysis, [2, 4] for joint queue length distribution, [3] for the sojourn time distribution, [12] for the fluid approximation, [35] for moments of performance measures, [16, 34] for diffusion approximation, [21, 36] for applications in inventory system, [1] for analysis on throughput maximization, [14] for asymptotic analysis for departure process, [38] for large deviations of queue length and departure processes, and [27, 28] for staffing levels to cope with time-varying arrivals in tandem queues.

LIL is a classical result in probability theory which characterizes the asymptotic variability of stochastic processs. The earliest LIL result was developed to treat the one-dimensional standard *Brownian motion* (BM) $W(t)$: *with probability one* (w.p.1),

$$\sup_{0\leq s\leq t} |W(s)| = O(\sqrt{t\log\log t}), \tag{1.1}$$

where the function $O(\cdot)$ means that $f(t) = O(g(t))$ if $\limsup_{t\to\infty} |f(t)/g(t)| \leq M$ for some $M > 0$. In the context of queueing systems, LILs in form of (1.1) can be used to characterize the asymptotic variability for important queueing performance measures around their means: For the queue length $Q(t)$ in a $GI/G/1$ queue, according to Theorem 3.4.2 in Chen and Mandelbaum [6] for *generalized Jackson networks* (GJN), it follows that

$$\sup_{0\leq s\leq t} |Q(s) - \bar{Q}(s)| = O(\sqrt{t\log\log t}) \tag{1.2}$$

w.p.1, where $\bar{Q}(s)$ is the fluid function acting as the asymptotic mean value for $Q(t)$. Readers can also refer to Theorem 6.11 in Chen and Yao [8] for a simple proof of (1.2).

The LIL in (1.1) gives a relatively rough estimate of the asymptotic variability because it only specifies the order in form of $O(\cdot)$. A more refined LIL result should be able to unravel the $O(\cdot)$ by establishing an explicit constant to better quantify the asymptotic variability. Indeed, according to Theorem 1.3.1* in [10], we have

$$\limsup_{t\to\infty} \frac{\sup_{0\leq s\leq t} |W(s))|}{\sqrt{t\log\log t}} = \sqrt{2}, \quad \text{w.p.1.}, \tag{1.3}$$

for the standard BM, where the limit "$\sqrt{2}$" on the right-hand side more explicitly quantifies the BM's stochastic variability (around its mean value 0) asymptotically as t grows large. Hereby, we refer to the LIL in form of (1.1) as the *implicit* LIL and that in form of (1.3) as the *explicit* LIL. See [15, 29, 30, 32] for explicit LIL results in multiclass priority queues and multi-server station GJNs under strict heavy traffic conditions. Because the analysis of explicit LILs is in general more involved, most LIL results in the extant literature are in the implicit form (see Chen and Yao [8] for a review). As noted by [8] (Remark 6.14, p.145; Remark 7.18, p.181), "It would be interesting to know what the limit

$$\limsup_{t\to\infty} \frac{\sup_{0\le s\le t} |Q(s) - \bar{Q}(s)|}{\sqrt{t \log\log t}} \tag{1.4}$$

is", and "the same questions may be asked for the workload process and the busy time process". In this paper, we try to partially answer this question by establishing the explicit LIL in a tandem queueing system.

1.2 A Strong Approximation Approach

A primary method to obtain the implicit LIL is via the *continuous mapping theorem* (CMT). For example, the implicit LIL for the $GI/G/1$ queue (Theorem 6.11 in [8]) is established by applying CMT to the implicit LILs of arrival and service processes. Because both processes are modeled by renewal processes, the implicit LIL in (1.2) can be directly concluded using the implicit LIL of a renewal process. See Panel (a) in Figure 1. Unfortunately, this approach does not help yield the explicit LIL.

In this paper, we establish the explicit LILs using the *strong approximation* (SA) methods, which are approximating functions constructed by fluid limits and BMs. For example, a renewal process $N(t)$ having interarrival times with finite rth moment, mean $1/\alpha > 0$ and standard deviation $\sigma > 0$ can be approximated by its SA $\widetilde{N}(t) \equiv \alpha t + \alpha^{3/2} W(t)$. Specifically, we have $\sup_{0\le s\le t} |\widetilde{N}(s) - N(s)| = o(t^{1/r})$, w.p.1 for some $r > 2$; See [10, 11]. Also see [7, 13, 18, 39, 40, 41] for SAs of queueing networks. Here, we say that $f(t) = o(g(t))$ if $\lim_{t\to\infty} |f(t)/g(t)| = 0$. We follow two steps. First, we show that the explicit LIL limit in

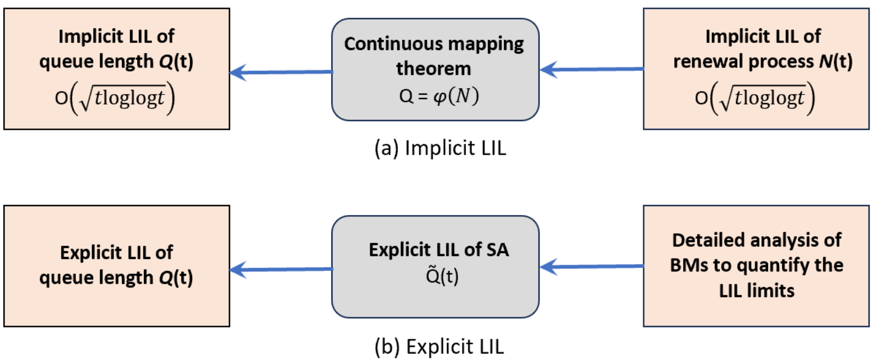

Figure 1: Implicit LIL vs. explicit LIL

form of (1.4) is equivalent, in the LIL sense, to that with $Q(t)$ replaced by its SA $\tilde{Q}(t)$, a function of the fluid limits and BMs, which is a continuous approximation of the discrete-valued $Q(t)$. Next, we derive the explicit LIL limits by carefully analyzing the SAs. A key step is to treat all the internal arrival processes because in our tandem queueing model, the departures of an upstream station serve as the arrivals of its immediate downstream station. In addition, our SA analysis depends on the traffic intensity ρ of the stations. For an underloaded station with $\rho < 1$ (overloaded with $\rho > 1$), the SA of its departure process is characterized solely by the arrival (service) times (see Theorem 3.1); while when the station is critically loaded with $\rho = 1$, we show that the SA draws from the variabilities in both the service and arrival times (Theorem 3.2).

1.3 Contributions and Organization

- We study a tandem queueing model and establish the LILs for performance metrics including the queue length, workload, busy time, idle time and departure processes. We give their explicit LIL limits as functions of the means and variances of the interarrival times and service times.

- We establish our LIL results using the fluid limits and SA approximations of the tandem queue. We first show that the LILs of the queueing metrics are equivalent to the those of the SAs. Next, we establish the LIL limits by properly treating the BM terms in the SA. Our analysis is highly dependent of the value of the traffic intensity ρ of the stations in consideration; we give fundamentally different treatments for the stages that are underloaded ($\rho < 1$), overloaded ($\rho > 1$) and critically loaded ($\rho = 1$).

- Our LIL results reveal clear-cut insights on how the variabilities received from upstream stages can be propagated to the downsteam echelons in the tandem queue model.

  i. An **underloaded** stage simply **transfers** the entirety of received upstream variability to the downstream stages; its own service-time variability makes no impact on any succeeding echelons.
  ii. An **overloaded** stage plays an overriding role by **blocking** the variability received from upstream stages; it resets the propagation process by feeding its successive stages with its own service-time variability alone.
  iii. A **critically loaded** stage is a middle ground between an underloaded and overloaded stage: On the one hand, it **inherits** the variability received from upstream stages; on the other hand, it **modifies** the variability with its own service-time variability and then passes it forward to the downstream stages.

Organization of the paper. In Section 2, we formally introduce the K-stage tandem queue and give the dynamic equations for the performance measures. In Section 3, we define the LILs for the performance measures and present the main results, namely, Theorem 3.1

and Theorem 3.2; we also discuss some implications of these results. In Section 4, we prove Theorems 3.1 and 3.2 using the strong approximation approach. We present several preliminary results as technical lemmas to reveal some structural results for the LIL limits. In Section 5, we give concluding remarks.

Summary of notations. We close this section by summarizing all notations used throughout the paper. All the random variables and stochastic processes are defined on the probability space $(\Omega, \mathcal{F}, \mathsf{P})$. E is the expectation operator. Let $\mathbb{R}(\mathbb{R}_+)$ be the space of (non-negative) real numbers. Let $\mathbb{D}^k(H)$ be the space of k-dimensional right-continuous function with left limit defined on $H \subset \mathbb{R}$, and $\mathbb{C}^k(H) \subset \mathbb{D}^k(H)$ is the space of continuous functions. Let $\mathbb{D}_0^k(H) \equiv \{x \in \mathbb{D}^k(H) : x(0) \geq 0\}$ and $\mathbb{C}_0^k(H) \equiv \{x \in \mathbb{C}^k(H) : x(0) \geq 0\}$ with $\equiv$ being a definition. For $k = 1$, let $\mathbb{D} \equiv \mathbb{D}^1$ and $\mathbb{C} \equiv \mathbb{C}^1$. For $f \in \mathbb{C}_0([0,\infty))$ and $f_n \in \mathbb{D}_0([0,\infty))$, $n = 1, 2, \ldots$, if $\sup_{0 \leq s \leq t} |f_n(s) - f(s)| \to 0$ for all $t \geq 0$ as $n \to \infty$, we say $f_n \to f$ uniformly on compact sets (u.o.c.). Suppose that $\mathbf{1}$ is an indicator function satisfying $\mathbf{1}_C(t) = 1$ if $t \in C$ and 0 otherwise. For real numbers x and y in $\mathbb{R}$, denote $x \vee y \equiv \max\{x, y\}$, $x \wedge y \equiv \min\{x, y\}$ and $(x)^+ \equiv \max\{x, 0\}$. For one-dimensional real function $f(t)$, define the supremum norm $||f||_t \equiv \sup_{0 \leq s \leq t} |f(s)|$ for all $t \geq 0$. Let $\varphi(t) \equiv \sqrt{2t \log \log t}$ for $t > \mathrm{e}$ (Euler constant). Let $\stackrel{\mathrm{d}}{=}$ denote equality in distribution. Let $\sum_{k=i}^{j} a_k \equiv 0$ if $i > j$ for all number sequence a_k.

2 The Tandem Queueing Model

We consider a queueing system comprised of K single-server stations, $K \geq 2$. All customers arrive at the first station for their first service via an exogenous arrival process; they proceed to the $(k+1)^{\text{st}}$ station after service completion at the k^{th} station, $k = 1, 2, \ldots, K-1$, and leave the system after the service completion at the K^{th} station. We say that a customer is in stage k if she is waiting in queue or in service at station k, where there is an infinite buffer for customers waiting for service. At every stage, the *first-come first-served* (FCFS) service discipline is enforced and the server is work-conserving, i.e., no server can be idle if there are customers waiting for service.

Primitive assumption. Let $v_0(n)$ be the interarrival time between the n^{th} and $(n+1)^{\text{st}}$ customers of the external arrival process, and $v_k(n)$ be the n^{th} service time at stage k, $k = 1, 2, \ldots, K$. Let $v_k = \{v_k(n), n = 1, 2, \ldots\}$ for $k = 0, 1, \ldots, K$. Suppose that v_k, $0 \leq k \leq K$, are $K+1$ mutually independent sequences of non-negative and iid random variables having means $\mathsf{E}[v_k(1)] \equiv 1/\mu_k$, variances $Var[v_k(1)]$ and *squared coefficients of variation* (SCVs) $c_k^2 \equiv Var[v_k(1)]/(\mathsf{E}[v_k(1)])^2$, respectively. Let

$$\mathcal{D} \equiv \left(\mu_k, c_k^2, 0 \leq k \leq K\right) \tag{2.1}$$

be the primitive data. For $k = 0, 1, \ldots, K$, define the partial sums,

$$V_k(n) \equiv \sum_{i=1}^{n} v_k(i), \quad n = 1, 2, \ldots, \tag{2.2}$$

and their corresponding renewal processes

$$S_k(t) \equiv \max\{n \geq 0 : V_k(n) \leq t\}, \quad t \geq 0, \tag{2.3}$$

where $S_0(t)$ counts the total number of external arrivals in $[0,t]$ and $S_k(t)$ counts the total number of potential service completions at stage k in $[0,t]$, $k = 1,2,\ldots,K$, provided that the server there is always busy. Define the traffic intensity at stage k:

$$\rho_k \equiv \frac{\min\{\mu_0, \mu_1, \ldots, \mu_{k-1}\}}{\mu_k}, \quad k = 1,2,\ldots,K. \tag{2.4}$$

We say that stage k is *underloaded* if $\rho_k < 1$, *critically loaded* if $\rho_k = 1$, and *overloaded* if $\rho_k > 1$.

Performance measures. For $k = 1,2,\ldots,K$, let $Q_k(t)$ be the stage-k queue length which is the total number of customers at stage k at time t. Suppose that the system is initially empty with $Q_k(0) = 0$ for all k. Let $Z_k(t)$ be the stage-k workload at time t, which is the total amount of work (measured in time) required to be processed until that station becomes empty, assuming there is no future arrivals after time t. Let $T_k(t)$ be the busy time of the stage-k server which is the cumulative time the server is busy by time t, namely,

$$T_k(t) = \int_0^t \mathbf{1}_{\{Q_k(s)>0\}} \mathsf{d}s. \tag{2.5}$$

Let $T_0(t) \equiv t$ and $I_k(t)$ be the stage-k idle time by time t, namely,

$$I_k(t) = t - T_k(t). \tag{2.6}$$

Denote $D_k(t) \equiv S_k(T_k(t))$ as the actual stage-k departure or service-completion process that counts the total number of service completions of in $[0,t]$. We have the following dynamic equations for these performance functions: for $k = 1,2,\ldots,K$, we have

$$Q_k(t) = S_{k-1}(T_{k-1}(t)) - S_k(T_k(t)) \geq 0, \tag{2.7}$$

$$Z_k(t) = V_k(S_{k-1}(T_{k-1}(t))) - T_k(t), \tag{2.8}$$

$$0 = \int_0^t Q_k(s) \mathsf{d} I_k(s), \tag{2.9}$$

where (2.7) holds by flow conservation, (2.8) holds because $V_k(S_{k-1}(T_{k-1}(t)))$ represents the total amount of work of all arrivals at the stage-k queue in $[0,t]$, and (2.9) stipulates that the idle time $I_k(t)$ increases only when $Q_k(t) = 0$ according to the non-idling rule.

Fluid approximations. For the performance measures above, the fluid approximations can be used to characterize their mean values and serve as the centering terms when defining the LILs. To properly introduce the fluid approximations, we define the following fluid-scaled processes:

$$\bar{Q}_k^{(n)}(t) \equiv \frac{1}{n} Q_k(nt), \quad \bar{Z}_k^{(n)}(t) \equiv \frac{1}{n} Z_k(nt), \quad \bar{T}_k^{(n)}(t) \equiv \frac{1}{n} T_k(nt),$$
$$\bar{I}_k^{(n)}(t) \equiv \frac{1}{n} I_k(nt), \quad \bar{D}_k^{(n)}(t) \equiv \frac{1}{n} D_k(nt), \quad k = 1,2,\ldots,K,$$

where n is the scaler. Next, we introduce fluid approximations as the limits of the *functional strong law of large numbers*. Readers can refer to Theorem 4.1 in [12] for details.

Lemma 2.1 (Fluid approximation)**.** *Suppose that the system is initially empty, we have, for $k = 1, 2, \ldots, K$,*

$$\left(\bar{Q}_k^{(n)}, \bar{Z}_k^{(n)}, \bar{T}_k^{(n)}, \bar{I}_k^{(n)}, \bar{D}_k^{(n)}\right) \to \left(\bar{Q}_k, \bar{Z}_k, \bar{T}_k, \bar{I}_k, \bar{D}_k\right) \equiv \bar{\mathbb{X}}_k, \quad \text{u.o.c., w.p.1,} \quad \text{as } n \to \infty,$$

where, for all $t \geq 0$, the fluid functions specified below:

$$\begin{aligned} &\bar{Q}_k(t) \equiv \Phi(\bar{X}_k)(t) \geq 0, \quad \bar{Z}_k(t) \equiv \frac{\bar{Q}_k(t)}{\mu_k}, \quad \bar{I}_k(t) \equiv \frac{1}{\mu_k}\Psi(\bar{X}_k)(t), \quad \bar{T}_k(t) \equiv t - \bar{I}_k(t), \\ &\bar{X}_k(t) \equiv \mu_{k-1}\bar{T}_{k-1}(t) - \mu_k t, \quad \bar{D}_k(t) \equiv \mu_k \bar{T}_k(t), \quad \bar{T}_0(t) \equiv t, \end{aligned} \tag{2.10}$$

and functions Φ and Ψ are defined on the space $\mathbb{D}_0([0,\infty))$ as below:

$$\Psi(x)(t) \equiv \sup_{0 \leq s \leq t} \{-x(s)\}^+ \quad \text{and} \quad \Phi(x)(t) \equiv x(t) + \sup_{0 \leq s \leq t} \{-x(s)\}^+. \tag{2.11}$$

Remark 2.1 (Oblique reflection mapping)**.** *The mapping (Ψ, Φ) is known as the one dimensional Oblique reflection mapping (ORM) and is Lipschitz continuous under the uniform topology, see Theorem 6.1 in [8]. In Lemma 2.1, the assumption that the system is initially empty implies that $\bar{X}_k(0) = 0$, then, for $\bar{X}_k$,*

$$\Psi(\bar{X}_k)(t) \equiv \sup_{0 \leq s \leq t} \{-\bar{X}_k(s)\} \quad \text{and} \quad \Phi(\bar{X}_k)(t) \equiv \bar{X}_k(t) + \sup_{0 \leq s \leq t} \{-\bar{X}_k(s)\}. \tag{2.12}$$

3 Main Results

In this section, we will establish the explicit LIL results for all performance measures $(Q_k, Z_k, T_k, I_k, D_k)$ and identify their LIL limits as simple analytic functions in terms of the primitive data $\mathcal{D}$ in (2.1). For simplicity, we omit the "explicit" and simply call them LILs for the rest of the paper.

We develop the following LILs for the tandem queueing system:

$$\mathcal{X}_k^* \equiv \left(Q_k^*, Z_k^*, T_k^*, I_k^*, D_k^*\right), \quad k = 1, 2, \ldots, K, \tag{3.1}$$

where

$$\begin{aligned} &Q_k^* \equiv \limsup_{L\to\infty} \frac{1}{\varphi(L)} \left|\left|Q_k - \bar{Q}_k\right|\right|_L, \quad Z_k^* \equiv \limsup_{L\to\infty} \frac{1}{\varphi(L)} \left|\left|Z_k - \bar{Z}_k\right|\right|_L, \\ &T_k^* \equiv \limsup_{L\to\infty} \frac{1}{\varphi(L)} \left|\left|T_k - \bar{T}_k\right|\right|_L, \quad I_k^* \equiv \limsup_{L\to\infty} \frac{1}{\varphi(L)} \left|\left|I_k - \bar{I}_k\right|\right|_L, \\ &D_k^* \equiv \limsup_{L\to\infty} \frac{1}{\varphi(L)} \left|\left|D_k - \bar{D}_k\right|\right|_L. \end{aligned} \tag{3.2}$$

We remark that the convergence in (3.1) is termwise convergence instead of joint convergence.

We make the following assumptions on the moments of $v_k(1)$ and enforce this for the rest of the paper. For some $r > 2$, we have

$$\mathsf{E}\left[v_k(1)^r\right] < \infty \quad \text{for all} \quad k = 0, 1, \ldots, K. \tag{3.3}$$

3.1 No Critical Loading

We first treat the relatively easy case without any critically loaded stage, i.e., all stages are either underloaded or overloaded. We report the LIL limits for all performance functions in Theorem 3.1 and provide in-depth discussions to gain insights into these results. We relegate the proof to Section 4. For the ease of notation, define the following constants:

$$C_{i,j} \equiv \sqrt{c_i^2 + c_j^2} \qquad \text{and} \qquad \widehat{C}_{i,j} \equiv c_i \vee c_j, \qquad 0 \le i, j \le K. \tag{3.4}$$

Theorem 3.1 (No critical loading)**.** *Given data $\mathcal{D}$ in* (2.1)*, we have the LILs in three cases:*

1. (***The full underloading case***) *If all stages are underloaded, that is, $\rho_k < 1$ for all $k = 1, 2, \ldots, K$, we have*

$$\mathcal{X}_k^* = \left(0, 0, \frac{\sqrt{\mu_0} C_{0,k}}{\mu_k}, \frac{\sqrt{\mu_0} C_{0,k}}{\mu_k}, \mu_0^{1/2} c_0\right). \tag{3.5}$$

2. (***The full overloading case***) *If all stages are overloaded, that is, $\rho_k > 1$ for all $k = 1, 2, \ldots, K$, we have*

$$\mathcal{X}_k^* = \left(\sqrt{\mu_{k-1} c_{k-1}^2 + \mu_k c_k^2}, \frac{\sqrt{\mu_{k-1}} C_{k-1,k}}{\mu_k}, 0, 0, \mu_k^{1/2} c_k\right). \tag{3.6}$$

3. (***The hybrid case***) *If some stages are underloaded and the others are overloaded, i.e., $\rho_k > 1$ for $k = k'_1, k'_2, \ldots, k'_{l_0}$, with $0 < k'_1 < k'_2 < \cdots < k'_{l_0} \le K$ and $l_0 < K$, and $\rho_k < 1$ for $k \ne k'_1, k'_2, \ldots, k'_{l_0}$, then, for $l = 1, 2, \ldots, l_0$:*

$$\mathcal{X}_{k'_l}^* = \left(\sqrt{\mu_{k'_{l-1}} c_{k'_{l-1}}^2 + \mu_{k'_l} c_{k'_l}^2}, \frac{\sqrt{\mu_{k'_{l-1}}} C_{k'_{l-1}, k'_l}}{\mu_{k'_l}}, 0, 0, \mu_{k'_l}^{1/2} c_{k'_l}\right). \tag{3.7}$$

The LILs for $k = 1, 2, \ldots, k'_1 - 1$ satisfy (3.5)*. For $k = k'_l + 1, k'_l + 2, \ldots, k'_{l+1} - 1$ and $l = 1, 2, \ldots, l_0$,*

$$\mathcal{X}_k^* = \left(0, 0, \frac{\sqrt{\mu_{k'_l}(c_{k'_l}^2 + c_k^2)}}{\mu_k}, \frac{\sqrt{\mu_{k'_l}} C_{k'_l, k}}{\mu_k}, \mu_{k'_l}^{1/2} c_{k'_l}\right). \tag{3.8}$$

To give insights into the mathematical results of Theorem 3.1, we next provide detailed discussions of the three cases in Theorem 3.1 via three remarks.

Remark 3.1 (The full underloading case: K underloaded $G/G/1$ queues)**.** *When all stages are underloaded with $\rho_k < 1$, $k = 1, 2, \ldots, K$, the overall system is in light traffic. The Markov process describing the system dynamics has a stability distribution [12]. Hence, the LIL limit for the queue length is zero, so is that of the workload due to Little's law. We next discuss the implications of the asymptotic variability of the busy time, idle time and departure process. In light traffic, the server capacity is not fully utilized, so the variability* *in the departure process at any stage is independent of the service-time variability and is determined only by the variability in its arrival process, described by the variance parameter* $\mu_0^{1/2} c_0$ *in* (3.5). *Hence, in the LIL sense, every stage behaves like a* $G/G/1$ *queue fed by the external arrival process* $S_0(t)$. *A similar implication can be drawn from the LIL limits of busy and idle times, which are comprised of* c_0^2 *and* c_k^2 *and have nothing to do with the variability parameters in any upstream queues* c_{k-1}^2 *for* $k = 2, 3, \ldots, K$. *Indeed, the LIL result at any stage is identical to that in an underloaded* $G/G/1$ *queue with an arrival process distributed as* S_0. *In summary, the service variability of an underloaded stage makes no impact on the remaining echelon; instead, it* ***passes forward*** *the entirety of the variability received from upstream stages to its succeeding downstream stages.*

Remark 3.2 (The full overloading case: K overloaded $G/G/1$ queues)**.** *When every stage is overloaded with* $\rho_k > 1$, $k = 1, 2, \ldots, K$, *the entire system is in heavy traffic. Because all servers are almost surely always busy asymptotically, the busy time* $T_k(t) \approx t$ *and the idle time* $I_k(t) \approx 0$, *so they hardly have any variability, which explains the zero LIL limits of* T_k^* *and* I_k^*. *Because there are almost infinitely many customers waiting at the queue, the variability of the departure depends only on its own service variability, described by the term* $\mu_k^{1/2} c_k$. *Similar to the case of* $G/G/1$, *the variabilities of the stage-*k *workload and queue length are determined by the stochastic fluctuations in its service times and (internal) arrival process, which correspond to the departure from its immediate upstream stage (i.e., stage* $k-1$*). Hence, in the LIL sense, it is equivalent to an overloaded* $G/G/1$ *queue having a renewal arrival process with rate* μ_{k-1} *and variability parameter* c_{k-1} *(because the stage-*k *server is asymptotically always busy and continues to produce service completions in a renewal fashion).*

Remark 3.3 (The hybrid case: Overloaded (underloaded) stages override (pass forward) previous memories)**.** *According to Remarks 3.1 and 3.2, we can see that, in the LIL sense, a stage-*k *queue often behaves like a* $G/G/1$ *queue of which the variability is impacted by its own service variability and the variability inherited from some upstream stages. For the hybrid case, the impact on the arrival side traces back to the latest upstream stage that is overloaded instead of its immediate upstream stage. Unlike an underloaded stage that simply transfers its upstream variabilities to its downsteam stages, this overloaded stage plays an* ***overriding*** *role in the network echelon; it eliminates the effects from all preceding stages and resets the forward propagation process by feeding the downstream stages with its own service variability alone. First, the LIL limit of an overloaded stage* k_l' *is similar to the full overloading case* (3.6) *except that, it is not its immediate upstream stage* $k_l' - 1$, *but its previous overloaded upstream stage* k_{l-1}' *that accounts for the variability on the arrival side. The underloaded stages between* k_{l-1}' *and* k_l' *(if any) have no impact on stage* k_l'. *Next, an underloaded queue* k *between two overloaded stages* k_l' *and* k_{l+1}' *(i.e.,* $k_l' < k < k_{l+1}'$*) behaves like an underloaded* $G/G/1$ *queue having renewal arrivals with rate* $\mu_{k_l'}$ *and variability parameter* $c_{k_l'}$. *Its LIL limit is similar to* (3.5) *with* μ_0 *and* c_0 *replaced by* $\mu_{k_l'}$ *and* $c_{k_l'}$.

3.2 One Critically Loaded Stage

The LIL analysis becomes much more complex when some stages are critically loaded with $\rho_k = 1$. As an initial attempt, we hereby focus on a tandem queue model with exactly one critically loaded stage. See Section 4.4 for detailed explanations on the technical challenges to treating more than one critically loaded stages. In what follows, we first present the LIL limits in Theorem 3.2 and then draw useful implications from these results.

Suppose that stage k_0 is critically loaded with $\rho_{k_0} = 1$, $1 \leq k_0 \leq K$. As we will show next, two overloaded stages (if it exists) play particularly critical roles: (i) *the last overloaded stage preceding stage* k_0 (if $k_0 > 1$), and (ii) *the first overloaded stage succeeding stage* k_0 (if $k_0 < K$). When $k_0 > 1$, let $\underline{k}_0$ be the index of the last overloaded stage (if any) preceding stage k_0, i.e.,

$$\underline{k}_0 \equiv \max\{k < k_0 : \rho_k > 1\}, \tag{3.9}$$

and set $\underline{k}_0 \equiv 0$ if all preceding stages are underloaded, i.e., $\rho_k < 1$ for all $k < k_0$. When $k_0 < K$, let $\bar{k}_0$ be the index of the first overloaded stage (if any) succeeding stage k_0, i.e.,

$$\bar{k}_0 \equiv \min\{k > k_0 : \rho_k > 1\}, \tag{3.10}$$

and set $\bar{k}_0 \equiv \infty$ if all succeeding stages are underloaded, i.e., $\rho_k < 1$ for all $k > k_0$.

Theorem 3.2 (One critically loaded stage)**.** *Given the model data* (2.1)*, we have the following explicit LILs.*

1. (***Stages preceding stage*** k_0) *The LIL limits for all stages preceding stage k_0 are specified by results in Theorem 3.1.*

2. (***Stage*** k_0) *The LIL limits for k_0 satisfies*

$$\mathcal{X}^*_{k_0} = \left(\sqrt{\mu_{\underline{k}_0}}C_{\underline{k}_0,k_0}, \frac{C_{\underline{k}_0,k_0}}{\sqrt{\mu_{\underline{k}_0}}}, \frac{C_{\underline{k}_0,k_0}}{\sqrt{\mu_{\underline{k}_0}}}, \frac{C_{\underline{k}_0,k_0}}{\sqrt{\mu_{\underline{k}_0}}}, \mu_{\underline{k}_0}^{1/2}\widehat{C}_{\underline{k}_0,k_0}\right). \tag{3.11}$$

3. (***Stages succeeding stage*** k_0) *For $k > k_0$, we specifies $\mathcal{X}^*_k$ in cases below:*

   a. *For stage $k = k_0 + 1, \ldots, (\bar{k}_0 \wedge K) - 1$, we have*

$$\mathcal{X}^*_k = \left(0, 0, \frac{\sqrt{\mu_{\underline{k}_0}(\widehat{C}^2_{\underline{k}_0,k_0} + c^2_k)}}{\mu_k}, \frac{\sqrt{\mu_{\underline{k}_0}(\widehat{C}^2_{\underline{k}_0,k_0} + c^2_k)}}{\mu_k}, \mu_{\underline{k}_0}^{1/2}\widehat{C}_{\underline{k}_0,k_0}\right). \tag{3.12}$$

   b. *For stage $\bar{k}_0$ (if $\bar{k}_0 \leq K$), we have*

$$\mathcal{X}^*_{\bar{k}_0} = \left(\sqrt{\mu_{k_0}\widehat{C}^2_{\underline{k}_0,k_0} + \mu_{\bar{k}_0}c^2_{\bar{k}_0}}, \frac{\sqrt{\mu_{k_0}(\widehat{C}^2_{\underline{k}_0,k_0} + c^2_{\bar{k}_0})}}{\mu_{\bar{k}_0}}, 0, 0, \mu_{\bar{k}_0}^{1/2}c_{\bar{k}_0}\right). \tag{3.13}$$

c. *For stages* $k = \bar{k}_0 + 1, \ldots, K$ *(if* $\bar{k}_0 < K$*), their LIL limits are equivalent to a no-critical-loading tandem queue (Theorem 3.1) having* $K - \bar{k}_0$ *stages and an external renewal arrival process with parameters* $\mu_{\bar{k}_0}$ *and* $c_{\bar{k}_0}$.

Remark 3.4 (The critically loaded stage)**.** *First, stage* k_0 *has no impact on the its preceding stages* $1, \ldots, k_0 - 1$*, of which the LIL limits are already characterized by Theorem 3.1. We next discuss the LIL for* k_0*. Consistent with results in Theorem 3.1, we know that stage* k_0*'s variability on the arrival side should inherit from the last overloaded queue preceding stage* k_0*, namely, stage* $\underline{k}_0$ *(if any). On the other hand, the structure of the LIL limits for stage* k_0 *is significantly different from those of underloaded and overloaded stages. First, the critical loading condition enables all performance functions (i.e., queue length, workload, and busy times) to preserve certain level of stochastic variabilities, none of the LIL limits in* (3.11) *is 0 (which is in stark contrast to the zero workload LIL in the underloaded case and the zero busy-time LIL in the overloaded case). Next, the LIL limit of the departure process* $\widehat{C}_{\underline{k}_0, k_0}$ *is the mixture of the variability parameter inherited from the upstream stages* $c_{\underline{k}_0}$ *and its own service-time variability parameter* c_{k_o} *(which is contrast to an underloaded stage which transfers the received upstream variability only and to an overloaded stage which blocks all upstream variability).*

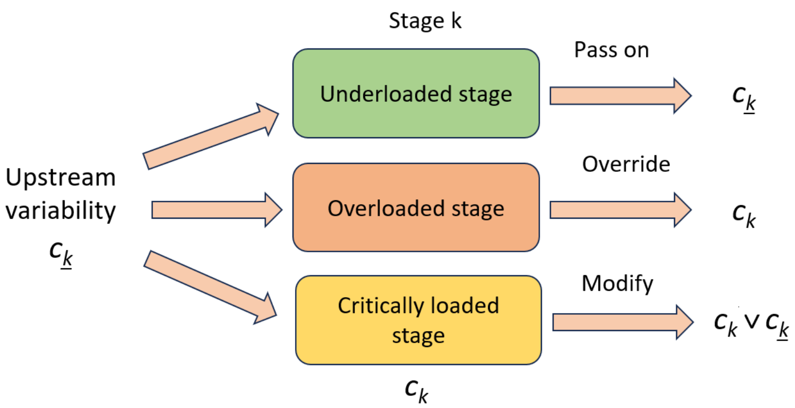

Figure 2: Different roles of underloaded, overloaded and critically loaded stages in the tandem queueing system. Here $\underline{k}$ is the last overloaded stage preceding stage k.

Remark 3.5 (Stages succeeding the critically loaded stage)**.** *As mentioned in Remark 3.3, an overloaded stage plays an overriding role in the tandem structure: it eliminates the "memory" of variabilities from previous stages and resets the remaining downstream echelons by fueling all subsequent stages with its own service variability. Because stage* $\bar{k}_0$ *is the first overloaded stage succeeding* k_0*, it does not allow the impact from previous stages to pass onto any later stages. Since server* $\bar{k}_0$ *is almost surely always busy, its departure process corresponds to a renewal process with parameters* $\mu_{\bar{k}_0}$ *and* $c_{\bar{k}_0}$*, which essentially turn the remaining stages into a* $(K - \bar{k}_0)$*-stage tandem queue model having an external renewal arrival with parameters* $\mu_{\bar{k}_0}$ *and* $c_{\bar{k}_0}$ *(this explains Case 3.c in Theorem 3.2). To understand Cases 3.a and 3.b, note that the variability of the departure process from the critically* *loaded stage (i.e., stage k_0) is $\widehat{C}_{\underline{k}_0,k_0}$ (as discussed in Remark 3.4) and each underloaded queue simply transfers any upstream variabilities received to next stages, all underloaded stages between k_0 and $\bar{k}_0$ are, in the LIL sense, equivalent to an underloaded $G/G/1$ queue having renewal arrivals with parameters $\mu_{\underline{k}_0}$ and $\widehat{C}_{\underline{k}_0,k_0}$, which explains* (3.12)*; similarly, the overloaded stage $\bar{k}_0$ is, in the LIL sense, equivalent to an overloaded $G/G/1$ queue having renewal arrivals with parameters $\mu_{\underline{k}_0}$ and $\widehat{C}_{\underline{k}_0,k_0}$, which explains* (3.13)*. See Figure 2 for a summary of the distinct roles of underloaded, overloaded and critically loaded stages.*

4 Proofs of Main Results

We prove Theorem 3.1 and Theorem 3.2 in this section. In Subsection 4.1, we give some technical lemmas which are building blocks of our proofs; we relegate their proofs to the Appendix. We prove Theorems 3.1 and 3.2 in Subsections 4.2 and 4.3, respectively.

4.1 Preliminary Results

Lemma 4.1 (Strong approximation). *If* (3.3) *holds, then, for some $r > 2$ and for all $k = 1, 2, \ldots, K$,*

$$\begin{aligned}&\left|\left|Q_k - \widetilde{Q}_k\right|\right|_L = o(L^{1/r}), \quad \left|\left|Z_k - \widetilde{Z}_k\right|\right|_L = o(L^{1/r}), \quad \left|\left|T_k - \widetilde{T}_k\right|\right|_L = o(L^{1/r}),\\ &\left|\left|I_k - \widetilde{I}_k\right|\right|_L = o(L^{1/r}), \quad \left|\left|D_k - \widetilde{D}_k\right|\right|_L = o(L^{1/r}), \quad \textit{w.p.1}, \end{aligned} \tag{4.1}$$

where $\widetilde{I}_0(t) \equiv 0$,

$$\begin{aligned}&(\widetilde{Q}_k, \widetilde{Y}_k) \equiv (\Phi, \Psi)(\widetilde{X}_k), \quad \widetilde{I}_k(t) \equiv \frac{1}{\mu}\widetilde{Y}_k(t), \quad \widetilde{T}_k(t) \equiv t - \widetilde{I}_k(t),\\ &\widetilde{X}_k(t) \equiv (\mu_{k-1} - \mu_k)t - \mu_{k-1}\widetilde{I}_{k-1}(t) + \mu_{k-1}^{1/2}c_{k-1}W_{k-1}(\bar{T}_{k-1}(t)) - \mu_k^{1/2}c_kW_k(\bar{T}_k(t)),\\ &\widetilde{Z}_k(t) \equiv \frac{1}{\mu_k}\left[\widetilde{Q}_k(t) + \mu_k^{1/2}c_k(W_k(\bar{T}_k(t)) - W_k(\rho_k t))\right],\\ &\widetilde{D}_k(t) \equiv \mu_k\widetilde{T}_k(t) + \mu_k^{1/2}c_kW_k(\bar{T}_k(t)), \end{aligned} \tag{4.2}$$

W_0 and W_k are mutually independent standard BMs associated with the exogenous arrival process and service process at stage-k, respectively, and Φ and Ψ are defined in (2.11)*.*

Lemma 4.1 is the tandem queue special case of Theorem 7.19 in [8], where we specify the routing matrix of the queueing network as

$$P = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ 0 & 0 & 0 & \cdots & 0 \end{pmatrix}.$$

See [8] for the proof of the SA results with a more general routing matrix P.

According to Lemma 4.1, the performance functions (which satisfy the dynamic equations (2.5)–(2.9)) can now be approximated by functions of BMs. In addition, (4.1) indicates that the errors of these approximations are $o(L^{1/r})$. To establish LILs of Q, Z, T, I and D, we will first develop the LILs of their SA counterparts specified in Lemma 4.1. Define the LILs of the SAs as:

$$\widetilde{\mathcal{X}}_k^* \equiv \left(\widetilde{Q}_k^*, \widetilde{Z}_k^*, \widetilde{T}_k^*, \widetilde{I}_k^*, \widetilde{D}_k^*\right), \quad k = 1, 2, \ldots, K, \tag{4.3}$$

where

$$\begin{aligned} \widetilde{Q}_k^* &\equiv \limsup_{L\to\infty} \frac{1}{\varphi(L)} \left|\left|\widetilde{Q}_k - \bar{Q}_k\right|\right|_L, \quad \widetilde{Z}_k^* \equiv \limsup_{L\to\infty} \frac{1}{\varphi(L)} \left|\left|\widetilde{Z}_k - \bar{Z}_k\right|\right|_L, \\ \widetilde{T}_k^* &\equiv \limsup_{L\to\infty} \frac{1}{\varphi(L)} \left|\left|\widetilde{T}_k - \bar{T}_k\right|\right|_L, \quad \widetilde{I}_k^* \equiv \limsup_{L\to\infty} \frac{1}{\varphi(L)} \left|\left|\widetilde{I}_k - \bar{I}_k\right|\right|_L, \\ \widetilde{D}_k^* &\equiv \limsup_{L\to\infty} \frac{1}{\varphi(L)} \left|\left|\widetilde{D}_k - \bar{D}_k\right|\right|_L. \end{aligned}$$

Our proofs build on the following lemmas: Lemma 4.2 guarantees that the LILs of the original performance functions coincide with the LILs of their SAs. Hence, it remains to estabilish the LILs $\widetilde{\mathcal{X}}_k^*$ in (4.3). Lemmas 4.3, 4.4 and 4.5 are useful building blocks to compute $\widetilde{\mathcal{X}}_k^*$. They also reveal some structural properties of the LIL limits for the three cases $\rho_k < 1$, $\rho_k > 1$ and $\rho_k = 1$, respectively. The proofs of all lemmas are relegated to the Appendix.

Lemma 4.2. *Suppose that* (3.3) *holds, we have* $\mathcal{X}_k^* = \widetilde{\mathcal{X}}_k^*$ *w.p.1 for all* $k = 1, 2, \ldots, K$.

Lemma 4.3. *If* $\rho_k < 1$ *for some stage-*k*, then,* $\widetilde{Q}_k^* = \widetilde{Z}_k^* = 0$ *w.p.1,* $k = 1, 2, \ldots, K$.

Lemma 4.4. *If* $\rho_k > 1$ *for some stage-*k *in the case the tandem system has no more than one critically loaded stage, then* $\widetilde{I}_k^* = \widetilde{T}_k^* = 0$ *w.p.1,* $k = 1, 2, \ldots, K$.

Lemma 4.5. *If* $\rho_k = 1$ *for some stage-*k *in the case the tandem system has no more than one critically loaded stage, then* $\widetilde{Q}_k^* = \mu_k \widetilde{Z}_k^*$ *and* $\widetilde{Z}_k^* = \widetilde{T}_k^* = \widetilde{I}_k^*$ *w.p.1,* $k = 1, 2, \ldots, K$.

See Figure 3 for a schematic illustration of the flows of the proofs.

4.2 Proof of Theorem 3.1

4.2.1 The full underloading case

Since $\rho_k < 1$ for all $k = 1, 2, \ldots, K$, we have $\mu_0 < \min\{\mu_1, \mu_2, \cdots, \mu_K\}$ and $\rho_k = \mu_0/\mu_k$, the corresponding fluid solution:

$$\bar{\mathbb{X}}_k(t) \quad = \quad (0, 0, \rho_k t, (1-\rho_k)t, \mu_0 t), \quad k = 1, 2, \ldots, K. \tag{4.4}$$

By Lemma 4.3, we note that the LIL $\widetilde{Q}_k^* = \widetilde{Z}_k^* = 0$ w.p.1 for all $k = 1, 2, \ldots, K$.

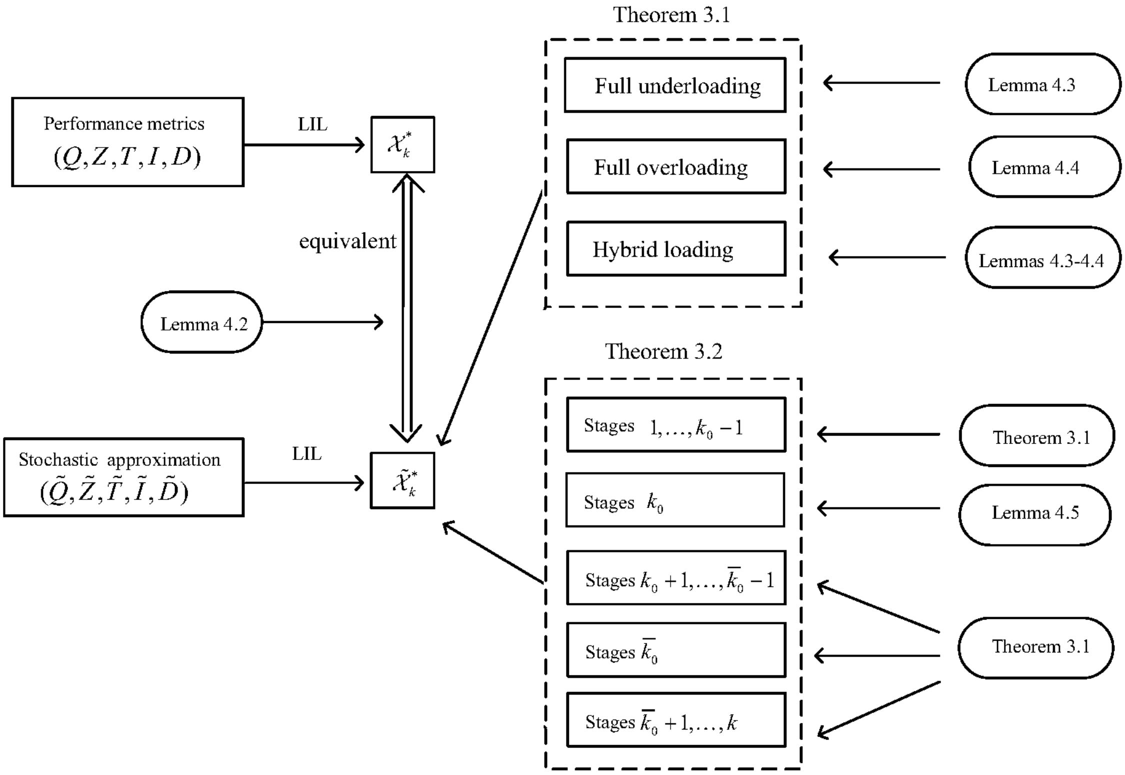

Figure 3: A schematic illustration of the proofs.

To treat the LIL of the idle time process, we have, by (4.2) and (4.4),

$$\begin{aligned}\widetilde{I}_k(t) - \bar{I}_k(t) &= \frac{1}{\mu_k}\left[\widetilde{Y}_k(t) - \bar{Y}_k(t)\right] \\ &= \frac{1}{\mu_k}\widetilde{Q}_k(t) - \frac{1}{\mu_k}\left[\widetilde{X}_k(t) - \bar{X}_k(t)\right],\end{aligned}$$

and because $\bar{X}_k(t) = (\mu_0 - \mu_k)t$ by $\rho_k < 1$ for all $k = 1, 2, \ldots, K$,

$$\begin{aligned}&\widetilde{X}_k(t) - \bar{X}_k(t) \\ =\;& \mu_{k-1}\widetilde{T}_{k-1}(t) - \mu_0 t + \mu_{k-1}^{1/2}c_{k-1}W_{k-1}(\bar{T}_{k-1}(t)) - \mu_k^{1/2}c_kW_k(\bar{T}_k(t)) \\ =\;& -\widetilde{Y}_{k-1}(t) - \bar{X}_{k-1}(t) + \mu_{k-1}^{1/2}c_{k-1}W_{k-1}(\bar{T}_{k-1}(t)) - \mu_k^{1/2}c_kW_k(\bar{T}_k(t)) \\ =\;& \left[\widetilde{X}_{k-1}(t) - \bar{X}_{k-1}(t)\right] - \widetilde{Q}_{k-1}(t) + \mu_{k-1}^{1/2}c_{k-1}W_{k-1}(\bar{T}_{k-1}(t)) - \mu_k^{1/2}c_kW_k(\bar{T}_k(t)) \\ =\;& \left[\widetilde{X}_1(t) - \bar{X}_1(t)\right] - \sum_{l=1}^{k-1}\widetilde{Q}_l(t) + \mu_1^{1/2}c_1W_1(\bar{T}_1(t)) - \mu_k^{1/2}c_kW_k(\bar{T}_k(t)) \\ =\;& -\sum_{l=1}^{k-1}\widetilde{Q}_l(t) + \mu_0^{1/2}c_0W_0(t) - \mu_k^{1/2}c_kW_k(\bar{T}_k(t)). \qquad (4.5)\end{aligned}$$

Hence, we have

$$\begin{aligned}\widetilde{I}_k(t)-\bar{I}_k(t) &= \frac{1}{\mu_k}\left[\widetilde{Y}_k(t)-\bar{Y}_k(t)\right]\\ &= \frac{1}{\mu_k}\sum_{l=1}^{k}\widetilde{Q}_l(t)-\frac{1}{\mu_k}\left[\mu_0^{1/2}c_0W_0(t)-\mu_k^{1/2}c_kW_k(\bar{T}_k(t))\right].\end{aligned}$$

Notice that $\widetilde{Q}_k^*=0$ w.p.1, and $\bar{T}_k(t)=(\mu_0/\mu_k)t$ for all $t\geq 0$, so $\mu_0^{1/2}c_0W_0(t)-\mu_k^{1/2}c_kW_k(\bar{T}_k(t))$ is a driftless BM with variance parameter $\mu_0(c_0^2+c_k^2)$. As a result, the LIL $\widetilde{I}_k^*=\sqrt{\mu_0}C_{0,k}/\mu_k$.

For the LIL of the busy time process, by (2.10) and (4.2), for all $k=1,2,\ldots,K$ and $t\geq 0$, we have

$$\widetilde{T}_k(t)-\bar{T}_k(t)=\bar{I}_k(t)-\widetilde{I}_k(t), \tag{4.6}$$

which always holds regardless of the value of the traffic intensity ρ_k. As a result, by (4.3), it follows that the LIL $\widetilde{T}_k^*=\sqrt{\mu_0}C_{0,k}/\mu_k$.

To develop the LIL of the departure process, we have, by (2.10), (4.2) and (4.5),

$$\begin{aligned}\widetilde{D}_k(t)-\bar{D}_k(t) &= \mu_k\left[\widetilde{T}_k(t)-\bar{T}_k(t)\right]+\mu_k^{1/2}c_kW_k(\bar{T}_k(t))\\ &= -\mu_k\left[\widetilde{I}_k(t)-\bar{I}_k(t)\right]+\mu_k^{1/2}c_kW_k(\bar{T}_k(t))\\ &= -\widetilde{Q}_k(t)+\left[\widetilde{X}_k(t)-\bar{X}_k(t)\right]+\mu_k^{1/2}c_kW_k(\bar{T}_k(t))\\ &= -\sum_{l=1}^{k}\widetilde{Q}_l(t)+\mu_0^{1/2}c_0W_0(t).\end{aligned}$$

This, together with $\widetilde{Q}_k^*=0$ w.p.1, implies that the LIL $\widetilde{D}_k^*=\mu_0^{1/2}c_0$ w.p.1.

Hence, invoking Lemma 4.2 we have established all the LILs in (3.5).

4.2.2 The full overloading case

Suppose that $\rho_k>1$ for all $k=1,2,\ldots,K$. We have $\mu_0>\mu_1>\cdots>\mu_K$ and $\rho_k=\mu_{k-1}/\mu_k$, the corresponding fluid solution:

$$\bar{\mathbb{X}}_k(t) = ((\mu_{k-1}-\mu_k)t,(\rho_k-1)t,t,0,\mu_kt), \quad k=1,2,\ldots,K. \tag{4.7}$$

By (2.10) and (4.2), $\bar{X}_k(t)=(\mu_{k-1}-\mu_k)t$ for all $k=1,2,\ldots,K$ and $t\geq 0$.

First, by Lemma 4.4, we note that $\widetilde{I}_k^*=\widetilde{T}_k^*=0$ w.p.1 for all $k=1,2,\ldots,K$. For the LIL of the queue length process, $k=1,2,\ldots,K$, by (2.10) and (4.2), we have $\bar{Q}_k(t)=\bar{X}_k(t)=(\mu_{k-1}-\mu_k)t$ for all $t\geq 0$, and

$$\begin{aligned}\widetilde{Q}_k(t)-\bar{Q}_k(t) &= \widetilde{X}_k(t)-\bar{X}_k(t)+\widetilde{Y}_k(t)\\ &= \mu_k\widetilde{I}_k(t)-\mu_{k-1}\widetilde{I}_{k-1}(t)+\mu_{k-1}^{1/2}c_{k-1}W_{k-1}(t)-\mu_k^{1/2}c_kW_k(t).\end{aligned} \tag{4.8}$$

Since $\widetilde{I}_k^*=0$ w.p.1 for all $k=1,2,\ldots,K$, we have the LIL $\widetilde{Q}_k^*=\sqrt{\mu_{k-1}c_{k-1}^2+\mu_kc_k^2}$ w.p.1 for all $k=1,2,\ldots,K$.

To treat the LIL of the workload process, we have, by (2.10) and (4.2), $\bar{Z}_k(t) = \bar{Q}_k(t)/\mu_k = (\rho_k - 1)t$ for all $t \geq 0$. This, together with (4.8), implies that

$$\begin{aligned}
\widetilde{Z}_k(t) - \bar{Z}_k(t) &= \frac{1}{\mu_k}\left\{\left[\widetilde{Q}_k(t) - \bar{Q}_k(t)\right] + \mu_k^{1/2} c_k (W_k(t) - W_k(\rho_k t))\right\} \\
&= \frac{1}{\mu_k}\left[\widetilde{X}_k(t) - \bar{X}_k(t) + \widetilde{Y}_k(t) + \mu_k^{1/2} c_k (W_k(t) - W_k(\rho_k t))\right] \\
&= \widetilde{I}_k(t) - \frac{\mu_{k-1}}{\mu_k}\widetilde{I}_{k-1}(t) + \frac{1}{\mu_k}\left[\mu_{k-1}^{1/2} c_{k-1} W_{k-1}(t) - \mu_k^{1/2} c_k W_k(\rho_k t)\right].
\end{aligned}$$

Notice that, by Lemma 4.4, $\widetilde{I}_k^* = 0$ w.p.1, we have the LIL $\widetilde{Z}_k^* = \sqrt{\mu_{k-1}} C_{k-1,k}/\mu_k$ w.p.1 for all $k = 1, 2, \ldots, K$.

For the LIL of the departure process, by (2.10) and (4.2),

$$\widetilde{D}_k(t) - \bar{D}_k(t) = \mu_k\left[\widetilde{T}_k(t) - \bar{T}_k(t)\right] + \mu_k^{1/2} c_k W_k(t) = -\mu_k \widetilde{I}_k(t) + \mu_k^{1/2} c_k W_k(t).$$

Since $\widetilde{I}_k^* = 0$ w.p.1, we have the LIL $\widetilde{D}_k^* = \mu_k^{1/2} c_k$ w.p.1 for all $k = 1, 2. \ldots, K$.

Hence, invoking Lemma 4.2, we have proved the LILs in (3.6).

4.2.3 The hybrid case

Recall that $k_1', k_2', \ldots, k_{l_0}' : 0 < k_1' \leq k_2' \leq \cdots \leq k_{l_0}' \leq K$ satisfy that $\rho_k > 1$ for all $k = k_1', k_2', \ldots, k_{l_0}'$ and $\rho_k < 1$ for all $k \neq k_1', k_2', \ldots, k_{l_0}'$, by (2.4) we have,

$$\begin{aligned}
&\mu_{k_i'} < \mu_{k_{i-1}'} < \mu_k, \quad k = k_{i-1}' + 1, k_{i-1}' + 2, \ldots, k_i' - 1, \quad i = 1, 2, \ldots, l_0, \\
&\mu_{k_{l_0}'} < \mu_k, \quad k = k_{l_0'} + 1, k_{l_0'} + 2, \ldots, K.
\end{aligned} \tag{4.9}$$

Next, we proceed to find the LILs for $k = k_1', k_2', \ldots, k_{l_0}'$ and then for $k \neq k_1', k_2', \ldots, k_{l_0}'$. According to Lemma 4.4, we have $\widetilde{I}_k^* = \widetilde{T}_k^* = 0$ w.p.1 for all $k : \rho_k > 1$. To develop the LIL of $\widetilde{Q}_k(t)$, $k = k_1', k_2', \ldots, k_{l_0}'$, we have, by (2.10) and (4.2),

$$\begin{aligned}
\widetilde{X}_{k_l'} &= -\widetilde{Y}_{k_l'-1}(t) + (\mu_{k_l'-1} - \mu_{k_l'})t + \mu_{k_l'-1}^{1/2} c_{k_l'-1} W_{k_l'-1}(\bar{T}_{k_l'-1}(t)) - \mu_{k_l'}^{1/2} c_{k_l'} W_{k_l'}(\bar{T}_{k_l'}(t)) \\
&= \widetilde{X}_{k_l'-1}(t) - \widetilde{Q}_{k_l'-1}(t) + (\mu_{k_l'-1} - \mu_{k_l'})t + \mu_{k_l'-1}^{1/2} c_{k_l'-1} W_{k_l'-1}(\bar{T}_{k_l'-1}(t)) \\
&\quad -\mu_{k_l'}^{1/2} c_{k_l'} W_{k_l'}(\bar{T}_{k_l'}(t)) \qquad (4.10)\\
&= -\mu_{k_{l-1}'} \widetilde{I}_{k_{l-1}'}(t) - \sum_{i=k_{l-1}'+1}^{k_l'-1} \widetilde{Q}_i(t) + (\mu_{k_{l-1}'} - \mu_{k_l'})t + \mu_{k_{l-1}'}^{1/2} c_{k_{l-1}'} W_{k_{l-1}'}(\bar{T}_{k_{l-1}'}(t)) \\
&\quad -\mu_{k_l'}^{1/2} c_{k_l'} W_{k_l'}(\bar{T}_{k_l'}(t)), \quad l = 1, 2, \ldots, l_0,
\end{aligned}$$

where $k_0' \equiv 0$. Notice that $\widetilde{Q}_{k_l'}(t) = \widetilde{X}_{k_l'}(t) + \widetilde{Y}_{k_l'}(t)$ in (4.2), we have

$$\begin{aligned}
\widetilde{Q}_{k_l'}(t) - \bar{Q}_{k_l'}(t) &= -\mu_{k_{l-1}'} \widetilde{I}_{k_{l-1}'}(t) + \mu_{k_l'} \widetilde{I}_{k_l'}(t) - \sum_{i=k_{l-1}'+1}^{k_l'-1} \widetilde{Q}_i(t) \\
&\quad +\mu_{k_{l-1}'}^{1/2} c_{k_{l-1}'} W_{k_{l-1}'}(\bar{T}_{k_{l-1}'}(t)) - \mu_{k_l'}^{1/2} c_{k_l'} W_{k_l'}(\bar{T}_{k_l'}(t)).
\end{aligned} \tag{4.11}$$

Notice that the LIL $\widetilde{I}_k^* = 0$ w.p.1 for all $k = k_1', k_2', \ldots, k_{l_0}'$, the LIL $\widetilde{Q}_k^* = 0$ w.p.1 for all $k \neq k_1', k_2', \ldots, k_{l_0}'$, w.p.1. Then, for all $l = 1, 2, \ldots, l_0$, the LIL $\widetilde{Q}_{k_l'}^* = \sqrt{\mu_{k_{l-1}'} c_{k_{l-1}'}^2 + \mu_{k_l'} c_{k_l'}^2}$, w.p.1, because the driftless BM $\mu_{k_{l-1}'}^{1/2} c_{k_{l-1}'} W_{k_{l-1}'}(\bar{T}_{k_{l-1}'}(t)) - \mu_{k_l'}^{1/2} c_{k_l'} W_{k_l'}(\bar{T}_{k_l'}(t))$ has variance parameter $\mu_{k_{l-1}'} c_{k_{l-1}'}^2 + \mu_{k_l'} c_{k_l'}^2$.

To treat the LIL of $\widetilde{Z}_k(t)$, $k = k_1', k_2', \ldots, k_{l_0}'$, we have, by (4.2) and (4.11),

$$\begin{aligned}
\widetilde{Z}_{k_l'}(t) - \bar{Z}_{k_l'}(t) &= \frac{1}{\mu_{k_l'}} \left[\widetilde{Q}_{k_l'}(t) - \bar{Q}_{k_l'}(t) + \mu_{k_l'}^{1/2} c_{k_l'} (W_{k_l'}(\bar{T}_{k_l'}(t)) - W_{k_l'}(\rho_{k_l'} t)) \right] \\
&= \frac{1}{\mu_{k_l'}} \Bigg[-\mu_{k_{l-1}'} \widetilde{I}_{k_{l-1}'}(t) + \mu_{k_l'} \widetilde{I}_{k_l'}(t) - \sum_{i=k_{l-1}'+1}^{k_l'-1} \widetilde{Q}_i(t) \\
&\qquad + \mu_{k_{l-1}'}^{1/2} c_{k_{l-1}'} W_{k_{l-1}'}(\bar{T}_{k_{l-1}'}(t)) - \mu_{k_l'}^{1/2} c_{k_l'} W_{k_l'}(\rho_{k_l'} t) \Bigg]
\end{aligned}$$

for $l = 1, 2, \ldots, l_0$. Notice that $\rho_{k_l'} = \mu_{k_{l-1}'} / \mu_{k_l'}$ by (2.4), and $\bar{T}_{k_l'}(t) = t$ for all $t \geq 0$ and $l = 1, 2, \ldots, l_0$, then $\mu_{k_{l-1}'}^{1/2} c_{k_{l-1}'} W_{k_{l-1}'}(\bar{T}_{k_{l-1}'}(t)) - \mu_{k_l'}^{1/2} c_{k_l'} W_{k_l'}(\rho_{k_l'} t)$ in the above equality is a driftless BM with variance parameter $\mu_{k_{l-1}'}(c_{k_{l-1}'}^2 + c_{k_l'}^2)$. Hence, we have $\widetilde{Z}_{k_l'}^* = \sqrt{\mu_{k_{l-1}'}} C_{k_{l-1}', k_l'} / \mu_{k_l'}$ w.p.1 for all $l = 1, 2, \ldots, l_0$.

For the LIL of $\widetilde{D}_k(t)$, $k = k_1', k_2', \ldots, k_{l_0}'$, we have, by (2.10) and (4.2),

$$\widetilde{D}_{k_l'}(t) - \bar{D}_{k_l'}(t) = \mu_{k_l'} \left[\widetilde{T}_{k_l'}(t) - \bar{T}_{k_l'}(t) \right] + \mu_{k_l'}^{1/2} c_{k_l'} W_{k_l'}(\bar{T}_{k_l'}(t)).$$

By Lemma 4.4, $\widetilde{T}_{k_l'}^* = 0$ w.p.1 for all $l = 1, 2, \ldots, l_0$, we have the LIL $\widetilde{D}_{k_l'}^* = \mu_{k_l'}^{1/2} c_{k_l'}$ w.p.1 for all $l = 1, 2, \ldots, l_0$.

Invoking Lemma 4.2, we can obtain all the LILs in (3.7) for stages $k_1', k_2', \ldots, k_{l_0}'$. We next proceed to establish the LILs for stages $k \neq k_1', k_2', \ldots, k_{l_0}'$. First, we note that all the LILs satisfy (3.5) for $k = 1, 2, \ldots, k_1' - 1$, because the first $k_1' - 1$ stages form an independent tandem queue in light traffic as in the full underloading case. Next, we consider the stages $k = k_l' + 1, k_l' + 2, \ldots, k_{l+1}' - 1$ with $l = 1, 2, \ldots, l_0$, where $k_{l_0+1}' \equiv K$. For all $k = k_l' + 1, k_l' + 2, \ldots, k_{l+1}' - 1$, we note by Lemma 2.1 and the traffic intensity (2.4) that $\rho_k = \mu_{k_l'} / \mu_k$, $\bar{T}_k(t) = \rho_k t$, $\bar{Q}_k(t) = 0$ and $\bar{X}_k(t) = (\mu_{k_l'} - \mu_k) t < 0$ for all $t \geq 0$. By Lemma 4.3, the LIL $\widetilde{Q}_k^* = 0$ w.p.1 for all $k = k_l' + 1, k_l' + 2, \ldots, k_{l+1}' - 1$. This, along with (4.2), implies the LIL $\widetilde{Z}_k^* = 0$ w.p.1 for all $k = k_l' + 1, k_l' + 2, \ldots, k_{l+1}' - 1$. By (2.10) and (4.2), we write

$$\widetilde{X}_k(t) - \bar{X}_k(t) = -\mu_{k_l'} \widetilde{I}_{k_l'}(t) - \sum_{i=k_l'+1}^{k-1} \widetilde{Q}_i(t) + \mu_{k_l'}^{1/2} c_{k_l'} W_{k_l'}(\bar{T}_{k_l'}(t)) - \mu_k^{1/2} c_k W_k(\bar{T}_k(t)). \tag{4.12}$$

For the LIL of $I_k(t)$, $k = k_l' + 1, k_l' + 2, \ldots, k_{l+1}' - 1$, we have,

$$\begin{aligned}
\widetilde{I}_k(t) - \bar{I}_k(t) &= \frac{1}{\mu_k} \left[\widetilde{Y}_k(t) - \bar{Y}_k(t) \right] = \frac{1}{\mu_k} \widetilde{Q}_k(t) - \frac{1}{\mu_k} \left[\widetilde{X}_k(t) - \bar{X}_k(t) \right] \\
&= \frac{\mu_{k_l'}}{\mu_k} \widetilde{I}_{k_l'}(t) + \frac{1}{\mu_k} \sum_{i=k_l'+1}^{k} \widetilde{Q}_i(t) - \frac{1}{\mu_k} \left[\mu_{k_l'}^{1/2} c_{k_l'} W_{k_l'}(\bar{T}_{k_l'}(t)) - \mu_k^{1/2} c_k W_k(\bar{T}_k(t)) \right].
\end{aligned}$$

By Lemmas 4.3 and 4.4, the LIL $\widetilde{I}^*_{k'_l} = \widetilde{Q}^*_k = 0$ w.p.1 for all $k = k'_l + 1, k'_l + 2, \ldots, k'_{l+1} - 1$, we have the LIL $\widetilde{I}^*_k = \sqrt{\mu_{k'_l}} C_{k'_l,k}/\mu_k$ w.p.1 for all $k = k'_l + 1, k'_l + 2, \ldots, k'_{l+1} - 1$, because $\mu^{1/2}_{k'_l} c_{k'_l} W_{k'_l}(\bar{T}_{k'_l}(t)) - \mu^{1/2}_k c_k W_k(\bar{T}_k(t))$ is a driftless BM with variance parameter $\mu_{k'_l}(c^2_{k'_l} + c^2_k)$.

For the LIL of $\widetilde{T}_k(t)$, $k = k'_l + 1, k'_l + 2, \ldots, k'_{l+1} - 1$, it follows from (4.6) that the LIL $\widetilde{T}^*_k = \sqrt{\mu_{k'_l}} C_{k'_l,k}/\mu_k$ w.p.1 for all $k = k'_l + 1, k'_l + 2, \ldots, k'_{l+1} - 1$.

To develop the LIL of $\widetilde{D}_k(t)$, $k = k'_l + 1, k'_l + 2, \ldots, k'_{l+1} - 1$, we have, by (4.2) and (4.12),

$$
\begin{aligned}
\widetilde{D}_k(t) - \bar{D}_k(t) &= \mu_k \left[\widetilde{T}_k(t) - \bar{T}_k(t)\right] + \mu^{1/2}_k c_k W_k(\bar{T}_k(t)) \\
&= -\left[\widetilde{Y}_k(t) - \bar{Y}_k(t)\right] + \mu^{1/2}_k c_k W_k(\bar{T}_k(t)) \\
&= -\widetilde{Q}_k(t) + \left[\widetilde{X}_k(t) - \bar{X}_k(t)\right] + \mu^{1/2}_k c_k W_k(\bar{T}_k(t)) \\
&= -\mu_{k'_l} \widetilde{I}_{k'_l}(t) - \sum_{i=k'_l+1}^{k} \widetilde{Q}_i(t) + \mu^{1/2}_{k'_l} c_{k'_l} W_{k'_l}(\bar{T}_{k'_l}(t)),
\end{aligned}
$$

where $\bar{T}_{k'_l}(t) = t$ for all $t \geq 0$. By Lemmas 4.3 and 4.4, the LILs $\widetilde{I}^*_{k'_l} = \widetilde{Q}^*_k = 0$ w.p.1 for all $k = k'_l + 1, k'_l + 2, \ldots, k'_{l+1} - 1$, we have the LIL $\widetilde{D}^*_k = \mu^{1/2}_{k'_l} c_{k'_l}$ w.p.1 for all $k = k'_l + 1, k'_l + 2, \ldots, k'_{l+1} - 1$. All LILs for class $k'_{l_0+1} \equiv K$ can be established similarly.

In summary, we have established all LILs results in (3.5), (3.6), (3.7) and (3.8). □

4.3 Proof of Theorem 3.2

4.3.1 Stages up to k_0 (Cases 1 and 2)

Stages preceding stage k_0. Since the first $k_0 - 1$ stages form an independent $(k_0 - 1)$-stage tandem model, their LILs are already characterized by Theorem 3.1. Specifically, for $k < k_0$, $\mathcal{X}^*_k$ satisfies (3.5) if k is in the full underloading case, it satisfies (3.6) if k is in the full overloading case, and it satisfies (3.7) and (3.8) if k is in the hybrid case.

Stage k_0. Since $\rho_{k_0} = 1$, we note that $\mu_{k_0} = \mu_{\underline{k}_0}$ by (3.9) and the fluid solution:

$$
\bar{\mathbb{X}}_{k_0}(t) = (0, 0, t, 0, \mu_{k_0} t) \tag{4.13}
$$

and $\bar{T}_{\underline{k}}(t) = t$ for all $t \geq 0$ by (2.10). By (4.2), it follows that

$$
\begin{aligned}
&\widetilde{X}_{k_0}(t) \\
=\ & -\widetilde{Y}_{k_0-1}(t) + (\mu_{k_0-1} - \mu_{k_0})t + \mu^{1/2}_{k_0-1} c_{k_0-1} W_{k_0-1}(\bar{T}_{k_0-1}(t)) - \mu^{1/2}_{k_0} c_{k_0} W_{k_0}(t) \\
=\ & \widetilde{X}_{k_0-1}(t) - \widetilde{Q}_{k_0-1}(t) + (\mu_{k_0-1} - \mu_{k_0})t + \mu^{1/2}_{k_0-1} c_{k_0-1} W_{k_0-1}(\bar{T}_{k_0-1}(t)) - \mu^{1/2}_{k_0} c_{k_0} W_{k_0}(t) \\
=\ & \widetilde{X}_{\underline{k}+1}(t) - \sum_{i=\underline{k}_0+1}^{k_0-1} \widetilde{Q}_i(t) + (\mu_{\underline{k}_0+1} - \mu_{k_0})t + \mu^{1/2}_{\underline{k}_0+1} c_{\underline{k}_0+1} W_{\underline{k}_0+1}(\bar{T}_{\underline{k}_0+1}(t)) - \mu^{1/2}_{k_0} c_{k_0} W_{k_0}(t) \\
=\ & -\widetilde{Y}_{\underline{k}_0}(t) - \sum_{i=\underline{k}_0+1}^{k_0-1} \widetilde{Q}_i(t) + \mu^{1/2}_{\underline{k}_0} c_{\underline{k}_0} W_{\underline{k}_0}(t) - \mu^{1/2}_{k_0} c_{k_0} W_{k_0}(t).
\end{aligned} \tag{4.14}
$$

To treat the LIL of the idle time process, we note that $\bar{X}_{k_0}(t) = \bar{I}_{k_0}(t) = 0$ and $\bar{T}_{k_0}(t) = t$ for all $t \geq 0$. Notice that by Lemma 4.3 the LIL $\widetilde{Q}_k^* = 0$ w.p.1 for all $k = \underline{k}_0+1, 2, \ldots, k_0-1$, and by Lemma 4.4 the LIL $\widetilde{I}_{\underline{k}_0}^* = 0$ w.p.1, we have the LIL, w.p.1,

$$
\begin{aligned}
\widetilde{I}_{k_0}^* &= \frac{1}{\mu_{k_0}} \limsup_{L\to\infty} \frac{1}{\varphi(L)} \left\{ \sup_{0\leq t\leq L} \widetilde{Y}_{k_0}(t) \right\} \\
&= \frac{1}{\mu_{k_0}} \limsup_{L\to\infty} \frac{1}{\varphi(L)} \left\{ \sup_{0\leq t\leq L} \left\{ \sup_{0\leq s\leq t} \left[-\widetilde{X}_{k_0}(s) \right] \right\} \right\} = \frac{1}{\mu_{k_0}} \limsup_{L\to\infty} \frac{1}{\varphi(L)} \left\{ \sup_{0\leq t\leq L} \left[-\widetilde{X}_{k_0}(t) \right] \right\} \\
&= \frac{1}{\mu_{k_0}} \limsup_{L\to\infty} \frac{1}{\varphi(L)} \left\{ \sup_{0\leq t\leq L} \left[-\mu_{\underline{k}_0}^{1/2} c_{\underline{k}_0} W_{\underline{k}_0}(t) + \mu_{k_0}^{1/2} c_{k_0} W_{k_0}(t) \right] \right\} = \frac{C_{\underline{k}_0,k_0}}{\sqrt{\mu_{\underline{k}_0}}},
\end{aligned} \tag{4.15}
$$

where the second equality holds by (2.11), the fifth equality holds because $\mu_{\underline{k}_0} = \mu_{k_0}$ and $-\mu_{\underline{k}_0}^{1/2} c_{\underline{k}_0} W_{\underline{k}_0}(t) + \mu_{k_0}^{1/2} c_{k_0} W_{k_0}(t)$ is a driftless BM with variance parameter $\mu_{\underline{k}_0} C_{\underline{k}_0,k_0}^2$.

The LILs of $\widetilde{Z}_{k_0}(t)$ and $\widetilde{T}_{k_0}(t)$ equal to $\widetilde{I}_{k_0}^*$ by Lemma 4.5, that is, $\widetilde{T}_{k_0}^* = \widetilde{Z}_{k_0}^* = C_{\underline{k}_0,k_0}/\sqrt{\mu_{\underline{k}_0}}$ w.p.1.

To develop the LIL of the queue length process, we note in advance that the Lipschitz constant for Φ is 2. To see this, for any given $\widetilde{X}'$ and $\widetilde{X}''$, we have

$$
\begin{aligned}
\left|\left| \Phi(\widetilde{X}') - \Phi(\widetilde{X}'') \right|\right|_L &\leq \sup_{0\leq t\leq L} \left\{ \left| \widetilde{X}'(t) - \widetilde{X}''(t) \right| + \left| \sup_{0\leq s\leq t} \{-\widetilde{X}'(s)\}^+ - \sup_{0\leq s\leq t} \{-\widetilde{X}''(s)\}^+ \right| \right\} \\
&\leq \left|\left| \widetilde{X}' - \widetilde{X}'' \right|\right|_L + \sup_{0\leq t\leq L} \sup_{0\leq s\leq t} \left| \widetilde{X}'(s) - \widetilde{X}''(s) \right| \\
&= 2 \left|\left| \widetilde{X}' - \widetilde{X}'' \right|\right|_L .
\end{aligned} \tag{4.16}
$$

By (4.2), $\widetilde{Q}_{k_0}(t) = \Phi(\widetilde{X}_{k_0})(t)$. For all $L > 0$, with (4.14) we have

$$
\begin{aligned}
& \sup_{0\leq t\leq L} \left| \Phi(\widetilde{X}_{k_0})(t) - \Phi(\mu_{\underline{k}_0}^{1/2} c_{\underline{k}_0} W_{\underline{k}_0} - \mu_{k_0}^{1/2} c_{k_0} W_{k_0})(t) \right| \\
\leq \; & 2 \sup_{0\leq t\leq L} \left| \widetilde{X}_{k_0}(t) - (\mu_{\underline{k}_0}^{1/2} c_{\underline{k}_0} W_{\underline{k}_0}(t) - \mu_{k_0}^{1/2} c_{k_0} W_{k_0}(t)) \right| = 2 \left(\widetilde{Y}_{\underline{k}_0}(t) + \sup_{0\leq t\leq L} \sum_{i=1}^{k_0-1} \widetilde{Q}_i(t) \right),
\end{aligned}
$$

where the inequality holds by (4.16). By Lemma 4.3 and Lemma 4.4, we have

$$
\begin{aligned}
& \limsup_{L\to\infty} \frac{1}{\varphi(L)} \left\{ \sup_{0\leq t\leq L} \left| \Phi(\widetilde{X}_{k_0})(t) - \Phi(\mu_{\underline{k}_0}^{1/2} c_{\underline{k}_0} W_{\underline{k}_0}(t) - \mu_{k_0}^{1/2} c_{k_0} W_{k_0}(t)) \right| \right\} \\
& \leq 2 \limsup_{L\to\infty} \frac{\widetilde{Y}_{\underline{k}_0}(t)}{\varphi(L)} + 2 \limsup_{L\to\infty} \frac{1}{\varphi(L)} \left\{ \sup_{0\leq t\leq L} \sum_{i=1}^{k_0-1} \widetilde{Q}_i(t) \right\} = 0, \quad \text{w.p.1.}
\end{aligned} \tag{4.17}
$$

Hence, w.p.1,

$$
\begin{aligned}
\widetilde{Q}^*_{k_0} &= \limsup_{L\to\infty} \frac{1}{\varphi(L)} \left\{ \sup_{0\le t\le L} \left| \Phi(\widetilde{X}_{k_0})(t) \right| \right\} \\
&= \limsup_{L\to\infty} \frac{1}{\varphi(L)} \left\{ \sup_{0\le t\le L} \left| \Phi(\mu_{\underline{k}}^{1/2} c_{\underline{k}_0} W_{\underline{k}_0}(t) - \mu_{k_0}^{1/2} c_{k_0} W_{k_0}(t)) \right| \right\} \\
&= \limsup_{L\to\infty} \frac{1}{\varphi(L)} \left\{ \sup_{0\le t\le L} \left| \mu_{\underline{k}_0}^{1/2} c_{\underline{k}_0} W_{\underline{k}_0}(t) - \mu_{k_0}^{1/2} c_{k_0} W_{k_0}(t) \right| \right\} = \sqrt{\mu_{\underline{k}_0}} C_{\underline{k}_0, k_0}, \qquad (4.18)
\end{aligned}
$$

where the second equality holds by (4.17), the third equality holds because

$$
\Phi(\mu_{\underline{k}_0}^{1/2} c_{\underline{k}_0} W_{\underline{k}_0}(t) - \mu_{k_0}^{1/2} c_{k_0} W_{k_0}(t))(t) \stackrel{\mathrm{d}}{=} \left| \mu_{\underline{k}_0}^{1/2} c_{\underline{k}_0} W_{\underline{k}_0}(t) - \mu_{k_0}^{1/2} c_{k_0} W_{k_0}(t) \right|,
$$

the fourth equality holds because $\mu_{\underline{k}_0}^{1/2} c_{\underline{k}_0} W_{\underline{k}_0}(t) - \mu_{k_0}^{1/2} c_{k_0} W_{k_0}(t)$ is a driftless BM with variance parameter $\mu_{\underline{k}_0} C^2_{\underline{k}_0, k_0}$. Readers can refer to Theorem 5.1 on p.119 in [31] for a similar result for a standard BM.

To treat the LIL of the departure process, by (2.10), (4.2) and (4.14), we have

$$
\begin{aligned}
& \widetilde{D}_{k_0}(t) - \bar{D}_{k_0}(t) \qquad (4.19) \\
= \; & \mu_{k_0} \left[\widetilde{T}_{k_0}(t) - \bar{T}_{k_0}(t) \right] + \mu_{k_0}^{1/2} c_{k_0} W_{k_0}(\bar{T}_{k_0}(t)) \\
= \; & -\widetilde{Y}_{k_0}(t) + \mu_{k_0}^{1/2} c_{k_0} W_{k_0}(t) = -\sup_{0\le s\le t} \left[-\widetilde{X}_{k_0}(s) \right] + \mu_{k_0}^{1/2} c_{k_0} W_{k_0}(t) \\
= \; & -\sup_{0\le s\le t} \left[\widetilde{Y}_{\underline{k}_0}(s) + \sum_{i=\underline{k}_0+1}^{k_0-1} \widetilde{Q}_i(s) - \mu_{\underline{k}_0}^{1/2} c_{\underline{k}_0} W_{\underline{k}_0}(s) + \mu_{k_0}^{1/2} c_{k_0} W_{k_0}(s) \right] + \mu_{k_0}^{1/2} c_{k_0} W_{k_0}(t) \\
\stackrel{\mathrm{d}}{=} \; & -\sup_{0\le s\le t} \left[\widetilde{Y}_{\underline{k}_0}(s) + \sum_{i=\underline{k}_0+1}^{k_0-1} \widetilde{Q}_i(s) - \mu_{\underline{k}_0}^{1/2} c_{\underline{k}_0} W_{\underline{k}_0}(s) - \mu_{k_0}^{1/2} c_{k_0} W_{k_0}(t-s) \right]. \qquad (4.20)
\end{aligned}
$$

By Lemmas 4.3 and 4.4, $\widetilde{I}^*_{\underline{k}_0} = \widetilde{Q}^*_k = 0$ w.p.1 for all $k = \underline{k}_0 + 1, 2, \ldots, k_0 - 1$ w.p.1, we have, w.p.1,

$$
\begin{aligned}
\widetilde{D}^*_{k_0} &= \limsup_{L\to\infty} \frac{1}{\varphi(L)} \left\{ \sup_{0\le t\le L} \sup_{0\le s\le t} \left| \mu_{\underline{k}_0}^{1/2} c_{\underline{k}_0} W_{\underline{k}_0}(s) + \mu_{k_0}^{1/2} c_{k_0} W_{k_0}(t-s) \right| \right\} \\
&= \sqrt{\mu_{\underline{k}_0}} \widehat{C}_{\underline{k}_0, k_0}, \qquad (4.21)
\end{aligned}
$$

where the second equality is from Lemma 4 in [15]. Invoking Lemma 4.2, we have proved all LILs results in (3.11).

4.3.2 Stages succeeding stage k_0 (Case 3)

Stages between k_0 and $\bar{k}_0$ (Case 3.a). First, by the definition of $\bar{k}_0$ in (3.10) we note that $\rho_k < 1$ for $k = k_0 + 1, \ldots, (\bar{k}_0 \wedge K) - 1$. Hence, we have the following fluid solution: for $k = k_0 + 1, \ldots, (\bar{k}_0 \wedge K) - 1$,

$$
\bar{\mathbb{X}}_k(t) = (0, 0, \rho_k t, (1-\rho_k)t, \mu_{k_0} t) \quad \text{with} \quad \rho_k = \frac{\mu_{k_0}}{\mu_k}. \qquad (4.22)
$$

For $l = 1, 2, \dots, (\bar{k}_0 \wedge K - 1) - k_0$, by (4.22) we note that $\bar{Q}_{k_0+l}(t) = 0$ for all $t \geq 0$. By (2.10) and (4.2), we have

$$\begin{aligned}
& \widetilde{X}_{k_0+l}(t) - \bar{X}_{k_0+l}(t) \\
= \; & \mu_{k_0+l-1}\left[\widetilde{T}_{k_0+l-1}(t) - \bar{T}_{k_0+l-1}(t)\right] \\
& + \mu_{k_0+l-1}^{1/2} c_{k_0+l-1} W_{k_0+l-1}(\bar{T}_{k_0+l-1}(t)) - \mu_{k_0+l}^{1/2} c_{k_0+l} W_{k_0+l}(\bar{T}_{k_0+l}(t)) \\
= \; & \left[\widetilde{X}_{k_0+l-1}(t) - \bar{X}_{k_0+l-1}(t)\right] - \widetilde{Q}_{k_0+l-1}(t) \\
& + \mu_{k_0+l-1}^{1/2} c_{k_0+l-1} W_{k_0+l-1}(\bar{T}_{k_0+l-1}(t)) - \mu_{k_0+l}^{1/2} c_{k_0+l} W_{k_0+l}(\bar{T}_{k_0+l}(t)) \\
= \; & \left[\widetilde{X}_{k_0+l-2}(t) - \bar{X}_{k_0+l-2}(t)\right] - \widetilde{Q}_{k_0+l-2}(t) - \widetilde{Q}_{k_0+l-1}(t) \\
& + \mu_{k_0+l-2}^{1/2} c_{k_0+l-2} W_{k_0+l-2}(\bar{T}_{k_0+l-2}(t)) - \mu_{k_0+l}^{1/2} c_{k_0+l} W_{k_0+l}(\bar{T}_{k_0+l}(t)) \\
= \; & \left[\widetilde{X}_{k_0+1}(t) - \bar{X}_{k_0+1}(t)\right] - \sum_{i=1}^{l-1} \widetilde{Q}_{k_0+i}(t) \qquad (4.23) \\
& + \mu_{k_0+1}^{1/2} c_{k_0+1} W_{k_0+1}(\bar{T}_{k_0+1}(t)) - \mu_{k_0+l}^{1/2} c_{k_0+l} W_{k_0+l}(\bar{T}_{k_0+l}(t)) \\
= \; & -\mu_{k_0} \widetilde{I}_{k_0}(t) - \sum_{i=1}^{l-1} \widetilde{Q}_{k_0+i}(t) + \mu_{k_0}^{1/2} c_{k_0} W_{k_0}(\bar{T}_{k_0}(t)) - \mu_{k_0+l}^{1/2} c_{k_0+l} W_{k_0+l}(\bar{T}_{k_0+l}(t)),
\end{aligned}$$

where $\bar{T}_{k_0}(t) = t$ for all $t \geq 0$. Invoking (2.11) and (4.14), we have

$$\begin{aligned}
& \widetilde{X}_{k_0+l}(t) - \bar{X}_{k_0+l}(t) \\
= \; & -\sup_{0 \leq u \leq t} \left\{ \widetilde{Y}_{\underline{k}_0}(u) + \sum_{i=\underline{k}_0+1}^{k_0-1} \widetilde{Q}_i(u) - \mu_{\underline{k}_0}^{1/2} c_{\underline{k}_0} W_{\underline{k}_0}(u) + \mu_{k_0}^{1/2} c_{k_0} W_{k_0}(u) \right\} \qquad (4.24) \\
& - \sum_{i=1}^{l-1} \widetilde{Q}_{k_0+i}(t) + \mu_{k_0}^{1/2} c_{k_0} W_{k_0}(t) - \mu_{k_0+l}^{1/2} c_{k_0+l} W_{k_0+l}(\bar{T}_{k_0+l}(t)).
\end{aligned}$$

To develop the LIL of $\widetilde{I}_k(t)$ for $l = 1, 2, \dots, (\bar{k}_0 \wedge K - 1) - k_0$, we have, by (2.10) and (4.2),

$$\begin{aligned}
\widetilde{I}_{k_0+l}(t) - \bar{I}_{k_0+l}(t) &= \frac{1}{\mu_{k_0+l}}\left[\widetilde{Q}_{k_0+l}(t) - \widetilde{X}_{k_0+l}(t)\right] - \left(t - \bar{T}_{k_0+l}(t)\right) \\
&= \frac{1}{\mu_{k_0+l}} \widetilde{Q}_{k_0+l}(t) - \frac{1}{\mu_{k_0+l}}\left[\widetilde{X}_{k_0+l}(t) - \mu_{k_0+l}\left(\bar{T}_{k_0+l}(t) - t\right)\right] \\
&= \frac{1}{\mu_{k_0+l}} \widetilde{Q}_{k_0+l}(t) - \frac{1}{\mu_{k_0+l}}\left[\widetilde{X}_{k_0+l}(t) - \bar{X}_{k_0+l}(t)\right]. \qquad (4.25)
\end{aligned}$$

Combining Lemma 2.1, (4.2), (4.24) and (4.25), we have,

$$
\begin{aligned}
\widetilde{I}^*_{k_0+l} &= \limsup_{L\to\infty} \frac{1}{\varphi(L)} \left|\left|\widetilde{I}_{k_0+l} - \bar{I}_{k_0+l}\right|\right|_L \\
&= \limsup_{L\to\infty} \frac{1}{\mu_{k_0+l}\varphi(L)} \left|\left|\widetilde{Q}_{k_0+l} - \left(\widetilde{X}_{k_0+l} - \bar{X}_{k_0+l}\right)\right|\right|_L \\
&= \frac{1}{\mu_{k_0+l}} \limsup_{L\to\infty} \frac{1}{\varphi(L)} \left|\left|\widetilde{X}_{k_0+l} - \bar{X}_{k_0+l}\right|\right|_L \\
&= \frac{1}{\mu_{k_0+l}} \limsup_{L\to\infty} \frac{1}{\varphi(L)} \left\{ \sup_{0\le t\le L} \left| \begin{array}{l} -\sup\limits_{0\le u\le t} \left\{ \widetilde{Y}_{\underline{k}_0}(u) + \sum\limits_{i=\underline{k}_0+1}^{k_0-1} \widetilde{Q}_i(u) - \mu_{\underline{k}_0}^{1/2} c_{\underline{k}_0} W_{\underline{k}_0}(u) \right. \\ \left. +\mu_{k_0}^{1/2} c_{k_0} W_{k_0}(u) \right\} + \mu_{k_0}^{1/2} c_{k_0} W_{k_0}(t) \\ -\sum\limits_{i=1}^{l-1} \widetilde{Q}_{k_0+i}(t) - \mu_{k_0+l}^{1/2} c_{k_0+l} W_{k_0+l}(T_{k_0+l}(t)) \end{array} \right| \right\} \\
&= \frac{1}{\mu_{k_0+l}} \limsup_{L\to\infty} \frac{1}{\varphi(L)} \left\{ \sup_{0\le t\le L} \left| \begin{array}{l} -\sup\limits_{0\le u\le t} \left\{ -\mu_{\underline{k}_0}^{1/2} c_{\underline{k}_0} W_{\underline{k}_0}(u) + \mu_{k_0}^{1/2} c_{k_0} W_{k_0}(u) \right\} \\ +\mu_{k_0}^{1/2} c_{k_0} W_{k_0}(t) - \mu_{k_0+l}^{1/2} c_{k_0+l} W_{k_0+l}(T_{k_0+1}(t)) \end{array} \right| \right\} \quad (4.26) \\
&= \frac{1}{\mu_{k_0+l}} \limsup_{L\to\infty} \frac{1}{\varphi(L)} \left\{ \sup_{0\le t\le L} \left| \begin{array}{l} \sup\limits_{0\le u\le t} \left\{ -\mu_{\underline{k}_0}^{1/2} c_{\underline{k}_0} W_{\underline{k}_0}(u) + \mu_{k_0}^{1/2} c_{k_0} W_{k_0}(t-u) \right\} \\ +\mu_{k_0+l}^{1/2} c_{k_0+l} W_{k_0+l}(T_{k_0+l}(t)) \end{array} \right| \right\} \\
&= \frac{1}{\mu_{k_0+l}} \limsup_{L\to\infty} \frac{1}{\varphi(L)} \left\{ \sup_{0\le t\le L} \left| \begin{array}{l} \sup\limits_{0\le u\le t} \left\{ -\mu_{\underline{k}_0}^{1/2} c_{\underline{k}_0} W_{\underline{k}_0}(u) + \mu_{k_0}^{1/2} c_{k_0} W_{k_0}(t-u) \right\} \\ +\mu_{k_0}^{1/2} c_{k_0+l} W_{k_0+l}(t) \end{array} \right| \right\},
\end{aligned}
$$

where the fifth equality holds because $\widetilde{Y}^*_{\underline{k}_0} = \widetilde{Q}^*_{k_0+l} = 0$ due to Lemmas 4.3 and 4.4. The numerator of the seventh equality in (4.26) is

$$
\begin{aligned}
&\sup_{0\le t\le L} \left| \sup_{0\le\theta\le 1} \left\{ -\mu_{\underline{k}_0}^{1/2} c_{\underline{k}_0} W_{\underline{k}_0}(\theta t) + \mu_{k_0}^{1/2} c_{k_0} W_{k_0}((1-\theta)t) + \mu_{k_0}^{1/2} c_{k_0+l} W_{k_0+l}(t) \right\} \right| \\
\stackrel{\mathrm{d}}{=}\ &\sup_{0\le t\le L} \left| \sup_{0\le\theta\le 1} \sqrt{\mu_{\underline{k}_0} c^2_{\underline{k}_0}\theta + \mu_{k_0} c^2_{k_0}(1-\theta) + \mu_{k_0} c^2_{k_0+l}}\, W(t) \right|,
\end{aligned}
$$

where $W(t)$ is a one-dimensional standard BM, and the equality in distribution holds because $W_{\underline{k}_0}(t)$, $W_{k_0}(t)$ and $W_{k_0+l}(t)$ are mutually independent BMs. This, along with (4.26), implies that, w.p.1,

$$
\widetilde{I}^*_{k_0+l} = \frac{\sup\limits_{0\le\theta\le 1} \sqrt{\mu_{\underline{k}_0} c^2_{\underline{k}_0}\theta + \mu_{k_0} c^2_{k_0}(1-\theta) + \mu_{k_0} c^2_{k_0+l}}}{\mu_{k_0+l}} = \frac{\sqrt{\mu_{\underline{k}_0}(\widehat{C}^2_{\underline{k}_0,k_0} + c^2_{k_0+l})}}{\mu_{k_0+l}}, \qquad (4.27)
$$

where the equality holds because $\mu_{\underline{k}_0} = \mu_{k_0}$.

For the LIL of the busy time process, it follows from (4.6) that the LIL $\widetilde{T}^*_{k_0+l} = \frac{1}{\mu_{k_0+l}}\sqrt{\mu_{\underline{k}_0}(\widehat{C}^2_{\underline{k}_0,k_0} + c^2_{k_0+l})}$ w.p.1 for $l = 1, 2, \ldots, (\bar{k}_0 \wedge K - 1) - k_0$.

For the LIL of $\widetilde{D}_{k_0+l}(t)$ for $l = 1, 2, \ldots, (\bar{k}_0 \wedge K - 1) - k_0$, we have

$$\begin{aligned}
\widetilde{D}_{k_0+l}(t) - \bar{D}_{k_0+l}(t) &= \mu_{k_0+l}\left[\widetilde{T}_{k_0+l}(t) - \bar{T}_{\mu_{k_0+l}}(t)\right] + \mu_{k_0+l}^{1/2} c_{k_0+l} W_{k_0+l}(\bar{T}_{k_0+l}(t)) \\
&= -\widetilde{Q}_{k_0+l}(t) + \left[\widetilde{X}_{k_0+l}(t) - \bar{X}_{k_0+l}(t)\right] + \mu_{k_0+l}^{1/2} c_{k_0+l} W_{k_0+l}(\bar{T}_{k_0+l}(t)) \\
&= -\sup_{0\le u\le t}\left\{\widetilde{Y}_{\underline{k}_0}(u) + \sum_{i=\underline{k}_0+1}^{k_0-1}\widetilde{Q}_i(u) - \mu_{\underline{k}_0}^{1/2} c_{\underline{k}_0} W_{\underline{k}_0}(u) + \mu_{k_0}^{1/2} c_{k_0} W_{k_0}(u)\right\} \\
&\qquad - \sum_{i=1}^{l}\widetilde{Q}_{k_0+i}(t) + \mu_{k_0}^{1/2} c_{k_0} W_{k_0}(\bar{T}_{k_0}(t)), \qquad (4.28)
\end{aligned}$$

where the third equality holds by (4.24) and the fluid $\bar{Q}_{k_0+l}(t) = 0$ and $\bar{T}_{k_0}(t) = t$ for all $t \ge 0$. By Lemmas 4.3 and 4.4, $\widetilde{Y}^*_{\underline{k}_0} = \widetilde{Q}^*_{k_0+l} = 0$ for $l = 1, 2, \ldots, (\bar{k}_0 \wedge K - 1) - k_0$. By (4.28), we have, w.p.1,

$$\begin{aligned}
\widetilde{D}^*_{k_0+l} &= \limsup_{L\to\infty}\frac{1}{\varphi(L)}\left|\left|\widetilde{D}_{k_0+l} - \bar{D}_{k_0+l}\right|\right|_L \\
&= \limsup_{L\to\infty}\frac{1}{\varphi(L)}\left\{\sup_{0\le t\le L}\left|\sup_{0\le u\le t}\left\{-\mu_{\underline{k}_0}^{1/2} c_{\underline{k}_0} W_{\underline{k}_0}(u) + \mu_{k_0}^{1/2} c_{k_0} W_{k_0}(u)\right\} - \mu_{k_0}^{1/2} c_{k_0} W_{k_0}(\underline{k}_0)\right|\right\} \\
&= \limsup_{L\to\infty}\frac{1}{\varphi(L)}\left\{\sup_{0\le t\le L}\left|\sup_{0\le u\le t}\left\{-\mu_{\underline{k}_0}^{1/2} c_{\underline{k}_0} W_{\underline{k}_0}(u) - \mu_{k_0}^{1/2} c_{k_0} W_{k_0}(t-u)\right\}\right|\right\} = \mu_{\underline{k}_0}^{1/2}\widehat{C}_{\underline{k}_0,k_0},
\end{aligned}$$

where the second equality holds similarly to (4.26), and the fourth equality holds by Lemma 4 in [15] and $\mu_{\underline{k}_0} = \mu_{k_0}$. Again, invoking Lemma 4.2 yields all LIL results in (3.12).

Stage $\bar{k}_0$ (Case 3.b). Notice that $\rho_{k_0} = 1$, $\rho_{\bar{k}_0} = \mu_{k_0}/\mu_{\bar{k}_0} > 1$ and $\rho_k < 1$ for all $k = k_0+1, k_0+2, \ldots, \bar{k}_0 - 1$, by (2.10) we have the following fluid solution: for all $t \ge 0$,

$$\bar{\mathbb{X}}_{\bar{k}_0}(t) = \left((\mu_{k_0} - \mu_{\bar{k}_0})t, (\rho_{\bar{k}_0} - 1)t, t, 0, \mu_{\bar{k}_0} t\right), \qquad (4.29)$$

and $\bar{X}_{\bar{k}_0}(t) = \bar{Q}_{\bar{k}_0}(t)$. By Lemma 4.4, we first note that $\widetilde{I}^*_{\bar{k}_0} = \widetilde{T}^*_{\bar{k}_0} = 0$ w.p.1.

To establish the LIL of the queue length process, by (2.10), (4.2) and (4.29), we have

$$\begin{aligned}
& \widetilde{Q}_{\bar{k}_0}(t) - \bar{Q}_{\bar{k}_0}(t) \\
= \;& (\mu_{\bar{k}_0-1} - \mu_{k_0})t - \widetilde{Y}_{\bar{k}_0-1}(t) + \mu_{\bar{k}_0-1}^{1/2} c_{\bar{k}_0-1} W_{\bar{k}_0-1}(\bar{T}_{\bar{k}_0-1}(t)) - \mu_{\bar{k}_0}^{1/2} c_{\bar{k}_0} W_{\bar{k}_0}(t) + \widetilde{Y}_{\bar{k}_0}(t) \\
= \;& (\mu_{\bar{k}_0-1} - \mu_{k_0})t - \widetilde{Q}_{\bar{k}_0-1}(t) + \widetilde{X}_{\bar{k}_0-1}(t) + \mu_{\bar{k}_0-1}^{1/2} c_{\bar{k}_0-1} W_{\bar{k}_0-1}(\bar{T}_{\bar{k}_0-1}(t)) \\
& -\mu_{\bar{k}_0}^{1/2} c_{\bar{k}_0} W_{\bar{k}_0}(t) + \widetilde{Y}_{\bar{k}_0}(t) \\
= \;& (\mu_{\bar{k}_0-2} - \mu_{k_0})t - \widetilde{Q}_{\bar{k}_0-1}(t) - \widetilde{Q}_{\bar{k}_0-2}(t) + \widetilde{X}_{\bar{k}_0-2}(t) + \mu_{\bar{k}_0-2}^{1/2} c_{\bar{k}_0-2} W_{\bar{k}_0-2}(\bar{T}_{\bar{k}_0-2}(t)) \\
& -\mu_{\bar{k}_0}^{1/2} c_{\bar{k}_0} W_{\bar{k}_0}(t) + \widetilde{Y}_{\bar{k}_0}(t) \\
= \;& -\sum_{i=k_0+1}^{\bar{k}_0-1}\widetilde{Q}_i(t) - \widetilde{Y}_{k_0}(t) + \mu_{k_0}^{1/2} c_{k_0} W_{k_0}(t) - \mu_{\bar{k}_0}^{1/2} c_{\bar{k}_0} W_{\bar{k}_0}(t) + \widetilde{Y}_{\bar{k}_0}(t). \qquad (4.30)
\end{aligned}$$

Then, by (4.14) and (4.30), we have

$$
\begin{aligned}
&\widetilde{Q}_{\bar{k}_0}(t) - \bar{Q}_{\bar{k}_0}(t) \\
&= - \sum_{i=k_0+1}^{\bar{k}_0-1} \widetilde{Q}_i(t) + \mu_{k_0}^{1/2} c_{k_0} W_{k_0}(t) - \mu_{\bar{k}_0}^{1/2} c_{\bar{k}_0} W_{\bar{k}_0}(t) + \widetilde{Y}_{\bar{k}_0}(t) \\
&\quad - \sup_{0\le u\le t} \left\{ \widetilde{Y}_{\underline{k}_0}(u) + \sum_{i=\underline{k}_0+1}^{k_0-1} \widetilde{Q}_i(u) - \mu_{\underline{k}_0}^{1/2} c_{\underline{k}_0} W_{\underline{k}_0}(u) + \mu_{k_0}^{1/2} c_{k_0} W_{k_0}(u) \right\} \\
&= - \sum_{i=k_0+1}^{\bar{k}_0-1} \widetilde{Q}_i(t) + \widetilde{Y}_{\bar{k}_0}(t) - \mu_{\bar{k}_0}^{1/2} c_{\bar{k}_0} W_{\bar{k}_0}(t) \\
&\quad - \sup_{0\le u\le t} \left\{ \widetilde{Y}_{\underline{k}_0}(u) + \sum_{i=\underline{k}_0+1}^{k_0-1} \widetilde{Q}_i(u) - \mu_{\underline{k}_0}^{1/2} c_{\underline{k}_0} W_{\underline{k}_0}(u) - \mu_{k_0}^{1/2} c_{k_0} W_{k_0}(t-u) \right\}.
\end{aligned}
$$

Hence, $\widetilde{Q}^*_{k_0+1}$

$$
\begin{aligned}
&= \limsup_{L\to\infty} \frac{1}{\varphi(L)} \left|\left| \widetilde{Q}_{k_0+1} - \bar{Q}_{k_0+1} \right|\right|_L \\
&= \limsup_{n\to\infty} \frac{1}{\varphi(L)} \left\{ \sup_{0\le t\le L} \left| \begin{array}{c} \sup\limits_{0\le u\le t} \left\{ \widetilde{Y}_{\underline{k}_0}(u) + \sum\limits_{i=\underline{k}_0+1}^{k_0-1} \widetilde{Q}_i(u) - \mu_{\underline{k}_0}^{1/2} c_{\underline{k}_0} W_{\underline{k}_0}(u) - \mu_{k_0}^{1/2} c_{k_0} W_{k_0}(t-u) \right\} \\ + \sum\limits_{i=k_0+1}^{\bar{k}_0-1} \widetilde{Q}_i(t) - \widetilde{Y}_{\bar{k}_0}(t) + \mu_{\bar{k}_0}^{1/2} c_{\bar{k}_0} W_{\bar{k}_0}(t) \end{array} \right| \right\} \\
&= \limsup_{L\to\infty} \frac{1}{\varphi(L)} \left\{ \sup_{0\le t\le L} \left| \sup_{0\le u\le t} \left[-\mu_{\underline{k}_0}^{1/2} c_{\underline{k}_0} W_{\underline{k}_0}(u) - \mu_{k_0}^{1/2} c_{k_0} W_{k_0}(t-u) \right] + \mu_{\bar{k}_0}^{1/2} c_{\bar{k}_0} W_{\bar{k}_0}(t) \right| \right\} \\
&= \sqrt{\mu_{k_0} \widehat{C}^2_{\underline{k}_0,k_0} + \mu_{\bar{k}_0} c^2_{\bar{k}_0}}, \quad \text{w.p.1}, \qquad (4.31)
\end{aligned}
$$

where the third equality holds because, by Lemma 4.3 and Lemma 4.4, $\widetilde{Q}^*_i = \widetilde{I}^*_{\underline{k}_0} = \widetilde{I}^*_{\bar{k}_0} = 0$ w.p.1 for all $i = \underline{k}_0 + 1, \underline{k}_0 + 2, \ldots, k_0 - 1$ and $k_0 + 1, k_0 + 2, \ldots, \bar{k}_0 - 1$, and the fourth equality holds similarly to (4.27) given $\mu_{\underline{k}_0} = \mu_{k_0}$.

To show the LIL of the workload process, we have, combing Lemma 2.1, (4.2), (4.14) and (4.30),

$$
\begin{aligned}
& \widetilde{Z}_{\bar{k}_0}(t) - \bar{Z}_{\bar{k}_0}(t) \\
= \; & \frac{1}{\mu_{\bar{k}_0}} \left[\widetilde{Q}_{\bar{k}_0}(t) - \bar{Q}_{\bar{k}_0}(t) + \mu_{\bar{k}_0}^{1/2} c_{\bar{k}_0} \left(W_{\bar{k}_0}(\bar{T}_{\bar{k}_0}(t)) - W_{\bar{k}_0}(\rho_{\bar{k}_0} t) \right) \right] \\
= \; & \frac{1}{\mu_{\bar{k}_0}} \left[- \sum_{i=k_0+1}^{\bar{k}_0 - 1} \widetilde{Q}_i(t) - \widetilde{Y}_{k_0}(t) + \mu_{k_0}^{1/2} c_{k_0} W_{k_0}(t) - \mu_{\bar{k}_0}^{1/2} c_{\bar{k}_0} W_{\bar{k}_0}(\rho_{\bar{k}_0} t) + \widetilde{Y}_{\bar{k}_0}(t) \right] \\
= \; & \frac{1}{\mu_{\bar{k}_0}} \left[- \sum_{i=k_0+1}^{\bar{k}_0 - 1} \widetilde{Q}_i(t) + \mu_{k_0}^{1/2} c_{k_0} W_{k_0}(t) - \mu_{\bar{k}_0}^{1/2} c_{\bar{k}_0} W_{\bar{k}_0}(\rho_{\bar{k}_0} t) + \widetilde{Y}_{\bar{k}_0}(t) \right] \\
& - \frac{1}{\mu_{\bar{k}_0}} \sup_{0 \le u \le t} \left\{ \widetilde{Y}_{\underline{k}_0}(u) + \sum_{i=\underline{k}_0+1}^{k_0 - 1} \widetilde{Q}_i(u) - \mu_{\underline{k}_0}^{1/2} c_{\underline{k}_0} W_{\underline{k}_0}(u) + \mu_{k_0}^{1/2} c_{k_0} W_{k_0}(u) \right\}. \quad (4.32)
\end{aligned}
$$

Next, similar to (4.31), we have $\widetilde{Z}^*_{\bar{k}_0}$

$$
\begin{aligned}
& = \limsup_{L \to \infty} \frac{1}{\mu_{\bar{k}_0} \varphi(L)} \left|\left| \widetilde{Z}_{\bar{k}_0} - \bar{Z}_{\bar{k}_0} \right|\right|_L \\
& = \limsup_{n \to \infty} \frac{1}{\mu_{\bar{k}_0} \varphi(L)} \left\{ \sup_{0 \le t \le L} \left| \begin{array}{l} \sum_{i=k_0+1}^{\bar{k}_0 - 1} \widetilde{Q}_i(t) - \mu_{k_0}^{1/2} c_{k_0} W_{k_0}(t) + \mu_{\bar{k}_0}^{1/2} c_{\bar{k}_0} W_{\bar{k}_0}(\rho_{\bar{k}_0} t) + \widetilde{Y}_{\bar{k}_0}(t) \\ + \sup_{0 \le u \le t} \left\{ \widetilde{Y}_{\underline{k}_0}(u) + \sum_{i=\underline{k}_0+1}^{k_0 - 1} \widetilde{Q}_i(u) - \mu_{\underline{k}_0}^{1/2} c_{\underline{k}_0} W_{\underline{k}_0}(u) + \mu_{k_0}^{1/2} c_{k_0} W_{k_0}(u) \right\} \end{array} \right| \right\} \\
& = \limsup_{n \to \infty} \frac{1}{\mu_{\bar{k}_0} \varphi(L)} \left\{ \sup_{0 \le t \le L} \left| \mu_{\bar{k}_0}^{1/2} c_{\bar{k}_0} W_{\bar{k}_0}(\rho_{\bar{k}_0} t) + \sup_{0 \le u \le t} \left\{ -\mu_{\underline{k}_0}^{1/2} c_{\underline{k}_0} W_{\underline{k}_0}(u) - \mu_{k_0}^{1/2} c_{k_0} W_{k_0}(t-u) \right\} \right| \right\} \\
& = \frac{\sqrt{\mu_{k_0} (\widehat{C}^2_{\underline{k}_0, k_0} + c^2_{\bar{k}_0})}}{\mu_{\bar{k}_0}}, \quad \text{w.p.1},
\end{aligned}
$$

where the last equality holds similarly to (4.27) with $\mu_{\bar{k}_0}^{1/2} c_{\bar{k}_0} W_{\bar{k}_0}(\rho_{\bar{k}_0} t)$ having variance parameter $\mu_{k_0} c^2_{\bar{k}_0}$.

To treat the LIL of the departure process, by Lemma 2.1 and (4.2), we have

$$
\begin{aligned}
\widetilde{D}_{\bar{k}_0}(t) - \bar{D}_{\bar{k}_0}(t) & = \mu_{\bar{k}_0} \left[\widetilde{T}_{\bar{k}_0}(t) - \bar{T}_{\bar{k}_0}(t) \right] + \mu_{\bar{k}_0}^{1/2} c_{\bar{k}_0} W_{\bar{k}_0}(\bar{T}_{\bar{k}_0}(t)) \\
& = -\mu_{\bar{k}_0} \left[\widetilde{I}_{\bar{k}_0}(t) - \bar{I}_{\bar{k}_0}(t) \right] + \mu_{\bar{k}_0}^{1/2} c_{\bar{k}_0} W_{\bar{k}_0}(t).
\end{aligned}
$$

This, together with the LIL $I^*_{\bar{k}_0} = 0$ (Lemma 4.4), implies that, w.p.1,

$$
\widetilde{D}^*_{\bar{k}_0} = \limsup_{L \to \infty} \frac{1}{\varphi(L)} \left|\left| -\mu_{\bar{k}_0} \widetilde{I}_{\bar{k}_0} + \mu_{\bar{k}_0}^{1/2} c_{\bar{k}_0} W_{\bar{k}_0} \right|\right|_L = \limsup_{L \to \infty} \frac{1}{\varphi(L)} \left|\left| \mu_{\bar{k}_0}^{1/2} c_{\bar{k}_0} W_{\bar{k}_0} \right|\right|_L = \mu_{\bar{k}_0}^{1/2} c_{\bar{k}_0}.
$$

Invoking Lemma 4.2, we have completed the proof for all LIL results in (3.13).

Stages succeeding $\bar{k}_0$ (Case 3.c). Note that the departure process from stage $\bar{k}_0$ corresponds to the arrival process for the downstream echelon which consists of stages $\bar{k}_0+1,\ldots,K$. Because stage $\bar{k}_0$ is overloaded w.p.1., the server is always busy so that its service completion process is a renewal process with parameters $\mu_{\bar{k}_0}$ and $c_{\bar{k}_0}$, the downstream segment of the system is equivalent to a tandem model having $K-\bar{k}_0$ stages fed by an "external" renewal arrival process with parameters $\mu_{\bar{k}_0}$ and $c_{\bar{k}_0}$. Hence, the LIL results of stages $\bar{k}_0+1,\ldots,K$ are characterized by Case 3 in Theorem 3.1. □

4.4 Multiple critically loaded stages: Technical challenges

In Theorem 3.2, we assume that there is only one critically loaded stage in the tandem queue. Next, we discuss the challenges in treating a tandem queue having more than one critically loaded stages. Suppose two stages $k_0' < k_0$ are critically loaded, i.e., $\rho_{k_0} = \rho_{k_0'} = 1$. For simplicity, assume $\rho_k < 1$ for all $k \neq k_0, k_0'$. The main challenge lies in the difficulty in analyzing the multi-fold composition of the function $\Phi(\cdot)$ (i.e., $\Phi(\Phi(\cdot))$ and $\Phi(\Phi(\Phi(\cdot)))$), because one critically loaded stage will give rise to one-fold $\Phi(\cdot)$. We can use the SA method to find the LIL $\widetilde{Q}^*_{k_0'}$ when k_0' is the first critically loaded stage. However, it is the critically loaded stage k_0' that makes the downstream analysis more challenging because $\widetilde{Q}_{k_0'}(t) = \Phi(\widetilde{X}_{k_0'})(t)$ must be transformed into Brownian motions in order to facilitate the subsequent analysis. In other words, the established LIL $\widetilde{Q}^*_{k_0'}$ does not directly help establish LILs for the performance measures of the subsequent stages. To put this into perspective, we discuss the LIL of the queue length of stage k_0: $\widetilde{Q}^*_{k_0}$.

For stage k_0' and k_0, following (4.14), we have

$$\widetilde{X}_{k_0'}(t) = -\sum_{i=1}^{k_0'-1} \widetilde{Q}_i(t) + \mu_0^{1/2} c_0 W_0(t) - \mu_{k_0'}^{1/2} c_{k_0'} W_{k_0'}(t), \tag{4.33}$$

$$\begin{aligned}
\widetilde{X}_{k_0}(t) &= -\sum_{i=1, i\neq k_0'}^{k_0-1} \widetilde{Q}_i(t) - \widetilde{Q}_{k_0'}(t) + \mu_0^{1/2} c_0 W_0(t) - \mu_{k_0}^{1/2} c_{k_0} W_{k_0}(t) \\
&= -\sum_{i=1, i\neq k_0'}^{k_0-1} \widetilde{Q}_i(t) - \Phi(\widetilde{X}_{k_0'})(t) + \mu_0^{1/2} c_0 W_0(t) - \mu_{k_0}^{1/2} c_{k_0} W_{k_0}(t). \qquad (4.34)
\end{aligned}$$

Notice that $k_0' < k_0$ and $\rho_{k_0} = \rho_{k_0'} = 1$ and $\rho_k < 1$ for all $k \neq k_0, k_0'$, it follows from (3.11) that $Q^*_{k_0'} = \sqrt{\mu_0} C_{0,k_0'}$ w.p.1. Since $\widetilde{Q}_{k_0}(t) = \Phi(\widetilde{X}_{k_0})(t)$ and $\widetilde{X}_{k_0}(t)$ is a function of

$\widetilde{Q}_{k_0'}(t) = \Phi(\widetilde{X}_{k_0'})(t)$, we have $\widetilde{Q}^*_{k_0}$

$$
\begin{aligned}
&= \limsup_{L\to\infty} \frac{1}{\varphi(L)} \left\{ \sup_{0\le t\le L} \Phi(\widetilde{X}_{k_0})(t) \right\} \\
&= \limsup_{L\to\infty} \frac{1}{\varphi(L)} \left\{ \sup_{0\le t\le L} \left\{ \widetilde{X}_{k_0}(t) + \sup_{0\le s\le t} \left[-\widetilde{X}_{k_0}(s) \right] \right\} \right\} \\
&= \limsup_{L\to\infty} \frac{1}{\varphi(L)} \left\{ \sup_{0\le t\le L} \left\{ \begin{array}{c} -\Phi(\widetilde{X}_{k_0'})(t) + \mu_0^{1/2} c_0 W_0(t) - \mu_{k_0}^{1/2} c_{k_0} W_{k_0}(t) \\ + \sup\limits_{0\le s\le t} \left[\Phi(\widetilde{X}_{k_0'})(s) - \mu_0^{1/2} c_0 W_0(s) + \mu_{k_0}^{1/2} c_{k_0} W_{k_0}(s) \right] \end{array} \right\} \right\} \\
&= \limsup_{L\to\infty} \frac{1}{\varphi(L)} \left\{ \sup_{0\le t\le L} \left\{ -\Phi(\widetilde{X}_{k_0'})(t) + \sup_{0\le s\le t} \left[\begin{array}{c} \Phi(\widetilde{X}_{k_0'})(s) + \\ \mu_0^{1/2} c_0 W_0(t-s) - \mu_{k_0}^{1/2} c_{k_0} W_{k_0}(t-s) \end{array} \right] \right\} \right\} \\
&= \limsup_{L\to\infty} \frac{1}{\varphi(L)} \left\{ \sup_{0\le t\le L} \left\{ \begin{array}{c} -\left[\widetilde{X}_{k_0'}(t) + \sup\limits_{0\le u\le t} [-\widetilde{X}_{k_0'}(u)] \right] \\ + \sup\limits_{0\le s\le t} \left[\begin{array}{c} \widetilde{X}_{k_0'}(s) + \sup\limits_{0\le v\le s} [-\widetilde{X}_{k_0'}(v)] \\ + \mu_0^{1/2} c_0 W_0(t-s) - \mu_{k_0}^{1/2} c_{k_0} W_{k_0}(t-s) \end{array} \right] \end{array} \right\} \right\} \\
&= \limsup_{L\to\infty} \frac{1}{\varphi(L)} \left\{ \sup_{0\le t\le L} \left\{ \begin{array}{c} -\sup\limits_{0\le u\le t} \left[-\widetilde{X}_{k_0'}(u) \right] + \\ \sup\limits_{0\le s\le t} \left[\begin{array}{c} \sup\limits_{0\le v\le s} \left[-\widetilde{X}_{k_0'}(v) \right] \\ + \mu_{k_0'}^{1/2} c_{k_0'} W_{k_0'}(t-s) - \mu_{k_0}^{1/2} c_{k_0} W_{k_0}(t-s) \end{array} \right] \end{array} \right\} \right\} \\
&= \limsup_{L\to\infty} \frac{1}{\varphi(L)} \left\{ \sup_{0\le t\le L} \left\{ \begin{array}{c} -\sup\limits_{0\le u\le t} \left[-\mu_0^{1/2} c_0 W_0(u) + \mu_{k_0'}^{1/2} c_{k_0'} W_{k_0'}(u) \right] \\ + \sup\limits_{0\le s\le t} \left[\begin{array}{c} \sup\limits_{0\le v\le s} \left[-\mu_0^{1/2} c_0 W_0(v) + \mu_{k_0'}^{1/2} c_{k_0'} W_{k_0'}(v) \right] \\ + \mu_{k_0'}^{1/2} c_{k_0'} W_{k_0'}(t-s) - \mu_{k_0}^{1/2} c_{k_0} W_{k_0}(t-s) \end{array} \right] \end{array} \right\} \right\}
\end{aligned}
\tag{4.35}
$$

where $\widetilde{X}_{k_0}(t)$ is in (4.33), $\widetilde{X}_{k_0}(t)$ is in (4.34), the third and the seventh equalities hold because $\widetilde{Q}^*_k = 0$ w.p.1 for $k < k_0$ and $k \neq k_0'$. In the last equality in (4.35), it says that it is the decomposition of the multi-fold compositions of "sup" functions of Brownian motions that complicates the analysis.

5 Conclusion

In this paper, we establish the LIL for a multi-stage tandem queueing system. Our LIL results, called explicit LILs, give explicit limits for the asymptotic variabilities of the system, which refine the implicit LIL results in [8]. Using the SAs and fluid limits, we develop the LIL limits for the following performance measures: the queue length, workload, busy time, idle time and departure processes. In our main results, we separately treat the cases with and without critically loaded stages (i.e., $\rho_k = 1$). In our first main result (i.e., Theorem 3.1), we show that, in the LIL sense, an underloaded (overloaded) stage is equivalent to an underloaded (overloaded) $G/G/1$ queue having a renewal arrival process with parameters corresponding to those of the last preceding stage. On the other hand, a tandem system having critically loaded stages are much more difficult to analyze. In Theorem 3.2, we report the LILs of a tandem queue having one critically loaded stage.

In addition, our LIL results reveal clear implications on how the variabilities received from upstream stages can be propagated to the downsteam echelons in the tandem queue model: (i) An underloaded stage simply transfers the entirety of received upstream variability to the downstream stages; its own service-time variability makes no impact on any succeeding echelons. (ii) An overloaded stage plays an overriding role by blocking the variability received from upstream stages; it resets the propagation process by feeding its successive stages with its own service-time variability alone. (iii) A critically loaded stage is a middle ground between an underloaded and overloaded stage: On the one hand, it inherits the variability received from upstream stages; on the other hand, it modifies the variability with its own service-time variability and then passes it forward to the downstream stages. One future direction is to study the case having multiple stages that are critically loaded. As indicated in Section 4.4, this may requires some new methodologies. Another future research is to extend our SA-based LIL analysis to the queueing networks having more general topological structure. This requires the analysis of more complex SAs (Theorem 7.19 in [8]) that consist of multi-dimensional BMs.

Acknowledgements: This work is supported by the National Nature Science Foundation of China under Grants No. 11871116 and 11971074, and by the Fundamental Research Funds for the Central Universities (No. 2023ZCJH02).

References

[1] Andradóttir S., Ayhan H. (2005). Throughput maximization for tandem lines with two stations and flexible servers. Operations Research, 53(3): 516–531.

[2] Blanc J.P.C., Iasnogorodski R., Nain P. (1988). Analysis of the $M/Gi/1 \to /M/1$ queueing model. Queueing System, 3(2): 129–156.

[3] Boxma O.J., Daduna, H. (2014). The cyclic queue and the tandem queue. Queueing Systems, 77(3): 275–295.

[4] Boxma O.J. (1984). $M/G/\infty$ tandem queues. Stochastic Processes and their Applications, 18(1): 153–164.

[5] Cao P., He S., Huang S., Liu Y. (2021). To pool or not to pool: Queueing design for large-scale service systems. Operations Research, 69(6): 1866–1885.

[6] Chen H., Mandelbaum A. (1994). Hierarchical modeling of stochastic network, part II: strong approximations. Stochastic Modeling and Analysis of Manufacturing Systems, Yao, D.D. (Ed.), 107–131.

[7] Chen H., Shen X. (2000). Strong approximations for multiclass feedforward queueing networks. Annals of Applied Probability, 10(3): 828–876.

[8] Chen H., Yao D.D. (2001). Fundamentals of Queueing Networks, Springer-Verlag, New York.

[9] Cheng D.W., Righter R. (1995). On the order of tandem queues. Queueing Systems, 21: 143–160.

[10] Csörgő M., Révész P. (1981). Strong Approximations in Probability and Statistics, Academic Press, New York.

[11] Csörgő M., Horváth L. (1993). Weighted Approximations in Probability and Statistics, Wiley, New York.

[12] Dai J.G. (1995). On the positive Harris recurrence for multiclass queueing networks: a unified approach via fluid limit models. Annals of Applied Probability, 5(1): 49–77.

[13] Glynn P.W. (1998). Strong approximations in queueing theory. In Asymptotic methods in probability and statistics, edited by Szyszkowicz B, North-Holland, 135–150.

[14] Glynn P.W., Whitt W. (1991). Departures from many queues in series. Annals of Applied Probability, 1(4): 546–572.

[15] Guo Y., Liu Y. (2015). A law of iterated logarithm for multiclass queues with preemptive priority service discipline. Queueing Systems, 79(3): 251–291.

[16] Harrison, M.J. (1978). The diffusion approximation for tandem queues in heavy traffic. Advances in Applied Probability, 10(4): 886–905.

[17] He B., Liu Y., Whitt W. (2016). Staffing a service system with non-poisson nonstationary arrivals. Probability in the Engineering and Informational Sciences, 30: 593–621.

[18] Horváth L. (1990). Strong approximations of open queueing network. Mathematics of Operations Research, 17: 487–508.

[19] Iglehart G.L. (1971). Multiple channel queues in heavy traffic: IV. Law of the iterated logarithm. Z.Wahrscheinlichkeitstheorie verw. Geb., 17: 168–180.

[20] Le L., Hossain E. (2008). Tandem queue models with applications to QoS routing in multihop wireless networks. IEEE transactions on mobile computing, 8: 1025–1040.

[21] Lee Y., Zipkin P. (1992). Tandem queues with planned inventories. Operations Research, 40(5): 936–947.

[22] Lévy P. Théorie de l'addition des variables aléatories. Gauthier-Villars, Paris, 1937.

[23] Lévy P. Procesus stochastique et mouvement Brownien. Gauthier-Villars, Paris, 1948.

[24] Liu Y. (2018). Staffing to stabilize the tail probability of delay in service systems with time-varying demand. Operations Research, 66(2): 514–534.

[25] Liu Y., Sun X., Hovey K. (2022). Scheduling to differentiate service in a multiclass service system. Operations Research, 70(1): 527–544.

[26] Liu Y., Whitt W. (2012). Stabilizing customer abandonment in many-server queues with time-varying arrivals. Operations Research, 60(6): 1551–1564.

[27] Liu Y., Whitt W. (2014). Stabilizing performance in networks of queues with time-varying arrival rates. Probability in the Engineering and Informational Sciences, 28(4): 419–449.

[28] Liu Y., Whitt W. (2017). Stabilizing performance in a service system with time-varying arrivals and customer feedback. European Journal of Operational Research, 256(2): 473–486.

[29] Minkevičius S., Steišūnas S. (2003). A law of the iterated logarithm for global values of waiting time in multiphase queues. Statistics and Probability Letters, 61(4): 359–371.

[30] Minkevičius S. (2014). On the law of the iterated logarithm in multiserver open queueing networks. Stochastics, 86(1): 46–59.

[31] Mörters P., Peres Y. Brownian Motion, Cambridge University Press, 2010.

[32] Sakalauskas L.L., Minkevičius, S. (2000). On the law of the iterated logarithm in open queueing networks. European Journal of Operational Research, 120(3): 632–640.

[33] Prabhu N.U. (1967). Transient behaviour of a tandem queue. Management Science 13(9): 631–639.

[34] Reed J., Yechiali U. (2013). Queues in tandem with customer deadlines and retrials. Queueing Systems, 73(1): 1–34.

[35] Scheller-Wolf A., Sigman K. (1998). Moments in tandem queues. Operations Research, 46(3): 378–380.

[36] Song J.S., Xiao L., Zhang H.Q., Zipkin P. (2021). Smart policies for multisource inventory systems and general tandem queues with order tracking and expediting. Operations Research, 70(4): 1953–2596.

[37] Sun X., Liu Y. (2021). Staffing many-server queues with autoregressive inputs. Naval Research Logistics, 68(3): 312–326.

[38] Zajic T. (1998). Rough asymptotics for tandem non-homogeneous $M/G/\infty$ queues via Poissonized empirical processes. Queueing Systems, 29(2): 161–174.

[39] Zhang H., Hsu G. (1992). Strong approximations for priority queues: Head-of-the-line-first discipline. Queueing Systems, 10: 213–233.

[40] Zhang H., Hsu G., Wang R. (1990). Strong approximitions for multiple channel queues in heavy traffic. Journal of Applied Probability, 28: 658–670.

[41] Zhang H. (1997). Strong approximations for irreducible closed queueing networks. Advances in Applied Probability, 29: 498–522.

APPENDIX

A Proofs of Lemma 4.2

Because all proofs are similar, we only give a proof for the queue length process. First, we note that $T^{1/r} = o(\varphi(T))$ for $r > 2$, and

$$\left| \frac{\left\|\widetilde{Q}_k - \bar{Q}_k\right\|_L}{\varphi(L)} - \frac{\left\|Q_k - \widetilde{Q}_k\right\|_L}{\varphi(L)} \right| \leq \frac{\left\|Q_k - \bar{Q}_k\right\|_L}{\varphi(L)} \leq \frac{\left\|\widetilde{Q}_k - \bar{Q}_k\right\|_L}{\varphi(L)} + \frac{\left\|Q_k - \widetilde{Q}_k\right\|_L}{\varphi(L)}.$$

Lemma 4.1 implies that $\lim\limits_{L\to\infty} \left\|Q_k - \widetilde{Q}_k\right\|_L / \varphi(L) = 0$ w.p.1. As a result, $\left\|Q_k - \bar{Q}_k\right\|_L / \varphi(L)$ and $\left\|Q_k - \bar{Q}_k\right\|_L / \varphi(L)$ have the same limits if they exist as $L \to \infty$. In other words, $Q_k^* = \widetilde{Q}_k^*$ for $k = 1, 2, \ldots, K$. The proofs for the other performance metrics follow similar arguments.

Hence, $\mathcal{X}_k^* = \widetilde{\mathcal{X}}_k^*$ w.p.1. $\square$

B Proof of Lemma 4.3

We prove Lemma 4.3 based on Lemmas B.1 and B.2, which are presented in advance and are proved after Lemma 4.3.

Lemma B.1. *If $\rho_k < 1$ for all $k = 1, 2, \ldots, K$, then*

$$\mathsf{P}\left\{\sup_{0\leq t\leq T} \widetilde{Q}_k(t) \geq z\right\} \leq k \exp\left\{-\gamma_k z\right\}, \tag{B.1}$$

where

$$\gamma_1 = \frac{\mu_1 - \mu_0}{\mu_0(c_0^2 + c_1^2)}, \quad \gamma_k = \min\left\{\frac{\mu_k - \mu_0}{\mu_0(c_0^2 + c_k^2)}, \frac{\gamma_1}{2(k-1)}, \frac{\gamma_2}{2(k-1)}, \ldots, \frac{\gamma_{k-1}}{2(k-1)}\right\} \tag{B.2}$$

for all $k = 2, 3, \ldots, K$.

Lemma B.2. *If there exist k_0 and l_0 such that $\rho_{k_0} \geq 1$, $\rho_{k_0+l} < 1$ for all $l = 1, 2, \ldots, l_0$ and $\rho_{k_0+l_0+1} \geq 1$ and $k_0 + l_0 + 1 \leq K$, then for all $k \geq k_0$ such that $\rho_k < 1$, we have*

$$\mathsf{P}\left\{\sup_{0\leq t\leq T} \widetilde{Q}_{k_0+l}(t) \geq z\right\} \leq l \exp\left\{-\gamma_l' z\right\}, \tag{B.3}$$

where

$$\gamma_1' = \frac{\mu_{k_0+1} - \mu_{k_0}}{\mu_{k_0}(c_{k_0}^2 + c_{k_0+1}^2)}, \tag{B.4}$$

$$\gamma_l' = \min\left\{\frac{\mu_{k_0+1} - \mu_{k_0}}{\mu_{k_0}(c_{k_0}^2 + c_{k_0+1}^2)}, \frac{\gamma_1'}{2(l-1)}, \frac{\gamma_2'}{2(l-1)}, \ldots, \frac{\gamma_{l-1}'}{2(l-1)}\right\} \tag{B.5}$$

for all $l = 2, 3, \ldots, l_0$.

Next, we prove Lemma 4.3.

Proof of Lemma 4.3 Given k, we prove the result in two cases: (i) $\rho_i < 1$ for all $i = 1, 2, \ldots, k$. By Lemma B.1, letting $z = 2\log T/\gamma_k$ in (B.1) yields

$$\mathsf{P}\left\{\sup_{0\le t\le T} \widetilde{Q}_k(t)/\log T \ge \frac{2}{\gamma_k}\right\} = k\frac{1}{T^2}.$$

This and Borel-Cantelli Lemma imply that $\sup_{0\le t\le T} \widetilde{Q}_k(t) = O(\log T)$. (ii) there exists a $k' < k$ such that $\rho_{k'} \ge 1$ and $\rho_i < 1$ for all $i = k'+1, k'+2, \ldots, k$. For this case, by Lemma B.2, the proof is similar with case (i) and is omitted.

As a result, $\widetilde{Q}_k^* = 0$ w.p.1 when $\rho_k < 1$. So, $\widetilde{Z}_k^* = 0$ w.p.1 under $\rho_k < 1$ because, by (2.10) and (4.2), $\mu_k \widetilde{Z}_k(t) = \widetilde{Q}_k(t)$ and $\mu_k \bar{Z}_k(t) = \bar{Q}_k(t)$ for k: $\rho_k < 1$. □

B.1 Proof of Lemma B.1

Since $\rho_k < 1$ for all $k = 1, 2, \ldots, K$, we have $\rho_k = \mu_0/\mu_k$ and $\mu_0 < \mu_k$ for all $k = 1, 2, \ldots, K$.

At first, we prove (B.1) holds with $k = 1$. By (4.2),

$$\widetilde{X}_1(t) \quad = \quad (\mu_0 - \mu_1)t + \mu_0^{1/2} c_0 W_0(t) - \mu_1^{1/2} c_1 W_1(\rho_1 t)$$

is a reflected BM with negative drift $(\mu_0 - \mu_1)$ and variance parameter $\mu_0(c_0^2 + c_1^2)$, then it follows from Theorem 6.3 in [8] that

$$\mathsf{P}\left\{\sup_{0\le t\le T} \widetilde{Q}_1(t) \ge z\right\} \le \exp\left\{-\frac{2(\mu_1 - \mu_0)}{\mu_0(c_0^2 + c_1^2)} z\right\} = \exp\left\{-\gamma_1 z\right\}, \tag{B.6}$$

that is, (B.1) holds with $k = 1$.

In order to prove (B.1) holds with $k = 2, 3, \ldots, K$, we first consider the following probability, for all $k = 2, 3, \ldots, K$, by (4.2),

$$\begin{aligned}
\mathsf{P}\left\{\sup_{0\le t\le T} \widetilde{Q}_k(t) \ge z\right\} &= \mathsf{P}\left\{\sup_{0\le t\le T}\left[\widetilde{X}_k(t) + \widetilde{Y}_k(t)\right] \ge z\right\} \\
&= \mathsf{P}\left\{\sup_{0\le t\le T}\left[\widetilde{X}_k(t) + \sup_{0\le s\le t}\left[-\widetilde{X}_k(s)\right]\right] \ge z\right\} \\
&= \mathsf{P}\left\{\sup_{0\le t\le T}\sup_{0\le s\le t}\left[\widetilde{X}_k(t) - \widetilde{X}_k(s)\right] \ge z\right\}.
\end{aligned} \tag{B.7}$$

Since, for all $0 \le s \le t$,

$$\begin{aligned}
\widetilde{X}_k(t) - \widetilde{X}_k(s) &= \mu_{k-1}\left[\widetilde{T}_{k-1}(t) - \widetilde{T}_{k-1}(s)\right] - \mu_k(t-s) \\
&\quad + \mu_{k-1}^{1/2} c_{k-1}\left[W_{k-1}(\bar{T}_{k-1}(t)) - W_{k-1}(\bar{T}_{k-1}(s))\right] \\
&\quad - \mu_k^{1/2} c_k \left[W_k(\bar{T}_k(t)) - W_k(\bar{T}_k(s))\right],
\end{aligned} \tag{B.8}$$

where

$$\begin{aligned}\mu_{k-1}\left[\widetilde{T}_{k-1}(t)-\widetilde{T}_{k-1}(s)\right] &= \mu_{k-1}(t-s)-\left[\widetilde{Y}_{k-1}(t)-\widetilde{Y}_{k-1}(s)\right]\\ &= \mu_{k-1}(t-s)-\left[\widetilde{Q}_{k-1}(t)-\widetilde{Q}_{k-1}(s)\right]+\left[\widetilde{X}_{k-1}(t)-\widetilde{X}_{k-1}(s)\right],\end{aligned}$$

we have

$$\begin{aligned}\widetilde{X}_k(t)-\widetilde{X}_k(s) &= \mu_{k-1}(t-s)-\mu_k(t-s)-\left[\widetilde{Q}_{k-1}(t)-\widetilde{Q}_{k-1}(s)\right]\\ &\quad+\mu_{k-1}^{1/2}c_{k-1}\left[W_{k-1}(\bar{T}_{k-1}(t))-W_{k-1}(\bar{T}_{k-1}(s))\right]\\ &\quad-\mu_k^{1/2}c_k\left[W_k(\bar{T}_k(t))-W_k(\bar{T}_k(s))\right]\\ &\quad+\left[\widetilde{X}_{k-1}(t)-\widetilde{X}_{k-1}(s)\right].\end{aligned}\tag{B.9}$$

So, by iteration, we have, for all $0 \le s \le t$,

$$\begin{aligned}\widetilde{X}_k(t)-\widetilde{X}_k(s) &= \mu_1(t-s)-\mu_k(t-s)-\sum_{l=1}^{k-1}\left[\widetilde{Q}_l(t)-\widetilde{Q}_l(s)\right]+\left[\widetilde{X}_1(t)-\widetilde{X}_1(s)\right]\\ &\quad+\mu_1^{1/2}c_1\left[W_1(\widetilde{T}_1(t))-W_1(\widetilde{T}_1(s))\right]-\mu_k^{1/2}c_k\left[W_k(\bar{T}_k(t))-W_k(\bar{T}_k(s))\right]\\ &= (\mu_0-\mu_k)(t-s)-\sum_{l=1}^{k-1}\left[\widetilde{Q}_l(t)-\widetilde{Q}_l(s)\right]\\ &\quad+\mu_0^{1/2}c_0\left[W_0(t)-W_0(s)\right]-\mu_k^{1/2}c_k\left[W_k(\bar{T}_k(t))-W_k(\bar{T}_k(s))\right]\\ &\overset{\mathrm{d}}{=} (\mu_0-\mu_k)(t-s)-\sum_{l=1}^{k-1}\left[\widetilde{Q}_l(t)-\widetilde{Q}_l(s)\right]\\ &\quad+\mu_0^{1/2}c_0\left[W_0(t)-W_0(s)\right]-\mu_0^{1/2}c_k\left[W_k(t)-W_k(s)\right]\\ &\le (\mu_0-\mu_k)(t-s)+\sum_{l=1}^{k-1}\widetilde{Q}_l(s)+\mu_0^{1/2}c_0W_0(t-s)-\mu_0^{1/2}c_kW_k(t-s).\end{aligned}$$

Following (B.7), we have

$$
\begin{aligned}
&\mathsf{P}\left\{\sup_{0\le t\le T}\widetilde{Q}_k(t)\ge z\right\}\\
\le\ &\mathsf{P}\left\{\sup_{0\le t\le T}\sup_{0\le s\le t}\left[(\mu_0-\mu_k)(t-s)+\mu_0^{1/2}c_0W_0(t-s)-\mu_0^{1/2}c_kW_k(t-s)\right]\ge\frac{z}{2}\right\}\\
&+\mathsf{P}\left\{\sup_{0\le t\le T}\sup_{0\le s\le t}\sum_{l=1}^{k-1}\widetilde{Q}_l(s)\ge\frac{z}{2}\right\}\\
=\ &\mathsf{P}\left\{\sup_{0\le t\le T}\left[(\mu_0-\mu_k)t+\mu_0^{1/2}c_0W_0(t)-\mu_0^{1/2}c_kW_k(t)\right]\ge\frac{z}{2}\right\} \qquad\qquad (B.10)\\
&+\mathsf{P}\left\{\sup_{0\le t\le T}\sum_{l=1}^{k-1}\widetilde{Q}_l(t)\ge\frac{z}{2}\right\}\\
\le\ &\exp\left\{-\frac{2(\mu_k-\mu_0)}{\mu_0(c_0^2+c_k^2)}\frac{z}{2}\right\}+\mathsf{P}\left\{\sum_{l=1}^{k-1}\sup_{0\le t\le T}\widetilde{Q}_l(t)\ge\frac{z}{2}\right\}\\
\le\ &\exp\left\{-\frac{2(\mu_k-\mu_0)}{\mu_0(c_0^2+c_k^2)}\frac{z}{2}\right\}+\sum_{l=1}^{k-1}\mathsf{P}\left\{\sup_{0\le t\le T}\widetilde{Q}_l(t)\ge\frac{z}{2(k-1)}\right\},
\end{aligned}
$$

where the second inequality holds because $(\mu_0-\mu_k)t+\mu_0^{1/2}c_0W_0(t)-\mu_0^{1/2}c_kW_k(t)$ is a BM with negative drift $(\mu_0-\mu_k)$ and variance parameter $\mu_0(c_0^2+c_k^2)$.

Next, we inductively prove that (B.1) holds for $k=2,3,\ldots,K$. We first prove (B.1) for $k=2$. By (B.10),

$$
\begin{aligned}
\mathsf{P}\left\{\sup_{0\le t\le T}\widetilde{Q}_2(t)\ge z\right\} &\le \exp\left\{-\frac{2(\mu_2-\mu_0)}{\mu_0(c_0^2+c_2^2)}\frac{z}{2}\right\}+\mathsf{P}\left\{\sup_{0\le t\le T}\widetilde{Q}_1(t)\ge\frac{z}{2}\right\}\\
&\le \exp\left\{-\frac{2(\mu_2-\mu_0)}{\mu_0(c_0^2+c_2^2)}\frac{z}{2}\right\}+\exp\left\{-\frac{\gamma_1}{2}z\right\}\\
&\le 2\exp\left\{-\gamma_2 t\right\},
\end{aligned}
$$

where γ_2 is defined in (B.2).

Suppose that (B.1) holds for all $k=2,3,\ldots,j-1$, we next show that (B.1) holds for $k=j\le K$. By the inductive hypothesis and (B.10),

$$
\begin{aligned}
\mathsf{P}\left\{\sup_{0\le t\le T}\widetilde{Q}_j(t)\ge z\right\} &\le \exp\left\{-\frac{2(\mu_j-\mu_0)}{\mu_0(c_0^2+c_j^2)}\frac{z}{2}\right\}+\sum_{l=1}^{j-1}\mathsf{P}\left\{\sup_{0\le t\le T}\widetilde{Q}_l(t)\ge\frac{z}{2(j-1)}\right\}\\
&\le \exp\left\{-\frac{2(\mu_j-\mu_0)}{\mu_0(c_0^2+c_j^2)}\frac{z}{2}\right\}+\sum_{l=1}^{j-1}\exp\left\{-\frac{\gamma_l}{2(j-1)}z\right\}\\
&\le j\exp\left\{-\gamma_j t\right\},
\end{aligned}
$$

where γ_j is defined in (B.2). So, (B.1) inductively holds for $k=2,3,\ldots,K$. □

B.2 Proof of Lemma B.2

We inductively prove this result. At first, we prove that (B.3) holds for $k = k_0 + 1$. By (B.8),

$$\begin{aligned}\widetilde{X}_{k_0+1}(t) - \widetilde{X}_{k_0+1}(s) &= (\mu_{k_0} - \mu_{k_0+1})(t-s) - \left[\widetilde{Y}_{k_0}(t) - \widetilde{Y}_{k_0}(s)\right] \\ &\quad + \mu_{k_0}^{1/2} c_{k_0} \left[W_{k_0}(\bar{T}_{k_0}(t)) - W_{k_0}(\bar{T}_{k_0}(s))\right] \\ &\quad - \mu_{k_0+1}^{1/2} c_{k_0+1} \left[W_{k_0+1}(\bar{T}_{k_0+1}(t)) - W_{k_0+1}(\bar{T}_{k_0+1}(s))\right].\end{aligned}$$

By (2.4), we have

$$\rho_{k_0} = \frac{\min\{\mu_0, \mu_1, \ldots, \mu_{k_0-1}\}}{\mu_{k_0}} \geq 1,$$

then $\min\{\mu_0, \mu_1, \ldots, \mu_{k_0-1}\} \geq \mu_{k_0}$. This, along with the fact that

$$\rho_{k_0+1} = \frac{\min\{\mu_0, \mu_1, \ldots, \mu_{k_0}\}}{\mu_{k_0+1}} = \frac{\mu_{k_0}}{\mu_{k_0+1}} < 1,$$

implies that $\mu_{k_0} < \mu_{k_0+1}$. By (2.12), we have, for all $0 \leq s \leq t$, $\widetilde{Y}_{k-1}(t) - \widetilde{Y}_{k-1}(s) \geq 0$ for all $k = 1, 2, \ldots, K$. Notice that $\bar{T}_{k_0}(t) = t$ and $\bar{T}_{k_0+1}(t) = \rho_{k_0+1} t$ for all $t \geq 0$, then

$$\begin{aligned}&\widetilde{X}_{k_0+1}(t) - \widetilde{X}_{k_0+1}(s) \\ \leq\ & (\mu_{k_0} - \mu_{k_0+1})(t-s) + \mu_{k_0}^{1/2} c_{k_0} \left[W_{k_0}(t) - W_{k_0}(s)\right] \\ & - \mu_{k_0+1}^{1/2} c_{k_0+1} \left[W_{k_0+1}(\rho_{k_0+1}(t)) - W_{k_0+1}(\rho_{k_0+1}(s))\right] \\ \stackrel{\text{d}}{=}\ & (\mu_{k_0} - \mu_{k_0+1})(t-s) + \mu_{k_0}^{1/2} c_{k_0} W_{k_0}(t-s) - \mu_{k_0}^{1/2} c_{k_0+1} W_{k_0+1}(t-s).\end{aligned}$$

As (B.10),

$$\begin{aligned}&\mathsf{P}\left\{\sup_{0\leq t\leq T} \widetilde{Q}_{k_0+1}(t) \geq z\right\} \\ \leq\ & \mathsf{P}\left\{\sup_{0\leq t\leq T} \left[(\mu_{k_0} - \mu_{k_0+1})t + \mu_{k_0}^{1/2} c_{k_0} W_{k_0}(t) - \mu_{k_0}^{1/2} c_{k_0+1} W_{k_0+1}(t)\right] \geq z\right\} \\ \leq\ & \exp\left\{-\frac{\mu_{k_0+1} - \mu_{k_0}}{\mu_{k_0}(c_{k_0}^2 + c_{k_0+1}^2)} z\right\} = \exp\left\{-\gamma_1' z\right\},\end{aligned}$$

where γ_1 is defined in (B.4), the second inequality is from Theorem 6.3 in [8] because $(\mu_{k_0} - \mu_{k_0+1})t + \mu_{k_0}^{1/2} c_{k_0} W_{k_0}(t) - \mu_{k_0}^{1/2} c_{k_0+1} W_{k_0+1}(t)$ is a BM with negative drift $(\mu_{k_0} - \mu_{k_0+1})$ and variance parameter $\mu_{k_0}(c_{k_0}^2 + c_{k_0+1}^2)$. So, (B.3) holds for $l = 1$.

Next, we prove (B.3) for $l = 2, 3, \ldots, l_0$. By (B.9), for $l = 2, 3, \ldots, l_0$,

$$\begin{aligned}
& \widetilde{X}_{k_0+l}(t) - \widetilde{X}_{k_0+l}(s) \\
= \; & \mu_{k_0+l-1}(t-s) - \mu_{k_0+l}(t-s) - \left[\widetilde{Q}_{k_0+l-1}(t) - \widetilde{Q}_{k_0+l-1}(s)\right] \\
& + \mu_{k_0+l-1}^{1/2} c_{k_0+l-1} \left[W_{k_0+l-1}(\bar{T}_{k_0+l-1}(t)) - W_{k_0+l-1}(\bar{T}_{k_0+l-1}(s))\right] \\
& - \mu_{k_0+l}^{1/2} c_{k_0+l} \left[W_{k_0+l}(\bar{T}_{k_0+l}(t)) - W_{k_0+l}(\bar{T}_{k_0+l}(s))\right] \\
& + \left[\widetilde{X}_{k_0+l-1}(t) - \widetilde{X}_{k_0+l-1}(s)\right] \\
= \; & \mu_{k_0+1}(t-s) - \mu_{k_0+l}(t-s) - \sum_{i=k_0+1}^{k_0+l-1} \left[\widetilde{Q}_i(t) - \widetilde{Q}_i(s)\right] \\
& + \mu_{k_0+1}^{1/2} c_{k_0+1} \left[W_{k_0+1}(\bar{T}_{k_0+1}(t)) - W_{k_0+1}(\bar{T}_{k_0+1}(s))\right] \\
& - \mu_{k_0+l}^{1/2} c_{k_0+l} \left[W_{k_0+l}(\bar{T}_{k_0+l}(t)) - W_{k_0+l}(\bar{T}_{k_0+l}(s))\right] \\
& + \left[\widetilde{X}_{k_0+1}(t) - \widetilde{X}_{k_0+1}(s)\right],
\end{aligned}$$

where, by (4.2),

$$\begin{aligned}
\widetilde{X}_{k_0+1}(t) - \widetilde{X}_{k_0+1}(s) \; = \; & \mu_{k_0} \left[\widetilde{T}_{k_0}(t) - \widetilde{T}_{k_0}(s)\right] - \mu_{k_0+1}(t-s) \\
& + \mu_{k_0}^{1/2} c_{k_0} \left[W_{k_0}(\bar{T}_{k_0}(t)) - W_{k_0}(\bar{T}_{k_0}(s))\right] \\
& - \mu_{k_0+1}^{1/2} c_{k_0+1} \left[W_{k_0+1}(\bar{T}_{k_0+1}(t)) - W_{k_0+1}(\bar{T}_{k_0+1}(s))\right] \\
= \; & \mu_{k_0}(t-s) - \mu_{k_0+1}(t-s) - \left[\widetilde{Y}_{k_0}(t) - \widetilde{Y}_{k_0}(s)\right] \\
& + \mu_{k_0}^{1/2} c_{k_0} \left[W_{k_0}(\bar{T}_{k_0}(t)) - W_{k_0}(\bar{T}_{k_0}(s))\right] \\
& - \mu_{k_0+1}^{1/2} c_{k_0+1} \left[W_{k_0+1}(\bar{T}_{k_0+1}(t)) - W_{k_0+1}(\bar{T}_{k_0+1}(s))\right].
\end{aligned}$$

Since $\widetilde{Y}_{k_0}(t) - \widetilde{Y}_{k_0}(s) \geq 0$ and $\widetilde{Q}_k(t) \geq 0$ for all k and $0 \leq s \leq t$, we have

$$\begin{aligned}
\widetilde{X}_{k_0+l}(t) - \widetilde{X}_{k_0+l}(s) \; \leq \; & (\mu_{k_0} - \mu_{k_0+l})(t-s) + \sum_{i=k_0+1}^{k_0+l-1} \widetilde{Q}_i(s) \\
& + \mu_{k_0}^{1/2} c_{k_0} \left[W_{k_0}(\bar{T}_{k_0}(t)) - W_{k_0}(\bar{T}_{k_0}(s))\right] \\
& - \mu_{k_0+l}^{1/2} c_{k_0+l} \left[W_{k_0+l}(\bar{T}_{k_0+l}(t)) - W_{k_0+l}(\bar{T}_{k_0+l}(s))\right] \\
\stackrel{\mathrm{d}}{=} \; & (\mu_{k_0} - \mu_{k_0+l})(t-s) + \sum_{i=k_0+1}^{k_0+l-1} \widetilde{Q}_i(s) \\
& + \mu_{k_0}^{1/2} c_{k_0} W_{k_0}(t-s) - \mu_{k_0}^{1/2} c_{k_0+l} W_{k_0+l}(t-s),
\end{aligned}$$

where the equality in distribution holds because $\bar{T}_{k_0}(t) = t$ and $\bar{T}_{k_0+l}(t) = \rho_{k_0+l} t =$

$(\mu_{k_0}/\mu_{k_0+l})t$ for all $t \geq 0$. So,

$$
\begin{aligned}
& \mathsf{P}\left\{\sup_{0\leq t\leq T} \widetilde{Q}_{k_0+l}(t) \geq z\right\} \\
\leq\ & \mathsf{P}\left\{\sup_{0\leq t\leq T}\sup_{0\leq s\leq t}\left[(\mu_{k_0}-\mu_{k_0+l})(t-s)+\mu_{k_0}^{1/2}c_{k_0}W_{k_0}(t-s)-\mu_{k_0}^{1/2}c_{k_0+l}W_{k_0+l}(t-s)\right] \geq \frac{z}{2}\right\} \\
& +\mathsf{P}\left\{\sup_{0\leq t\leq T}\sum_{i=k_0+1}^{k_0+l-1}\widetilde{Q}_i(t) \geq \frac{z}{2}\right\} \\
=\ & \mathsf{P}\left\{\sup_{0\leq t\leq T}\left[(\mu_{k_0}-\mu_{k_0+l})t+\mu_{k_0}^{1/2}c_{k_0}W_{k_0}(t)-\mu_{k_0}^{1/2}c_{k_0+l}W_{k_0+l}(t)\right] \geq \frac{z}{2}\right\} \\
& +\sum_{i=k_0+1}^{k_0+l-1}\mathsf{P}\left\{\sup_{0\leq t\leq T}\widetilde{Q}_i(t) \geq \frac{z}{2(l-1)}\right\} \\
\leq\ & \exp\left\{-\frac{2(\mu_{k_0+l}-\mu_{k_0})}{\mu_{k_0}(c_{k_0}^2+c_{k_0+l}^2)}\frac{z}{2}\right\}+\sum_{i=k_0+1}^{k_0+l-1}\mathsf{P}\left\{\sup_{0\leq t\leq T}\widetilde{Q}_i(t) \geq \frac{z}{2(l-1)}\right\}.
\end{aligned}
\tag{B.11}
$$

Next we inductively prove (B.3) holds for all $l = 2, 3, \ldots, l_0$. For $l = 2$,

$$
\begin{aligned}
\mathsf{P}\left\{\sup_{0\leq t\leq T}\widetilde{Q}_{k_0+2}(t) \geq z\right\} &\leq \exp\left\{-\frac{2(\mu_{k_0+l}-\mu_{k_0})}{\mu_{k_0}(c_{k_0}^2+c_{k_0+l}^2)}\frac{z}{2}\right\}+\mathsf{P}\left\{\sup_{0\leq t\leq T}\widetilde{Q}_{k_0+1}(t) \geq \frac{z}{2}\right\} \\
&\leq \exp\left\{-\frac{\mu_{k_0+l}-\mu_{k_0}}{\mu_{k_0}(c_{k_0}^2+c_{k_0+l}^2)}z\right\}+\exp\left\{-\frac{\gamma_1'}{2}z\right\} \\
&\leq 2\exp\left\{-\gamma_2' z\right\},
\end{aligned}
$$

where γ_2' is defined in (B.4). That is, (B.3) holds for $l = 2$.

Suppose that (B.3) holds for $l = 2, 3, \ldots, j-1$, we next show that (B.1) holds for j, $j \leq l_0$. By the inductive hypothesis and (B.11),

$$
\begin{aligned}
\mathsf{P}\left\{\sup_{0\leq t\leq T}\widetilde{Q}_{k_0+j}(t) \geq z\right\} &\leq \exp\left\{-\frac{\mu_{k_0+j}-\mu_{k_0}}{\mu_{k_0}(c_{k_0}^2+c_{k_0+j}^2)}z\right\}+\sum_{i=k_0+1}^{k_0+j-1}\exp\left\{-\frac{\gamma_1'}{2(j-1)}z\right\} \\
&\leq j\exp\left\{-\gamma_j' z\right\},
\end{aligned}
$$

where γ_l' is defined in (B.4). So, (B.3) holds for $l = 2, 3, \ldots, l_0$.

In summary, (B.3) holds. □

C Proof of Lemma 4.4

By (4.6), it suffices to prove that

$$
\widetilde{I}_k^* = 0, \quad \text{w.p.1} \tag{C.1}
$$

for $k : \rho_k > 1$

No critical loading. Suppose that $\rho_k > 1$ for all $k = k_1, k_2, \ldots, k_{l_0}$ and $\rho_k < 1$ for all $k \neq k_1, k_2, \ldots, k_{l_0}$. When $l_0 = K$ ($l_0 = 0$), it is the full overloading (underloading) case. By (2.10) and (4.2), we note that, for all $t \geq 0$, $\bar{T}_k(t) = t$ for $k = k_1, k_2, \ldots, k_{l_0}$ and $\bar{T}_k(t) = \rho_k t$ for $k \neq k_1, k_2, \ldots, k_{l_0}$. Next, we inductively prove that (C.1) holds for all $k = k_1, k_2, \ldots, k_{l_0}$.

At first, we prove (C.1) for k_1. If $k_1 = 1$, by (4.2), $\lim_{t\to\infty} \widetilde{X}_1(t) = +\infty$ w.p.1, because $\widetilde{X}_1(t)$ is a BM with positive drift $\mu_0 - \mu_1 > 0$. Since $\widetilde{I}_1(t) = \widetilde{Y}_1(t)/\mu_1 = \sup_{0\leq s\leq t} \left[-\widetilde{X}_1(s)\right]^+ /\mu_1$, we have $\sup_{t\geq 0} \widetilde{I}_1(t) < +\infty$ w.p.1. So, (C.1) holds for $k = 1$. If $k_1 > 1$, by (2.10) and (4.2), then

$$\begin{aligned}
\widetilde{X}_{k_1}(t) &= \widetilde{X}_{k_1-1}(t) - \widetilde{Q}_{k_1-1}(t) + (\mu_{k_1-1} - \mu_{k_1})t \\
&\quad + \mu_{k_1-1}^{1/2} c_{k_1-1} W_{k_1-1}(\bar{T}_{k_1-1}(t)) - \mu_{k_1}^{1/2} c_{k_1} W_{k_1}(\bar{T}_{k_1}(t)) \\
&= \widetilde{X}_1(t) - \sum_{i=1}^{k_1-1} \widetilde{Q}_i(t) + (\mu_1 - \mu_{k_1})t + \mu_1^{1/2} c_1 W_1(\bar{T}_1(t)) - \mu_{k_1}^{1/2} c_{k_1} W_{k_1}(\bar{T}_{k_1}(t)) \\
&= -\sum_{i=1}^{k_1-1} \widetilde{Q}_i(t) + (\mu_0 - \mu_{k_1})t + \mu_0^{1/2} c_0 W_0(t) - \mu_{k_1}^{1/2} c_{k_1} W_{k_1}(\bar{T}_{k_1}(t)).
\end{aligned} \tag{C.2}$$

By Lemma 4.3, for all $k = 1, 2, \ldots, k_1 - 1$, $\widetilde{Q}_k^* = 0$ w.p.1 because $\rho_k < 1$. By (4.9), it follows $\lim_{t\to\infty} \widetilde{X}_{k_1}(t) = +\infty$ w.p.1. By (2.11), $\sup_{t\geq 0} \widetilde{Y}_{k_1}(t) < \infty$ w.p.1. So, (C.1) holds for $k = k_1$.

Suppose that (C.1) holds for all $k = k_1, k_2, \ldots, k_l$, we next prove that (C.1) holds for $k = k_{l+1}$. By (2.10) and (4.2), we have

$$\begin{aligned}
\widetilde{X}_{k_{l+1}}(t) &= \widetilde{X}_{k_{l+1}-1}(t) - \widetilde{Q}_{k_{l+1}-1}(t) + (\mu_{k_{l+1}-1} - \mu_{k_{l+1}})t \\
&\quad + \mu_{k_{l+1}-1}^{1/2} c_{k_{l+1}-1} W_{k_{l+1}-1}(\bar{T}_{k_{l+1}-1}(t)) - \mu_{k_{l+1}}^{1/2} c_{k_{l+1}} W_{k_{l+1}}(\bar{T}_{k_{l+1}}(t)) \\
&= -\mu_{k_l} \widetilde{I}_{k_l}(t) - \sum_{i=k_l+1}^{k_{l+1}-1} \widetilde{Q}_i(t) + (\mu_{k_l} - \mu_{k_{l+1}})t \\
&\quad + \mu_{k_l}^{1/2} c_{k_l} W_{k_l}(\bar{T}_{k_l}(t)) - \mu_{k_{l+1}}^{1/2} c_{k_{l+1}} W_{k_{l+1}}(\bar{T}_{k_{l+1}}(t)).
\end{aligned} \tag{C.3}$$

By Lemma 4.3, $\widetilde{Q}_k^* = 0$ w.p.1 for all $k \neq k_1, k_2, \ldots, k_{l_0}$, and $\bar{T}_{k_l}(t) = \bar{T}_{k_{l+1}}(t) = t$ for all $t \geq 0$. By (4.9) and the inductive hypothesis, we have $\lim_{t\to\infty} \widetilde{X}_{k_{l+1}}(t) = +\infty$ w.p.1, and further $\sup_{t\geq 0} \widetilde{I}_{k_{l+1}}(t) < \infty$ w.p.1 by (2.11). So, (C.1) holds for $k = k_{l+1}$.

Hence, (C.1) holds for all $k = k_1, k_2, \ldots, k_{l_0}$.

One critically loaded stage Accordingly, $\rho_{k_0} = 1$ and $\rho_k \neq 1$ for all $k \neq k_0$. Consider the stage $\bar{k}_0$, by (2.10) and (4.2), we have

$$\begin{aligned}
& \widetilde{X}_{\bar{k}_0}(t) - \bar{X}_{\bar{k}_0}(t) \\
= \ & (\mu_{\bar{k}_0-1} - \mu_{\bar{k}_0})t - \widetilde{Y}_{\bar{k}_0-1}(t) + \mu^{1/2}_{\bar{k}_0-1} c_{\bar{k}_0-1} W_{\bar{k}_0-1}(\bar{T}_{\bar{k}_0-1}(t)) - \mu^{1/2}_{\bar{k}_0} c_{\bar{k}_0} W_{\bar{k}_0}(t) \\
= \ & (\mu_{\bar{k}_0-1} - \mu_{\bar{k}_0})t - \widetilde{Q}_{\bar{k}_0-1}(t) + \widetilde{X}_{\bar{k}_0-1}(t) + \mu^{1/2}_{\bar{k}_0-1} c_{\bar{k}_0-1} W_{\bar{k}_0-1}(\bar{T}_{\bar{k}_0-1}(t)) \\
& -\mu^{1/2}_{\bar{k}_0} c_{\bar{k}_0} W_{\bar{k}_0}(t) \\
= \ & (\mu_{\bar{k}_0-2} - \mu_{\bar{k}_0})t - \widetilde{Q}_{\bar{k}_0-1}(t) - \widetilde{Q}_{\bar{k}_0-2}(t) + \widetilde{X}_{\bar{k}_0-2}(t) + \mu^{1/2}_{\bar{k}_0-2} c_{\bar{k}_0-2} W_{\bar{k}_0-2}(\bar{T}_{\bar{k}_0-2}(t)) \\
& -\mu^{1/2}_{\bar{k}_0} c_{\bar{k}_0} W_{\bar{k}_0}(t) \\
= \ & -\sum_{i=k_0+1}^{\bar{k}_0-1} \widetilde{Q}_i(t) - \widetilde{Y}_{k_0}(t) + \mu^{1/2}_{k_0} c_{k_0} W_{k_0}(t) - \mu^{1/2}_{\bar{k}_0} c_{\bar{k}_0} W_{\bar{k}_0}(t).
\end{aligned} \tag{C.4}$$

If $k_0 = 1$, then $\widetilde{X}_{\bar{k}_0}(t) = \bar{X}_{\bar{k}_0}(t) - \sum_{i=2}^{\bar{k}_0-1} \widetilde{Q}_i(t) - \widetilde{Y}_1(t) + \mu_1^{1/2} c_1 W_1(t) - \mu^{1/2}_{\bar{k}_0} c_{\bar{k}_0} W_{\bar{k}_0}(t)$ where, by (2.10), $\bar{X}_{\bar{k}_0}(t) = (\mu_0 - \mu_{\bar{k}_0})t > 0$ and $\widetilde{Y}_1(t) = \Psi(\widetilde{X}_1)(t)$ with $\widetilde{X}_1(t) = \mu_0^{1/2} c_0 W_0(t) - \mu_1^{1/2} c_1 W_1(t)$. Notice that, by Lemma 4.3, $\widetilde{Q}^*_i = 0$ w.p.1 for all $i = 2, 3, \ldots, \bar{k}_0 - 1$, $\mu_1^{1/2} c_1 W_1(t) - \mu^{1/2}_{\bar{k}_0} c_{\bar{k}_0} W_{\bar{k}_0}(t)$ and $\mu_0^{1/2} c_0 W_0(t) - \mu_1^{1/2} c_1 W_1(t)$ are two driftless BMs. This, together with $\mu_0 - \mu_{\bar{k}_0} > 0$, implies that $\lim_{t\to\infty} \widetilde{X}_{\bar{k}_0}(t) = +\infty$ w.p.1. As a result, $\sup_{t\geq 0} \widetilde{I}_{\bar{k}_0}(t) < +\infty$ w.p.1 by (2.11). So, (C.1) holds for $k = \bar{k}_0$.

If $k_0 > 1$ and $\rho_k = \mu_0/\mu_k < 1$ for all $k = 1, 2, \ldots, k_0 - 1$, then (C.4) holds where $\widetilde{Y}_{k_0}(t) = \Psi(\widetilde{X}_{k_0})(t)$ with $\widetilde{X}_{k_0}(t) = -\sum_{i=1}^{k_0-1} \widetilde{Q}_i(t) + \mu_0^{1/2} c_0 W_0(t) - \mu^{1/2}_{k_0} c_{k_0} W_{k_0}(t)$ by (4.2). By Lemma 4.3, $\widetilde{Q}^*_i = 0$ w.p.1 for all $i = 1, 2, \ldots, k_0 - 1$ and $i = k_0 + 1, k_0 + 2, \ldots, \bar{k}_0 - 1$. Notice that both $\mu^{1/2}_{k_0} c_{k_0} W_{k_0}(t) - \mu^{1/2}_{\bar{k}_0} c_{\bar{k}_0} W_{\bar{k}_0}(t)$ and $\mu_0^{1/2} c_0 W_0(t) - \mu^{1/2}_{k_0} c_{k_0} W_{k_0}(t)$ are two driftless BMs, this, together with $\bar{X}_{\bar{k}_0}(t) = (\mu_0 - \mu_{\bar{k}_0})t > 0$ for all $t \geq 0$, implies that $\lim_{t\to\infty} \widetilde{X}_{\bar{k}_0}(t) = +\infty$ w.p.1. Similar with the last case of $k_0 = 1$, (C.1) holds for $k = \bar{k}_0$.

If $k_0 > 1$ and there is a $k : 1 \leq k < k_0$ such that $\rho_k > 1$, then, by the case of the **No critical loading**, (C.1) holds for $1 \leq k \leq \bar{k}_0 - 1$: $\rho_k > 1$. By (2.10), $\bar{X}_{\bar{k}_0}(t) = (\mu_{k_0} - \mu_{\bar{k}_0})t > 0$ for all $t \geq 0$. Then, by (4.14) and (C.4), we have

$$\begin{aligned}
\widetilde{X}_{\bar{k}_0}(t) = (\mu_{k_0} - \mu_{\bar{k}_0})t - \sum_{i=k_0+1}^{\bar{k}_0-1} \widetilde{Q}_i(t) + \mu^{1/2}_{k_0} c_{k_0} W_{k_0}(t) - \mu^{1/2}_{\bar{k}_0} c_{\bar{k}_0} W_{\bar{k}_0}(t) \\
- \sup_{0\leq u\leq t} \left\{ \widetilde{Y}_{\underline{k}_0}(u) + \sum_{i=\underline{k}_0+1}^{k_0-1} \widetilde{Q}_i(u) - \mu^{1/2}_{\underline{k}_0} c_{\underline{k}_0} W_{\underline{k}_0}(u) + \mu^{1/2}_{k_0} c_{k_0} W_{k_0}(u) \right\}.
\end{aligned}$$

By Lemma 4.3, $\widetilde{Q}^*_i = 0$ w.p.1 for all $i = \underline{k}_0 + 1, \underline{k}_0 + 2, \ldots, k_0 - 1$ and $i = k_0 + 1, k_0 + 2, \ldots, \bar{k}_0 - 1$. Since $\mu^{1/2}_{k_0} c_{k_0} W_{k_0}(t) - \mu^{1/2}_{\bar{k}_0} c_{\bar{k}_0} W_{\bar{k}_0}(t)$ and $-\mu^{1/2}_{\underline{k}_0} c_{\underline{k}_0} W_{\underline{k}_0}(t) + \mu^{1/2}_{k_0} c_{k_0} W_{k_0}(t)$ are two driftless BMs, and $\widetilde{I}^*_{\underline{k}_0} = 0$ w.p.1 because $\rho_{\underline{k}_0} > 1$ and the case of the **No critical**

loading, notice that $\bar{X}_{\bar{k}_0}(t) = (\mu_0 - \mu_{\bar{k}_0})t > 0$ for all $t \geq 0$, we get (C.1) holds for $k = \bar{k}_0$ similarly with the last case of $k_0 = 1$.

So far, (C.1) holds for all $k \leq \bar{k}_0 : \rho_k > 1$. For all $k > \bar{k}_0 : \rho_k > 1$, we note that because stage $\bar{k}_0$ is overloaded w.p.1., the server is always busy so that its service completion process, in the LIL sense, is a renewal process with parameters $\mu_{\bar{k}_0}$ and $c_{\bar{k}_0}$, the downstream segment of the system is equivalent to a tandem model having $K - \bar{k}_0$ stages fed by an "external" renewal arrival process with parameters $\mu_{\bar{k}_0}$ and $c_{\bar{k}_0}$. So, we can get (C.1) similarly with the case of of the **No critical loading**. □

D Proof of Lemma 4.5

Let $\rho_{k_0} = 1$. At first, by (2.10), we note that $\bar{Q}_{k_0}(t) = \mu_{k_0}\bar{Z}_{k_0}(t)$, $\widetilde{Q}_{k_0}(t) = \mu_{k_0}\widetilde{Z}_{k_0}(t)$. By (4.3), it follows that $\widetilde{Q}^*_{k_0} = \mu_{k_0}\widetilde{Z}^*_{k_0}$ w.p.1. So, by Lemma 4.2, $Q^*_{k_0} = \mu_{k_0}Z^*_{k_0}$ w.p.1. That is, the Little's law in the LIL version holds w.p.1.

Next, we prove $\widetilde{Z}^*_{k_0} = \widetilde{T}^*_{k_0} = \widetilde{I}^*_{k_0}$. By (4.6), it suffices to prove the first equality.

If $k_0 = 1$, by (2.10) and (4.2), then, for all $t \geq 0$,

$$\widetilde{Q}_1(t) = \Phi(\widetilde{X}_1)(t) = \mu_1\widetilde{Z}_1(t), \quad \widetilde{X}_1(t) = \mu_0^{1/2}c_0W_0(t) - \mu_1^{1/2}c_1W_1(t).$$

Notice that $\widetilde{Y}_1(t) = \mu_1\widetilde{I}_1(t) = \Psi(\widetilde{X}_1)(t)$, we have

$$\begin{aligned}
\widetilde{I}^*_1 &= \frac{1}{\mu_1}\limsup_{L\to\infty}\frac{1}{\varphi(L)}\left\{\sup_{0\leq t\leq L}\sup_{0\leq s\leq t}\left[-\widetilde{X}_1(s)\right]\right\} = \frac{1}{\mu_1}\limsup_{L\to\infty}\frac{1}{\varphi(L)}\left\{\sup_{0\leq t\leq L}\left[-\widetilde{X}_1(t)\right]\right\} \\
&= \frac{1}{\mu_1}\limsup_{L\to\infty}\frac{1}{\varphi(L)}\left\{\sup_{0\leq t\leq L}\left|\widetilde{X}_1(t)\right|\right\} = \frac{1}{\mu_1}\limsup_{L\to\infty}\frac{1}{\varphi(L)}\left\{\sup_{0\leq t\leq L}\Phi(\widetilde{X}_1)(t)\right\} \\
&= \widetilde{Z}^*_1,
\end{aligned} \tag{D.1}$$

where, in fact, $\widetilde{I}^*_1 = \widetilde{Z}^*_1 = C_{0,1}/\mu_1$ w.p.1.

If $k_0 > 1$, similar with (4.14), we have

$$\widetilde{X}_{k_0}(t) = -\widetilde{Y}_{\underline{k}_0}(t) - \sum_{i=\underline{k}_0+1}^{k_0-1}\widetilde{Q}_i(t) + \mu_{\underline{k}_0}^{1/2}c_{\underline{k}_0}W_{\underline{k}_0}(t) - \mu_{k_0}^{1/2}c_{k_0}W_{k_0}(t). \tag{D.2}$$

By Lemma 4.3 and Lemma 4.4, we have $\widetilde{I}^*_{\underline{k}_0} = \widetilde{Q}^*_i = 0$ w.p.1 for all $i = \underline{k}_0+1, \underline{k}_0+2, \ldots, k_0-$

1, it follows that, similar with (D.1),

$$\begin{aligned}
\widetilde{I}^*_{k_0} &= \frac{1}{\mu_{k_0}} \limsup_{L\to\infty} \frac{1}{\varphi(L)} \left\{ \sup_{0\le t\le L} \left| \Psi(\widetilde{X}_{k_0})(t) \right| \right\} \\
&= \frac{1}{\mu_{k_0}} \limsup_{L\to\infty} \frac{1}{\varphi(L)} \left\{ \sup_{0\le t\le L} \left| \mu^{1/2}_{\underline{k}_0} c_{\underline{k}_0} W_{\underline{k}_0}(t) - \mu^{1/2}_{k_0} c_{k_0} W_{k_0}(t) \right| \right\} \\
&= \frac{1}{\mu_{k_0}} \limsup_{L\to\infty} \frac{1}{\varphi(L)} \left\{ \sup_{0\le t\le L} \left| \Phi(\mu^{1/2}_{\underline{k}_0} c_{\underline{k}_0} W_{\underline{k}_0} - \mu^{1/2}_{k_0} c_{k_0} W_{k_0})(t) \right| \right\} \\
&= \frac{1}{\mu_{k_0}} \limsup_{L\to\infty} \frac{1}{\varphi(L)} \left\{ \sup_{0\le t\le L} \left| \Phi(\widetilde{X}_{k_0})(t) \right| \right\} \\
&= \frac{1}{\mu_{k_0}} \widetilde{Q}^*_{k_0} = \widetilde{Z}^*_{k_0}.
\end{aligned}$$

□