

Functional law of the iterated logarithm for multi-server queues with batch arrivals and customer feedback

Yongjiang Guo, Yunan Liu & Renhu Pei

Annals of Operations Research

ISSN 0254-5330

Volume 264

Combined 1-2

Ann Oper Res (2018) 264:157-191

DOI 10.1007/s10479-017-2529-9



Your article is protected by copyright and all rights are held exclusively by Springer Science +Business Media New York. This e-offprint is for personal use only and shall not be self-archived in electronic repositories. If you wish to self-archive your article, please use the accepted manuscript version for posting on your own website. You may further deposit the accepted manuscript version in any repository, provided it is only made publicly available 12 months after official publication or later and provided acknowledgement is given to the original source of publication and a link is inserted to the published article on Springer's website. The link must be accompanied by the following text: "The final publication is available at link.springer.com".

Functional law of the iterated logarithm for multi-server queues with batch arrivals and customer feedback

Yongjiang Guo¹ · Yunan Liu² · Renhu Pei¹

Published online: 19 May 2017
© Springer Science+Business Media New York 2017

Abstract A functional law of the iterated logarithm (FLIL) and its corresponding law of the iterated logarithm (LIL) are established for a multi-server queue with batch arrivals and customer feedback. The FLIL and LIL, which quantify the magnitude of asymptotic fluctuations of the stochastic processes around their mean values, are developed in three cases: underloaded, critically loaded and overloaded, for five performance measures: queue length, workload, busy time, idle time and departure process. Both FLIL and LIL are proved using an approach based on strong approximations.

Keywords Functional law of the iterated logarithm · Multi-server queue · Batch arrival · Customer feedback · Nonexponential service times · Strong approximation

1 Introduction

We develop a functional law of the iterated logarithm (FLIL) and its corresponding law of the iterated logarithm (LIL) for the multi-server $GI^B/GI^F/N$ queue, which has a renewal batch arrival process with independent and identically distributed (i.i.d.) batch sizes (the GI^B), i.i.d. non-exponential service times (the second GI), N servers in parallel, and customer feedback after service completion in a Bernoulli fashion (the superscript F).

Many asymptotic results have been developed for queueing networks having the batch arrivals. For single-server queueing models, there is a large volume of literature on the asymptotic analysis, including stationary probability distribution analysis (Machihara 1999; Ommeren 1990), fluid approximation (Chen and Shanthikumar 1994; Dai 1995), diffusion

✉ Yunan Liu
yliu48@ncsu.edu

Yongjiang Guo
yongerguo@bupt.edu.cn

¹ School of Science, Beijing University of Posts and Telecommunications, Beijing 100876, China

² Industrial and Systems Engineering, North Carolina State University, Raleigh, NC 27607, USA

approximation in heavy traffic (Glynn and Whitt 1987; Pang and Whitt 2012), LIL limits (Guo and Liu 2015; Minkevičius and Steišūnas 2003; Sakalauskas 2000), strong approximation (Chen and Mandelbaum 1994; Chen and Shen 2000; Glynn and Whitt 1991a, b; Horváth 1992; Mandelbaum et al. 1998; Whitt 1983; Zhang 1997; Zhang et al. 1990; Zhang and Hsu 1992), stationary optimal policies under a linear cost structure (Lee and Srinivasan 1989), etc. Related heavy-traffic results for multi-server queues include the diffusion approximation (Chen and Shanthikumar 1994) and the LIL limits (Minkevičius 2014) for generalized Jackson network in strictly heavy traffic. See Iglehart (1971) for the FLIL and LIL limits of multi-channel queue models. Queueing models having after-service feedback have also proven useful in modeling real service systems, see Dai (1995) for stochastic queueing networks with customer feedback, Yom-Tov and Mandelbaum (2014) for the Markovian Erlang-R model and Liu and Whitt (2017) for time-varying staffing recommendations to stabilize performance in queues with feedback. Readers are referred to Borovkov (1984), Chen and Yao (2001) and Whitt (2002) for a whole asymptotic analysis for queueing networks.

FLILs and LILs. The earliest FLIL results were developed for the standard *Brownian motion* (BM). Let W be a one-dimensional standard BM. By considering a sequence of scaled BMs indexed by n , $W_n(t) \equiv W(nt)/\sqrt{n \log \log n}$, $0 \leq t \leq 1$, $n \geq 3$, Strassen (1964) showed that, with probability one (w.p.1), the sequence $\{W_n, n \geq 3\}$ is *relatively compact*, that is, every subsequence of W_n has a convergent subsubsequence, and the limits of all convergent subsubsequences are contained in a compact set:

$$\mathcal{E} = \left\{ x \in C^1[0, 1] : x(0) = 0, \int_0^1 \left(\frac{d}{dt} x(t) \right)^2 dt \leq 1 \right\},$$

where $C^1[0, 1]$ is the space of the one-dimensional continuous functions on $[0, 1]$. Intuitively, the set \mathcal{E} is the space of absolutely continuous functions with a controlled total variation. Iglehart (1971) adapted Strassen’s approach to establish the FLIL for a multi-channel queueing systems; also see Glynn and Whitt (1986, 1987, 1988) for a Little’s law version of FLIL.

The earliest LIL result was developed for the standard BM by Lévy (1937) and Lévy (1948), that is,

$$\limsup_{t \rightarrow \infty} \frac{W(t)}{\sqrt{2t \log \log t}} = - \liminf_{t \rightarrow \infty} \frac{W(t)}{\sqrt{2t \log \log t}} = 1, \quad \text{w.p.1.} \tag{1}$$

The version of LIL in (1) is called the *strong form* because it provides an explicit value [the “1” on the right-hand side of (1)] to quantify the asymptotic rate of the increasing variability for a standard BM. Motivated by the Lévy’s works, other LIL results have later been developed for queueing systems, including single-server priority queues (Guo and Liu 2015), multi-channel queues (Iglehart 1971), strictly overloaded tandem queueing model (Minkevičius and Steišūnas 2003) and generalized Jackson networks (Minkevičius 2014; Sakalauskas 2000). In contrast to the strong form in (1), Chen and Yao (2001) developed a *weak form* of LIL for the queue length process Q (centered by its fluid function \bar{Q}) of the $GI/GI/1$ queue: they showed that $\sup_{0 \leq t \leq T} |Q(t) - \bar{Q}(t)|$ is in the same order of the function $\sqrt{T \log \log T}$ as $T \rightarrow \infty$. This result is called the *weak form* because the value of the LIL limit [as in (1)] was not identified. Also see Chen and Mandelbaum (1994), Chen and Shen (2000) and Chen and Yao (2001) for other results on weak-form LILs for generalized Jackson network and feedforward queueing network.

Our contributions We summarize our contributions below:

- (i) First, we establish a strong-form FLIL and LIL for key performance functions of the $GI^B/GI^F/N$ queue including the queue length, workload, busy time, idle time and departure processes.
- (ii) Second, our results cover all three regimes defined in terms of the traffic intensity ρ : (i) *underloaded* (UL) with $\rho < 1$, (ii) *critically loaded* (CL) with $\rho = 1$ and (iii) *overloaded* (OL) with $\rho > 1$ (unlike most previous works which merely focused on either UL or OL state).
- (iii) Third, in terms of the model input parameters (e.g., arrival rate, feedback probability and moments of service times), we identify the FLIL and LIL limits as simple analytic expressions. The FLIL limits are compact sets which are explicitly defined by the model parameters; and the LIL limits are closed-form functions of the model parameters. These simple FLIL and LIL limits provide structural insights and interesting results. For instance, Theorem 3 shows that Little’s law always holds in the UL and CL cases but may fail in the OL case; Little’s Law holds in the OL case only when the variance of service times are zero. To gain insights into our explicit FLIL and LIL limits, we provide detailed discussions in Sect. 4 and numerical examples (see Sect. 6).
- (iv) Finally, different from previous methods for LIL and FLIL [e.g., using probability inequalities (Minkevičius and Steišūnas 2003; Minkevičius 2014; Sakalauskas 2000) and construction of renewal processes (Iglehart 1971)], we adopt a new *strong approximation* (SA) based approach; we believe that this new approach may help stimulate future research (e.g., facilitating proofs of asymptotic results of other queueing models). We next give more details of our SA-based approach.

The strong approximation approach. To illustrate the framework of SA, consider a renewal process $\{N(t), t \geq 0\}$ with rate $\lambda > 0$ and interrenewal-time variance $\sigma^2 < \infty$. Define $\bar{N}(t) = \lambda t$ and $\tilde{N}(t) = \lambda t + \lambda^{3/2}\sigma W(t)$, where W is the one-dimensional standard BM. In fact, $\bar{N}(t)$ and $\tilde{N}(t)$ are the fluid limit and SA of $N(t)$ respectively. It follows from Horváth (1984a) and Horváth (1984b) that, if the r th moment of interarrival times is finite, $r > 2$, then

$$\sup_{0 \leq t \leq L} |N(t) - \tilde{N}(t)| = o(L^{1/r}), \quad \text{w.p.1,} \tag{2}$$

where the little function “ $o(\cdot)$ ” means that $f(t) = o(g(t))$ as $t \rightarrow \infty$ if $\lim_{t \rightarrow \infty} |f(t)/g(t)| = 0$. Equation (2) implies that the renewal process is approximated by a one-dimensional standard BM with an error $o(L^{1/r})$ with $r > 2$ if the r th moment of interarrival times is finite. For $t \in [0, 1]$, the FLIL of the renewal process $N(t)$ can be obtained by taking $n \rightarrow \infty$ in the sequence:

$$\frac{N(nt) - \bar{N}(nt)}{\sqrt{2n \log \log n}} = \frac{N(nt) - \tilde{N}(nt)}{\sqrt{2n \log \log n}} + \frac{\tilde{N}(nt) - \bar{N}(nt)}{\sqrt{2n \log \log n}} = \frac{\lambda^{3/2}\sigma W(nt)}{\sqrt{2n \log \log n}} + o(1),$$

where the second equality holds by (2). Hence, the FLIL of renewal process is now transformed to the FLIL problem of the corresponding scaled BM. SAs have been developed for various stochastic processes, such as random walks (Csörgő et al. 1987) and renewal-related processes (Csörgő and Révész 1981; Csörgő and Horváth 1993). There is a body of literature using the SA to study queueing models, including the $GI/GI/1$ queue (Chen and Yao 2001), $GI/GI/\infty$ queue (Glynn and Whitt 1991a), multiple channel queue (Zhang et al. 1990), tandem-queue network (Glynn and Whitt 1991b), generalized Jackson network (Chen and Mandelbaum 1994; Horváth 1992; Zhang 1997), non-preemptive priority queue (Zhang

and Hsu 1992), time-dependent Markovian network queues (Mandelbaum and Massey 1995; Mandelbaum et al. 1998) and feedforward queueing networks (Chen and Shen 2000; Chen and Yao 2001).

Our SA-based approach follows four steps: first, we need to establish the fluid limits and strong approximations for the desired performance functions (e.g., the queue length and workload processes). Second, we connect the FLILs of these performance functions to the FLILs of their corresponding SAs; these SAs are usually in forms of some continuous functions of BMs. Next, we directly treat the BM related processes to obtain closed-form FLIL limits. Finally, we obtain the LIL limits from their corresponding FLIL limits. One major advantage of the SA-based approach is that, once the SAs of the desired performance functions are established, we are able to take advantage of existing FLIL results for BMs. In this case, developing FLILs for BM-related functions is much easier than treating FLILs for the discrete performance functions (e.g., queue length process). Nevertheless, there are two major difficulties in this SA-based approach: it is in general not easy to develop the heavy-traffic fluids and strong approximations for the desired performance functions. Commonly used methods include the continuous mapping approach (Chen and Shen 2000; Chen and Yao 2001; Zhang 1997; Zhang et al. 1990) and probability inequality approach (Horváth 1992; Zhang and Hsu 1992). In addition, obtaining the LIL limits from the corresponding FLIL results may not be straightforward.

Organization of the paper In Sect. 2, we introduce the $GI^B/GI^F/N$ model, define key performance functions and give useful preliminary results. In Sect. 3, we review the fluid limits of the $GI^B/GI^F/N$ queue which are building blocks for the FLIL and LIL. In Sect. 4, we present our main FLIL and LIL results (Theorems 2, 3, 4, 5). In Sect. 5, we prove the main results and other supporting results. In Sect. 6, we give concrete numerical examples to gain insights into the main results. Finally, we draw conclusions in Sect. 7. Additional supporting materials, including additional numerical examples and omitted proofs, appear in “Appendix” section.

Notations We close this section by summarizing all notations. All random variables and processes are defined on a common probability space $(\Omega, \mathcal{F}, \mathbb{P})$. We let $E(X)$ and $\text{Var}(X)$ be the mean and variance for X . We write $X \stackrel{d}{=} Y$ if X and Y have the same distributions. For any positive integer k , we denote by \mathbb{R}^k and \mathbb{R}_+^k the sets of the k -dimensional real numbers and nonnegative real numbers. Vector $e \in \mathbb{R}^k$ is a column vector with all of its entries being ones. We denote $\mathbb{R} = \mathbb{R}^1$ and $\mathbb{R}_+ = \mathbb{R}_+^1$. For $a \in \mathbb{R}$, $[a]^+ \equiv \max\{a, 0\}$. For any $t \in \mathbb{R}$, denote by $[t]$ the maximal integer no more than t . Let $\mathbb{D}^k[a, b]$ be the space of k -dimensional right continuous functions on $[a, b]$ having left limits on $(a, b]$, endowed the Skorohod topology, see Ethier and Kurtz (1986). Let $\mathbb{C}^k[a, b]$ be the subset of continuous paths in $\mathbb{D}^k[a, b]$. Denote $\mathbb{D} \equiv \mathbb{D}^1$ and $\mathbb{C} \equiv \mathbb{C}^1$. We say that $f_n \rightrightarrows \mathcal{K}_f$ w.p.1 if $\{f_n, n \geq 1\}$ is relatively compact (i.e., every subsequence has a convergent subsubsequence) and the set of all limit points is the compact set \mathcal{K}_f . Let $\|f\|_T \equiv \sup_{0 \leq t \leq T} |f(t)|$ be the uniform norm of f . We say $f_n \rightarrow f$ uniformly on compact set (u.o.c.) if $\|f_n - f\|_T \rightarrow 0$, as $n \rightarrow \infty$. We say $f(t) = O(g(t))$ as $t \rightarrow \infty$ if $\limsup_{t \rightarrow \infty} |f(t)/g(t)| \leq M$ for some $M > 0$. We use “ \equiv ” to denote a definition. We summarize all acronyms in Table 2 of “Appendix” section.

2 The $GI^B/GI^F/N$ queueing model

Consider a multi-server $GI^B/GI^F/N$ queue with batch arrivals and customer feedback. The servers are indexed by $1, 2, \dots, N$. External customers arrive in batches and are served in the

order of arrival. Customers in the same batch are served in an arbitrary order. If a customer finds at least one available server, it starts service with the available server having the smallest index. If all servers are busy, the arrival waits in queue. When a customer finishes service, with probability $0 \leq p < 1$ it returns for more service (joining the end of the queue) and with probability $1 - p$ it leaves the system. We assume a work-conserving service discipline, i.e., no server is allowed to be idle if there is a customer waiting in queue.

Let $u(n)$ be interarrival time between the $(n - 1)$ th and n th batches, $v_j(n)$ be the service time of the n th served customer by server j , and $\xi(n)$ be the size of the n th arrived batch, $n = 1, 2, \dots$. Suppose that $u \equiv \{u(n), n = 1, 2, \dots\}$, $v_j \equiv \{v_j(n), n = 1, 2, \dots\}$ and $\xi \equiv \{\xi(n), n = 1, 2, \dots\}$ are mutually independent i.i.d. sequences of non-negative random variables, having means $E[u(1)] \equiv 1/\alpha$, $E[v_j(1)] \equiv 1/\mu$ and $E[\xi(1)] \equiv m$, and squared coefficients of variation (SCV) $c_a^2 \equiv Var[u(1)]/(E[u(1)])^2$, $c_s^2 \equiv Var[v_j(1)]/(E[v_j(1)])^2$ and $c_b^2 \equiv Var[\xi(1)]/(E[\xi(1)])^2$. Throughout the rest of the paper, we assume that, for some $r > 2$,

$$E[u(1)^r] < \infty, \quad E[v_j(1)^r] < \infty, \quad E[\xi(1)^r] < \infty, \quad \text{for all } j = 1, 2, \dots, N. \quad (3)$$

We use the i.i.d. sequence $\gamma \equiv \{\gamma_j(n), j = 1, 2, \dots, N, n = 1, 2, \dots\}$ to characterize the Bernoulli feedback. After a customer completing service by server j , let $\gamma_j(n) = 1$ if the customer decides to revisit the system, and let $\gamma_j(n) = 0$ if the customer leaves the system. We assume that the sequence γ is independent of u, v, ξ , and has mean $E[\gamma_j(1)] \equiv p \in [0, 1)$.

Define the partial sums of the interarrival times, service times, batches and routings:

$$U(n) \equiv \sum_{k=1}^n u(k), \quad V_j(n) \equiv \sum_{k=1}^n v_j(k), \quad B(n) \equiv \sum_{k=1}^n \xi(k), \quad \Gamma_j(n) \equiv \sum_{k=1}^n \gamma_j(k), \quad (4)$$

$n = 1, 2, \dots$, and two renewal processes:

$$A(t) \equiv \max\{n \geq 0 : U(n) \leq t\} \quad \text{and} \quad S_j(t) \equiv \max\{n \geq 0 : V_j(n) \leq t\}, \quad (5)$$

where $A(t)$ counts the total number of batch arrivals in $(0, t]$ and $S_j(t)$ counts the number of customers server j can potentially serve in $(0, t]$ (assuming server j is always busy).

Define the traffic intensity:

$$\rho \equiv \frac{\lambda}{N\mu} \quad \text{with} \quad \lambda = m\alpha + pN\mu. \quad (6)$$

We say that the system is UL when $\rho < 1$, CL when $\rho = 1$ and OL when $\rho > 1$.

Performance functions Let $Q(t)$ be the total number of customers in the system at time t and $Z(t)$ be the workload at time t , that is, the time until the system first becomes empty assuming no future arrivals after time t . Let $T_j(t)$ indicate the cumulative busy time of server j during $[0, t]$ and $I_j(t) \equiv t - T_j(t)$ indicates the cumulative idle time of server j during $[0, t]$. Let $T(t) \equiv \sum_{j=1}^N T_j(t)$ and $I(t) \equiv \sum_{j=1}^N I_j(t) = Nt - T(t)$ be the total busy time and idle time for all N servers in $[0, t]$. Let $D(t)$ denote the total number of departures by time t .

System equations We assume the system is empty initially, i.e., $Q(0) = 0$. Flow conservation implies that

$$Q(t) = B(A(t)) - \sum_{j=1}^N S_j(T_j(t)) + \sum_{j=1}^N \Gamma_j(S_j(T_j(t))), \quad (7)$$

where, by time t , $B(A(t))$ is the total number of customer arrivals, $S_j(T_j(t))$ is the total number of service completions at server j and $\Gamma_j(S_j(T_j(t)))$ is the total number of customers that are feedback from server j .

To relate the queue length process to idle times, we rewrite (7) as

$$Q(t) = X(t) + Y(t), \tag{8}$$

where the auxiliary processes:

$$Y(t) \equiv (1 - p)\mu I(t) = (1 - p)\mu \sum_{j=1}^N I_j(t) \geq 0, \tag{9}$$

$$X(t) \equiv \theta_N t + [B(A(t)) - m\alpha t] + \sum_{j=1}^N [\Gamma_j(S_j(T_j(t))) - pS_j(T_j(t))] + p \sum_{j=1}^N [S_j(T_j(t)) - \mu T_j(t)] - \sum_{j=1}^N [S_j(T_j(t)) - \mu T_j(t)], \tag{10}$$

and $\theta_N \equiv m\alpha - (1 - p)N\mu$. We note that, for any $j = 1, 2, \dots, N$, $I_j(t)$ cannot increase at time t if $Q(t) \geq N$ under the work-conserving service discipline. That is, when $Q(t) \geq N$, $dY(t) = 0$. On the other hand, if $Q(t) < N$, then $dY(t) \geq 0$. This implies that

$$(N - Q(t))dY(t) \geq 0. \tag{11}$$

More discussion on the queue length and idle time see Sect. B.2.

When $N = 1$, the multi-dimensional system dynamical equations simplify to:

$$\begin{aligned} Q(t) &\equiv B(A(t)) - D(t) \equiv X(t) + Y(t) \geq 0, \quad D(t) \equiv S(T(t)) - \Gamma(S(T(t))), \\ X(t) &\equiv \theta_1 t + [B(A(t)) - mA(t)] + m[A(t) - \alpha t] \\ &\quad + [\Gamma(S(T(t))) - pS(T(t))] - (1 - p)[S(T(t)) - \mu T(t)], \\ Y(t) &\equiv (1 - p)\mu I(t), \quad \int_0^t Q(t)dY(t) = 0, \\ Z(t) &\equiv V(B(A(t)) + \Gamma(S(T(t)))) - T(t) = V(Q(t) + S(T(t))) - T(t), \end{aligned} \tag{12}$$

where $\Gamma \equiv \Gamma_1$, $V \equiv V_1$, $S \equiv S_1$. In this case an equivalent representation for $Q(t)$ and $Y(t)$ is

$$Q(t) = \Phi(X)(t) \quad \text{and} \quad Y(t) = \Psi(X)(t), \tag{13}$$

where two functions (Ψ, Φ) , defined as

$$\Psi(x)(t) = \sup_{0 \leq s \leq t} [-x(s)]^+ \quad \text{and} \quad \Phi(x)(t) = x(t) + \Psi(x)(t), \tag{14}$$

are known as the one-dimensional *oblique reflection mapping* (ORM) (see Sect. B.1 for an alternative definition of ORM).

The objective of the rest of the paper is to establish the FLIL and LIL for performance processes (Q, Z, T, I, D) and identify the their corresponding FLIL and LIL limits as functions of the model input data

$$\mathcal{D} \equiv (\alpha, \mu, m, p, c_a^2, c_s^2, c_b^2, N). \tag{15}$$

3 Fluid limits

Since the forms of the FLIL and LIL involve the performance measures centered by their corresponding fluid functions, we now give the corresponding fluid limits. We first define the fluid-scaled processes:

$$\begin{aligned} \bar{Q}^{(n)}(t) &\equiv \frac{1}{n}Q(nt), & \bar{Y}^{(n)}(t) &\equiv \frac{1}{n}Y(nt), & \bar{X}^{(n)}(t) &\equiv \frac{1}{n}X(nt), & \bar{Z}^{(n)}(t) &\equiv \frac{1}{n}Z(nt), \\ \bar{T}^{(n)}(t) &\equiv \frac{1}{n}T(nt), & \bar{I}^{(n)}(t) &\equiv \frac{1}{n}I(nt), & \bar{D}^{(n)}(t) &\equiv \frac{1}{n}D(nt), & \bar{T}_j^{(n)}(t) &\equiv \frac{1}{n}T_j(nt), \\ \bar{I}_j^{(n)}(t) &\equiv \frac{1}{n}I_j(nt), & \bar{D}_j^{(n)}(t) &\equiv \frac{1}{n}D_j(nt), & & & & j = 1, 2, \dots, N. \end{aligned}$$

We next give the *functional strong law of large numbers* (FSLLN) results for the $GI^B/GI^F/N$ model. The proof is similar with [Chen and Yao \(2001\)](#) and [Dai \(1995\)](#). We give the proof in Sect. C.1 of “Appendix” section.

Theorem 1 (FSLLN for $GI^B/GI^F/N$) *Assume the system is initially empty. If $E[u(1)] < \infty$, $E[v_j(1)] < \infty$ and $E[\xi(1)] < \infty$, $j = 1, 2, \dots, N$, then,*

(i) *If $N \geq 1$ and $\rho \geq 1$, then, as $n \rightarrow \infty$,*

$$\left(\bar{Q}^{(n)}, \bar{X}^{(n)}, \bar{T}_j^{(n)}, \bar{I}_j^{(n)}, \bar{D}^{(n)}, \bar{D}_j^{(n)}\right) \rightarrow (\bar{Q}, \bar{X}, \bar{T}_j, \bar{I}_j, \bar{D}, \bar{D}_j) \quad \text{u.o.c., w.p.1,}$$

where $\bar{Q}(t) = \bar{X}(t) = \theta_N t$, $\bar{T}_j(t) = t - \bar{I}_j(t) = t$, $\bar{T}(t) = Nt$, $\bar{I}(t) = 0$, $\bar{D}_j(t) = (1 - p)\mu t$ and $\bar{D}(t) = \sum_{j=1}^N \bar{D}_j(t)$.

(ii) *If $N = 1$, then, as $n \rightarrow \infty$,*

$$\left(\bar{Q}^{(n)}, \bar{X}^{(n)}, \bar{Y}^{(n)}, \bar{Z}^{(n)}, \bar{T}^{(n)}, \bar{I}^{(n)}, \bar{D}^{(n)}\right) \rightarrow (\bar{Q}, \bar{X}, \bar{Y}, \bar{Z}, \bar{T}, \bar{I}, \bar{D}), \quad \text{u.o.c., w.p.1,}$$

where $\bar{Q}, \bar{X}, \bar{Y}, \bar{Z}, \bar{T}, \bar{I}, \bar{D}$ satisfy, for $t \geq 0$,

$$\begin{aligned} (\bar{Q}, \bar{Y}) &= (\Phi, \Psi)(\bar{X}), & \bar{X}(t) &= \theta_1 t, & \bar{I}(t) &= \frac{1}{(1 - p)\mu} \bar{Y}(t), \\ \bar{Z}(t) &= \frac{1}{\mu} \bar{Q}(t), & \bar{T}(t) &= t - \bar{I}(t), & \bar{D}(t) &= (1 - p)\mu \bar{T}(t). \end{aligned} \tag{16}$$

4 Main results

In this section, we establish the LIL and FLIL for the $GI^B/GI^F/N$ queue. we first define the LIL and FLIL scalings in Sect. 4.1; next in Sect. 4.2 we develop the FLIL and LIL results in the UL, CL and OL cases. We give all proofs in Sect. 5.

4.1 The LIL and FLIL scalings

LIL scaling and limits Using the fluid limits given in Sect. 3, we now define the LIL scaling and limits. Let

$$Q_{\text{sup}}^* \equiv \limsup_{t \rightarrow \infty} \frac{Q(t) - \bar{Q}(t)}{\varphi(t)} \quad \text{and} \quad Q_{\text{inf}}^* \equiv \liminf_{t \rightarrow \infty} \frac{Q(t) - \bar{Q}(t)}{\varphi(t)}. \tag{17}$$

where $\varphi(t) = \sqrt{2t \log \log t}$ for all $t > e$ (Euler's constant). Similarly, we define the upper and lower LIL-scaled processes: $Z_{\text{sup}}^*, Z_{\text{inf}}^*, T_{\text{sup}}^*, T_{\text{inf}}^*, I_{\text{sup}}^*, I_{\text{inf}}^*, D_{\text{sup}}^*, D_{\text{inf}}^*, T_{j,\text{sup}}^*, T_{j,\text{inf}}^*$ and $I_{j,\text{sup}}^*, I_{j,\text{inf}}^*, j = 1, 2, \dots, N$. We will express all the LIL limits in terms of the input data \mathcal{D} in (15).

FLIL scaling and limits For any $t \in [0, 1]$ and $n = 3, 4, \dots$, define

$$Q^n(t) \equiv \frac{Q(nt) - \bar{Q}(nt)}{\varphi(n)}. \tag{18}$$

Similarly we define the FLIL-scaled processes: $X^n(t), Y^n(t), I^n(t), T^n(t), Z^n(t), D^n(t), T_j^n(t)$ and $I_j^n(t)$ in the same token of (18), $j = 1, 2, \dots, N$. We will develop the FLIL results by showing that

$$(Q^n, Z^n, I^n, T^n, D^n) \rightrightarrows (\mathcal{K}_Q, \mathcal{K}_Z, \mathcal{K}_I, \mathcal{K}_T, \mathcal{K}_D) \equiv \mathcal{K}^*, \quad \text{w.p.1}, \tag{19}$$

and identifying the compact sets $\mathcal{K}_Q, \mathcal{K}_Z, \mathcal{K}_I, \mathcal{K}_T, \mathcal{K}_D$ characterized by the input data \mathcal{D} in (15) and the compact set \mathcal{G}_k defined as

$$\mathcal{G}_k(\delta) \equiv \left\{ x \in \mathbb{C}^k[0, 1] : x(0) = 0, \int_0^1 [\dot{x}(t)]^2 dt \leq \delta^2 \right\}, \quad \delta > 0, \tag{20}$$

where the square denotes inner product, and $\dot{x}(t)$ denotes the derivative of $x(t)$ which exists almost everywhere with respect to Lebesgue measure.

We remark that $\mathcal{G}_k(\delta)$ is the space of continuous functions having a δ -controlled quadratic variation. We give some examples to gain insights. For instance, $x_1(t) = \delta at \in \mathcal{G}(\delta)$ for $0 < a \leq 1$ because $x_1(0) = 0$ and $\int_0^1 [\dot{x}_1(t)]^2 dt = a^2 \delta^2 \leq \delta^2$; $x_2(t) \equiv (\delta b_1 t, \delta b_2 t)' \in \mathcal{G}_2(\delta)$ for b_1, b_2 such that $0 < b_1^2 + b_2^2 \leq 1$, because $x_2(0) = 0$ and $\int_0^1 [\dot{x}_2(t)]^2 dt = \delta^2(b_1^2 + b_2^2) \leq \delta^2$.

4.2 The LIL and FLIL limits

We now give our main results. For the OL and CL $GI^B/GI^F/N$ model with $N \geq 1$, we give the FLIL and LIL results for the queue length, busy time and idle time process. For the case $N = 1$, we give the FLIL and LIL results for all performance functions in OL, UL and CL cases. We give all proofs in Sect. 5.

Theorem 2 (FLIL for $GI^B/GI^F/N$) For $N \geq 1$, we have

$$\mathcal{K}_Q = \begin{cases} \Phi(\mathcal{G}(\sigma_o(N))), & \rho = 1, \\ \mathcal{G}(\sigma_o(N)), & \rho > 1, \end{cases} \tag{21}$$

where $\mathcal{G} \equiv \mathcal{G}_1$ is defined in (20) and

$$\sigma_o^2(N) = m^2 \alpha (c_a^2 + c_b^2) + p(1-p)\mu + (1-p)^2 N \mu c_s^2. \tag{22}$$

If $\rho = 1$, then for all $t \in [0, 1]$,

$$(1-p)\mu \sum_{j=1}^N I_j^n(t) \rightrightarrows \Psi(\mathcal{G}(\sigma_o(N))), \quad \text{w.p.1}. \tag{23}$$

If $\rho > 1$, then $I_j^n(t) \rightrightarrows 0$ and $T_j^n(t) \rightrightarrows 0$ w.p.1 for all $t \in [0, 1]$ and $j = 1, 2, \dots, N$.

Remark 1 (Understanding FLIL for $GI^B/GI^F/N$) When $\rho \geq 1$, the FLIL limit for the queue length is mainly determined by total variability $\sigma_o(N)$, including the variabilities of the arrival, batch, routing and service distributions. When $\rho = 1$, the weighted idle time FLIL limit in (23) is dependent on $\sigma_o(N)$ through continuous mapping Ψ . When $\rho > 1$, the FLILs for busy and idle time both degenerate to a single point zero, because all the servers are busy almost all time so that both busy and idle times are almost linear in time t . By the definition of the relatively compact, if $f_n \Rightarrow \mathcal{K}_f$ w.p.1 and the set \mathcal{K}_f contains a single point, then $f_n \rightarrow \mathcal{K}_f$ w.p.1. Therefore, the FLIL here implies the FSLLN, that is, $I_j^n(t) \rightarrow 0$ and $T_j^n(t) \rightarrow 0$ w.p.1 as $n \rightarrow \infty$.

Theorem 3 (FLIL for $GI^B/GI^F/1$) *If $N = 1$, then the convergence (19) holds with $\mathcal{K}^* =$*

$$\left\{ \begin{array}{ll} \left\{ \left(0, 0, \frac{\sigma_{uc}}{(1-p)\mu}x, -\frac{\sigma_{uc}}{(1-p)\mu}x, \sigma_{D,u}x \right) : x \in \mathcal{G}(1) \right\}, & \text{if } \rho < 1, \\ \left\{ \left(\Phi(x), \frac{1}{\mu}\Phi(x), \frac{1}{(1-p)\mu}\Psi(x), -\frac{1}{(1-p)\mu}\Psi(x), h(y) \right) : x \in \mathcal{G}(\sigma_o), y \in \mathcal{G}_2(\sigma_o) \right\}, & \text{if } \rho = 1, \\ \left\{ \left(\sigma_o x, \frac{\sigma_{Z,o}}{\mu}x, 0, 0, (1-p)\sigma_{D,o}x \right) : x \in \mathcal{G}(1) \right\}, & \text{if } \rho > 1, \end{array} \right. \quad (24)$$

where h is a continuous function mapping: $\mathbb{C} \times \mathbb{C} \rightarrow \mathbb{C}$, defined as

$$h(x, y)(t) = y(t) + \inf_{0 \leq s \leq t} [x(s) - y(s)], \quad (25)$$

$\sigma_o \equiv \sigma_o(1) > 0$ defined in (22), and

$$\begin{aligned} \sigma_{uc} &\equiv \sqrt{\alpha m^2 (c_a^2 + c_b^2) + (1-p)m\alpha c_s^2 + pm\alpha}, \\ \sigma_{D,u} &\equiv \sqrt{\alpha m^2 (c_a^2 + c_b^2)}, \quad \sigma_{D,o} \equiv \sqrt{\mu c_s^2 + \frac{p\mu}{1-p}}, \\ \sigma_{Z,o} &\equiv \sqrt{\alpha m^2 (c_a^2 + c_b^2) + c_s^2 (p\mu^{1/2} - (m\alpha + p\mu)^{1/2})^2 + p(1-p)\mu}. \end{aligned} \quad (26)$$

Remark 2 (Little’s law in FLIL) Unlike in the fluid limit where Little’s Law always holds (e.g., $\bar{Q}(t) = \mu \bar{Z}(t)$ in (16)), Little’s law in FLIL, such as, $\mathcal{K}_Q = \mu \mathcal{K}_Z$, continues to hold in the UL and CL cases, but fails in the OL case. In general Little’s law fails in FLIL because FLIL describes the asymptotic stochastic fluctuations around the means rather than the mean values (that are characterized by the fluid model). In fact, in (24) we have $\mathcal{K}_Q = \mu \mathcal{K}_Z$ in the OL case if we let $c_s = 0$ (little’s law holds in the OL case only when the service times are deterministic).

Remark 3 (FLIL for $GI/GI/1$) Setting $m = 1, c_b = 0$ and $p = 0$ yields the FLIL for the $GI/GI/1$ special case. The FLIL parameters in (26) simplify to the following:

$$\begin{aligned} \sigma_{uc} = \sigma_{Z,o} &= \sqrt{\alpha (c_a^2 + c_s^2)}, \quad \sigma_{D,u} = \sqrt{\alpha c_a^2}, \quad \sigma_o = \sqrt{\alpha c_a^2 + \mu c_s^2}, \\ \sigma_{D,o} &= \sqrt{\mu c_s^2}. \end{aligned} \quad (27)$$

Here for the $GI/GI/1$ queue, the FLIL of the departure $\mathcal{K}_D = h(\mathcal{G}_2(\sigma_o))$ coincides with the result in Theorem 4.2 in Iglehart (1971).

We next give the LIL results for the $GI^B/GI^F/N$ model.

Theorem 4 (LIL for $GI^B/GI^F/N$)

- (i) If $\rho = 1$, then $Q_{sup}^* = \sigma_o(N)$, $Q_{inf}^* = 0$.
- (ii) If $\rho > 1$, then $Q_{sup}^* = -Q_{inf}^* = \sigma_o(N)$, $T_{sup}^* = T_{inf}^* = I_{sup}^* = I_{inf}^* = 0$ and $T_{j,sup}^* = T_{j,inf}^* = I_{j,sup}^* = I_{j,inf}^* = 0$ for all $j = 1, 2, \dots, N$.

Remark 4 (Understanding the $GI^B/GI^F/N$ LIL limits) According to the FLIL scaling in (17), the LIL limits given in Theorem 4 imply the fluid limits in (i) in Theorem 1, because $\varphi(n)/n \rightarrow 0$ as $n \rightarrow \infty$. The LIL limits of T, I, T_j and I_j given in (ii) in Theorem 4 are consistent with their FLIL limits in Theorem 2, because when both the lim sup and lim inf are zero, the limit is zero. In the CL case, $Q_{inf}^* = 0$ implies that the scaled queue length $Q(t)/\varphi(t)$ approaches 0 infinitely often on every sample path. However in the OL case, the queue length has both positive and negative fluctuations around its fluid limit.

Theorem 5 (LIL for $GI^B/GI^F/1$) If $N = 1$, then the LIL limits are given below:

- (i) If $\rho < 1$, then

$$Q_{sup}^* = Q_{inf}^* = Z_{sup}^* = Z_{inf}^* = 0, \quad D_{sup}^* = -D_{inf}^* = \sigma_{D,u},$$

$$T_{sup}^* = -T_{inf}^* = I_{sup}^* = -I_{inf}^* = \frac{\sigma_{uc}}{(1-p)\mu}. \tag{28}$$

- (ii) If $\rho = 1$, then

$$Q_{sup}^* = \mu Z_{sup}^* = \sigma_o, \quad I_{sup}^* = -T_{inf}^* = \frac{\sigma_o}{(1-p)\mu},$$

$$Q_{inf}^* = Z_{inf}^* = I_{inf}^* = T_{sup}^* = 0. \tag{29}$$

- (iii) If $\rho > 1$, then

$$Q_{sup}^* = -Q_{inf}^* = \sigma_o, \quad Z_{sup}^* = -Z_{inf}^* = \frac{\sigma_{Z,o}}{\mu},$$

$$T_{sup}^* = T_{inf}^* = I_{sup}^* = I_{inf}^* = 0, \quad D_{sup}^* = -D_{inf}^* = (1-p)\sigma_{D,o}. \tag{30}$$

Remark 5 (Little’s law in LIL) Similar to the observations in Remark 2, the LIL Little’s Law for Q and Z holds in the UL and CL cases, i.e., $Q_{sup}^* = \mu Z_{sup}^*$ and $Q_{inf}^* = \mu Z_{inf}^*$, but fails in the OL case, i.e., $Q_{sup}^* \neq \mu Z_{sup}^*$ and $Q_{inf}^* \neq \mu Z_{inf}^*$. Note that the workload Z keeps track of the total amount of unfinished service times while the queue length Q only counts the number of unfinished customers. Although, as time goes to infinity in an OL system, many customers will never be served so their service variability will play no role in the FLIL limit of Q . Nevertheless, their service variability still make an impact to the workload Z (this explains why Little’s law holds in the OL case (30) only if we set $c_s = 0$).

Remark 6 (LIL for $GI/GI/1$) Setting $m = 1, c_b = 0$ and $p = 0$ yields the $GI/GI/1$ special case. In particular, (28)–(30) hold with $\sigma_{uc}, \sigma_{D,u}, \sigma_{Z,o}, \sigma_{D,o}$ and σ_o defined in (27). For example,

$$Q_{sup}^* = \begin{cases} 0, & \text{if } \rho < 1, \\ \sqrt{\alpha(c_a^2 + c_s^2)}, & \text{if } \rho = 1, \\ \sqrt{\alpha c_a^2 + \mu c_s^2}, & \text{if } \rho > 1, \end{cases} \quad Q_{inf}^* = \begin{cases} 0, & \text{if } \rho \leq 1, \\ -\sqrt{\alpha c_a^2 + \mu c_s^2}, & \text{if } \rho > 1, \end{cases}$$

$$Z_{sup}^* = \begin{cases} 0, & \text{if } \rho < 1, \\ \frac{1}{\mu} \sqrt{\alpha(c_a^2 + c_s^2)}, & \text{if } \rho \geq 1, \end{cases} \quad Z_{inf}^* = \begin{cases} 0, & \text{if } \rho \leq 1, \\ -\frac{1}{\mu} \sqrt{\alpha(c_a^2 + c_s^2)}, & \text{if } \rho > 1. \end{cases}$$

In Sect. D of “Appendix” section, we conduct sensitivity analysis for the FLIL and LIL limits: we study the impact of all model input parameters, e.g., the mean values, variabilities, upon the FLIL and LIL limits. Also see Sects. 6 and E in “Appendix” section for related numerical studies.

5 Proofs

In this section, we prove our main results. We first provide two preliminary results (i) FLIL for BM (Strassen 1964) (Lemma 1) and (ii) strong approximations for $GI^B/GI^F/N$ (Theorem 6). Next, we apply these results to prove Theorems 2, 3, 4 and 5.

Strassen (1964) firstly developed the FLIL for the k -dimensional standard BM.

Lemma 1 (Strassen’s FLIL for BM Strassen 1964) *Let $W_n(t) \equiv W(nt)/\sqrt{n \log \log n}$, $n \geq 3$, $t \in [0, 1]$ where W is a k -dimensional standard BM, then $W_n \Rightarrow \mathcal{G}_k(1)$ w.p.1, $k = 1, 2, \dots$*

We next give the SAs and leave the proof in Sect. C.2 of “Appendix” section.

Theorem 6 (Strong approximations for $GI^B/GI^F/N$) *Suppose (3) is satisfied.*

(i) *If $N \geq 1$ and $\rho \geq 1$, then, for some $r > 2$, w.p.1,*

$$\sup_{0 \leq t \leq L} |Q(t) - \tilde{Q}(t)| = o(L^{1/r}), \quad \sup_{0 \leq t \leq L} |(1-p)\mu I(t) - \Psi(\tilde{X})(t)| = o(L^{1/r}), \tag{31}$$

where $\tilde{Q}(t) = \Phi(\tilde{X})(t)$, and

$$\begin{aligned} \tilde{X}(t) = & \bar{X}(t) + m\alpha^{1/2}c_a W_a(t) + mc_b W_b(\alpha t) - (1-p)\mu^{1/2} \sum_{j=1}^N c_{s,j} W_{s,j}(t) \\ & + \sqrt{p(1-p)} \sum_{j=1}^N W_{f,j}(\mu t), \end{aligned} \tag{32}$$

and $W_b, W_a, W_{s,j}$ and $W_{f,j}$ are mutually independent BMs associated with the batches, arrival, service and feedback, $j = 1, 2, \dots, N$.

(ii) *If $N = 1$, then*

$$\begin{aligned} \sup_{0 \leq t \leq L} |Q(t) - \tilde{Q}(t)| = o(L^{1/r}), \quad \sup_{0 \leq t \leq L} |T(t) - \tilde{T}(t)| = o(L^{1/r}), \\ \sup_{0 \leq t \leq L} |Z(t) - \tilde{Z}(t)| = o(L^{1/r}), \quad \sup_{0 \leq t \leq L} |I(t) - \tilde{I}(t)| = o(L^{1/r}), \\ \sup_{0 \leq t \leq L} |D(t) - \tilde{D}(t)| = o(L^{1/r}), \quad \text{for some } r > 2, \end{aligned} \tag{33}$$

where $W_s \equiv W_{s,1}$, $W_f \equiv W_{f,1}$, and $\tilde{T}(t)$ is given in (16),

$$\begin{aligned} (\tilde{Q}, \tilde{Y}) = & (\Phi, \Psi)(\tilde{X}), \quad \tilde{I}(t) = \frac{1}{(1-p)\mu} \tilde{Y}(t), \quad \tilde{T}(t) = t - \tilde{I}(t), \\ \tilde{X}(t) = & \bar{X}(t) + m\alpha^{1/2}c_a W_a(t) + mc_b W_b(\alpha t) - (1-p)\mu^{1/2}c_s W_s(\tilde{T}(t)) \\ & + \sqrt{p(1-p)}W_f(\mu\tilde{T}(t)), \end{aligned}$$

$$\begin{aligned} \tilde{Z}(t) &= \frac{1}{\mu} [\tilde{Q}(t) + \mu^{1/2} c_s (W_s(\bar{T}(t)) - W_s(\rho t))], \\ \tilde{D}(t) &= (1 - p) [\mu \tilde{T}(t) + \mu^{1/2} c_s W_s(\bar{T}(t))] - \sqrt{p(1 - p)} W_f(\mu \bar{T}(t)). \end{aligned} \quad (34)$$

Remark 7 $\tilde{X}(t) - \bar{X}(t)$ is a driftless BM with variance parameter $\sigma_o^2(N)$ defined in (22) in Theorem 2 if $N \geq 1$, and σ^2 defined in Theorem 3 if $N = 1$.

5.1 Proof of Theorem 2

For all $t \in [0, 1]$ and $n = 3, 4, \dots$, define

$$\tilde{Q}^n(t) = \frac{\tilde{Q}(nt) - \bar{Q}(nt)}{\varphi(n)}, \quad (35)$$

where \bar{Q} and \tilde{Q} are defined in Theorem 1 and Lemma 6, respectively. By Lemma 6, since $L^{1/r} = o(\varphi(L))$ for all $r > 2$, we have, for all $t \in [0, 1]$, w.p.1,

$$\lim_{n \rightarrow \infty} \frac{|Q(nt) - \tilde{Q}(nt)|}{\varphi(n)} \leq \limsup_{n \rightarrow \infty} \frac{\sup_{0 \leq s \leq n} |Q(s) - \tilde{Q}(s)|}{\varphi(n)} = \limsup_{n \rightarrow \infty} \frac{o(n^{1/r})}{\varphi(n)} = 0.$$

So, for all $t \in [0, 1]$,

$$\lim_{n \rightarrow \infty} \frac{Q(nt) - \tilde{Q}(nt)}{\varphi(n)} = 0, \quad \text{w.p.1.} \quad (36)$$

Note that, by (18),

$$Q^n(t) = \frac{Q(nt) - \tilde{Q}(nt)}{\varphi(n)} + \tilde{Q}^n(t).$$

This, and (36), implies that it suffices to prove $\tilde{Q}^n \Rightarrow \mathcal{K}_Q$ if one tries to prove $Q^n \Rightarrow \mathcal{K}_Q$.

Similarly we define $\tilde{X}^n(t)$ in the same token of (35). It is only needed to prove $\tilde{X}^n \Rightarrow \mathcal{K}_X$ if one tries to prove $X^n \Rightarrow \mathcal{K}_X$, where \mathcal{K}_X is some compact set of absolutely continuous functions.

The CL case. If $\rho = 1$, by Theorem 1, $\bar{X}(t) = 0$ for all $t \geq 0$, by Lemma 6, $\tilde{X}(t)$ is a driftless BM with variance parameter $\sigma_o^2(N)$. Then, for all $t \in [0, 1]$, $\tilde{X}^n(t) \Rightarrow \mathcal{G}(\sigma_o(N))$ w.p.1. Since Φ is a continuous mapping under uniform topology, for all $t \in [0, 1]$, $\tilde{Q}^n(t) \Rightarrow \Phi(\mathcal{G}(\sigma_o(N)))$ w.p.1.

To prove (23), we note that

$$(1 - p)\mu I^n(t) = \frac{1}{\varphi(n)} \left[(1 - p)\mu \sum_{j=1}^N I_j(nt) - \Psi(\tilde{X})(nt) \right] + \frac{1}{\varphi(n)} \Psi(\tilde{X})(nt).$$

For all $t \in [0, 1]$, by (31),

$$\left| \frac{1}{\varphi(n)} \left[(1 - p)\mu \sum_{j=1}^N I_j(nt) - \Psi(\tilde{X})(nt) \right] \right| \leq \frac{n^{1/r}}{\varphi(n)} \rightarrow 0 \quad \text{as } n \rightarrow \infty,$$

and by Theorem 7 (Strassen’s CMT),

$$\frac{1}{\varphi(n)} \Psi(\tilde{X})(nt) \rightrightarrows \Psi(\mathcal{G}(\sigma_o(N)))$$

because $\tilde{X}^n(t) \rightrightarrows \mathcal{G}(\sigma_o(N))$ w.p.1 for all $t \in [0, 1]$. So, (23) holds.

The OL case. If $\rho > 1$, then by (32) and Theorem 1 $\tilde{X}(t)$ is a BM with positive drift $\theta_N > 0$, which implies that $\lim_{t \rightarrow \infty} \tilde{X}(t) = +\infty$ w.p.1. So, by the definition of continuous mapping Ψ , $\sup_{t \geq 0} \Psi(\tilde{X})(t) < \infty$ w.p.1. As a result, for all $t \in [0, 1]$, $\Psi(\tilde{X})(nt)/\varphi(n) \rightarrow 0$ w.p.1. as $n \rightarrow \infty$. This, and (31), implies that $(1 - p)\mu \sum_{j=1}^N I_j^n(t) \rightarrow 0$ w.p.1 as $n \rightarrow \infty$. Notice that $(1 - p)\mu > 0$ and $I_j(t) \geq 0$, then for $j = 1, 2, \dots, N$ and $t \in [0, 1]$, $I_j^n(t) \rightarrow 0$ w.p.1 as $n \rightarrow \infty$. Since $T_j(t) = t - I_j(t)$, for $j = 1, 2, \dots, N$ and $t \in [0, 1]$, $T_j^n(t) \rightarrow 0$ w.p.1 as $n \rightarrow \infty$. For the FLIL of Q , since $\rho > 1$, we have $\bar{Q}(t) = \tilde{X}(t)$ and then $\tilde{Q}(t) - \bar{Q}(t) = \tilde{X}(t) - \bar{X}(t) + \Psi(\tilde{X})(t)$. This follows that, for all $t \in [0, 1]$,

$$\tilde{Q}^n(t) = \tilde{X}^n(t) + \frac{\Psi(\tilde{X})(nt)}{\varphi(n)} \rightrightarrows \mathcal{G}(\sigma_o(N)), \quad \text{w.p.1,}$$

because that $\tilde{X}(t) - \bar{X}(t)$ is a driftless BM with variance parameter $\sigma_o^2(N)$. □

5.2 Proof of Theorem 3

As in the proof of Theorem 2, we define $\tilde{Y}^n(t), \tilde{I}^n(t), \tilde{B}^n(t), \tilde{Z}^n(t), \tilde{D}^n(t)$ in the same token of (35), and we transfer the original problem (19) for the $GI^B/GI/1$ queue into the following problem:

$$(\tilde{Q}^n, \tilde{Z}^n, \tilde{I}^n, \tilde{B}^n, \tilde{D}^n) \rightrightarrows \mathcal{K}^*, \quad \text{w.p.1,} \tag{37}$$

where \mathcal{K}^* is given in (19).

The UL case. If $\rho < 1$, then, $\rho = m\alpha/((1 - p)\mu)$, and by (16),

$$(\bar{Q}, \bar{Z}, \bar{X}, \bar{Y}, \bar{T}, \bar{I}, \bar{D})(t) = (0, 0, \theta_1 t, -\theta_1 t, \rho t, (1 - \rho)t, m\alpha t), \tag{38}$$

and by (34), $\tilde{X}(t)$ is a BM with negative drift $\theta_1 < 0$ and variance parameter σ_{uc}^2 . So, \tilde{Q} is a reflected BM (i.e., $\Phi(\tilde{X})$) with negative drift $\theta_1 < 0$. Therefore, it follows from Theorem 6.3 in Chen and Yao (2001) that

$$\sup_{0 \leq t \leq L} \tilde{Q}(t) = O(\log L), \quad \text{w.p.1.}$$

Because $O(\log L)/\varphi(L) \rightarrow 0$ as $L \rightarrow \infty$, we have

$$\frac{\sup_{0 \leq t \leq L} \tilde{Q}(t)}{\varphi(L)} \rightarrow 0, \quad \text{w.p.1, as } L \rightarrow \infty.$$

Since $\bar{Q}(t) = 0$ for any $t \geq 0$, we have, for all $t \in [0, 1]$,

$$\tilde{Q}^n(t) \rightarrow 0 \quad \text{w.p.1, as } n \rightarrow \infty, \tag{39}$$

and furthermore, for all $t \in [0, 1]$, $\tilde{Q}^n(t) \rightrightarrows 0$ w.p.1 as $n \rightarrow \infty$. For the FLIL of workload Z , since $\tilde{Z}^n(t) = \tilde{Q}^n(t)/\mu$ from (34), we have, for all $t \in [0, 1]$, $\tilde{Z}^n(t) = \tilde{Q}^n(t)/\mu \rightrightarrows 0$ w.p.1. as $n \rightarrow \infty$.

For the FLIL of the idle time I , since $\tilde{Q}(t) = \tilde{X}(t) + \tilde{Y}(t)$, we have

$$\tilde{Y}(t) - \bar{Y}(t) = \tilde{Q}(t) - \tilde{X}(t) - \bar{Y}(t) = \tilde{Q}(t) - [\tilde{X}(t) - \bar{X}(t)].$$

This, and Remark 7, implies that, for all $t \in [0, 1]$,

$$\tilde{Y}^n(t) = \tilde{Q}^n(t) - \frac{\tilde{X}(nt) - \bar{X}(nt)}{\varphi(n)} \Rightarrow \mathcal{G}(\sigma_{uc}), \quad \text{w.p.1,}$$

because $\tilde{Q}^n(t) \Rightarrow 0$, where σ_{uc} is defined in (26). So, for all $t \in [0, 1]$,

$$\tilde{I}^n(t) = \frac{1}{\mu(1-p)} \tilde{Y}^n(t) \Rightarrow \frac{1}{\mu(1-p)} \mathcal{G}(\sigma_{uc}), \quad \text{w.p.1.}$$

For the FLIL of the busy time T , since $\tilde{T}(t) - \bar{T}(t) = \tilde{I}(t) - \bar{I}(t)$, it follows that, for all $t \in [0, 1]$,

$$\tilde{T}^n(t) = -\tilde{I}^n(t) \Rightarrow -\frac{1}{\mu(1-p)} \mathcal{G}(\sigma_{uc}), \quad \text{w.p.1.}$$

For the FLIL of departure D , since, by (16) and (34),

$$\begin{aligned} \tilde{D}(t) - \bar{D}(t) &= (1-p) \{ \mu [\tilde{T}(t) - \bar{T}(t)] + \mu^{1/2} c_s W_s(\rho t) \} - \sqrt{p(1-p)} W_f(\mu \bar{T}(t)) \\ &= (1-p) \{ \mu [\tilde{I}(t) - \bar{I}(t)] + \mu^{1/2} c_s W_s(\rho t) \} - \sqrt{p(1-p)} W_f(\mu \rho t) \\ &= [\tilde{Q}(t) - \bar{Q}(t)] + [\tilde{X}(t) - \bar{X}(t)] + (1-p) \mu^{1/2} c_s W_s(\rho t) - \sqrt{p(1-p)} W_f(\mu \rho t) \\ &= -\tilde{Q}(t) + m\alpha^{1/2} c_a W_a(t) + m c_b W_b(\alpha t), \end{aligned}$$

we have, for all $t \in [0, 1]$,

$$\tilde{D}^n(t) = -\tilde{Q}^n(t) + \frac{m\alpha^{1/2} c_a W_a(nt) + m c_b W_b(\alpha nt)}{\varphi(n)} \Rightarrow \mathcal{G} \left(\sqrt{\alpha m^2 (c_a^2 + c_b^2)} \right), \quad \text{w.p.1,}$$

because $\tilde{Q}^n(t) \rightarrow 0$ w.p.1. and $m\alpha^{1/2} c_a W_a(t) + m c_b W_b(\alpha t)$ is a driftless BM with variance $m^2 \alpha (c_a^2 + c_b^2)$.

The CL case. If $\rho = 1$, then $m\alpha = (1-p)\mu$ and

$$(\bar{Q}, \bar{Z}, \bar{X}, \bar{Y}, \bar{T}, \bar{I}, \bar{D})(t) = (0, 0, 0, 0, t, 0, (1-p)\mu t), \quad (40)$$

and as in the proof of Theorem 2, for all $t \in [0, 1]$,

$$\tilde{X}^n(t) = \frac{\tilde{X}(nt)}{\varphi(n)} \Rightarrow \mathcal{G}(\sigma_o), \quad \text{w.p.1,}$$

where σ_o is defined in Theorem 3. By continuous mapping theorem (CMT) B.3, for all $t \in [0, 1]$, w.p.1,

$$\tilde{Q}^n(t) \Rightarrow \Phi(\mathcal{G}(\sigma_o)) \text{ and } \tilde{Y}^n(t) \Rightarrow \Psi(\mathcal{G}(\sigma_o)).$$

For the FLILs of I , T and Z , we have, for all $t \in [0, 1]$, w.p.1,

$$\begin{aligned} \tilde{I}^n(t) &= \frac{1}{\mu(1-p)} \tilde{Y}^n(t) \Rightarrow \frac{1}{\mu(1-p)} \Psi(\mathcal{G}(\sigma_o)), \\ \tilde{T}^n(t) &= -\tilde{I}^n(t) \Rightarrow -\frac{1}{\mu(1-p)} \Psi(\mathcal{G}(\sigma_o)), \\ \tilde{Z}^n(t) &= \frac{1}{\mu} \tilde{Q}^n(t) \Rightarrow \frac{1}{\mu} \Phi(\mathcal{G}(\sigma_o)). \end{aligned}$$

For the FLIL of D , we note that, from (34),

$$\begin{aligned} & \tilde{D}(t) - \bar{D}(t) \\ &= (1 - p) [-\mu\tilde{I}(t) + \mu^{1/2}c_s W_s(t)] - \sqrt{p(1 - p)}W_f(\mu\bar{T}(t)) \\ &= - \sup_{0 \leq s \leq t} [-\tilde{X}(s)] + (1 - p)\mu^{1/2}c_s W_s(t) - \sqrt{p(1 - p)}W_f(\mu t) \\ &= \inf_{0 \leq s \leq t} [\tilde{X}(s)] + (1 - p)\mu^{1/2}c_s W_s(t) - \sqrt{p(1 - p)}W_f(\mu t) \\ &= \inf_{0 \leq s \leq t} \left[m\alpha^{1/2}c_a W_a(s) + mc_b W_b(\alpha s) - (1 - p)\mu^{1/2}c_s W_s(s) + \sqrt{p(1 - p)}W_f(\mu s) \right] \\ &\quad + (1 - p)\mu^{1/2}c_s W_s(t) - \sqrt{p(1 - p)}W_f(\mu t) \\ &= h(m\alpha^{1/2}c_a W_a(t) + mc_b W_b(\alpha t), (1 - p)\mu^{1/2}c_s W_s(t) - \sqrt{p(1 - p)}W_f(\mu t)). \end{aligned}$$

By CMT B.3, for all $t \in [0, 1]$,

$$\begin{aligned} & \tilde{D}^n(t) \\ &= \frac{h(m\alpha^{1/2}c_a W_a(nt) + mc_b W_b(\alpha nt), (1 - p)\mu^{1/2}c_s W_s(nt) - \sqrt{p(1 - p)}W_f(\mu nt))}{\varphi(n)} \\ &\Rightarrow h(\mathcal{G}_2(\sigma_o)), \quad \text{w.p.1,} \end{aligned}$$

because, by Corollary 2.2 in Iglehart (1971), for all $t \in [0, 1]$, w.p.1,

$$\begin{aligned} & \left(\frac{m\alpha^{1/2}c_a W_a(nt) + mc_b W_b(\alpha nt)}{\varphi(n)}, \frac{(1 - p)\mu^{1/2}c_s W_s(nt) - \sqrt{p(1 - p)}W_f(\mu nt)}{\varphi(n)} \right) \\ &\Rightarrow \mathcal{G}_2(\sigma_o). \end{aligned}$$

The OL case. If $\rho > 1$, then by (6), $\rho = (m\alpha + p\mu)/\mu$ and

$$(\bar{Q}, \bar{Z}, \bar{X}, \bar{Y}, \bar{T}, \bar{I}, \bar{D})(t) = (\theta_1 t, (\rho - 1)t, \theta_1 t, 0, t, 0, (1 - p)\mu t), \quad (41)$$

where $\bar{Q}(t) = \theta_1 t = \mu(\rho - 1)t = \mu\bar{Z}(t)$. As in the proof of Theorem 2, $\sup_{t \geq 0} \tilde{Y}(t) = \sup_{t \geq 0} \Psi(\tilde{X})(t) < \infty$ w.p.1., and $\tilde{Y}^n(t) \rightarrow 0$ w.p.1. as $n \rightarrow \infty$. Hence $\tilde{I}^n(t) = \tilde{Y}^n(t)/[(1 - p)\mu] \Rightarrow 0$ as $n \rightarrow \infty$. For the FLIL of T , since $\tilde{T}(t) - \bar{T}(t) = -\tilde{I}(t)$, we have, for all $t \in [0, 1]$, $\tilde{T}^n(t) = -\tilde{I}^n(t) \Rightarrow 0$ w.p.1.

For the FLIL of Q , as in the proof of Theorem 2, $\tilde{Q}(t) - \bar{Q}(t) = \tilde{X}(t) - \bar{X}(t) + \tilde{Y}(t)$, and, for all $t \in [0, 1]$,

$$\tilde{Q}^n(t) = \frac{\tilde{X}(nt)}{\varphi(n)} + \tilde{Y}^n(t) \Rightarrow \mathcal{G}(\sigma_o), \quad \text{w.p.1.}$$

For the FLIL of Z , since, by (16) and (34),

$$\begin{aligned} & \tilde{Z}(t) - \bar{Z}(t) \\ &= \frac{1}{\mu} [\tilde{Q}(t) + \mu^{1/2}c_s (W_s(t) - W_s(\rho t))] - \frac{1}{\mu}\bar{Q}(t) \\ &= \frac{1}{\mu} [\tilde{X}(t) + \tilde{Y}(t) - \bar{X}(t) + \mu^{1/2}c_s (W_s(t) - W_s(\rho t))] \equiv \frac{1}{\mu}W_{Z,o}(t) + \frac{1}{\mu}\tilde{Y}(t), \end{aligned}$$

where

$$W_{Z,o}(t) \equiv m\alpha^{1/2}c_a W_a(t) + mc_b W_b(\alpha t) + p\mu^{1/2}c_s W_s(t) - \mu^{1/2}c_s W_s(\rho t) + \sqrt{p(1-p)}W_f(\mu t).$$

Notice that $p\mu^{1/2}c_s W_s(t) - \mu^{1/2}c_s W_s(\rho t) \stackrel{d}{=} c_s(p\mu^{1/2} - (m\alpha + p\mu)^{1/2})W_s(t)$, then $W_{Z,o}(t)$ is a driftless BM with variance parameter $\sigma_{Z,o}^2$ defined in (26). This, and $\tilde{Y}^n(t)/\mu \Rightarrow 0$ w.p.1 for all $t \in [0, 1]$, implies that, for all $t \in [0, 1]$,

$$\tilde{Z}^n(t) = \frac{1}{\mu} \frac{W_{Z,o}(nt)}{\varphi(n)} + \frac{1}{\mu} \tilde{Y}^n(t) \Rightarrow \frac{1}{\mu} \mathcal{G}(\sigma_{Z,o}), \quad \text{w.p.1.}$$

For the FLIL of D , since

$$\begin{aligned} \tilde{D}(t) - \bar{D}(t) &= (1-p) [\mu (\tilde{T}(t) - \bar{T}(t)) + \mu^{1/2}c_s W_s(\bar{T}(t))] - \sqrt{p(1-p)}W_f(\mu \bar{T}(t)) \\ &= (1-p) \left[-\mu \tilde{T}(t) + \mu^{1/2}c_s W_s(t) - \sqrt{\frac{p}{1-p}}W_f(\mu t) \right], \end{aligned}$$

we have, for all $t \in [0, 1]$,

$$\begin{aligned} \tilde{D}^n(t) &= (1-p) \left[-\mu \tilde{T}^n(t) + \frac{\mu^{1/2}c_s W_s(nt) - \sqrt{\frac{p}{1-p}}W_f(n\mu t)}{\varphi(n)} \right] \\ &\Rightarrow (1-p)\mathcal{G}(\sigma_{D,o}), \quad \text{w.p.1,} \end{aligned}$$

where $\sigma_{D,o}$ is defined in (26).

Hence, the result is proved. □

5.3 Proofs of Theorems 4 and 5

Since the proof of Theorem 4 is similar to that of Theorem 5, we only prove Theorem 5.

(i) Case $\rho < 1$. By (24), since $\mathcal{K}_Q = \mathcal{K}_Z = \{0\}$, we have $Q_{sup}^* = Q_{inf}^* = Z_{sup}^* = Z_{inf}^* = 0$. For I, T and D , we firstly observe that $\sup_{x \in \mathcal{G}(\delta)} x(1) = \delta$ and $\inf_{x \in \mathcal{G}(\delta)} x(1) = -\delta$ for any $\delta > 0$, where the supremum and infimum are actually attained for the functions $x(t) = \delta t$ and $x(t) = -\delta t$ respectively. This and (24) imply that (28) holds.

(ii) Case $\rho = 1$. By (24), $\mathcal{K}_Q = \Phi(\mathcal{G}(\sigma_o))$. As Corollary 3.1 and 3.2 in Iglehart (1971), we have

$$\sup_{x \in \Phi(\mathcal{G}(\sigma_o))} x(1) = \sigma_o \quad \text{and} \quad \inf_{x \in \Phi(\mathcal{G}(\sigma_o))} x(1) = 0,$$

where the supremum and the infimum are actually attained for the functions $x(t) = \sigma_o t$ and $x(t) = 0$ respectively. That is, $Q_{sup}^* = \sigma_o$ and $Q_{inf}^* = 0$. Notice that $\mathcal{K}_Z = \Phi(\mathcal{G}(\sigma_o))/\mu$ from (24), we get Z_{sup}^* and Z_{inf}^* in (29).

For I , we firstly note that $\mathcal{K}_I = (1/(1-p)\mu)\Psi(\mathcal{G}(\sigma_o))$. On the one hand, Strassen has shown that $\mathcal{G}(\delta)$ is compact in $\mathbb{C}^1[0, 1]$ for any $\delta > 0$ and that $|x(b) - x(a)| \leq \delta(b-a)^{1/2}$

for any $y \in \mathcal{G}(\delta)$ and $0 \leq a \leq b \leq 1$. Reader can also see Iglehart (1971). This follows $|y(b)| \leq \delta\sqrt{b} \leq \delta$ for any $y \in \mathcal{G}(\delta)$ and $0 \leq b \leq 1$, then,

$$\sup_{x \in \Psi(\mathcal{G}(\sigma_o))} x(1) = \sup_{y \in \mathcal{G}(\sigma_o)} \sup_{0 \leq s \leq 1} \{-y(s)\} = \sigma_o,$$

where the supremum is attained for the function $y(s) = -s$. On the other hand, notice $y(0) = 0$ for any $y \in \mathcal{G}(\delta)$, then

$$\inf_{x \in \Psi(\mathcal{G}(\sigma_o))} x(1) = \inf_{y \in \mathcal{G}(\sigma_o)} \sup_{0 \leq s \leq 1} \{-y(s)\} = 0,$$

where the infimum is attained for the function $y(s) = 0$. From above, $I_{sup}^* = \sigma_o / [(1 - p)\mu]$ and $I_{inf}^* = 0$.

For T , we note that $\mathcal{K}_T = -\mathcal{K}_I = -\Psi(\mathcal{G}(\sigma_o)) / [(1 - p)\mu]$, which follows the desired T_{sup}^* and T_{inf}^* in (29).

Case $\rho > 1$. The analysis is similar with the case $\rho < 1$ and omitted. □

6 Numerical examples

We now consider numerical examples to gain insights into our LIL limits. In particular, we study the sensitivity of our LIL limits to the model input parameters, including the number of servers N , the traffic intensity ρ and the variability parameters (c_a, c_b and c_s). Additional numerical examples appear in Sect. E of “Appendix” section.

Example 1 (Sensitivity to the traffic intensity ρ) We first consider a $GI^B/GI^F/1$ model with fixed variability parameter $c_a = c_b = c_s = 1$. We study the sensitivity to the traffic intensity ρ , by varying of parameters in (6): α, m, μ and p . The idea is to vary the value of ρ so that we can cover all UL, CL and OL cases.

We consider the following four cases:

- (i) Varying α with others fixed at $m = 1, \mu = 1, p = 1/2, c_a = c_b = c_s = 1$.
- (ii) Varying m with others fixed at $\alpha = 1/2, \mu = 1, p = 1/2, c_a = c_b = c_s = 1$.
- (iii) Varying μ with others fixed at $\alpha = 1/2, m = 1, p = 1/2, c_a = c_b = c_s = 1$.
- (iv) Varying p from 0 to 1 and let $\alpha = 1/2, m = 1, \mu = 1, c_a = c_b = c_s = 1$.

For case (i), we plot the key LIL limits as functions of α in Fig. 1. We observe that the LIL limits jump at $\alpha = 0.5$ which correspond to the CL case ($\rho = 1$). For cases (ii)–(iv), we plots the key LIL limits in Figs. 2, 3 and 4 in Sect. E of “Appendix” section.

Example 2 (Sensitivity to variability parameters) We next consider a $GI^B/GI^F/1$ model with α, m, μ, p fixed and study the sensitivity of the LIL limits to the variability parameters c_a, c_b, c_s . We consider the following two cases:

- (i) **Arrival variability** We fix $m = 2, \mu = 1, c_s = c_b = p = 0$ and study the dependence on c_a when the system is UL ($\alpha = 1/4$), CL ($\alpha = 1/2$) and OL ($\alpha = 1$).
- (ii) **Service variability** We fix $\alpha = 1/2, m = 2, c_a = c_b = p = 0$ and study the dependence on c_s when the system is OL ($\mu = 1/2$), CL ($\mu = 1$) and UL ($\mu = 2$).

For case (i) we give the LIL limits as functions of c_a in Table 1. We give the numerical results in Sect. E of “Appendix” section for case (ii).

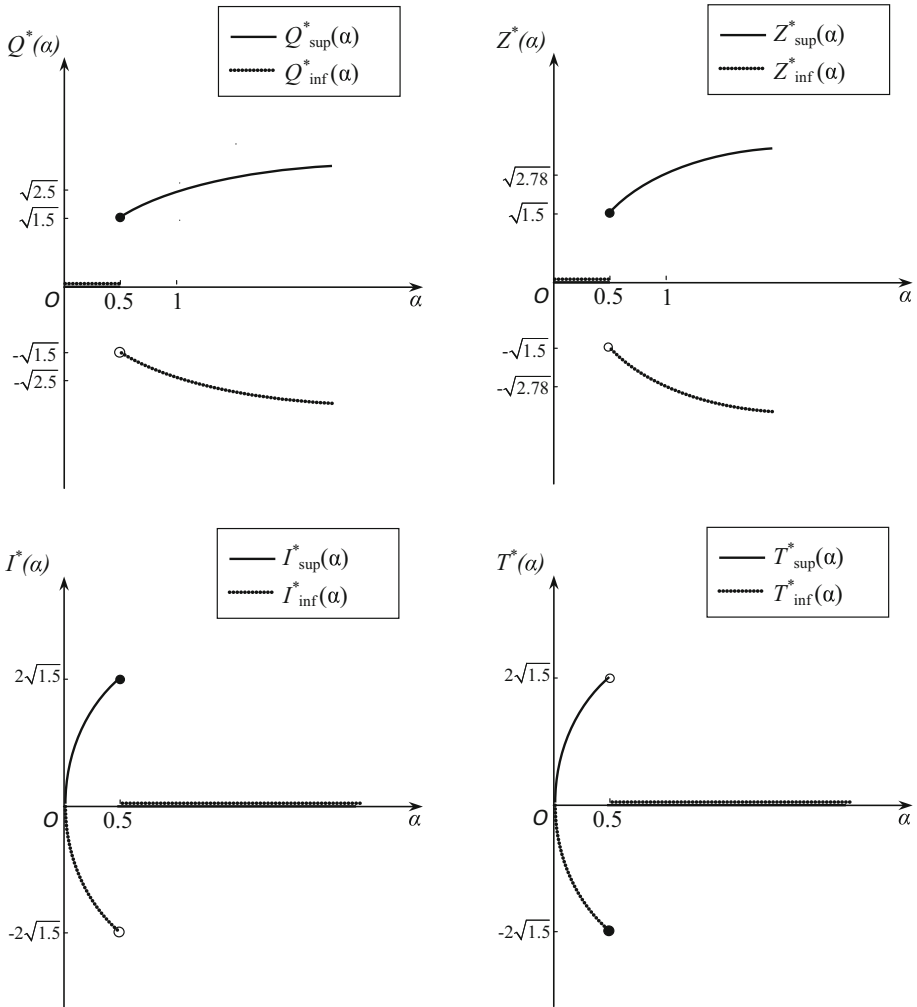


Fig. 1 LIL limits of (i) of Example 1 as functions of α , with $m = 1, \mu = 1, p = 1/2, c_a = c_b = c_s = 1$

Table 1 The superior and inferior LIL limits as functions of c_a for case (i) of Example 2

| α | $Q_{\text{sup}}^*(c_a)$ | $Z_{\text{sup}}^*(c_a)$ | $I_{\text{sup}}^*(c_a)$ | $T_{\text{sup}}^*(c_a)$ | $D_{\text{sup}}^*(c_a)$ |
|---------------|-------------------------|-------------------------|-------------------------|-------------------------|-------------------------|
| $\frac{1}{4}$ | 0 | 0 | c_a | c_a | c_a |
| $\frac{1}{2}$ | $\sqrt{2}c_a$ | $\sqrt{2}c_a$ | $\sqrt{2}c_a$ | 0 | 0 |
| 1 | $2c_a$ | $2c_a$ | 0 | 0 | 0 |
| α | $Q_{\text{inf}}^*(c_a)$ | $Z_{\text{inf}}^*(c_a)$ | $I_{\text{inf}}^*(c_a)$ | $T_{\text{inf}}^*(c_a)$ | $D_{\text{inf}}^*(c_a)$ |
| $\frac{1}{4}$ | 0 | 0 | $-c_a$ | $-c_a$ | $-c_a$ |
| $\frac{1}{2}$ | 0 | 0 | 0 | $-\sqrt{2}c_a$ | 0 |
| 1 | $-2c_a$ | $-2c_a$ | 0 | 0 | 0 |

Example 2 shows that the LIL limits are either zero or linear functions of variability parameters c_a and c_s . See Sects. D and E in “Appendix” section for more discussions on the sensitivity analysis to the input model parameters.

7 Conclusions

Using a strong approximation approach, we have developed a FLIL and its LIL for $GI^B/GI^F/N$ queue with Bernoulli feedback by focusing on five key performance processes: queue length, workload, busy time, idle time and departure processes. The FLIL and its LIL results cover all three important cases UL, CL and OL categorized by the traffic intensity. Refining the FSLLNs and the corresponding limiting fluid functions which are often used to approximate the mean values, the FLIL and LIL provide an estimate for the asymptotic rate of the increasing stochastic variability of these performance functions in functional set version and numerical version. We have identified these FLIL and LIL as explicit mathematical expressions of the first and second moments of the batch, batch-interarrival, service time and feedback. Comprehensive discussions and numerical experiments have been provided to gain insights of these FLIL and LIL limits.

Our SA-based approach can be a viable tool to establish FLIL and LIL limits for more general queueing systems, especially network models having a multi-dimensional state process, such as the generalized Jackson network, reentrant models and feedback systems (Liu and Whitt 2014b). To do so, the first step is to obtain the heavy-traffic fluid limits of the designated network queue model (Liu and Whitt 2014a). Based on the fluid model, we need to next develop the strong approximations (Mandelbaum and Massey 1995). Depending on the values of the traffic intensities in a queueing network, the SA results can be multi-dimensional reflected BMs with different drift (positive, zero or negative). We plan to next develop the FLIL and LIL results for the reentrant network queueing system.

Acknowledgements The first author acknowledge support from NSFC Grant 11471053. The second author also acknowledges support from NSF Grant CMMI 1362310.

Appendix

Overview. This appendix contains additional materials supplementing the main paper. In Sect. A, we summarize all acronyms used in this paper. In Sect. B, we give additional preliminary results, including one-dimensional ORM (Sect. B.1), discussions on queue length and idle times (Sect. B.2) and the continuous mapping theorem for FLIL (Sect. B.3). In Sect. C we give additional proofs, including the proofs of Theorem 1 (Sect. C.1) and Lemmas 6 (Sect. C.2). In Sect. D, we study the sensitivity of the FLIL and LIL limits to the input parameters. In Sect. E, additional numerical examples are given to supplement Sect. 6.

A Summary of the abbreviated words

We give a glossary of all acronyms used in the main paper in Table 2.

Table 2 Summary of all acronyms used in the main paper (in the alphabetic order)

| Acronyms | Meanings |
|----------|--|
| BM | Brownian motion |
| CL | Critically loaded |
| CMT | Continuous mapping theorem |
| FLIL | Functional law of the iterated logarithm |
| FSLLN | Functional strong law of large numbers |
| i.i.d. | Independent and identically distributed |
| LIL | Law of the iterated logarithm |
| OL | Overloaded |
| u.o.c. | Uniformly over compact sets |
| ORM | Oblique reflection mapping |
| SA | Strong approximation |
| SCV | Squared coefficient of variation |
| UL | Underloaded |
| w.p.1 | With probability 1 |

B Additional preliminary results

In this section, we give additional preliminary results that will be used in the proofs.

B.1 The oblique reflection mapping

We now provide an alternative definition of the ORM. Define $\mathbb{D}_b \equiv \{x \in \mathbb{D}[0, \infty) : x(0) \geq b\}$ for any $b \in \mathbb{R}$, and $\mathbb{D}_\uparrow \equiv \{x \in \mathbb{D}_0 : x \text{ is non-decreasing}\}$.

Definition 1 (The one-dimensional ORM Harrison 1985) For any function $x \in \mathbb{D}_0$, if there exists a unique pair of functions $(z, y) \in \mathbb{D}_0^2$ satisfying

- (i) $z(t) = x(t) + y(t) \geq 0$;
- (ii) y is nondecreasing and $y(0) = 0$;
- (iii) $\int_0^\infty z(t) \mathbf{d}y(t) = 0$,

then the map from x to (y, z) is called the one-dimensional *oblique reflection mapping*, denoted by $(z, y) = (\Phi, \Psi)(x)$.

B.2 Analysis of queue length and idle times

This section facilitates the treatment of (11). Define

$$\mathcal{H}_1 \equiv \{y \in \mathbb{D}_\uparrow : x(t) + y(t) \geq 0, \forall t \geq 0\},$$

$$\mathcal{H}_2 \equiv \{y \in \mathbb{D}_\uparrow : [x(t) + y(t)] \mathbf{d}y(t) \geq 0, \forall t \geq 0\}.$$

For $x \in \mathbb{D}_0$, it is proved in Reiman (1984) that, $\Psi(x)$ is the least element of \mathcal{H}_1 , that is, for any $y \in \mathcal{H}_1$, $y(t) \geq \Psi(x)(t)$; and in Chen and Shanthikumar (1994) that $\Psi(x)$ is the maximum element of \mathcal{H}_2 , that is, for any $y \in \mathcal{H}_2$, $y(t) \leq \Psi(x)(t)$.

Let $x \in \mathbb{D}_b$ with $b \in \mathbb{R}_+$, $y \in \mathbb{D}_\uparrow$, $z \in \mathbb{D}$, and suppose that

$$z = x + y \geq 0, \quad (b - z(t)) \mathbf{d}y(t) \geq 0, \quad \forall t \geq 0. \tag{42}$$

Then, together with Theorem 2.2 in [Chen and Shanthikumar \(1994\)](#), we have

$$\Psi(x)(t) \leq y(t) \leq \Psi(x - b)(t), \quad \forall t \geq 0. \tag{43}$$

Comparing (8) and (11) with (42) and (43), we have

$$\Psi(X)(t) \leq (1 - p)\mu \sum_{j=1}^N I_j(t) \leq \Psi(X - N)(t). \tag{44}$$

B.3 The continuous mapping theorem for FLIL

The following result is a Corollary of Theorem 3 in [Strassen \(1964\)](#).

Theorem 7 (Strassen’s CMT) *Let $\{x_n : n \geq 1\}$ be a relatively compact sequence in $\mathbb{C}^k[0, 1]$ endowed with the uniform norm and with the compact set \mathcal{G}_k as its set of limit points. If f is a continuous function on $\mathbb{C}^k[0, 1]$ into some metric space \mathbb{S} with Borel sets ψ , then the sequence $\{f(x_n) : n \geq 1\}$ is relatively compact in (\mathbb{S}, ψ) and the set of its limit points coincides with $f(\mathcal{G}_k)$, a compact set.*

C Additional proofs

C.1 Proof of Theorem 1

We only prove (i) because (ii) is similar to Sect. 6.3 in [Chen and Yao \(2001\)](#). Since $0 \leq \bar{T}_j^{(n)}(t) - \bar{T}_j^{(n)}(s) \leq t - s$ for any $0 \leq s \leq t$, $\{\bar{T}_j^{(n)}, n \geq 1\}$ is uniformly Lipschitz continuous and hence equicontinuous. Therefore, the proof of the u.o.c. convergence reduces to the proof that all convergent subsequences must converge to the same limit. That is, it is sufficient to show that $\bar{T}_j^{(n_l)}(t) \rightarrow \bar{T}_j(t)$ u.o.c. as $l \rightarrow \infty$, which implies that $\bar{T}_j^{(n)}(t) \rightarrow \bar{T}_j(t)$ u.o.c. as $n \rightarrow \infty$, where $\{n_l, l \geq 1\}$ is a sequence of integers convergent to infinity. To simplify the notation, we assume without loss of generality that $n_l = l$.

For the case $\rho \geq 1$, by FSLLN, we have $\bar{X}^{(n)}(t) \rightarrow \theta_N t \geq 0$ u.o.c. w.p.1 as $n \rightarrow \infty$, which implies that, as $n \rightarrow \infty$, w.p.1,

$$\Psi(\bar{X}^{(n)})(t) \rightarrow 0, \quad \text{u.o.c.}, \quad \text{and} \quad \Psi\left(\bar{X}^{(n)} - \frac{N}{n}\right)(t) \rightarrow 0, \quad \text{u.o.c.}$$

This, together with (44), yields the convergence below:

$$(1 - p)\mu \sum_{j=1}^N \bar{I}_j^{(n)}(t) \rightarrow 0 \equiv (1 - p)\mu \sum_{j=1}^N \bar{I}_j(t), \quad \text{u.o.c.}, \quad \text{w.p.1.}$$

Because $\bar{I}_i(t) \geq 0$ by definition, $\bar{I}_i(t) = 0$ for any $i = 1, 2, \dots, N$. This implies that $\bar{T}_i(t) = t$ for any i . Therefore, $\bar{I}(t) = 0$, $\bar{T}(t) = t$ and $\bar{Q}(t) = \theta_N t$. □

C.2 Proof of Theorem 6

Lemma 2 *Suppose $\rho \geq 1$. If (3) holds, then, for all j ,*

$$\sup_{0 \leq t \leq L} |T_j(t) - t| = O\left(\sqrt{L \log \log L}\right) \quad \text{w.p.1.}$$

Proof Since (3) holds, we have, w.p.1.

$$\begin{aligned} \sup_{0 \leq t \leq L} |A(t) - \alpha t| &= O\left(\sqrt{L \log \log L}\right), & \sup_{0 \leq t \leq L} |S_j(t) - \mu t| &= O\left(\sqrt{L \log \log L}\right), \\ \sup_{0 \leq t \leq L} |V_j(t) - t/\mu| &= O\left(\sqrt{L \log \log L}\right), & \sup_{0 \leq t \leq L} |B(t) - mt| &= O\left(\sqrt{L \log \log L}\right), \\ \sup_{0 \leq t \leq L} |\Gamma_j(t) - pt| &= O\left(\sqrt{L \log \log L}\right), & j &= 1, 2, \dots, N, \end{aligned}$$

where $V_j(t) = V_j(\lfloor t \rfloor)$, $B(t) = B(\lfloor t \rfloor)$ and $\Gamma_j(t) = \Gamma_j(\lfloor t \rfloor)$ for all $t \geq 0$.

If $\rho \geq 1$, then, by Lemma 1, $\bar{T}_j(t) = t - \bar{I}_j(t) = t$ for all j , and by (9),

$$\begin{aligned} &\sup_{0 \leq t \leq L} |X(t) - \bar{X}(t)| \\ &\leq \sup_{0 \leq t \leq L} |B(A(t)) - m\alpha t| + \sum_{j=1}^N \sup_{0 \leq t \leq L} |\Gamma_j(S_j(T_j(t))) - pS_j(T_j(t))| \\ &\quad + \sum_{j=1}^N p \sup_{0 \leq t \leq L} |S_j(T_j(t)) - \mu T_j(t)| + \sum_{j=1}^N \sup_{0 \leq t \leq L} |S_j(T_j(t)) - \mu T_j(t)| \\ &\leq \sup_{0 \leq t \leq L} |B(A(t)) - mA(t)| + m \sup_{0 \leq t \leq L} |A(t) - \alpha t| + \sum_{j=1}^N \sup_{0 \leq t \leq (\mu+1)L} |\Gamma_j(t) - pt| \\ &\quad + \sum_{j=1}^N p \sup_{0 \leq t \leq L} |S_j(t) - \mu t| + \sum_{j=1}^N \sup_{0 \leq t \leq L} |S_j(t) - \mu t| \\ &= O\left(\sqrt{L \log \log L}\right), \quad \text{w.p.1.} \end{aligned}$$

Because Ψ is continuous under uniform norm, then, w.p.1,

$$\begin{aligned} \sup_{0 \leq t \leq L} |\Psi(X)(t) - \Psi(\bar{X})(t)| &= O\left(\sqrt{L \log \log L}\right), \\ \sup_{0 \leq t \leq L} |\Psi(X - N)(t) - \Psi(\bar{X})(t)| &= O\left(\sqrt{L \log \log L}\right). \end{aligned}$$

Since $\Psi(\bar{X})(t) = \Psi(\theta_N t) = 0$, and with (44), we have

$$\sup_{0 \leq t \leq L} \left| (1 - p)\mu \sum_{j=1}^N I_j(t) \right| = O\left(\sqrt{L \log \log L}\right), \quad \text{w.p.1.}$$

Notice that $(1 - p)\mu > 0$, $I_i(t) \geq 0$ for all i , then, w.p.1,

$$\sup_{0 \leq t \leq L} |(1 - p)\mu I_i(t)| \leq \sup_{0 \leq t \leq L} \left| (1 - p)\mu \sum_{j=1}^N I_j(t) \right| = O\left(\sqrt{L \log \log L}\right),$$

which implies that $\sup_{0 \leq t \leq L} |I_i(t)| = O\left(\sqrt{L \log \log L}\right)$ w.p.1. This follows that, for all j ,

$$\sup_{0 \leq t \leq L} |T_j(t) - t| = \sup_{0 \leq t \leq L} |I_j(t)| = O\left(\sqrt{L \log \log L}\right), \quad \text{w.p.1.}$$

□

Proof of Theorem 6 By the SA in Csörgő and Révész (1981), w.p.1,

$$\begin{aligned} \sup_{0 \leq t \leq L} |B(t) - mt - mc_b W_b(t)| &= o(L^{1/r}), \\ \sup_{0 \leq t \leq L} |A(t) - \alpha t - \alpha^{1/2} c_a W_a(t)| &= o(L^{1/r}). \end{aligned} \tag{45}$$

By Lemma 2.3 (vi) of Chen and Shen (2000), we have, w.p.1,

$$\sup_{0 \leq t \leq L} |B(A(t)) - m\alpha t - m\alpha^{1/2} c_a W_a(t) - mc_b W_b(\alpha t)| = o(L^{1/r}). \tag{46}$$

(i) Suppose that $N \geq 1$ and $\rho \geq 1$. If $\rho \geq 1$, then $\bar{T}_j(t) = t$ for all j . With (32), we have

$$\begin{aligned} X(t) - \tilde{X}(t) &= [B(A(t)) - m\alpha t - m\alpha^{1/2} c_a W_a(t) - mc_b W_b(\alpha t)] \\ &\quad - \sum_{j=1}^N [S_j(T_j(t)) - \mu T_j(t) - \mu^{1/2} c_{s,j} W_{s,j}(T_j(t))] \\ &\quad - \sum_{j=1}^N \mu^{1/2} c_{s,j} [W_{s,j}(T_j(t)) - W_{s,j}(\bar{T}_j(t))] \\ &\quad + \sum_{j=1}^N [\Gamma_j(S_j(T_j(t))) - pS_j(T_j(t)) - \sqrt{p(1-p)} W_{f,j}(S_j(T_j(t)))] \\ &\quad + \sum_{j=1}^N \sqrt{p(1-p)} [W_{f,j}(S_j(T_j(t))) - W_{f,j}(\mu \bar{T}_j(t))] \\ &\quad + \sum_{j=1}^N p [S_j(T_j(t)) - \mu T_j(t) - \mu^{1/2} c_{s,j} W_{s,j}(T_j(t))] \\ &\quad + \sum_{j=1}^N p\mu^{1/2} c_{s,j} [W_{s,j}(T_j(t)) - W_{s,j}(\bar{T}_j(t))]. \end{aligned} \tag{47}$$

By Lemma 2 and Lemma 6.21 in Chen and Yao (2001), $\sup_{0 \leq t \leq L} |W_{s,j}(T_j(t)) - W_{s,j}(\bar{T}_j(t))| = o(L^{1/r})$ w.p.1. By Lemma 2, since

$$\begin{aligned} \sup_{0 \leq t \leq L} |S_k(T_k(t)) - \mu \bar{T}_k(t)| &\leq \sup_{0 \leq t \leq L} |S_k(T_k(t)) - \mu T_k(t)| + \mu \sup_{0 \leq t \leq L} |T_k(t) - \bar{T}_k(t)| \\ &= O(\sqrt{L \log \log L}), \quad \text{w.p.1,} \end{aligned}$$

we have

$$\sup_{0 \leq t \leq L} |W_{f,j}(S_j(T_j(t))) - W_{f,j}(\mu \bar{T}_j(t))| = o(L^{1/r}), \quad \text{w.p.1.}$$

Since $T_k(t) \leq t$ for all $t \geq 0$, then

$$\sup_{0 \leq t \leq L} |S_k(T_k(t)) - \mu T_k(t) - \mu^{1/2} c_{s,k} W_{s,k}(T_k(t))| = o(L^{1/r}), \quad \text{w.p.1.}$$

Using $S_k(T_k(t)) \leq S_k(t) \leq (\mu + 1)t$ w.p.1 for large t , we have

$$\sup_{0 \leq t \leq L} \left| \Gamma_j(S_j(T_j(t))) - pS_j(T_j(t)) - \sqrt{p(1-p)}W_{f,j}(S_j(T_j(t))) \right| = o(L^{1/r}), \quad \text{w.p.1.}$$

Hence, $\sup_{0 \leq t \leq L} |X(t) - \tilde{X}(t)| = o(L^{1/r})$ w.p.1.

Since Ψ is a continuous function under uniform norm, we have, w.p.1,

$$\begin{aligned} \sup_{0 \leq t \leq L} |\Psi(X)(t) - \Psi(\tilde{X})(t)| &= o(L^{1/r}) \quad \text{and} \\ \sup_{0 \leq t \leq L} |\Psi(X - N)(t) - \Psi(\tilde{X})(t)| &= o(L^{1/r}). \end{aligned}$$

This, together with (44), implies that $\sup_{0 \leq t \leq L} |(1-p)\mu I(t) - \Psi(\tilde{X})(t)| = o(L^{1/r})$. For the SA of Q , we note by (44) that $\Psi(X)(t) + X(t) \leq Q(t) \leq \Psi(X - N)(t) + X(t)$, and

$$\Phi(X)(t) \leq Q(t) \leq \Phi(X - N)(t) + N.$$

Since Ψ is a continuous function under uniform norm, similarly with the analysis for X , we have $\sup_{0 \leq t \leq L} |Q(t) - \tilde{Q}(t)| = o(L^{1/r})$. So (31) follows.

(ii) Next, we suppose $N = 1$. Like (47), we have

$$\begin{aligned} X(t) - \tilde{X}(t) &= [B(A(t)) - m\alpha t - m\alpha^{1/2}c_a W_a(t) - mc_b W_b(\alpha t)] \\ &\quad - (1-p)[S(T(t)) - \mu T(t) - \mu^{1/2}c_s W_s(T(t))] \\ &\quad - (1-p)\mu^{1/2}c_s [W_s(T(t)) - W_s(\tilde{T}(t))] \\ &\quad + [\Gamma(S(T(t))) - pS(T(t)) - \sqrt{p(1-p)}W_f(S(T(t)))] \\ &\quad + \sqrt{p(1-p)}[W_f(S(T(t))) - W_f(\mu\tilde{T}(t))], \end{aligned}$$

where $S(t) = S_1(t)$. Similarly with Theorem 6.11 in Chen and Yao (2001), we have $\sup_{0 \leq t \leq L} |T(t) - \tilde{T}(t)| = O(\sqrt{L \log \log L})$ w.p.1. This, and Lemma 6.21 in Chen and Yao (2001), gives $\sup_{0 \leq t \leq L} |W_s(T(t)) - W_s(\tilde{T}(t))| = o(L^{1/r})$ w.p.1. Since $\sup_{0 \leq t \leq L} |S(T(t)) - \mu\tilde{T}(t)| = O(\sqrt{L \log \log L})$ w.p.1, we have

$$\sup_{0 \leq t \leq L} |W_f(S(T(t))) - W_f(\mu\tilde{T}(t))| = o(L^{1/r}), \quad \text{w.p.1.}$$

Since $T(t) \leq t$ for all $t \geq 0$, then

$$\sup_{0 \leq t \leq L} |S(T(t)) - \mu T(t) - \mu^{1/2}c_s W_s(T(t))| = o(L^{1/r}), \quad \text{w.p.1.}$$

Using $S(T(t)) \leq S(t) \leq (\mu + 1)t$ w.p.1 for large t , we have

$$\sup_{0 \leq t \leq L} \left| \Gamma(S(T(t))) - pS(T(t)) - \sqrt{p(1-p)}W_f(S(T(t))) \right| = o(L^{1/r}), \quad \text{w.p.1.}$$

Hence, $\sup_{0 \leq t \leq L} |X(t) - \tilde{X}(t)| = o(L^{1/r})$ w.p.1. This yields the SAs of Q, Y, I, T (which is similar to Theorem 6.16 and 7.19 in Chen and Yao (2001)). For the SA of Z , we have, by (12),

$$\begin{aligned} Z(t) - \tilde{Z}(t) &= \left[V(Q(t) + S(T(t))) - \frac{1}{\mu}(Q(t) + S(T(t))) \right. \\ &\quad \left. + \frac{1}{\mu}\mu^{1/2}c_s W_s \left(\frac{1}{\mu}(Q(t) + S(T(t))) \right) \right] \end{aligned}$$

$$\begin{aligned}
 & + \frac{1}{\mu} [Q(t) - \tilde{Q}(t)] + \frac{1}{\mu} [S(T(t)) - \mu T(t) - \mu^{1/2} c_s W_s(T(t))] \\
 & + \frac{1}{\mu} \mu^{1/2} c_s [W_s(T(t)) - W_s(\tilde{T}(t))] \\
 & - \frac{1}{\mu} \mu^{1/2} c_s \left[W_s \left(\frac{1}{\mu} (B(A(t)) + \Gamma(S(T(t)))) \right) \right. \\
 & \quad \left. - W_s \left(\frac{1}{\mu} (m\alpha t + p\mu \tilde{T}(t)) \right) \right],
 \end{aligned}$$

where

$$W_s \left(\frac{1}{\mu} (m\alpha t + p\mu \tilde{T}(t)) \right) = W_s(\rho t),$$

because, if $\rho < 1$, then

$$\tilde{T}(t) = \rho t = \frac{m\alpha}{(1-p)\mu} t \quad \text{and} \quad \frac{1}{\mu} (m\alpha t + p\mu \tilde{T}(t)) = \frac{m\alpha}{\mu} \left(1 + \frac{p}{1-p} \right) t = \rho t,$$

and if $\rho \geq 1$,

$$\rho t = \frac{m\alpha + p\mu}{\mu} t, \quad \tilde{T}(t) = t \quad \text{and} \quad \frac{1}{\mu} (m\alpha t + p\mu \tilde{T}(t)) = \rho t.$$

Notice that, by FSLLN, for a large t ,

$$\begin{aligned}
 & B(A(t)) + \Gamma(S(T(t))) = Q(t) + S(T(t)) \leq (m\alpha + p\mu + 1)t, \quad \text{w.p.1,} \\
 & \sup_{0 \leq t \leq L} \left| \frac{1}{\mu} (B(A(t)) + \Gamma(S(T(t)))) - \frac{1}{\mu} (m\alpha t + p\mu \tilde{T}(t)) \right| = O \left(\sqrt{L \log \log L} \right) \quad \text{w.p.1.}
 \end{aligned}$$

The proof of the SA of Z is similar to that of X .

For the SA of D , by its definition we have

$$\begin{aligned}
 D(t) - \tilde{D}(t) &= (1-p) [S(T(t)) - \mu T(t) - \mu^{1/2} c_s W_s(T(t))] \\
 &+ (1-p)\mu [T(t) - \tilde{T}(t)] + (1-p)\mu^{1/2} c_s [W_s(T(t)) - W_s(\tilde{T}(t))] \\
 &- \left[\Gamma(S(T(t))) - pS(T(t)) - \sqrt{p(1-p)} W_f(S(T(t))) \right] \\
 &- \sqrt{p(1-p)} [W_f(S(T(t))) - W_f(\mu \tilde{T}(t))].
 \end{aligned}$$

Similar to the treatment of the SA of X , we have $\sup_{0 \leq t \leq L} |D(t) - \tilde{D}(t)| = o(L^{1/r})$. \square

D Sensitivity analysis for FLIL and LIL limits

In this section we conduct sensitivity analysis for the FLIL and LIL limits given in Theorem 3 as functions of all model parameters.

D.1 Sensitivity analysis of FLIL limits

For the function h in \mathcal{K}_D in the CL case, we note that, in the proof of Theorem 3 in Sect. 5, we use the function $h(x, y)$ with

$$x(t) = m\alpha^{1/2} c_a W_a(t) + mc_b W_b(\alpha t), \quad y(t) = (1-p)\mu^{1/2} c_s W_s(t) - \sqrt{p(1-p)} W_f(\mu t).$$

Then, we have

$$h(x, y)(t) \equiv \begin{cases} h_a(m\alpha^{1/2}c_a W_a(t)), & \text{if } c_b = c_s = p = 0, \\ h_a(mc_b W_b(\alpha t)), & \text{if } c_a = c_s = p = 0, \\ h_a(x(t)), & \text{if } c_s = p = 0, \\ h_s(\mu^{1/2}c_s W_s(t)), & \text{if } c_a = c_b = p = 0, \\ h_s(-\sqrt{p(1-p)}W_f(\mu t)), & \text{if } c_a = c_b = c_s = 0, \\ h_s(y(t)), & \text{if } c_a = c_b = 0, \end{cases}$$

where, for function x' ,

$$h_a(x')(t) \equiv \inf_{0 \leq s \leq t} [x'(s)], \quad h_s(x')(t) \equiv x'(t) + \inf_{0 \leq s \leq t} [-x'(s)].$$

Then, by Theorem 7 (Strassen’s CMT), w.p.1,

$$\frac{h(x, y)(nt)}{\varphi(n)} \Rightarrow \begin{cases} h_a(\alpha^{1/2}mc_a), & \text{if } c_b = c_s = p = 0, \\ h_a(\alpha^{1/2}mc_b), & \text{if } c_a = c_s = p = 0, \\ h_a\left(\sqrt{\alpha m^2(c_a^2 + c_b^2)}\right), & \text{if } c_s = p = 0, \\ h_s(\mu^{1/2}c_s), & \text{if } c_a = c_b = p = 0, \\ h_s(\sqrt{p(1-p)}\mu), & \text{if } c_a = c_b = c_s = 0, \\ h_s\left(\sqrt{(1-p)^2\mu c_s^2 + p(1-p)\mu}\right), & \text{if } c_a = c_b = 0. \end{cases}$$

The above functions are used to explain the FLIL limits for $GI^B/GI^F/1$ in the following cases.

For the FLILs of the $GI^B/GI^F/1$ queue in Theorem 3, we first note that all the FLILs mainly depend on all the variance parameters. In deed, if all variances are zero, that is, $c_a^2 = c_b^2 = c_s^2 = p = 0$, then $\mathcal{K}_Q = \mathcal{K}_Z = \mathcal{K}_I = \mathcal{K}_T = \mathcal{K}_D = \{0\}$, which implies that all process Q^n, Z^n, I^n, T^n, D^n converge to zero w.p.1.

Variability in batch arrivals

- (i) Let $c_a^2 = c_s^2 = p = 0$, that is, all deviations are assumed to be from the batches B , then

$$\sigma_{uc} = \sigma_o = \sigma_{D,u} = \sigma_{Z,o} = \sqrt{\alpha m^2 c_b^2}, \quad \sigma_{D,o} = 0, \tag{48}$$

and

$$\mathcal{K}^* = \begin{cases} \left\{ \left(0, 0, \frac{\sigma_{uc}}{\mu}x, -\frac{\sigma_{uc}}{\mu}x, \sigma_{uc}x \right) : x \in \mathcal{G}(1) \right\}, & \text{if } \rho < 1, \\ \left\{ \left(\Phi(x), \frac{1}{\mu}\Phi(x), \frac{1}{\mu}\Psi(x), -\frac{1}{\mu}\Psi(x), h_a(x) \right) : x \in \mathcal{G}(\sigma_{uc}) \right\}, & \text{if } \rho = 1, \tag{49} \\ \left\{ \left(\sigma_{uc}x, \frac{\sigma_{uc}}{\mu}x, 0, 0, 0 \right) : x \in \mathcal{G}(1) \right\}, & \text{if } \rho > 1. \end{cases}$$

- (ii) Let $c_b^2 = c_s^2 = p = 0$, that is, all deviations are assumed to be from the arrival process A , then \mathcal{K}^* satisfies (49) with

$$\sigma_{uc} = \sigma_o = \sigma_{D,u} = \sigma_{Z,o} = \sqrt{\alpha m^2 c_a^2}, \quad \sigma_{D,o} = 0. \tag{50}$$

(iii) The joint impact from the batch B and the arrival process A . Let $c_s^2 = p = 0$, then (49) holds with

$$\sigma_{uc} = \sigma_o = \sigma_{D,u} = \sigma_{Z,o} = \sqrt{\alpha m^2 (c_a^2 + c_b^2)}, \quad \sigma_{D,o} = 0. \tag{51}$$

Variability in service times Let $c_a^2 = c_b^2 = p = 0$, that is, the variability stems only from the random service process S , then

$$\sigma_{uc} = \sigma_{Z,o} = \sqrt{\alpha m c_s^2}, \quad \sigma_o = \sigma_{D,o} = \sqrt{\mu c_s^2}, \quad \sigma_{D,u} = 0, \tag{52}$$

and

$$\mathcal{K}^* = \begin{cases} \left\{ \left(0, 0, \frac{\sigma_{uc}}{\mu} x, -\frac{\sigma_{uc}}{\mu} x, 0 \right) : x \in \mathcal{G}(1) \right\}, & \text{if } \rho < 1, \\ \left\{ \left(\Phi(x), \frac{1}{\mu} \Phi(x), \frac{1}{\mu} \Psi(x), -\frac{1}{\mu} \Psi(x), h_s(x) \right) : x \in \mathcal{G}(\sigma_o) \right\}, & \text{if } \rho = 1, \\ \left\{ \left(\sigma_o x, \frac{\sigma_{uc}}{\mu} x, 0, 0, \sigma_o x \right) : x \in \mathcal{G}(1) \right\}, & \text{if } \rho > 1. \end{cases} \tag{53}$$

Variability in Bernoulli feedback Let $c_a^2 = c_b^2 = c_s^2 = 0$, that is, the variability stems only from the stochastic feedback process γ , then

$$\sigma_{uc} = \sqrt{p m \alpha}, \quad \sigma_{D,u} = 0, \quad (1-p)\sigma_{D,o} = \sigma_o = \sigma_{Z,o} = \sqrt{p(1-p)}\mu, \tag{54}$$

and $\mathcal{K}^* =$

$$\begin{cases} \left\{ \left(0, 0, \frac{\sigma_{uc}}{(1-p)\mu} x, -\frac{\sigma_{uc}}{(1-p)\mu} x, 0 \right) : x \in \mathcal{G}(1) \right\}, & \text{if } \rho < 1, \\ \left\{ \left(\Phi(x), \frac{1}{\mu} \Phi(x), \frac{1}{(1-p)\mu} \Psi(x), -\frac{1}{(1-p)\mu} \Psi(x), h_s(x) \right) : x \in \mathcal{G}(\sigma_o) \right\}, & \text{if } \rho = 1, \\ \left\{ \left(\sigma_o x, \frac{\sigma_o}{\mu} x, 0, 0, \sigma_o x \right) : x \in \mathcal{G}(1) \right\}, & \text{if } \rho > 1. \end{cases} \tag{55}$$

Variability in service times and feedback Let $c_a^2 = c_b^2 = 0$, that is, the variability comes from the service S and the feedback γ . We have the FLILs (24) with the following parameters:

$$\begin{aligned} \sigma_{uc} &= \sqrt{(1-p)m\alpha c_s^2 + pm\alpha}, \quad \sigma_{Z,o} = \sqrt{c_s^2(p\mu^{1/2} - (m\alpha + p\mu)^{1/2})^2 + p(1-p)\mu}, \\ \sigma_{D,u} &= 0, \quad \sigma_{D,o} = \sqrt{\mu c_s^2 + \frac{p\mu}{1-p}}, \quad \sigma_o = \sqrt{(1-p)^2 \mu c_s^2 + p(1-p)\mu}. \end{aligned} \tag{56}$$

From (49)–(56), we find that, different from processes B , A and joint $B(A)$, the service S , feedback Γ and joint $\Gamma(S)$ affect the FLIL in a different way.

D.2 Sensitivity analysis for LIL limits

Following Sect. D.1, we now consider the impact of parameters upon their corresponding LIL limits for $GI^B/GI^F/1$. Specifically, we study how variabilities in different model components affect the LIL limits. If $c_a^2 = c_b^2 = c_s^2 = p = 0$, that is, all precesses are deterministic, then all the superior and inferior limits in (28)–(30) are zeros. Hence, it is said that all the superior and inferior limits are determined by the variances.

Variability in batch arrivals If $c_a^2 = c_s^2 = p = 0$, that is, LIL is only affected by the random batch size B , then all superior and inferior limits in (28)–(30) hold with parameters given in (48). Similarly, if $c_b^2 = c_s^2 = p = 0$ ($c_s^2 = p = 0$), then the limits (28)–(30) hold with

parameters given in (50) [(51)]. These three cases describe the impact of the arrival process on the LIL limits. For example, under condition $c_s^2 = p = 0$, we have

$$Q_{sup}^* = \mu Z_{sup}^* = \begin{cases} 0, & \text{if } \rho < 1, \\ \sqrt{\alpha m^2 (c_a^2 + c_b^2)}, & \text{if } \rho = 1, \\ \sqrt{\alpha m^2 (c_a^2 + c_b^2)}, & \text{if } \rho > 1, \end{cases}$$

$$Q_{inf}^* = \mu Z_{inf}^* = \begin{cases} 0, & \text{if } \rho \leq 1, \\ -\sqrt{\alpha m^2 (c_a^2 + c_b^2)}, & \text{if } \rho > 1, \end{cases}$$

where we note that $\alpha m^2 (c_a^2 + c_b^2)$ is the variance parameter of the compound renewal process $B(A(t))$, and it entirely determines the LIL limits for Q and Z . If we furthermore assume that $c_a = 0$ or $c_b = 0$, then we get the LIL limits similar with the above respectively, that is, characterized by $\sqrt{\alpha m^2 c_b^2}$ or $\sqrt{\alpha m^2 c_a^2}$. In this sense, the impacts from the batch and the arrival are separable.

Variability in Bernoulli feedback If $c_a^2 = c_b^2 = c_s^2 = 0$, that is, LIL is only impacted by the random Bernoulli feedback process Γ , then all superior and inferior limits in (28)–(30) hold with parameters given in (54). For example,

$$Q_{sup}^* = \mu Z_{sup}^* = \begin{cases} 0, & \text{if } \rho < 1, \\ \sqrt{p m \alpha}, & \text{if } \rho = 1, \\ \sqrt{p(1-p)\mu}, & \text{if } \rho > 1, \end{cases}$$

$$Q_{inf}^* = \mu Z_{inf}^* = \begin{cases} 0, & \text{if } \rho \leq 1, \\ -\sqrt{p(1-p)\mu}, & \text{if } \rho > 1, \end{cases}$$

where we note that $p(1-p)$ is the variance parameter of the feedback Γ . We also note that, if $\rho = 1$, then $\lambda = m\alpha/(1-p)$ and $Q_{sup}^* = \mu Z_{sup}^* = \sqrt{\lambda p(1-p)}$. In this sense, we can say that the variance of the feedback $p(1-p)$ determines the LIL limits under $c_a^2 = c_b^2 = c_s^2 = 0$.

Variability in service times If $c_a^2 = c_b^2 = p = 0$, that is, LIL is only determined by the randomness in the service times S , then the limits (28)–(30) hold with parameters given in (52), for example,

$$Q_{sup}^* = \begin{cases} 0, & \text{if } \rho < 1, \\ \sqrt{\mu c_s^2}, & \text{if } \rho = 1, \\ \sqrt{\mu c_s^2}, & \text{if } \rho > 1, \end{cases} \quad Q_{inf}^* = \begin{cases} 0, & \text{if } \rho \leq 1, \\ -\sqrt{\mu c_s^2}, & \text{if } \rho > 1, \end{cases}$$

and

$$Z_{sup}^* = \begin{cases} 0, & \text{if } \rho < 1, \\ \frac{1}{\mu} \sqrt{\mu c_s^2}, & \text{if } \rho = 1, \\ \frac{1}{\mu} \sqrt{m \alpha c_s^2}, & \text{if } \rho > 1, \end{cases} \quad Z_{inf}^* = \begin{cases} 0, & \text{if } \rho \leq 1, \\ -\frac{1}{\mu} \sqrt{m \alpha c_s^2}, & \text{if } \rho > 1. \end{cases}$$

We note that these LIL limits are determined by the SCV c_s^2 , and the Little’s law fails in this case, i.e., $Q_{sup}^* \neq \mu Z_{sup}^*$ and $Q_{inf}^* \neq \mu Z_{inf}^*$ if $c_s \neq 0$.

Variability in service times and feedback If $c_a^2 = c_b^2 = 0$, that is, LIL is characterized by both the random service time S and the random feedback mechanism Γ , then the limits (28)–(30) hold with parameters given in (56), for example, since $\lambda = m\alpha/(1 - p)$ when $\rho = 1$, we have

$$Q_{sup}^* = \begin{cases} 0, & \text{if } \rho < 1, \\ \sqrt{(1 - p)m\alpha c_s^2 + \lambda p(1 - p)}, & \text{if } \rho = 1, \\ \sqrt{(1 - p)^2 \mu c_s^2 + p(1 - p)\mu}, & \text{if } \rho > 1, \end{cases}$$

$$Q_{inf}^* = \begin{cases} 0, & \text{if } \rho \leq 1, \\ -\sqrt{(1 - p)^2 \mu c_s^2 + p(1 - p)\mu}, & \text{if } \rho > 1, \end{cases}$$

and

$$Z_{sup}^* = \begin{cases} 0, & \text{if } \rho < 1, \\ \frac{1}{\mu} \sqrt{(1 - p)m\alpha c_s^2 + \lambda p(1 - p)}, & \text{if } \rho = 1, \\ \frac{1}{\mu} \sqrt{c_s^2 (p\mu^{1/2} - (m\alpha + p\mu)^{1/2})^2 + p(1 - p)\mu}, & \text{if } \rho > 1, \end{cases}$$

$$Z_{inf}^* = \begin{cases} 0, & \text{if } \rho \leq 1, \\ -\frac{1}{\mu} \sqrt{c_s^2 (p\mu^{1/2} - (m\alpha + p\mu)^{1/2})^2 + p(1 - p)\mu}, & \text{if } \rho > 1. \end{cases}$$

Compared with the impact from the batch and the arrival, the impact from the service and the feedback is not separable and is more complicated. This is so because the feedback adheres its own deviation to the service process S . We also note that the Little’s law fails here too, i.e., $Q_{sup}^* \neq \mu Z_{sup}^*$ and $Q_{inf}^* \neq \mu Z_{inf}^*$.

We conclude that the LIL limits are non-linear functions of the mean parameters but linear functions of the variability parameters.

E Additional numerical examples

E.1 A new example

Example 3 (Impact of N on the queue length LIL limits) We consider a sequence of the queueing models $\{GI^B/GI^F/N, N = 1, 2, \dots, 5\}$. Let $\alpha = 2.5, m = 1, \mu = 1, p =$

Table 3 Impact of N on the queue length LIL limits for Example 3

| N | 1 | 2 | 3 | 4 | 5 |
|-------------|---------------|-----------------|---------------|-----------------|--------------|
| ρ | 3 | $\frac{3.5}{2}$ | $\frac{4}{3}$ | $\frac{4.5}{4}$ | 1 |
| Q_{sup}^* | $\sqrt{5.5}$ | $\sqrt{6}$ | $\sqrt{6.5}$ | $\sqrt{7}$ | $\sqrt{7.5}$ |
| Q_{inf}^* | $-\sqrt{5.5}$ | $-\sqrt{6}$ | $-\sqrt{6.5}$ | $-\sqrt{7}$ | 0 |

0.5, $c_a = c_b = c_{s,j} = 1$. In this setting, we compute the values of the LIL limits ρ , Q_{sup}^* and Q_{inf}^* in Table 3; they are monotone functions in N .

In Example 3, we find that Q_{sup}^* and Q_{inf}^* are increasing and decreasing respectively in N except $Q_{\text{inf}}^* = 0$ in the CL case: $N = 5$.

E.2 More on Example 1

To supplement the main paper, here we plot the key LIL limits in Figs. 2, 3 and 4 for cases (ii)–(iv). We also provide the explicit functions for the LILs in all 4 cases.

- (i) Increase α from 0 and let $m = 1, \mu = 1, p = 1/2, c_a = c_b = c_s = 1$. In this case, $\rho = 2\alpha$ if $\alpha \leq 1/2$, and $\alpha + 1/2$ otherwise. Then, we get the superior and the inferior limits in Table 4, which gives Fig. 1.

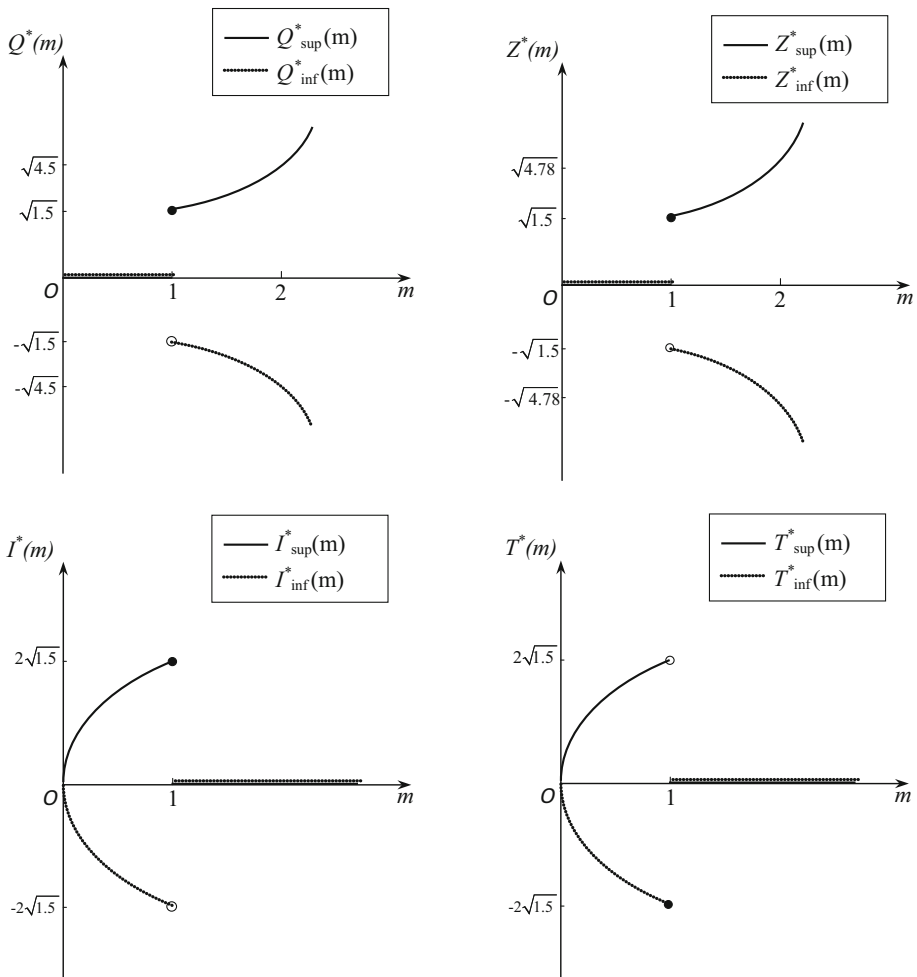


Fig. 2 LIL limits of (ii) of Example 1 as functions of m , with $\alpha = 1/2, \mu = 1, p = 1/2, c_a = c_b = c_s = 1$

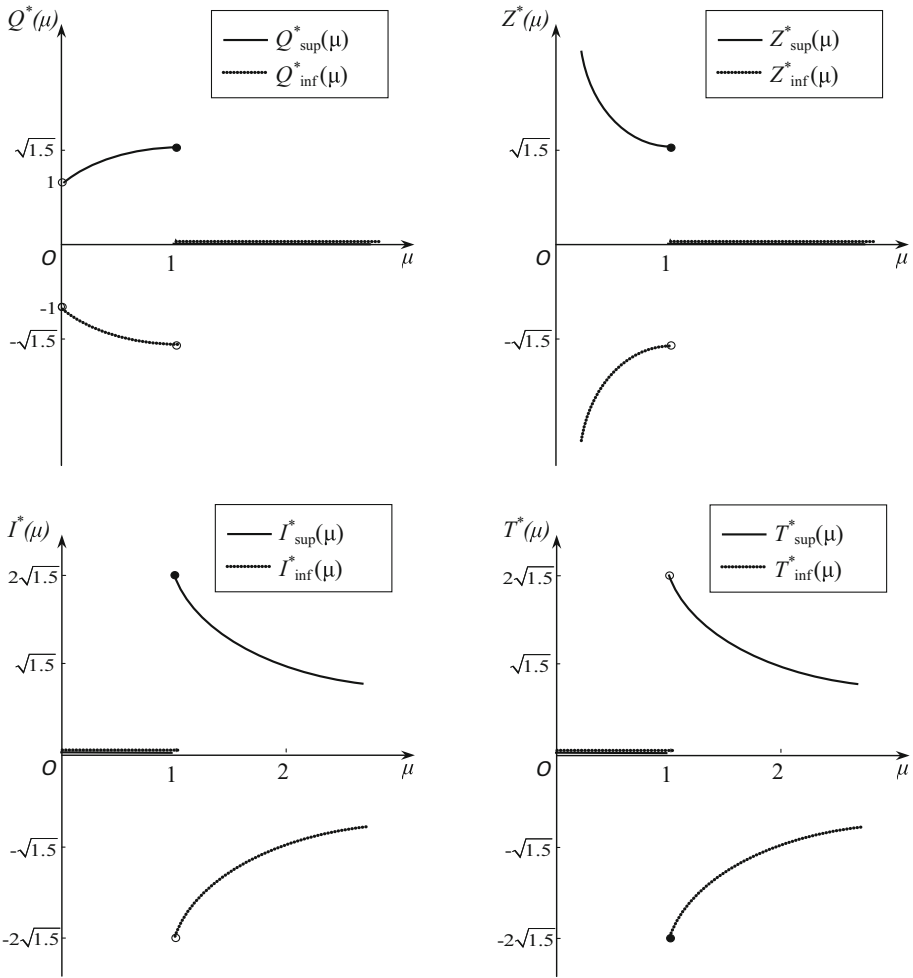


Fig. 3 LIL limits of (iii) of Example 1 as functions of μ , with $\alpha = 1/2, m = 1, p = 1/2, c_a = c_b = c_s = 1$

- (ii) Increase m from 0 and let $\alpha = 1/2, \mu = 1, p = 1/2, c_a = c_b = c_s = 1$. In this case, $\rho = m$ if $m \leq 1$ and $m/2 + 1/2$ otherwise. Then, we get the superior and the inferior limits in Table 5, which gives Fig. 2.
- (iii) Increase μ from 0 and let $\alpha = 1/2, m = 1, p = 1/2, c_a = c_b = c_s = 1$. In this case, $\rho = 1/\mu$ if $\mu \geq 1$ and $1/(2\mu) + 1/2$ otherwise. Then, we get the superior and the inferior limits in Table 6, which gives Fig. 3.
- (iv) Increase p from 0 to 1, and let $\alpha = 1/2, m = 1, \mu = 1, c_a = c_b = c_s = 1$. In this case, $\rho = 1/[2(1 - p)]$ if $p \leq 1/2$ and $1/2 + p$ otherwise. Then, we get the superior and the inferior limits in Table 7, which gives Fig. 4.

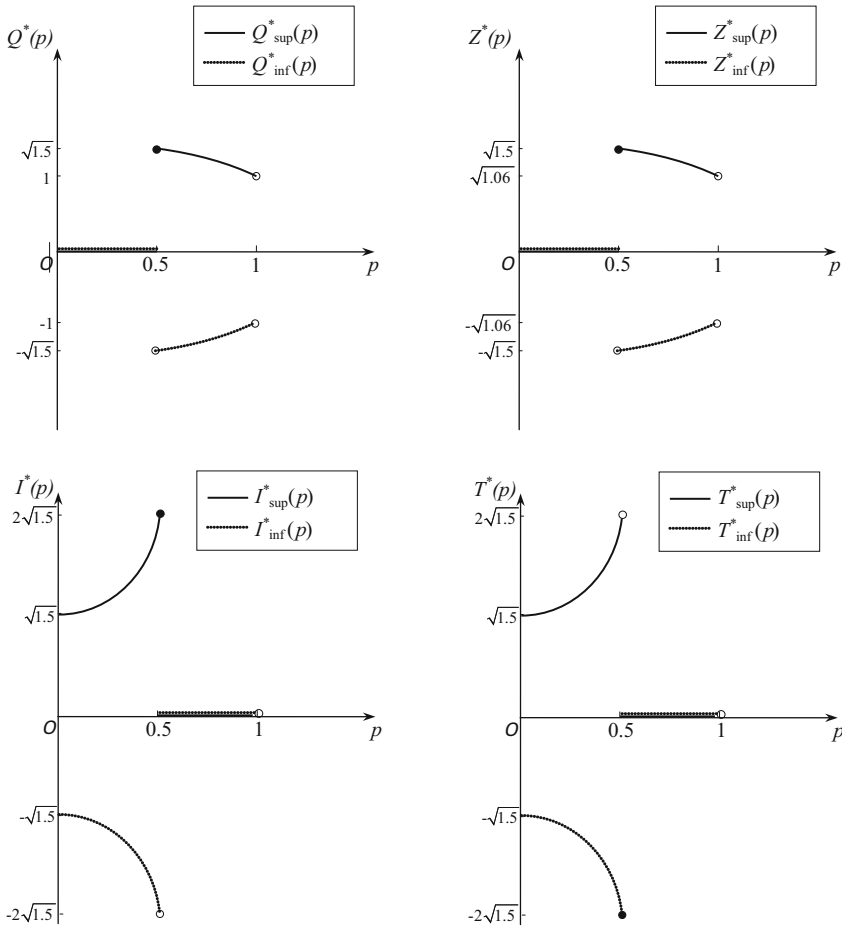


Fig. 4 LIL limits of (iv) of Example 1 as functions of p , with $\alpha = 1/2, m = 1, \mu = 1, c_a = c_b = c_s = 1$

Table 4 The LIL limits for (i) of Example 1 with α varying

| α | $< \frac{1}{2}$ | $= \frac{1}{2}$ | $> \frac{1}{2}$ |
|----------------------------|--------------------|------------------------|---|
| $Q_{\text{sup}}^*(\alpha)$ | 0 | $\sqrt{\frac{3}{2}}$ | $\sqrt{2\alpha + \frac{1}{2}}$ |
| $Z_{\text{sup}}^*(\alpha)$ | 0 | $\sqrt{\frac{3}{2}}$ | $\sqrt{3\alpha + 1 - \sqrt{\alpha + \frac{1}{2}}}$ |
| $I_{\text{sup}}^*(\alpha)$ | $2\sqrt{3\alpha}$ | $2\sqrt{\frac{3}{2}}$ | 0 |
| $T_{\text{sup}}^*(\alpha)$ | $2\sqrt{3\alpha}$ | 0 | 0 |
| $Q_{\text{inf}}^*(\alpha)$ | 0 | 0 | $-\sqrt{2\alpha + \frac{1}{2}}$ |
| $Z_{\text{inf}}^*(\alpha)$ | 0 | 0 | $-\sqrt{3\alpha + 1 - \sqrt{\alpha + \frac{1}{2}}}$ |
| $I_{\text{inf}}^*(\alpha)$ | $-2\sqrt{3\alpha}$ | 0 | 0 |
| $T_{\text{inf}}^*(\alpha)$ | $-2\sqrt{3\alpha}$ | $-2\sqrt{\frac{3}{2}}$ | 0 |

Table 5 The LIL limits for (ii) of Example 1 with m varying

| m | <1 | $=1$ | >1 |
|----------------|-------------------------------|------------------------|--|
| $Q_{sup}^*(m)$ | 0 | $\sqrt{\frac{3}{2}}$ | $\sqrt{m^2 + \frac{1}{2}}$ |
| $Z_{sup}^*(m)$ | 0 | $\sqrt{\frac{3}{2}}$ | $\sqrt{m^2 + \frac{1}{2}m + 1 - \sqrt{\frac{1}{2}m + \frac{1}{2}}}$ |
| $I_{sup}^*(m)$ | $2\sqrt{m^2 + \frac{1}{2}m}$ | $2\sqrt{\frac{3}{2}}$ | 0 |
| $T_{sup}^*(m)$ | $2\sqrt{m^2 + \frac{1}{2}m}$ | 0 | 0 |
| $Q_{inf}^*(m)$ | 0 | 0 | $-\sqrt{m^2 + \frac{1}{2}}$ |
| $Z_{inf}^*(m)$ | 0 | 0 | $-\sqrt{m^2 + \frac{1}{2}m + 1 - \sqrt{\frac{1}{2}m + \frac{1}{2}}}$ |
| $I_{inf}^*(m)$ | $-2\sqrt{m^2 + \frac{1}{2}m}$ | 0 | 0 |
| $T_{inf}^*(m)$ | $-2\sqrt{m^2 + \frac{1}{2}m}$ | $-2\sqrt{\frac{3}{2}}$ | 0 |

Table 6 The LIL limits for (iii) of Example 1 with μ varying

| μ | >1 | $=1$ | <1 |
|------------------|------------------------------------|------------------------|---|
| $Q_{sup}^*(\mu)$ | 0 | $\sqrt{\frac{3}{2}}$ | $\sqrt{1 + \frac{1}{2}\mu}$ |
| $Z_{sup}^*(\mu)$ | 0 | $\sqrt{\frac{3}{2}}$ | $\frac{1}{\mu}\sqrt{\mu - \sqrt{\frac{1}{2}\mu(1 + \mu)} + \frac{3}{2}}$ |
| $I_{sup}^*(\mu)$ | $\frac{2}{\mu}\sqrt{\frac{3}{2}}$ | $2\sqrt{\frac{3}{2}}$ | 0 |
| $T_{sup}^*(\mu)$ | $\frac{2}{\mu}\sqrt{\frac{3}{2}}$ | 0 | 0 |
| $Q_{inf}^*(\mu)$ | 0 | 0 | $-\sqrt{1 + \frac{1}{2}\mu}$ |
| $Z_{inf}^*(\mu)$ | 0 | 0 | $-\frac{1}{\mu}\sqrt{\mu - \sqrt{\frac{1}{2}\mu(1 + \mu)} + \frac{3}{2}}$ |
| $I_{inf}^*(\mu)$ | $-\frac{2}{\mu}\sqrt{\frac{3}{2}}$ | 0 | 0 |
| $T_{inf}^*(\mu)$ | $-\frac{2}{\mu}\sqrt{\frac{3}{2}}$ | $-2\sqrt{\frac{3}{2}}$ | 0 |

Table 7 The LIL limits for (iv) of Example 1 with p varying

| p | $<\frac{1}{2}$ | $=\frac{1}{2}$ | $>\frac{1}{2}$ |
|----------------|------------------------------------|------------------------|---|
| $Q_{sup}^*(p)$ | 0 | $\sqrt{\frac{3}{2}}$ | $\sqrt{2-p}$ |
| $Z_{sup}^*(p)$ | 0 | $\sqrt{\frac{3}{2}}$ | $\sqrt{2p - 2p\sqrt{\frac{1}{2} + p} + \frac{3}{2}}$ |
| $I_{sup}^*(p)$ | $\frac{1}{1-p}\sqrt{\frac{3}{2}}$ | $2\sqrt{\frac{3}{2}}$ | 0 |
| $T_{sup}^*(p)$ | $\frac{1}{1-p}\sqrt{\frac{3}{2}}$ | 0 | 0 |
| $Q_{inf}^*(p)$ | 0 | 0 | $-\sqrt{2-p}$ |
| $Z_{inf}^*(p)$ | 0 | 0 | $-\sqrt{2p - 2p\sqrt{\frac{1}{2} + p} + \frac{3}{2}}$ |
| $I_{inf}^*(p)$ | $-\frac{1}{1-p}\sqrt{\frac{3}{2}}$ | 0 | 0 |
| $T_{inf}^*(p)$ | $-\frac{1}{1-p}\sqrt{\frac{3}{2}}$ | $-2\sqrt{\frac{3}{2}}$ | 0 |

E.3 More on Example 2

To supplement Sect. 6, we provide numerical results for case (ii) of Example 2 in Table 8.

Table 8 The LIL limits for (ii) of Example 2 with c_s varying

| μ | $Q_{\text{sup}}^*(c_a)$ | $Z_{\text{sup}}^*(c_a)$ | $I_{\text{sup}}^*(c_a)$ | $T_{\text{sup}}^*(c_a)$ | $D_{\text{sup}}^*(c_a)$ |
|---------------|-------------------------|-------------------------|-------------------------|-------------------------|-------------------------|
| 2 | 0 | 0 | $\frac{1}{2}c_s$ | $\frac{1}{2}c_s$ | 0 |
| 1 | c_s | c_s | c_s | 0 | 0 |
| $\frac{1}{2}$ | $\sqrt{1/2}c_s$ | $2c_s$ | 0 | 0 | $\sqrt{1/2}c_s$ |
| μ | $Q_{\text{inf}}^*(c_a)$ | $Z_{\text{inf}}^*(c_a)$ | $I_{\text{inf}}^*(c_a)$ | $T_{\text{inf}}^*(c_a)$ | $D_{\text{inf}}^*(c_a)$ |
| 2 | 0 | 0 | $-\frac{1}{2}c_s$ | $-\frac{1}{2}c_s$ | 0 |
| 1 | 0 | 0 | 0 | $-c_s$ | 0 |
| $\frac{1}{2}$ | $-\sqrt{1/2}c_s$ | $-2c_s$ | 0 | 0 | $-\sqrt{1/2}c_s$ |

References

Borovkov, A. A. (1984). *Asymptotic methods in queueing theory*. New York: Wiley.

Chen, H., & Mandelbaum, A. (1994). Hierarchical modeling of stochastic network, part II: Strong approximations. In D. D. Yao (Ed.), *Stochastic modeling and analysis of manufacturing systems* (pp. 107–131).

Chen, H., & Shanthikumar, J. G. (1994). Fluid limits and diffusion approximations for networks of multi-server queues in heavy traffic. *Discrete Event Dynamic Systems*, 4, 269–291.

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

Chen, H., & Yao, D. D. (2001). *Fundamentals of queueing networks*. New York: Springer.

Csörgő, M., Deheuvels, P., & Horváth, L. (1987). An approximation of stopped sums with applications in queueing theory. *Advances in Applied Probability*, 19(3), 674–690.

Csörgő, M., & Horváth, L. (1993). *Weighted approximations in probability and statistics*. New York: Wiley.

Csörgő, M., & Révész, P. (1981). *Strong approximations in probability and statistics*. New York: Academic.

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.

Ethier, S. N., & Kurtz, T. G. (1986). *Markov processes: Characterization and convergence*. New York: Wiley.

Glynn, P. W., & Whitt, W. (1986). A central-limit-theorem version of $L = \lambda W$. *Queueing Systems*, 1(2), 191–215.

Glynn, P. W., & Whitt, W. (1987). Sufficient conditions for functional limit theorem versions of $L = \lambda W$. *Queueing Systems*, 1(3), 279–287.

Glynn, P. W., & Whitt, W. (1988). An LIL version of $L = \lambda W$. *Mathematics of Operations Research*, 13(4), 693–710.

Glynn, P. W., & Whitt, W. (1991a). A new view of the heavy-traffic limit for infinite-server queues. *Advances in Applied Probability*, 23(1), 188–209.

Glynn, P. W., & Whitt, W. (1991b). Departures from many queues in series. *Annals of Applied Probability*, 1(4), 546–572.

Guo, Y., & Liu, Y. (2015). A law of iterated logarithm for multiclass queues with preemptive priority service discipline. *Queueing Systems*, 79, 251–291.

Harrison, J. M. (1985). *Brownian motion and stochastic flow system*. New York: Wiley.

Horváth, L. (1984a). Strong approximation of renewal processes. *Stochastic Processes and Their Applications*, 18(1), 127–138.

Horváth, L. (1984b). Strong approximation of extended renewal processes. *The Annals of Probability*, 12(4), 1149–1166.

Horváth, L. (1992). Strong approximations of open queueing networks. *Mathematics of Operations Research*, 17(2), 487–508.

Iglehart, G. L. (1971). Multiple channel queues in heavy traffic: IV. Law of the iterated logarithm. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete*, 17, 168–180.

- Lee, H. S., & Srinivasan, M. M. (1989). Control policies for $M^X/G/1$ queueing system. *Management Science*, 35(6), 708–721.
- Lévy, P. (1937). *Théorie de l'addition des variables aléatoires*. Paris: Gauthier-Villars.
- Lévy, P. (1948). *Processus stochastique et mouvement Brownien*. Paris: Gauthier-Villars.
- Liu, Y., & Whitt, W. (2014a). Algorithms for time-varying networks of many-server fluid queues. *INFORMS Journal on Computing*, 26(1), 59–73.
- Liu, Y., & Whitt, W. (2014b). Stabilizing performance in networks of queues with time-varying arrival rates. *Probability in the Engineering and Informational Sciences*, 28(4), 419–449.
- 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.
- Machihara, F. (1999). A $BMAP/SM/1$ queue with service times depending on the arrival process. *Queueing Systems*, 33(4), 277–291.
- Mandelbaum, A., & Massey, W. A. (1995). Strong approximations for time-dependent queues. *Mathematics of Operations Research*, 20(1), 33–64.
- Mandelbaum, A., Massey, W. A., & Reiman, M. (1998). Strong approximations for Markovian service networks. *Queueing Systems*, 30, 149–201.
- Minkevičius, S. (2014). On the law of the iterated logarithm in multiserver open queueing networks. *Stochastics*, 86(1), 46–59.
- 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.
- Pang, G. D., & Whitt, W. (2012). Infinite-server queues with batch arrivals and dependent service times. *Probability in the Engineering and Information Sciences*, 26, 197–220.
- Reiman, M. I. (1984). Open queueing networks in heavy traffic. *Mathematic of Operations Research*, 9, 441–458.
- 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.
- Strassen, V. (1964). An invariance principle for the law of the iterated logarithm. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete*, 3(3), 211–226.
- Van Ommeren, J. C. W. (1990). Simple approximations for the batch-arrival $M^X/G/1$ queue. *Operations Research*, 38(4), 678–685.
- Whitt, W. (1983). Comparing batch delays and customer delays. *Bell System Technical Journal*, 62(7), 2001–2009.
- Whitt, W. (2002). *Stochastic-process limits*. New York: Springer.
- Yom-Tov, G., & Mandelbaum, A. (2014). Erlang R: A time-varying queue with reentrant customers, in support of healthcare staffing. *Manufacturing and Service Operations Management*, 16, 283–299.
- Zhang, H. (1997). Strong approximations of irreducible closed queueing networks. *Advances in Applied Probability*, 29(2), 498–522.
- Zhang, H., & Hsu, G. X. (1992). Strong approximations for priority queues: Head-of-the-line-first discipline. *Queueing Systems*, 10(3), 213–234.
- Zhang, H., Hsu, G. X., & Wang, R. X. (1990). Strong approximations for multiple channels in heavy traffic. *Journal of Applied Probability*, 27(3), 658–670.