Polling Systems in Heavy Traffic: A Bessel Process Limit

E. G. Coffman, Jr.,[†] *A. A. Puhalskii*,[#] and *M. I. Reiman*[†]

[†]AT&T Bell Laboratories, Murray Hill, NJ 07974

[#]Institute for Problems in Information Transmission, Moscow, Russia 101447

ABSTRACT

This paper studies the classical polling model under the exhaustive-service assumption; such models continue to be very useful in performance studies of computer/communication systems. The analysis here extends earlier work of the authors to the general case of nonzero switchover times. It shows that, under the standard heavy-traffic scaling, the total unfinished work in the system tends to a Bessel-type diffusion in the heavy-traffic limit. It verifies in addition that, with this change in the limiting unfinished-work process, the averaging principle established earlier by the authors carries over to the general model.

September 13, 1995

Polling Systems in Heavy Traffic: A Bessel Process Limit

E. G. Coffman, Jr.,[†] *A. A. Puhalskii,*[#] and *M. I. Reiman*[†]

[†]AT&T Bell Laboratories, Murray Hill, NJ 07974

[#]Institute for Problems in Information Transmission, Moscow, Russia 101447

1 Introduction

In classical polling models, $M \geq 2$ queues are visited by a single server in cyclic order. Such models have many applications in the performance analysis of communication systems, including token rings and packet switches, where a single-server resource (e.g., a communication link) is shared among many demands on the resource (e.g., traffic streams). An analysis of the 5ESS® switching system is a modern example [8] and a sequel to work on earlier switching systems [7]. Introductions to a massive literature addressing many different applications can be found in [9] and [13].

This paper focuses on polling with *exhaustive* service: the visit of the server to any given queue terminates only when no work remains to be done at that queue. We number the queues from 1 to M and assume they are served in that order. The time for the server to switch over (or move) from queue i to queue $i+1$ is nonzero in general, and is allowed to be random and to depend on i.

An exact analysis of exhaustive polling systems is quite difficult; hopes for explicit solutions are soon abandoned in favor of numerical methods and approximations. A recent study of asymptotic behavior derived from heavy-traffic (diffusion) limits has been a promising approach, one that leads to relatively simple formulas which in turn yield useful insights. The cornerstone of the theory is an averaging principle proved in [3] by the present authors. In a recent application of this principle, Reiman and Wein [12] studied set-up scheduling problems in two-class single-server queues.

A limitation of the results in [3] is the often untenable assumption of zero switchover times. The main contribution of this paper is a proof that the total unfinished work in the general two-queue system tends, in the heavy traffic limit, to a Bessel type diffusion rather than the reflected Brownian motion in the case of zero switchover times. We verify that, as a corollary, the averaging principle in [3] carries over to the general model. The remainder of this section describes the averaging principle and gives a heuristic argument leading to the new diffusion limit for exhaustive polling systems. Section 2 introduces notation and formulates our main results. A threshold queue very similar to the one in [3] is analyzed in Section 3. Results for this queueing system supply bounds for the polling system, as shown in Section 4. Further preliminaries are taken up in Section 5, where the tightness of a number of basic processes is proved. The development of sections 3–5 culminates in the proofs in Section 6 of our main results. A critical element in the proofs is a semimartingale representation of the unfinished work process which allows us to use general convergence results for semimartingales [6, 10]. Concluding remarks are given in Section 7.

Briefly, the mathematical model is as follows. Customers arrive at the ith queue in a renewal process with rate λ_i and interarrival-time variance σ_{ai}^2. The service rate parameter at the ith queue is μ_i and the service-time variance is σ_{si}^2. Let d_i be the mean switchover time from queue i to queue $i+1$. Define $\rho = \rho_1 + \cdots + \rho_M$, where $\rho_i = \lambda_i/\mu_i$ is the traffic intensity at queue i.

We first review the case of $M = 2$ queues and zero switchover times $d_i = 0, 1 \le i \le M$, as presented in [3]. Let $U_t, t \ge 0$, denote the total unfinished work (service time) in queues 1 and 2 at time t. Then since the process $(U_t, t \ge 0)$ is the same as the unfinished work process in the corresponding $\Sigma GI/G/1$ system, we can extend the heavy-traffic limit theorem of Iglehart and Whitt [4] as follows (see also [11]). Consider a sequence of systems indexed by n, and let ρ^n denote the traffic intensity of the nth system. The heavy-traffic limit stipulates that $\rho^n \to 1$ as $n \to \infty$ with

$$\sqrt{n}(\rho^n - 1) \equiv c^n \to c \text{ as } n \to \infty, \ -\infty < c < \infty.$$

(As in the standard set-up, we also assume that $\lambda_i^n \to \lambda_i > 0$, $(\sigma_{si}^n)^2 \to \sigma_{si}^2$ as $n \to \infty$, $i = 1, 2$. There is one more technical assumption that we defer until later; it implies that the Lindeberg condition holds.) For the scaled process $V_t^n = n^{-1/2} U_{nt}^n$, $n \ge 1$, $0 \le t \le 1$, under the above conditions, $V^n \xrightarrow{d} V$, as $n \to \infty$, where V is reflected Brownian motion with drift c and infinitesimal variance

$$\sigma^2 \equiv \sum_{i=1}^{M} \lambda_i (\sigma_{si}^2 + \rho_i^2 \sigma_{ai}^2) > 0.$$

The averaging principle proved in [3] deals with queue lengths; converted to unfinished work, the principle states that, for any continuous function $f : \mathbf{R}_+ \to \mathbf{R}$ and any $T > 0$, we have

$$\int_0^T f(V_t^{n,i}) dt \xrightarrow{d} \int_0^T \left(\int_0^1 f(uV_t) du \right) dt, \quad i = 1, 2, \tag{1.1}$$

where $V_t^{n,i}$ is the time scaled and normalized unfinished work at queue i and the symbol $\xrightarrow{d}$ denotes convergence in distribution . Extended to general M, the corresponding averaging principle for sojourn times $W_t(i)$ at queue i is given by

$$\int_0^T f(Y_t^{n,i}dt \xrightarrow{d} \int_0^T \int_0^1 f\left(\frac{1-\rho_i}{\varrho} V_t u\right) dudt, \quad 1 \le i \le M, \tag{1.2}$$

where $Y_t^{n,i} = W_{nt}^n(i)/\sqrt{n},\ n \ge 1,\ 0 \le t \le 1$, and $\varrho = \sum_{1 \le j < k \le M} \rho_j \rho_k$.

We now return to nonzero switchover times with the expected values $d_i,\ 1 \le i \le M$. While a similar averaging principle can be expected, the unfinished-work process is no longer the same as in the $\Sigma GI/G/1$ system, so the limit diffusion V may be different. To see what this limit process should be, we give the following heuristic argument.

Consider the same sequence of systems as before, assuming in addition to the previous conditions that $d_i^n \to d_i,\ 0 \le d_i < \infty$. As before, $V_t^n = n^{-1/2} U_{nt}^n$. The drift $c(x)$ of the limit process V at point x is the limit $\Delta \to 0,\ n \to \infty$ of

$$c_\Delta^n(x) = \Delta^{-1} E[V_{t+\Delta}^n - V_t^n | V_t^n = x > 0].$$

Work enters the system at rate ρ^n per unit time. We assume that Δ is small enough that V_t^n does not reach zero during $[t, t+\Delta]$. With nonzero switchover times, work leaves the system at a rate less than 1, which we calculate as follows. Let $r^n(x)$ denote the fraction of time the server spends doing useful work (not switching) when $V_t^n = x$. Then $r^n(x)$ is the rate at which work leaves the system. Since there are $O(\sqrt{n})$ cycles per unit of 'diffusion' time, we can write

$$r^n(x) = \frac{E[\text{useful work done over a cycle}]}{\text{E[duration of a cycle]}}.$$

For simplicity, let $M = 2$ and start the cycle at the moment the server switches to queue 1. On average, it takes time $\frac{\sqrt{n}\,x}{1-\rho_1}$ to empty queue 1, d_1 to switch to queue 2, $\frac{\sqrt{n}\,x}{1-\rho_2}$ to empty queue 2, and d_2 to switch back to queue 1. The useful work is $\omega(x) = \frac{\sqrt{n}\,x}{1-\rho_1} + \frac{\sqrt{n}\,x}{1-\rho_2}$, so we have $r^n(x) = \frac{\omega(x)}{\omega(x)+d_1+d_2}$. But in heavy traffic, $\rho_1 + \rho_2 = 1$, so a little algebra yields

$$r^n(x) = \frac{x}{x + d/\sqrt{n}},$$

where $d = \rho_1\rho_2(d_1 + d_2)$. Extending the calculation to general M gives the same result with d generalized to $d = \varrho(d_1 + \cdots + d_M)$. Now $c_\Delta^n(x) = \sqrt{n}[\rho^n - r^n(x)]$, so the limit $\Delta \to 0,\ n \to \infty$ yields

$$c_\Delta^n(x) \to c(x) \equiv c + d/x.$$

Note that the seemingly innocuous addition of a $O(1)$ switchover time to a cycle which takes $O(\sqrt{n})$ time (before normalization) produces a dramatic change in the form of the drift.

A heuristic calculation along the above lines shows that the infinitesimal variance is unaffected by the addition of switchover times. We are thus led to expect that $V^n \to V$, where the limit process V is a one-dimensional diffusion with state dependent drift $c(x)$, and constant variance σ^2. This fact is proved rigorously for $M = 2$. The limit process is a Bessel process with negative drift. When $2\varrho/\sigma^2 < 1$, the process can hit the origin, in which case it instantaneously reflects. When $c < 0$, V is positive recurrent and has a stationary distribution with density

$$\pi(x) = \frac{a(ax)^\beta e^{-ax}}{\Gamma(\beta+1)}, \ x \geq 0, \tag{1.3}$$

where $a = 2|c|/\sigma^2$, $\beta = 2d/\sigma^2$. This is the gamma density of order β and scale a.

We further verify that the averaging principles (1.1) and (1.2) hold for $M = 2$, with V the above Bessel process. Extended to general M, we have

$$\frac{1}{T}\int_0^T f(Y_t^{n,i})dt \xrightarrow{d} \frac{1}{T}\int_0^T \int_0^1 f\left(\frac{1-\rho_i}{\varrho}V_t u\right) du dt,$$

so if we let V_0 have the stationary distribution (1.3), then

$$E\frac{1}{T}\int_0^T \int_0^1 f\left(\frac{1-\rho_i}{\varrho}V_t u\right) du dt = \int_0^\infty \frac{a(ax)^\beta e^{-x}}{\Gamma(\beta+1)} \left[\int_0^1 f\left(\frac{1-\rho_i}{\varrho}ux\right) du\right] dx.$$

For example, if $f(x) = x$, we find that the limiting sojourn times have the means

$$\frac{\beta+1}{a}\frac{1-\rho_i}{2\varrho}.$$

2 Results

We begin with notational matters. In the standard set-up for heavy traffic limits, we consider a sequence of two-queue polling systems. For the n^{th} system, denote by

$\tau_i^{n,\ell} = \xi_1^{n,\ell} + \cdots + \xi_i^{n,\ell}, i \geq 1, \ell = 1, 2$, the time of the i^{th} arrival to the ℓ^{th} queue in terms of interarrival times $\xi_i^{n,\ell}$,

$\eta_i^{n,\ell}$, $i \geq 1, \ell = 1, 2$, the i^{th} service time in the ℓ^{th} queue,

$s_i^{n,\ell}$, $i \geq 1, \ell = 1, 2$, the i^{th} switchover time from the ℓ^{th} queue.

We assume that $\xi_i^{n,\ell}, i \geq 1,\ \eta_i^{n,\ell}, i \geq 1,\ s_i^{n,\ell}, i \geq 1,\ \ell = 1,2$, are independent i.i.d. sequences. As in the previous section, we introduce, for $n = 1,2,\ldots$ and $\ell = 1,2$,

$$\lambda_\ell^n = (E\xi_1^{n,\ell})^{-1}, \quad \mu_\ell^n = (E\eta_1^{n,\ell})^{-1}, \quad d_\ell^n = Es_1^{n,\ell} \ ,$$
$$\rho_\ell^n = \frac{\lambda_\ell^n}{\mu_\ell^n}, \qquad \rho^n = \rho_1^n + \rho_2^n \ .$$

Instead of dealing with the variances $(\sigma_{ai}^n)^2$ and $(\sigma_{si}^n)^2$, it is more convenient here to introduce

$$(\sigma_\ell^n)^2 = E(\eta_1^{n,\ell} - \rho_\ell^n \xi_1^{n,\ell})^2 \ , \ell = 1,2 \ .$$

As in Section 1, we assume the limits, as $n \to \infty$ for $\ell = 1,2$,

$$\lambda_\ell^n \to \lambda_\ell, \quad \mu_\ell^n \to \mu_\ell > 0, \quad \sigma_\ell^n \to \sigma_\ell \ , \tag{2.1}$$

$$d_\ell^n \to d_\ell \ , \tag{2.2}$$

and assume the heavy traffic condition

$$\lim_{n\to\infty} \sqrt{n}(\rho^n - 1) = c \ . \tag{2.3}$$

The Lindeberg conditions mentioned earlier are, for $\ell = 1,2$, $\epsilon > 0$,

$$\lim_{n\to\infty} E(\xi_1^{n,\ell})^2 \cdot 1(\xi_1^{n,\ell} > \epsilon\sqrt{n}) = 0 \ , \tag{2.4}$$

$$\lim_{n\to\infty} E(\eta_1^{n,\ell})^2 \cdot 1(\eta_1^{n,\ell} > \epsilon\sqrt{n}) = 0 \ , \tag{2.5}$$

$$\lim_{n\to\infty} E(s_1^{n,\ell})^2 \cdot 1(s_1^{n,\ell} > \epsilon\sqrt{n}) = 0 \ . \tag{2.6}$$

Also let

$$\sigma^2 = \lambda_1\sigma_1^2 + \lambda_2\sigma_2^2 > 0, \quad \rho_\ell = \lambda_\ell/\mu_\ell \ , \ \ell = 1,2.$$

Recall that U_t^n is the total unfinished work in the n^{th} system at time t, with U_0^n independent of $\{\xi_i^{n,\ell}, i \geq 1\}$, $\{\eta_i^{n,\ell}, i \geq 1\}$, and $\{s_i^{n,\ell}, i \geq 1\}$, $\ell = 1,2$, and that

$$V_t^n = \frac{1}{\sqrt{n}} U_{nt}^n, \quad t \geq 0 \ ,$$
$$V^n = (V_t^n, t \geq 0) \ .$$

Recalling that $d = \rho_1\rho_2(d_1 + d_2)$, let the process $X = (X_t, t \geq 0)$ solve the equation

$$dX_t = [2(d + c(X_t \vee 0)^{1/2}) + \sigma^2]dt + 2\sigma(X_t \vee 0)^{1/2} dW_t, \quad X_0 \geq 0 \ , \tag{2.7}$$

where $W = (W_t, t \geq 0)$ is a standard Brownian motion, and X_0 and W are independent. Next, define $V = (V_t, t \geq 0)$ as the diffusion process on $[0, \infty)$ with the generator

$$Lg(x) = \left(c + \frac{d}{x}\right) \frac{dg}{dx}(x) + \frac{1}{2}\sigma^2 \frac{d^2 g}{dx^2}(x) \ ,$$

where the domain of L is

$$D(L) = \{g \in C_K^2([0, \infty)) : g(x) = \tilde{g}(x^2) \text{ for some } \ \tilde{g} \in C_K^2([0, \infty))\} \ ,$$

$C_K^2([0, \infty))$ being the space of twice continuously differentiable functions on $[0, \infty)$ with compact support.

A proof of the following technical result is similar to the proof of the existence of the Bessel diffusion [5, Chapter 4, Examples 8.2 and 8.3].

Lemma 2.1 *For given X_0 and V_0, the processes X and V exist and are unique in law. If V_0 is distributed as $\sqrt{X_0}$, then the distributions of V and $\sqrt{X}$ coincide.*

In the main result below, and throughout the remainder of the paper, all processes are considered as random elements of the Skorohod space $D[0, \infty)$ (see, e.g., [6, 10]), and convergence in distribution for the processes is understood as weak convergence of the induced measures on $D[0, \infty)$. By $\xrightarrow{d}$ we denote convergence in distribution in an appropriate metric space.

Theorem 2.1 *Assume that $V_0^n \xrightarrow{d} V_0$, as $n \to \infty$, where V_0 is a nonnegative random variable. If conditions (2.1) - (2.6) hold, then*

$$V^n \xrightarrow{d} V.$$

As a consequence of Theorem 2.1, we get the averaging principle for unfinished work. Let $U_t^{n,\ell}, t \geq 0$, denote the unfinished work at time t at queue $\ell = 1, 2$, and define $V_t^{n,\ell} = U_{nt}^{n,\ell}/\sqrt{n}$, $V^{n,\ell} = (V_t^{n,\ell}, t \geq 0)$.

Theorem 2.2 *Let $f(x), x \geq 0$, be a real-valued continuous function. Then, under the conditions of Theorem 2.1, for $t > 0$,*

$$\int_0^t f(V_s^{n,\ell})\, ds \xrightarrow{d} \int_0^t \left(\int_0^1 f(uV_s)\, du\right) ds, \quad \ell = 1, 2.$$

To conclude this section, it is instructive to compare the result of Theorem 2.1 to a related Bessel process limit obtained by Yamada [14, Theorem 1]. Note that the process of total unfinished work satisfies the equation

$$U_t^n = U_0^n + S_t^n - \int_0^t 1(U_s^n > 0)\alpha_s^n\, ds, \tag{2.8}$$

where

$$S_t^n = S_t^{n,1} + S_t^{n,2},\ S_t^{n,\ell} = \sum_{i=1}^{A_t^{n,\ell}} \eta_i^{n,\ell}, \quad \ell = 1,2, \tag{2.9}$$

$A^{n,\ell} = (A_t^{n,\ell}, t \geq 0)$, $\ell = 1,2$, are the input processes, i.e.,

$$A_t^{n,\ell} = \max\left(j : \sum_{i=1}^{j} \xi_i^{n,\ell} \leq t \right),$$

and α_s^n is the indicator of the event that the server is not switching over (i.e., is serving) at time s.

According to (2.8), if $U_s^n > 0$, then the instantaneous rate at which work leaves the system is α_s^n. The heuristic argument of Section 1 shows that it is reasonable to replace α_s^n by $r^n(U_s^n)$, i.e., consider the process $\breve{U}^n = (\breve{U}_t^n, t \geq 0)$ defined as the solution to

$$\breve{U}_t^n = U_0^n + S_t^n - \int_0^t 1(\breve{U}_s^n > 0) r^n(\breve{U}_s^n)\, ds \tag{2.10}$$

as an approximation for U^n. Equation (2.10) is of the type studied by Yamada. The conditions of our Theorem 2.1 allow us, with some reservations, to apply his Theorem 1; the limit process that this gives us turns out to be the same as the one in Theorem 2.1.

This comparison justifies our guess that $r^n(U_s^n)$ can be substituted for α_s^n in (2.8). Moreover, it is plausible to conjecture that one can weaken the much more restrictive conditions of Yamada's result. Indeed, the techniques used in the proof of Theorem 2.1 can be applied to a proof of the following generalization of Yamada's result. In this generalization, we assume that $\breve{U}^n = (\breve{U}_t^n, t \geq 0)$ is a nonnegative process satisfying (2.10), where $r^n(x), x \geq 0$, is a nonnegative bounded function. We further let $\breve{V}_t^n = \breve{U}_{nt}^n/\sqrt{n}, t \geq 0$, $\breve{V}^n = (\breve{V}_t^n, t \geq 0)$ and $\overline{r}^n = \sup_{x\geq 0} r^n(x)$. The previous notation is preserved.

Theorem 2.3 *Assume that $r^n(x)$ satisfies the following conditions*

(r1) $\lim_{x,n\to\infty} x(\overline{r}^n - r^n(x)) = d$,

(r2) $\sup_{x,n} x(\overline{r}^n - r^n(x)) < \infty$.

Assume that, as $n \to \infty$, $\sqrt{n}(\overline{r}^n - \rho^n) \to c$ and conditions (2.1), (2.4) and (2.5) hold. If $\breve{V}_0^n \xrightarrow{d} V_0$, then $\breve{V}^n \xrightarrow{d} V$.

The main improvements over Yamada's result are that we do not need the input processes to be Poisson (Yamada conjectured that this extension holds, but did not give a proof); we can do without the condition $\rho^n \leq \overline{r}^n$; and we do not require the existence of the fourth moments of service times.

3 A Threshold Queue

This section studies a threshold queue similar to the one considered in Section 3 of [3]. The distinction is that the busy periods start when the unfinished work exceeds a level h, i.e., the threshold is for the unfinished work and not the queue length as in [3].

We use the notation of [3]. Consider a sequence of threshold queues indexed by n. The generic interarrival and service times are denoted by ξ^n and η^n respectively. The threshold for the unfinished work in the nth queue is $h^n = \sqrt{n}a^n$, where a^n is a given constant. We are assuming that

$$\sup_n E(\xi^n)^2 < \infty, \quad \sup_n E(\eta^n)^2 < \infty \ , \tag{3.1}$$

and, letting $\lambda^n = (E\xi^n)^{-1}$ and $\mu^n = (E\eta^n)^{-1}$, assume that

$$\lim_{n\to\infty} \lambda^n = \lambda > 0, \quad \lim_{n\to\infty} \mu^n = \mu > 0, \quad \lim_{n\to\infty} a^n = a > 0 \ , \quad \lambda < \mu \ . \tag{3.2}$$

As in [3], within each busy period, at most one of the interarrival periods is allowed to be *exceptional*, i.e., have a distribution other than that of ξ^n. Specifically, for each $i \geq 1$, we introduce a nonnegative random variable $\tilde{\xi}_i^n$ and an integer-valued random variable χ_i^n which correspond to the i^{th} cycle. If there are at least χ_i^n arrivals in the i^{th} busy period, then the $(\chi_i^n)^{th}$ arrival has an exceptional interarrival period whose duration is taken to be $\tilde{\xi}_i^n$. If the busy period has less than χ_i^n arrivals, no exceptional arrivals occur. We assume that there exists a family of sequences $\{\zeta_i^n(r), i \geq 1\}$, $r > 0$, of identically distributed nonnegative random variables such that

$$\frac{1}{\sqrt{n}}\zeta_1^n(r) \xrightarrow{\text{P}} 0 \ \text{ as } \ n \to \infty, \ r > 0, \ \ \lim_{r\to\infty} \overline{\lim_{n\to\infty}} \sum_{i=1}^{\lfloor t\sqrt{n}\rfloor} P(\tilde{\xi}_i^n > \zeta_i^n(r)) = 0, \ t > 0 \ , \tag{3.3}$$

and that the joint distribution of $\zeta_i^n(r)$, the normal interarrival times, and the service times in the i^{th} cycle does not depend on i. We allow for two interarrival times to be dependent if one is taken from a busy period of the ith cycle and the other is taken either from another cycle or from an accumulation period of the ith cycle. However, interarrival (except for the exceptional), as well as service, times within each accumulation or busy period are assumed to be mutually independent. We also assume that the time of the first arrival, which we denote by $\overline{\xi}_1^n$, may have a distribution different from that of the generic interarrival time, and that

$$\frac{\overline{\xi}_1^n}{\sqrt{n}} \xrightarrow{\text{P}} 0 \ . \tag{3.4}$$

Introduce $X^n(t) = Y^n(nt)/\sqrt{n}$, $t \geq 0$, where $Y^n(t)$ is the unfinished work at t, and assume that $X^n(0) = 0$. In what follows, $\xrightarrow{P}$ denotes convergence in probability.

The following result is well-known and will be used several times in the remainder of the paper (see [4] for a proof).

Lemma 3.1 *Let $\{\zeta_i^n, i \geq 1\}, n \geq 1$, be a triangular array of nonnegative i.i.d. random variables such that, for any $\epsilon > 0$,*

$$\lim_{n\to\infty} E(\zeta_1^n)^2 \cdot 1(\zeta_1^n > \epsilon\sqrt{n}) = 0.$$

Let $N^n, n \geq 1$, be nonnegative integer-valued random variables such that, for some $q > 0$,

$$\lim_{n\to\infty} P\left(\frac{N^n}{n} > q\right) = 0.$$

Then as $n \to \infty$,

$$\frac{1}{\sqrt{n}} \max_{1\leq i\leq N^n} \zeta_i^n \xrightarrow{P} 0.$$

Theorem 3.1 *Let $f(x)$, $x \in R_+$, denote a bounded continuous function. If conditions (3.1)–(3.4) hold, then for any $T > 0$*

$$\int_0^T f(X^n(t))dt \xrightarrow{\mathrm{P}} T\int_0^1 f(au)du \quad \text{as} \quad n \to \infty\ .$$

Proof. We proceed as in the proof of Theorem 3.1 in [3]. Define the times

$$\begin{aligned}
\gamma_0^n &= 0\ , \\
\alpha_i^n &= \inf(t > \gamma_{i-1}^n : X^n(t) > 0)\ , i \geq 1\ , \\
\beta_i^n &= \inf(t > \gamma_{i-1}^n : X^n(t) > a^n),\ i \geq 1\ , \\
\gamma_i^n &= \inf(t > \beta_i^n : X^n(t) = 0),\ \ i \geq 1\ .
\end{aligned} \tag{3.5}$$

Note that the β_i^n start and the γ_i^n terminate busy periods. We prove that

$$\gamma_{\lfloor\sqrt{n}t\rfloor}^n \xrightarrow{\mathrm{P}} a\left(\frac{1}{\lambda} + \frac{1}{\mu-\lambda}\right)\mu t \quad \text{as} \quad n \to \infty\ , \tag{3.6}$$

and

$$\int_0^{\gamma_{\lfloor\sqrt{n}t\rfloor}^n} f(X^n(s))ds \xrightarrow{\mathrm{P}} \mu t\left(\frac{1}{\lambda} + \frac{1}{\mu-\lambda}\right)\int_0^a f(u)du \quad \text{as} \quad n \to \infty\ , \tag{3.7}$$

which immediately give the assertion of the theorem.

For $i \geq 2$, denote by $\overline{\xi}_i^n$ the time between γ_{i-1}^n and the first arrival after γ_{i-1}^n, i.e., $\overline{\xi}_i^n = \alpha_i^n - \gamma_{i-1}^n$; and denote by $\{\xi_{i,k}^{n,1}, k \geq 1\}$ and $\{\xi_{i,k}^{n,2}, k \geq 1\}$ the i.i.d. sequences, with generic random variable ξ^n, from which normal interarrival times on $[\alpha_i^n, \beta_i^n]$ and $[\beta_i^n, \alpha_{i+1}^n]$, respectively, are taken. Similarly, let $\{\eta_{i,k}^{n,l}, k \geq 1\}$, $i \geq 1$, $l = 1, 2$, be sequences from which service times of requests arriving in $[\alpha_i^n, \beta_i^n]$ and $[\beta_i^n, \gamma_i^n]$, respectively, are drawn. Note that, by the conditions of the theorem, the distribution of $\{\zeta_i^n(r), \xi_{i,k}^{n,l},\ \eta_{i,k}^{n,l},\ l = 1, 2,\ k \geq 1\}$ does not depend on $i = 1, 2, \ldots$.

In a sense, the $\overline{\xi}_i^n$, $i \geq 2$, also represent exceptional interarrival times. By Lemma 3.1 in [3], we know that they satisfy conditions similar to those imposed on $\tilde{\xi}_i^n$, i.e.,

$$\lim_{r\to\infty} \overline{\lim_{n\to\infty}} \sum_{i=2}^{\lfloor t\sqrt{n}\rfloor} P(\overline{\xi}_i^n > \overline{\zeta}_i^n(r)) = 0, \quad t > 0\ , \tag{3.8}$$

where

$$\overline{\zeta}_i^n(r) = \max_{1\leq k\leq \lfloor r\sqrt{n}\rfloor} \xi_{i-1,k}^{n,2}, \quad i \geq 2, r > 0\ .$$

Moreover, as $n \to \infty$,

$$\frac{\overline{\zeta}_i^n(r)}{\sqrt{n}} \xrightarrow{\mathrm{P}} 0, \quad i \geq 2, r > 0\ . \tag{3.9}$$

Define for $i \geq 1$

$$A_i^{n,1}(t) = 1(\overline{\xi}_i^n \leq t) + \sum_{k=1}^{\infty} 1\left(\overline{\xi}_i^n + \sum_{j=1}^{k} \xi_{i,j}^{n,1} \leq t\right)\ , \tag{3.10}$$

$$\begin{aligned} A_i^{n,2}(t) = &\sum_{k=1}^{\chi_i^n-1} 1\left(\sum_{j=1}^{k} \xi_{i,j}^{n,2} \leq t\right) + 1\left(\sum_{j=1}^{\chi_i^n-1} \xi_{i,j}^{n,2} + \tilde{\xi}_i^n \leq t\right) \\ &+ \sum_{k=\chi_i^n}^{\infty} 1\left(\sum_{j=1}^{k} \xi_{i,j}^{n,2} + \tilde{\xi}_i^n \leq t\right)\ , \end{aligned} \tag{3.11}$$

$$S_i^{n,l}(k) = \sum_{j=1}^{k} \eta_{i,j}^{n,l},\ k = 1, 2, \ldots,\ l = 1, 2\ . \tag{3.12}$$

For homogeneity of notation, we further set $\overline{\zeta}_1^n(r) = \overline{\xi}_1^n$. As in [3], by (3.3) and (3.8), it is enough to prove (3.6) and (3.7) on the events

$$\Gamma^n(r) = \bigcap_{i=1}^{\lfloor t\sqrt{n}\rfloor} \{\tilde{\xi}_i^n \leq \zeta_i^n(r), \overline{\xi}_i^n \leq \overline{\zeta}_i^n(r)\}\ .$$

Define the interval lengths

$$u_i^n = \beta_i^n - \gamma_{i-1}^n, v_i^n = \gamma_i^n - \beta_i^n, i \geq 1\ ,$$

so that by (3.5) and (3.10)–(3.12)

$$\begin{aligned} u_i^n &= \inf(t > 0 : \quad \frac{1}{\sqrt{n}} S_i^{n,1}(A_i^{n,1}(nt)) > a^n)\ , \\ v_i^n &= \inf(t > 0 : \quad \frac{1}{\sqrt{n}}(nt - S_i^{n,2}(A_i^{n,2}(nt))) > \frac{1}{\sqrt{n}} S_i^{n,1}(A_i^{n,1}(nu_i^n)))\ , \end{aligned} \tag{3.13}$$

and

$$\gamma_i^n - \gamma_{i-1}^n = u_i^n + v_i^n \ . \tag{3.14}$$

In analogy with (3.10) and (3.11), define (since r is fixed, it is omitted in the new notation below)

$$\begin{aligned}
\overline{A}_i^{n,1}(t) &= 1(\overline{\zeta}_i^n(r) \leq t) + \sum_{k=1}^{\infty} 1\left(\overline{\zeta}_i^n(r) + \sum_{j=1}^{k} \xi_{i,j}^{n,1} \leq t\right) \ , \\
\underline{A}_i^{n,1}(t) &= 1 + \sum_{k=1}^{\infty} 1\left(\sum_{j=1}^{k} \xi_{i,j}^{n,1} \leq t\right) \ , \\
\overline{A}_i^{n,2}(t) &= 1 + \sum_{k=1}^{\infty} 1\left(\sum_{j=1}^{k} \xi_{i,j}^{n,2} \leq t\right) \ , \\
\underline{A}_i^{n,2}(t) &= \sum_{k=1}^{\infty} 1\left(\zeta_i^n(r) + \sum_{j=1}^{k} \xi_{i,j}^{n,2} \leq t\right) \ ,
\end{aligned} \tag{3.15}$$

and define as in (3.13) and (3.14)

$$\begin{aligned}
\overline{u}_i^n &= \inf(t > 0 : \ \frac{1}{\sqrt{n}} S_i^{n,1}(\overline{A}_i^{n,1}(nt)) > a^n) \ , \\
\underline{u}_i^n &= \inf(t > 0 : \ \frac{1}{\sqrt{n}} S_i^{n,1}(\underline{A}_i^{n,1}(nt)) > a^n) \ , \\
\overline{v}_i^n &= \inf(t > 0 : \ \frac{1}{\sqrt{n}}(nt - S_i^{n,2}(\overline{A}_i^{n,2}(nt))) > \frac{1}{\sqrt{n}} S_i^{n,1}(A_i^{n,1}(nu_i^n))) \ , \\
\underline{v}_i^n &= \inf(t > 0 : \ \frac{1}{\sqrt{n}}(nt - S_i^{n,2}(\underline{A}_i^{n,2}(nt))) > \frac{1}{\sqrt{n}} S_i^{n,1}(A_i^{n,1}(nu_i^n))) \ ,
\end{aligned} \tag{3.16}$$

and

$$\overline{\gamma}_i^n = \sum_{j=1}^{i}(\overline{u}_j^n + \overline{v}_j^n), \quad \underline{\gamma}_i^n = \sum_{j=1}^{i}(\underline{u}_j^n + \underline{v}_j^n), \quad i \geq 1 \ , \quad \overline{\gamma}_0^n = \underline{\gamma}_0^n = 0 \ . \tag{3.17}$$

Since $\overline{\xi}_i^n \leq \overline{\zeta}_i^n(r)$, $\tilde{\xi}_i^n \leq \zeta_i^n(r)$, $1 \leq i \leq t\sqrt{n}$, on $\Gamma^n(r)$, we have by (3.10), (3.11) and (3.15) that

$$\begin{aligned}
\overline{A}_i^{n,1}(t) &\leq A_i^{n,1}(t) \leq \underline{A}_i^{n,1}(t) \ , \\
\underline{A}_i^{n,2}(t) &\leq A_i^{n,2}(t) \leq \overline{A}_i^{n,2}(t) \ ,
\end{aligned} \tag{3.18}$$

on $\Gamma^n(r)$, and hence by (3.13) and (3.16), for $1 \leq i \leq t\sqrt{n}$,

$$\underline{u}_i^n \leq u_i^n \leq \overline{u}_i^n, \ \underline{v}_i^n \leq v_i^n \leq \overline{v}_i^n \ , \tag{3.19}$$

on $\Gamma^n(r)$, and then by (3.14) and (3.17), for $1 \leq i \leq t\sqrt{n}$,

$$\underline{\gamma}_i^n - \underline{\gamma}_{i-1}^n \leq \gamma_i^n - \gamma_{i-1}^n \leq \overline{\gamma}_i^n - \overline{\gamma}_{i-1}^n \ , \tag{3.20}$$

on $\Gamma^n(r)$.

Now we prove (3.6) for $\overline{\gamma}_{\lfloor\sqrt{n}t\rfloor}^n$ and $\underline{\gamma}_{\lfloor\sqrt{n}t\rfloor}^n$; this will imply (3.6) for $\gamma_{\lfloor\sqrt{n}t\rfloor}^n$ on $\Gamma^n(r)$. Consider only the upper bound process. The proof for $\underline{\gamma}_{\lfloor\sqrt{n}t\rfloor}^n$ is similar.

First, note that by (3.15),

$$
\begin{aligned}
\underline{A}_i^{n,1}(t) &= \inf\left(k : \sum_{j=1}^{k} \xi_{i,j}^{n,1} > t\right) , \\
\underline{A}_i^{n,2}(t) &= \inf\left(k : \zeta_i^n(r) + \sum_{j=1}^{k+1} \xi_{i,j}^{n,2} > t\right) .
\end{aligned}
$$

Since $\{\xi_{i,k}^{n,l}, k \geq 1\}, \{\eta_{i,k}^{n,l}, k \geq 1\}$, $l = 1, 2$, are i.i.d., we have by (3.1) and (3.2) that

$$
\frac{1}{\sqrt{n}} \sum_{k=1}^{\lfloor \sqrt{n}t \rfloor} \xi_{i,k}^{n,l} \xrightarrow{\mathrm{P}} \frac{t}{\lambda}, \quad \frac{1}{\sqrt{n}} \sum_{k=1}^{\lfloor \sqrt{n}t \rfloor} \eta_{i,k}^{n,l} \xrightarrow{\mathrm{P}} \frac{t}{\mu}, \ l = 1, 2,
$$

and hence, by (3.3), (3.12), and Lemma 2.1 in [3],

$$
\frac{1}{\sqrt{n}} \underline{A}_i^{n,l}(\sqrt{n}t) \xrightarrow{\mathrm{P}} \lambda t, \ \frac{1}{\sqrt{n}} S_i^{n,l}(\underline{A}_i^{n,l}(\sqrt{n}t)) \xrightarrow{\mathrm{P}} \frac{\lambda}{\mu} t, \ l = 1, 2. \tag{3.21}
$$

Similarly, using (3.9),

$$
\frac{1}{\sqrt{n}} \overline{A}_i^{n,l}(\sqrt{n}t) \xrightarrow{\mathrm{P}} \lambda t, \ \frac{1}{\sqrt{n}} S_i^{n,l}(\overline{A}_i^{n,l}(\sqrt{n}t)) \xrightarrow{\mathrm{P}} \frac{\lambda}{\mu} t \ , l = 1, 2. \tag{3.22}
$$

By Lemma 2.1 in [3] and (3.16), (3.2)

$$
\sqrt{n} \overline{u}_i^n \xrightarrow{\mathrm{P}} \frac{a\mu}{\lambda}, \ \sqrt{n} \underline{u}_i^n \xrightarrow{\mathrm{P}} \frac{a\mu}{\lambda}. \tag{3.23}
$$

Hence by (3.19), $\sqrt{n} u_i^n \xrightarrow{\mathrm{P}} \frac{a\mu}{\lambda}$, and by (3.18), $A_i^{n,1}(\sqrt{n}t)/\sqrt{n} \xrightarrow{\mathrm{P}} \lambda t$, so that, if we define

$$
w_i^n = \frac{1}{\sqrt{n}} S_i^{n,1}(A_i^{n,1}(n u_i^n)) - a^n, \tag{3.24}
$$

then by the inequality

$$
0 \leq w_i^n \leq \frac{1}{\sqrt{n}} \sup_{1 \leq j \leq A_i^{n,1}(n u_i^n)} \eta_{i,j}^{n,1} \tag{3.25}
$$

together with (3.1) and Lemma 3.1, we have $w_i^n \xrightarrow{\mathrm{P}} 0$. This gives us by Lemma 2.1 in [3] and (3.16), (3.22), (3.2),

$$
\sqrt{n} \overline{v}_i^n \xrightarrow{\mathrm{P}} \frac{a\mu}{\mu - \lambda}, \ i \geq 1. \tag{3.26}
$$

Then, by (3.17) and (3.23),

$$
\sqrt{n}(\overline{\gamma}_i^n - \overline{\gamma}_{i-1}^n) \xrightarrow{\mathrm{P}} a\mu\left(\frac{1}{\lambda} + \frac{1}{\mu - \lambda}\right), \ i \geq 1 \ . \tag{3.27}
$$

Since

$$
\overline{\gamma}_{\lfloor \sqrt{n}t \rfloor}^n = \frac{1}{\sqrt{n}} \sum_{k=1}^{\lfloor \sqrt{n}t \rfloor} \sqrt{n}(\overline{\gamma}_k^n - \overline{\gamma}_{k-1}^n) \ ,
$$

and since $(\overline{\gamma}_i^n - \overline{\gamma}_{i-1}^n,\ i \ge 1)$ are identically distributed by construction, we would have, in view of Lemma 2.4 in [3],

$$\overline{\gamma}^n_{\lfloor\sqrt{n}t\rfloor} \xrightarrow{\text{P}} a\mu\left(\frac{1}{\lambda} + \frac{1}{\mu-\lambda}\right) t \tag{3.28}$$

provided

$$\lim_{k\to\infty} \overline{\lim_{n\to\infty}} \sqrt{n} P\left(\sqrt{n}(\overline{\gamma}_1^n - \overline{\gamma}_0^n) > k\right) = 0\ . \tag{3.29}$$

By (3.17), this would follow from

$$\begin{aligned} \lim_{k\to\infty} \overline{\lim_{n\to\infty}} \sqrt{n} P(\sqrt{n}\overline{u}_1^n > k) &= 0\ , \\ \lim_{k\to\infty} \overline{\lim_{n\to\infty}} \sqrt{n} P(\sqrt{n}\overline{v}_1^n > k) &= 0\ . \end{aligned} \tag{3.30}$$

Consider the first limit. By (3.16)

$$\begin{aligned} P(\sqrt{n}\overline{u}_1^n > k) &= P(S_1^{n,1}(\overline{A}_1^{n,1}(\sqrt{n}k)) \le \sqrt{n}a^n) \\ &\le P(\overline{A}_1^{n,1}(\sqrt{n}k) \le \frac{1}{2}\lambda^n\sqrt{n}k) + P(S_1^{n,1}(\frac{1}{2}\lambda^n\sqrt{n}k) \le \sqrt{n}a^n)\ . \end{aligned} \tag{3.31}$$

By (3.12) we have, applying Chebyshev's inequality and (3.1) and (3.2), that, for any $\varepsilon > 0$,

$$\lim_{k\to\infty} \overline{\lim_{n\to\infty}} \sqrt{n} P(|S_1^{n,l}(k\sqrt{n}) - \frac{k}{\mu^n}\sqrt{n}| \ge k\sqrt{n}\varepsilon) = 0\ ,\ l = 1, 2. \tag{3.32}$$

By an analogue of (3.32) for interarrival times, and by (3.15),

$$\lim_{k\to\infty} \overline{\lim_{n\to\infty}} \sqrt{n} P(|\overline{A}_1^{n,l}(\sqrt{n}k) - \lambda^n k\sqrt{n}| \ge k\sqrt{n}\varepsilon) = 0\ ,\ l = 1, 2. \tag{3.33}$$

Relations (3.31)–(3.33) prove the first convergence in (3.30).

For the second convergence, we first prove that

$$\lim_{k\to\infty} \overline{\lim_{n\to\infty}} \sqrt{n} P\left(w_1^n > \frac{1}{3}\left(1 - \frac{\lambda^n}{\mu^n}\right) k\right) = 0. \tag{3.34}$$

By (3.25), (3.18) and (3.19),

$$\begin{aligned} P\left(w_1^n > \frac{1}{3}\left(1 - \frac{\lambda^n}{\mu^n}\right) k\right) &\le P(\sqrt{n}u_1^n > k) + P(A_1^{n,1}(\sqrt{n}k) > \mu^n\sqrt{n}k) \\ +P\left(\sup_{1\le j\le\lfloor\mu^n\sqrt{n}k\rfloor} \eta_{i,j}^{n,1} > \frac{1}{3}\left(1 - \frac{\lambda^n}{\mu^n}\right) k\sqrt{n}\right) &\le P(\sqrt{n}\overline{u}_1^n > k) \\ +P(\underline{A}_1^{n,1}(\sqrt{n}k) > \mu^n\sqrt{n}k) + \mu^n\sqrt{n}kP&\left(\eta_{1,1}^{n,1} > \frac{1}{3}\left(1 - \frac{\lambda^n}{\mu^n}\right) k\sqrt{n}\right). \end{aligned}$$

We have proved that $\lim_{k\to\infty} \varlimsup_{n\to\infty} \sqrt{n}P(\sqrt{n}\overline{u}_1^n > k) = 0$; by an analogue of (3.33) for $\underline{A}_1^{n,l}$, (3.2), and since $\lambda < \mu$, $\lim_{k\to\infty} \varlimsup_{n\to\infty} \sqrt{n}P(\underline{A}_1^{n,1}(\sqrt{n}k) > \mu^n\sqrt{n}k) = 0$. Finally, by Chebyshev's inequality,

$$\sqrt{n}kP\left(\eta_{1,1}^{n,1} > \frac{1}{3}\left(1 - \frac{\lambda^n}{\mu^n}\right)k\sqrt{n}\right) \le \frac{9}{(1-\lambda^n/\mu^n)^2 k\sqrt{n}}E(\eta_{1,1}^{n,1})^2,$$

and applying (3.1) and (3.2), we arrive at (3.34).

Going back to the proof of the second inequality in (3.30), write, by (3.16) and (3.24), in analogy with (3.31),

$$\begin{aligned}
P(\sqrt{n}\overline{v}_1^n > k) = P(\sup_{t\le k}(\sqrt{n}t - S_1^{n,2}(\overline{A}_1^{n,2}(\sqrt{n}t))) \le \sqrt{n}w_1^n + \sqrt{n}a^n)& \\
\le P(S_1^{n,2}(\overline{A}_1^{n,2}(\sqrt{n}k)) \ge \sqrt{n}(k - w_1^n) - \sqrt{n}a^n)& \\
\le P(\overline{A}_1^{n,2}(\sqrt{n}k) > \frac{1}{2}(\lambda^n + \mu^n)\sqrt{n}k) + P\left(w_1^n > \frac{1}{3}\left(1 - \frac{\lambda^n}{\mu^n}\right)k\right)& \\
+P\left(S_1^{n,2}\left(\frac{1}{2}(\lambda^n+\mu^n)\sqrt{n}k\right) \ge \sqrt{n}\left(\left(\frac{1}{3} + \frac{2}{3}\frac{\lambda^n}{\mu^n}\right)k - a^n\right)\right)&.
\end{aligned}$$

Putting together (3.32), (3.33) and (3.34) yields the second convergence in (3.30).

To prove (3.7) on $\Gamma^n(r)$, we apply Lemma 2.4 in [3], i.e., we prove that

$$\lim_{n\to\infty}\frac{1}{\sqrt{n}}\sum_{i=1}^{\lfloor\sqrt{n}t\rfloor} P\left(\left\{\left|\sqrt{n}\int_{\gamma_{i-1}^n}^{\gamma_i^n} f(X^n(s))ds - \mu\left(\frac{1}{\lambda} + \frac{1}{\mu-\lambda}\right)\int_0^a f(u)du\right| > \epsilon\right\} \cap\Gamma^n(r)\right) = 0, \epsilon > 0\ , \qquad (3.35)$$

and

$$\lim_{k\to\infty}\varlimsup_{n\to\infty}\sum_{i=1}^{\lfloor\sqrt{n}t\rfloor} P\left(\left\{\sqrt{n}\left|\int_{\gamma_{i-1}^n}^{\gamma_i^n} f(X^n(s))ds\right| > k\right\}\cap\Gamma^n(r)\right) = 0\ . \qquad (3.36)$$

Note that (3.36) is easy. For, by the right inequality in (3.20) and the boundedness of f, we have, letting $\|\cdot\|$ denote the sup norm,

$$\varlimsup_{n\to\infty}\sum_{i=1}^{\lfloor\sqrt{n}t\rfloor} P\left(\left\{\sqrt{n}\left|\int_{\gamma_{i-1}^n}^{\gamma_i^n} f(X^n(s))ds\right| > k\right\}\cap\Gamma^n(r)\right) \le \varlimsup_{n\to\infty}\sum_{i=1}^{\lfloor\sqrt{n}t\rfloor} P(\sqrt{n}(\overline{\gamma}_i^n - \overline{\gamma}_{i-1}^n)\|f\| > k)$$

which tends to 0 as $k\to\infty$ by (3.29) and by the fact that the $(\overline{\gamma}_i^n - \overline{\gamma}_{i-1}^n)$, $i \ge 1$, are identically distributed.

By (3.14), (3.35) would follow if

$$\lim_{n\to\infty}\frac{1}{\sqrt{n}}\sum_{i=1}^{\lfloor\sqrt{n}t\rfloor} P\left(\left\{\left|\sqrt{n}\int_0^{u_i^n} f(X^n(s+\gamma_{i-1}^n))ds - \frac{\mu}{\lambda}\int_0^a f(u)du\right| > \epsilon\right\}\cap\Gamma^n(r)\right) = 0\ ,$$

and

$$\lim_{n\to\infty} \frac{1}{\sqrt{n}} \sum_{i=1}^{\lfloor \sqrt{n} t \rfloor} P\left(\left\{ \left| \sqrt{n} \int_0^{v_i^n} f(X^n(s+\beta_i^n))ds - \frac{\mu}{\mu-\lambda} \int_0^a f(u)du \right| > \epsilon \right\} \cap \Gamma^n(r) \right) = 0 \ . \tag{3.37}$$

so,

$$P\left(\left\{\sup_{u\le\sqrt{n}v_i^n\wedge\frac{a\mu}{\mu-\lambda}}\left|X^n\left(\frac{u}{\sqrt{n}}+\beta_i^n\right)-(a-(1-\frac{\lambda}{\mu})u)\right|>\eta\right\}\cap\ \Gamma^n(r)\right)$$

$$\le\ P\left(\left\{\sup_{u\le\frac{a\mu}{\mu-\lambda}}\left|\frac{S_i^{n,2}(A_i^{n,2}(\sqrt{n}u))}{\sqrt{n}}-\frac{\lambda}{\mu}u\right|>\frac{\eta}{3}\right\}\cap\Gamma^n(r)\right)$$

$$+\ P(w_i^n>\frac{\eta}{3})+1\left(|a^n-a|>\frac{\eta}{3}\right)\ .$$

Since the distributions of $(\underline{A}_i^{n,2}(t),t\ge 0)$, $(\overline{A}_i^{n,2}(t),t\ge 0)$ and $(S_i^{n,2}(t),t\ge 0)$ do not depend on i, we conclude from the above, (3.2) and (3.18) that the left-hand side of (3.40) is not greater than

$$t\ \overline{\lim_{n\to\infty}}\,P\left(\sup_{u\le\frac{a\mu}{\mu-\lambda}}\left|\frac{1}{\sqrt{n}}S_1^{n,2}(\overline{A}_1^{n,2}(\sqrt{n}u))-\frac{\lambda}{\mu}u\right|>\frac{\eta}{3}\right)$$

$$+\ t\ \overline{\lim_{n\to\infty}}\,P\left(\sup_{u\le\frac{a\mu}{\mu-\lambda}}\left|\frac{1}{\sqrt{n}}S_1^{n,2}(\underline{A}_1^{n,2}(\sqrt{n}u))-\frac{\lambda}{\mu}u\right|>\frac{\eta}{3}\right)$$

which is zero by (3.21) and (3.22); (3.40) is proved.

Now on the event

$$\left\{\sup_{u\le\sqrt{n}v_i^n\wedge\frac{a\mu}{\mu-\lambda}}\left|X^n\left(\frac{u}{\sqrt{n}}+\beta_i^n\right)-(a-(1-\frac{\lambda}{\mu})u)\right|\le\eta\right\}\ ,$$

we have that $X^n\left(\frac{u}{\sqrt{n}}+\beta_i^n\right)\le a+\eta$, $u\in[0,\sqrt{n}v_i^n\wedge\frac{a\mu}{\mu-\lambda}]$, and therefore, for $u\in[0,\sqrt{n}v_i^n\wedge\frac{a\mu}{\mu-\lambda}]$,

$$\left|f\left(X^n\left(\frac{u}{\sqrt{n}}+\beta_i^n\right)\right)-f(a-(1-\frac{\lambda}{\mu})u)\right|\le\omega_f(\eta,a+\eta)\ ,$$

where $\omega_f(\delta,T)$ is the modulus of continuity of f on $[0,T]$ for partitions of diameter δ. This implies by the continuity of f that, for all η small enough and for all i,

$$\left\{\sup_{u\le\sqrt{n}v_i^n\wedge\frac{a\mu}{\mu-\lambda}}\left|X^n\left(\frac{u}{\sqrt{n}}+\beta_i^n\right)-(a-(1-\frac{\lambda}{\mu})u)\right|\le\eta\right\}$$

$$\subset\ \left\{\int_0^{\sqrt{n}v_i^n\wedge\frac{a\mu}{\mu-\lambda}}\left|f\left(X^n\left(\frac{u}{\sqrt{n}}+\beta_i^n\right)\right)-f(a-(1-\frac{\lambda}{\mu})u)\right|du\le\frac{\epsilon}{2}\right\}\ ,$$

so for η small enough

$$P\left(\left\{\int_0^{\sqrt{n}v_i^n\wedge\frac{a\mu}{\mu-\lambda}}\left|f\left(X^n\left(\frac{u}{\sqrt{n}}+\beta_i^n\right)\right)-f(a-(1-\frac{\lambda}{\mu})u)\right|du>\frac{\epsilon}{2}\right\}\cap\Gamma^n(r)\right)$$

$$\le\ P\left(\left\{\sup_{u\le\sqrt{n}v_i^n\wedge\frac{a\mu}{\mu-\lambda}}\left|X^n\left(\frac{u}{\sqrt{n}}+\beta_i^n\right)-(a-(1-\frac{\lambda}{\mu})u)\right|>\eta\right\}\cap\Gamma^n(r)\right)\ ,$$

and so (3.39) follows from (3.40). Thus (3.37), (3.35) and (3.7) are proved. This completes the proof of the theorem. ∎

4 Upper and Lower Bounds for Unfinished Work

In this section, having in view the averaging principle, we derive a limit theorem for the integral $\int_0^T f(V_t^{n,1})\,dt$, where $f(x)$ is a real-valued continuous function on the positive half-line, assuming that $V^n \xrightarrow{d} \tilde{V}$ with some continuous process $\tilde{V}$. This is carried out by providing suitable upper and lower bounds for the unfinished work at an individual queue in analogy with the contents of Section 4 in [3]. The main result is the following.

Theorem 4.1 *Assume that, in addition to the conditions of Theorem 2.1,* $V^n \xrightarrow{\mathrm{d}} \tilde{V}$*, where* $\tilde{V} = (\tilde{V}_t, t \geq 0)$ *is a nonnegative continuous process such that, for any* $T > 0$*,* $\int_0^T 1(\tilde{V}_t = 0)\,dt = 0$ *P-a.s. Then, for any continuous function* $f(x)$ *on* R_+*,*

$$\int_0^T f(V_t^{n,1})\,dt \xrightarrow{\mathrm{d}} \int_0^T \left(\int_0^1 f(u\tilde{V}_t)du\right) dt \ .$$

Proof. We first assume that $f(x)$ is bounded and nonnegative. We note that it is enough to prove that

$$\int_0^T f(V_t^{n,1}) \cdot 1(\delta \leq V_t^n \leq K)dt \xrightarrow{\mathrm{d}} \int_0^T \left(\int_0^1 f(u\tilde{V}_t)du\right) \cdot 1(\delta \leq \tilde{V}_t \leq K)dt \ , \tag{4.1}$$

for any δ and K, $0 < \delta < K$, such that

$$\int_0^T [1(\tilde{V}_t = \delta) + 1(\tilde{V}_t = K)]dt = 0 \quad P\text{–a.s.} \tag{4.2}$$

The argument is given in the proof of Theorem 2.1 in [3]. So, we prove (4.1) assuming (4.2). As in [3], choose $\epsilon \in \left(0, \frac{\delta}{2}\right)$ such that $N = (K - \delta)/\epsilon$ is an integer and, given $r(\epsilon) < \epsilon/2$, let, for $0 \leq i \leq N$,

$$\begin{aligned}
a_i(\epsilon) &= \delta + i\epsilon \ , \\
B_{r(\epsilon)}(\epsilon, i) &= (a_i(\epsilon) - r(\epsilon), a_i(\epsilon) + r(\epsilon)) \ , \\
C_{r(\epsilon)}(\epsilon, i) &= (0, a_i(\epsilon) - \epsilon + r(\epsilon)) \cup (a_i(\epsilon) + \epsilon - r(\epsilon), \infty) \ , \\
\zeta_0^n(\epsilon, i) &= 0 \ , \\
\tau_k^n(\epsilon, i) &= \inf(t > \zeta_{k-1}^n(\epsilon, i) : V_t^n \in B_{r(\epsilon)}(\epsilon, i)), \quad k \geq 1 \ , \\
\zeta_k^n(\epsilon, i) &= \inf(t > \tau_k^n(\epsilon, i) : V_t^n \in C_{r(\epsilon)}(\epsilon, i)), \quad k \geq 1 \ , \\
\zeta_0(\epsilon, i) &= 0 \ , \\
\tau_k(\epsilon, i) &= \inf(t > \zeta_{k-1}(\epsilon, i) : \tilde{V}_t \in B_{r(\epsilon)}(\epsilon, i)), \quad k \geq 1 \ , \\
\zeta_k(\epsilon, i) &= \inf(t > \tau_k(\epsilon, i) : \tilde{V}_t \in C_{r(\epsilon)}(\epsilon, i)), \quad k \geq 1 \ .
\end{aligned}$$

The argument of the proof of Lemma 4.1 in [3] applied to V^n and $\tilde{V}$ shows that

$$\tau_k(\epsilon,i) < \zeta_k(\epsilon,i) \ \ P\text{–a.s. on } \ \{\tau_k(\epsilon,i)<\infty\} \ ,$$

$$\lim_{k\to\infty} P\left(\min_{0\le i\le N} \zeta_k(\epsilon,i) \le T\right) = 0 \ , \tag{4.3}$$

and that $r(\epsilon)$ can be chosen so that, as $n\to\infty$,

$$(V^n, (\tau_k^n(\epsilon,i)\wedge T,\ \zeta_k^n(\epsilon,i)\wedge T)_{k\ge 1, 0\le i\le N}) \xrightarrow{\mathrm{d}} \left(\tilde{V}, (\tau_k(\epsilon,i)\wedge T,\ \zeta_k(\epsilon,i)\wedge T)_{k\ge 1, 0\le i\le N}\right) \ . \tag{4.4}$$

where convergence is in R^∞. In analogy with [3], denote by $n\kappa_j^{n,1}(i,k)$, $j\ge 0$, the successive times after $n\tau_k^n(\epsilon,i)$, when the unfinished work at the first queue becomes equal to 0. These are times when switchovers from queue 1 to queue 2 start. We denote the switchover times starting after $n\tau_k^n(\epsilon,i)$ by $s_0^{n,1}(i,k)$, $s_1^{n,1}(i,k),\dots$. Obviously, they are independent and distributed as $s_1^{n,1}$. We also let

$$\vartheta_{i,k}^{n,1} = \begin{cases} \min(j:\kappa_{j+1}^{n,1}(i,k) > \zeta_k^n(\epsilon,i)\wedge T), & \text{if } \kappa_0^{n,1}(i,k)\le \zeta_k^n(\epsilon,i)\wedge T, \\ 0, & \text{if } \kappa_0^{n,1}(i,k) > \zeta_k^n(\epsilon,i)\wedge T. \end{cases}$$

Analogously, we denote by $n\kappa_j^{n,2}(i,k)$, $j\ge 0$, the successive times after $n\tau_k^n(\epsilon,i)$, when the unfinished work at queue 2 becomes equal to 0, and we denote by $s_0^{n,2}(i,k)$, $s_1^{n,2}(i,k),\dots$ the switchover times from queue 2 to queue 1 which start at $n\kappa_0^{n,2}(i,k)$, $n\kappa_1^{n,2}(i,k),\dots$. We note again that $s_0^{n,2}(i,k)$, $s_1^{n,2}(i,k),\dots$ are independent and distributed as $s_1^{n,2}$. Also let

$$\vartheta_{i,k}^{n,2} = \begin{cases} \min(j:\kappa_{j+1}^{n,2}(i,k) > \zeta_k^n(\epsilon,i)\wedge T), & \text{if } \kappa_0^{n,2}(i,k)\le \zeta_k^n(\epsilon,i)\wedge T, \\ 0, & \text{if } \kappa_0^{n,2}(i,k) > \zeta_k^n(\epsilon,i)\wedge T. \end{cases}$$

For given i,k and ϵ, let $n\kappa_j^n, j\ge 0$, be successive times after $n(\tau_k^n(\epsilon,i)\wedge T)$ when the first queue empties, and let $n\theta_j^n, j\ge 1$, be successive times after $n\kappa_0^n$ when the server comes back to the first queue; obviously, $\kappa_{j-1}^n < \theta_j^n < \kappa_j^n, j\ge 1$. Next, define the first passage times

$$\begin{aligned} \psi_j^n &= \inf(t>\kappa_{j-1}^n : V_t^{n,1} > a_i(\epsilon)-\epsilon)\wedge \theta_j^n,\ j\ge 1, \\ \phi_j^n &= \inf(t>\theta_j^n : V_t^{n,1} \le V_{\psi_j^n}^{n,1})),\ j\ge 1. \end{aligned}$$

Note that $n\kappa_j^n$ is a service completion time for the first queue, and if $\kappa_j^n\le \zeta_k^n(\epsilon,i)<\infty$, then $n\psi_j^n$ is an arrival time for the first queue.

Let the arrivals on $[n\kappa_0^n,\infty)$ be numbered successively starting from 1. Let $\tilde{\xi}_1^n$ denote the time period between $n\kappa_0^n$ and the first arrival. Denote by $\tilde{\xi}_l^n, l\ge 2$, the times between the

$(l-1)^{\text{th}}$ and l^{th} of these arrivals. Obviously, $\{\tilde{\xi}_l^n, l \geq 2\}$ is a set of i.i.d. random variables with the distribution of the generic interarrival time for the first queue.

Let $\tilde{\chi}_j^{n,1}$ be the index of the arrival occurring at or just before $n\psi_j^n, j \geq 1$, and let $\tilde{\chi}_j^{n,2}$ be the index of the arrival occurring just after $n\phi_j^n, j \geq 1$. For $j \geq 1$, let v_j^n be the time period between $n\phi_j^n$ and the $\tilde{\chi}_j^{n,2\text{th}}$ arrival. Denote by $\{\tilde{\eta}_{j,l}^n, l \geq 1\}, j = 1, 2, \ldots$, independent copies of the sequence of service times at the first queue, which are also independent of $\{\tilde{\xi}_l^n, l \geq 1\}$. Again by the i.i.d. assumptions, we may assume that, for each $1 \leq j \leq \vartheta_{i,k}^{n,1}$, the service times for completions in $(\phi_j^n, \kappa_j^n]$ are $\tilde{\eta}_{j,1}^n, \tilde{\eta}_{j,2}^n, \ldots$.

Now consider the threshold queue with the threshold $h^n = \sqrt{n}(a_i(\epsilon) - \epsilon)$ which has the sequence

$$\{\tilde{\xi}_1^n, \tilde{\xi}_2^n, \ldots, \tilde{\xi}_{\tilde{\chi}_1^{n,1}}^n, v_1^n, \tilde{\xi}_{\tilde{\chi}_1^{n,2}+1}^n, \ldots, \tilde{\xi}_{\tilde{\chi}_2^{n,1}}^n, v_2^n, \tilde{\xi}_{\tilde{\chi}_2^{n,2}+1}^n, \ldots\}$$

of interarrival times; service times in the j^{th} busy period of this queue are $\tilde{\eta}_{j,l}^n, l = 1, 2, \ldots$. Denote by $\tilde{V}_t^{n,1}$ the normalized and time-scaled unfinished work at this queue. Also let $\tilde{\beta}_i^n$ and $\tilde{\gamma}_i^n$ be defined for this queue as β_i^n and γ_i^n respectively in (3.5).

Then the construction above yields

$$\tilde{V}_t^{n,1} = \begin{cases} V_{t-\tilde{\gamma}_{j-1}^n + \kappa_{j-1}^n}^{n,1}, & t \in [\tilde{\gamma}_{j-1}^n, \tilde{\beta}_j^n], \quad 1 \leq j \leq \vartheta_{i,k}^{n,1}, \\ V_{t-\tilde{\beta}_j^n + \phi_j^n}^{n,1}, & t \in [\tilde{\beta}_j^n, \tilde{\gamma}_j^n], \quad 1 \leq j \leq \vartheta_{i,k}^{n,1}. \end{cases} \tag{4.5}$$

$$\begin{aligned} \tilde{\gamma}_j^n &= \sum_{l=1}^{j} [(\psi_l^n - \kappa_{l-1}^n) + (\kappa_l^n - \phi_l^n)], \quad 1 \leq j \leq \vartheta_{i,k}^{n,1}, \quad \tilde{\gamma}_0^n = 0, \\ \tilde{\beta}_j^n &= \tilde{\gamma}_{j-1}^n + (\psi_j^n - \kappa_{j-1}^n), \quad 1 \leq j \leq \vartheta_{i,k}^{n,1}, \quad \tilde{\beta}_0^n = 0. \end{aligned} \tag{4.6}$$

The exceptional interarrival times for this queue are v_1^n, $v_2^n, \ldots$, i.e., these are the interarrival times of the first arrivals in busy periods.

Equalities (4.5), (4.6), the definition of $\vartheta_{i,k}^{n,1}$, and the assumption $f \geq 0$ show that

$$\int_0^{\tilde{\vartheta}^n} f(\tilde{V}_t^{n,1}) dt \leq \int_{\tau_k^n(\epsilon,i)\wedge T}^{\zeta_k^n(\epsilon,i)\wedge T} f(V_t^{n,1}) dt, \tag{4.7}$$

where $\tilde{\vartheta}^n = \tilde{\gamma}_{\vartheta_{i,k}^{n,1}}^n$.

Now we construct an upper bound process $\hat{V}^{n,1}$. Whereas we truncated the original process $V^{n,1}$ at the level $a_i(\epsilon) - \epsilon$ to obtain the lower bound process, here we will extend $V^{n,1}$ to level $a_i(\epsilon) + \epsilon$ to obtain an upper bound process. Introduce independent replicas $\{\hat{\xi}_{j,\ell}^{n,k}, \ell \geq 1\}$,

$j \geq 1$, $k = 1, 2$, of the interarrival time sequence at the first queue and independent replicas $\{\hat{\eta}^n_{j,\ell}, \ell \geq 1\}$, $j \geq 1$, of the service time sequence at the first queue.

Let

$$\varphi^n_j = \inf(t > \kappa^n_{j-1} : V^{n,1}_t > a_i(\varepsilon) + \varepsilon) \wedge \theta^n_j,\ j \geq 1,$$
$$\vartheta^n_j = \inf(t > \theta^n_j : V^{n,1}_t \leq V^{n,1}_{\varphi^n_j}),\ j \geq 1.$$

Note that if $V^{n,1}_{\theta^n_j} \leq a_i(\varepsilon) + \varepsilon$, then $\varphi^n_j = \vartheta^n_j = \theta^n_j$. Let $\hat{\chi}^n_j, j \geq 1$, index the arrival in the original queue occurring at or just after $n\varphi^n_j$ (recall that the numbering starts from the arrival after $n\kappa^n_0$), and let $\overline{\upsilon}^n_j, j \geq 1$, denote the time between $n\varphi^n_j$ and the $\hat{\chi}^{n\,\mathrm{th}}_j$ arrival. By definition, $\overline{\upsilon}^n_j \leq \tilde{\xi}^n_{\hat{\chi}^n_j}$. Also $\overline{\upsilon}^n_j = 0$ if $\varphi^n_j < \theta^n_j$.

Construct as follows a threshold queue with the threshold $h^n = \sqrt{n}(a_i(\varepsilon) + \varepsilon)$. In the first cycle the interarrival times in the accumulation period are taken from the sequence $\{\tilde{\xi}^n_1, \tilde{\xi}^n_2, \ldots, \tilde{\xi}^n_{\hat{\chi}^n_1}, \hat{\xi}^{n,1}_{1,1}, \hat{\xi}^{n,1}_{1,2}, \ldots\}$ (note that if $V^{n,1}_{\varphi^n_1} > a_i(\epsilon) + \epsilon$, then $\hat{\xi}^{n,1}_{1,1}, \hat{\xi}^{n,1}_{1,2}, \ldots$ are not used). Denoting the threshold queue normalized and time-scaled unfinished work at t by $\hat{V}^{n,1}_t$, define

$$\hat{\beta}^n_1 = \inf(t > 0 : \hat{V}^{n,1}_t > a_i(\epsilon) + \epsilon).$$

Then $n\hat{\beta}^n_1$ ends the first accumulation period. If $V^{n,1}_{\varphi^n_1} \leq a_i(\epsilon) + \epsilon$, which happens if $V^{n,1}_{\theta^n_1} \leq a_i(\epsilon) + \epsilon$, then after $n\hat{\beta}^n_1$ the service times are $\hat{\eta}^n_{1,1}, \hat{\eta}^n_{1,2}, \ldots$, and the initial interarrival times are taken from $\{\hat{\xi}^{n,2}_{1,\ell}, \ell \geq 1\}$ until time $n\check{\beta}^n_1$, where

$$\check{\beta}^n_1 = \inf(t > \hat{\beta}^n_1 : \hat{V}^{n,1}_t \leq V^{n,1}_{\vartheta^n_1}).$$

We then take $\hat{\xi}^{n,2}_{1,\overline{\chi}^n_1}$ to be the last random variable in the sequence $\{\hat{\xi}^{n,2}_{1,1}, \hat{\xi}^{n,2}_{1,2}, \ldots\}$ that was actually realized as an interarrival time in $[n\hat{\beta}^n_1, n\check{\beta}^n_1]$. In the case that $V^{n,1}_{\varphi^n_1} > a_i(\epsilon) + \epsilon$, which happens if $V^{n,1}_{\theta^n_1} > a_i(\epsilon) + \epsilon$, we define $\check{\beta}^n_1 = \hat{\beta}^n_1$ and set $\overline{\chi}^n_1 = \hat{\xi}^{n,2}_{1,\overline{\chi}^n_1} = 0$. In both cases, the first arrival after $n\check{\beta}^n_1$ is made to occur at time $n\check{\beta}^n_1 + \overline{\upsilon}^n_1$, so that its interarrival time $\hat{\upsilon}^n_1$ always satisfies $\hat{\upsilon}^n_1 \leq \hat{\xi}^{n,2}_{1,\overline{\chi}^n_1+1} + \overline{\upsilon}^n_1 \leq \hat{\xi}^{n,2}_{1,\overline{\chi}^n_1+1} + \tilde{\xi}^n_{\hat{\chi}^n_1}$. The subsequent interarrival times are $\tilde{\xi}^n_{\hat{\chi}^n_1+1}, \ldots, \tilde{\xi}^n_{\hat{\chi}^n_2}$, and the service times after $n\check{\beta}^n_1$ are the same as for $V^{n,1}$ after ϑ^n_1. The arrival terminating the interarrival time $\tilde{\xi}^n_{\hat{\chi}^n_2}$ corresponds to the arrival in the original queue occurring at or after $n\varphi^n_2$. After that arrival, the interarrival times are again taken to be $\hat{\xi}^{n,1}_{2,1}, \hat{\xi}^{n,1}_{2,2}, \ldots$ until the threshold has been exceeded (these times are not used if $V^{n,1}_{\varphi^n_2} > a_i(\epsilon) + \epsilon$). After this has happened at $n\hat{\beta}^n_2$, where

$$\hat{\beta}^n_2 = \inf(t > \check{\beta}^n_1 : \hat{V}^{n,1}_t > a_i(\epsilon) + \epsilon),$$

and until $n\check{\beta}_2^n$, where

$$\check{\beta}_2^n = \inf(t > \hat{\beta}_2^n : \hat{V}_t^{n,1} \leq V_{\vartheta_2^n}^{n,1}),$$

the service times are $\hat{\eta}_{2,1}^n, \hat{\eta}_{2,2}^n, \ldots$ and the interarrival times are $\hat{\xi}_{2,1}^{n,2}, \hat{\xi}_{2,2}^{n,2}, \ldots$ (as above these are not used if $V_{\varphi_2^n}^{n,1} > a_i(\epsilon) + \epsilon$ and hence $\hat{\beta}_2^n = \check{\beta}_2^n$)

where $\check{\chi}_j^n$ indexes the first arrival in the original queue after κ_{j-1}^n. The first interarrival time is again $\tilde{\xi}_1^n$. We conclude that Theorem 3.1 holds for $\tilde{V}^{n,1}$ and $\widehat{V}^{n,1}$.

Next, by (4.7) and (4.8) we have the bounds

$$
\begin{aligned}
\int_0^{\tilde{\vartheta}^n} f(\tilde{V}_t^{n,1})dt &\leq \int_{\tau_k^n(\epsilon,i)\wedge T}^{\zeta_k^n(\epsilon,i)\wedge T} f(V_t^{n,1})dt \\
&\leq \int_0^{\hat{\vartheta}^n} f(\hat{V}_t^{n,1})dt + [\kappa_0^n \wedge T - \tau_k^n(\epsilon,i)\wedge T]\|f\| \qquad (4.9)\\
&\qquad + [\zeta_k^n(\epsilon,i)\wedge T - \kappa^n_{\vartheta^{n,1}_{i,k}} \wedge T]\|f\|.
\end{aligned}
$$

Define

$$
\begin{aligned}
\tilde{\vartheta}_{i,k}^{n,1} &= \min(j : \tilde{\gamma}_{j+1}^n > \zeta_k^n(\epsilon,i)\wedge T - \tau_k^n(\epsilon,i)\wedge T),\\
\widehat{\vartheta}_{i,k}^{n,1} &= \min(j : \widehat{\gamma}_{j+1}^n > \zeta_k^n(\epsilon,i)\wedge T - \tau_k^n(\epsilon,i)\wedge T).
\end{aligned}
$$

Let $\tilde{\vartheta}_{i,k}^{n,1}(t) = \min(j \geq 0 : \tilde{\gamma}_{j+1}^n > t)$ and $\widehat{\vartheta}_{i,k}^{n,1}(t) = \min(j \geq 0 : \widehat{\gamma}_{j+1}^n > t)$. In the course of proving Theorem 3.1 we established (3.6). Since $\tilde{V}^{n,1}$ and $\widehat{V}^{n,1}$ meet the conditions of Theorem 3.1, we can write for these processes, in analogy with (3.6)

$$
\tilde{\gamma}^n_{\lfloor\sqrt{n}t\rfloor} \xrightarrow{\mathrm{P}} t(a_i(\epsilon)-\epsilon)\mu_1\left(\frac{1}{\lambda_1}+\frac{1}{\mu_1-\lambda_1}\right),\ \widehat{\gamma}^n_{\lfloor\sqrt{n}t\rfloor} \xrightarrow{\mathrm{P}} t(a_i(\epsilon)+\epsilon)\mu_1\left(\frac{1}{\lambda_1}+\frac{1}{\mu_1-\lambda_1}\right)\ .
$$

Then Lemma 2.1 in [3] yields

$$
\frac{\tilde{\vartheta}_{i,k}^{n,1}(t)}{\sqrt{n}} \xrightarrow{\mathrm{P}} \frac{t}{\mu_1(a_i(\epsilon)-\epsilon)}\left(\frac{1}{\lambda_1}+\frac{1}{\mu_1-\lambda_1}\right)^{-1},\ \frac{\widehat{\vartheta}_{i,k}^{n,1}(t)}{\sqrt{n}} \xrightarrow{\mathrm{P}} \frac{t}{\mu_1(a_i(\epsilon)+\epsilon)}\left(\frac{1}{\lambda_1}+\frac{1}{\mu_1-\lambda_1}\right)^{-1}. \tag{4.10}
$$

Lemma 2.2 in [3] then yields

$$
\tilde{\gamma}^n_{\widehat{\vartheta}_{i,k}^{n,1}(t)} \xrightarrow{\mathrm{P}} \frac{a_i(\epsilon)-\epsilon}{a_i(\epsilon)+\epsilon}t, \quad \widehat{\gamma}^n_{\tilde{\vartheta}_{i,k}^{n,1}(t)} \xrightarrow{\mathrm{P}} \frac{a_i(\epsilon)+\epsilon}{a_i(\epsilon)-\epsilon}t\ . \tag{4.11}
$$

Note also that

$$
\begin{aligned}
\tilde{\vartheta}_{i,k}^{n,1} &= \tilde{\vartheta}_{i,k}^{n,1}(\zeta_k^n(\epsilon,i)\wedge T - \tau_k^n(\epsilon,i)\wedge T), &(4.12)\\
\hat{\vartheta}_{i,k}^{n,1} &= \hat{\vartheta}_{i,k}^{n,1}(\zeta_k^n(\epsilon,i)\wedge T - \tau_k^n(\epsilon,i)\wedge T)\ , &(4.13)
\end{aligned}
$$

and

$$
\widehat{\vartheta}_{i,k}^{n,1} \leq \vartheta_{i,k}^{n,1} \leq \tilde{\vartheta}_{i,k}^{n,1}. \tag{4.14}
$$

Let $U_k^n(\epsilon,i)$ and $V_k^n(\epsilon,i)$ denote respectively the lower bound in (4.9) with $\tilde{\vartheta}^n$ $(=\tilde{\gamma}^n_{\vartheta^{n,1}_{i,k}})$ changed to $\tilde{\omega}^n = \tilde{\gamma}^n_{\widehat{\vartheta}^{n,1}_{i,k}}$, and the upper bound in (4.9) with $\widehat{\vartheta}^n$ $(=\widehat{\gamma}^n_{\vartheta^{n,1}_{i,k}})$ changed to $\widehat{\omega}^n = \widehat{\gamma}^n_{\tilde{\vartheta}^{n,1}_{i,k}}$. By (4.14),

$$U_k^n(\epsilon,i) \le \int_{\tau_k^n(\epsilon,i)\wedge T}^{\zeta_k^n(\epsilon,i)\wedge T} f(V_t^{n,1})dt \le V_k^n(\epsilon,i). \tag{4.15}$$

We now show that

$$U_k^n(\epsilon,i) \xrightarrow{\mathrm{d}} U_k(\epsilon,i), \quad V_k^n(\epsilon,i) \xrightarrow{\mathrm{d}} V_k(\epsilon,i), \quad k\ge 1, 0\le i\le N, \tag{4.16}$$

where

$$\begin{aligned} U_k(\epsilon,i) &= \frac{a_i(\epsilon)-\epsilon}{a_i(\epsilon)+\epsilon}(\zeta_k(\epsilon,i)\wedge T - \tau_k(\epsilon,i)\wedge T)\int_0^1 f(u(a_i(\epsilon)-\epsilon))du, \\ V_k(\epsilon,i) &= \frac{a_i(\epsilon)+\epsilon}{a_i(\epsilon)-\epsilon}(\zeta_k(\epsilon,i)\wedge T - \tau_k(\epsilon,i)\wedge T)\int_0^1 f(u(a_i(\epsilon)+\epsilon))du\ , \end{aligned} \tag{4.17}$$

By Theorem 3.1 and (4.11),

$$\begin{aligned} \int_0^{\tilde{\gamma}^n_{\widehat{\vartheta}^{n,1}_{i,k}}(t)} f(\tilde{V}_s^{n,1})ds &\xrightarrow{\mathrm{P}} \frac{a_i(\epsilon)-\epsilon}{a_i(\epsilon)+\epsilon}t\int_0^1 f(u(a_i(\epsilon)-\epsilon))du\ , \\ \int_0^{\widehat{\gamma}^n_{\tilde{\vartheta}^{n,1}_{i,k}}(t)} f(\widehat{V}_s^{n,1})ds &\xrightarrow{\mathrm{P}} \frac{a_i(\epsilon)+\epsilon}{a_i(\epsilon)-\epsilon}t\int_0^1 f(u(a_i(\epsilon)+\epsilon))du\ . \end{aligned} \tag{4.18}$$

In view of (4.12), Lemma 2.2 in [3] shows that (4.4) and (4.18) imply

$$\begin{aligned} \int_0^{\tilde{\omega}^n} f(\tilde{V}_t^{n,1})dt &\xrightarrow{\mathrm{d}} \frac{a_i(\epsilon)-\epsilon}{a_i(\epsilon)+\epsilon}(\zeta_k(\epsilon,i)\wedge T - \tau_k(\epsilon,i)\wedge T)\int_0^1 f(u(a_i(\epsilon)-\epsilon))du\ , \\ \int_0^{\widehat{\omega}^n} f(\widehat{V}_t^{n,1})dt &\xrightarrow{\mathrm{d}} \frac{a_i(\epsilon)+\epsilon}{a_i(\epsilon)-\epsilon}(\zeta_k(\epsilon,i)\wedge T - \tau_k(\epsilon,i)\wedge T)\int_0^1 f(u(a_i(\epsilon)+\epsilon))du\ . \end{aligned}$$

Since $|\kappa_0^n\wedge T - \tau_k^n(\epsilon,i)\wedge T| \xrightarrow{\mathrm{P}} 0$, and $|\zeta_k^n(\epsilon,i)\wedge T - \kappa^n_{\vartheta^{n,1}_{i,k}}\wedge T| \xrightarrow{\mathrm{P}} 0$ obviously hold, (4.16) is proved. Moreover, the same argument shows that

$$\begin{aligned} (V^n,(U_k^n(\epsilon,i))_{k\ge1,0\le i\le N}) &\xrightarrow{\mathrm{d}} (\tilde{V},(U_k(\epsilon,i))_{k\ge1,0\le i\le N}), \\ (V^n,(V_k^n(\epsilon,i))_{k\ge1,0\le i\le N}) &\xrightarrow{\mathrm{d}} (\tilde{V},(V_k(\epsilon,i))_{k\ge1,0\le i\le N})\ . \end{aligned} \tag{4.19}$$

Next, defining

$$U^n(\epsilon) = \sum_{k=1}^{\infty}\sum_{i=0}^{N} U_k^n(\epsilon,i),\ V^n(\epsilon) = \sum_{k=1}^{\infty}\sum_{i=0}^{N} V_k^n(\epsilon,i), \tag{4.20}$$

we need to prove that

$$(V^n,U^n(\epsilon)) \xrightarrow{\mathrm{d}} (\tilde{V},U(\epsilon)),\ (V^n,V^n(\epsilon)) \xrightarrow{\mathrm{d}} (\tilde{V},V(\epsilon)), \tag{4.21}$$

where

$$U(\epsilon) = \sum_{k=1}^{\infty} \sum_{i=0}^{N} U_k(\epsilon, i), \; V(\epsilon) = \sum_{k=1}^{\infty} \sum_{i=0}^{N} V_k(\epsilon, i). \tag{4.22}$$

We prove the first convergence result in (4.21); the proof of the second uses the same reasoning.

Since $U_k^n(\epsilon, i) = 0$ if $\tau_k^n(\epsilon, i) \geq T$, we have by (4.3) and (4.4), for $\eta > 0$,

$$\begin{aligned} &\overline{\lim}_{M\to\infty}\overline{\lim}_{n\to\infty} P\left(\left|\sum_{k=1}^{M}\sum_{i=0}^{N} U_k^n(\epsilon, i) - U^n(\epsilon)\right| > \eta\right) \\ &\leq \overline{\lim}_{M\to\infty}\overline{\lim}_{n\to\infty} P\left(\min_{0\leq i\leq N} \zeta_M^n(\epsilon, i) \wedge (T+1) < T\right) \\ &\leq \overline{\lim}_{M\to\infty} P\left(\min_{0\leq i\leq N} \zeta_M(\epsilon, i) \wedge (T+1) \leq T\right) = 0. \end{aligned} \tag{4.23}$$

Analogously,

$$\sum_{k=1}^{M}\sum_{i=0}^{N} U_k(\epsilon, i) \xrightarrow{\mathrm{P}} U(\epsilon) \quad (M \to \infty)\ . \tag{4.24}$$

Next, by (4.19) and the continuous mapping theorem, we have

$$\left(V^n, \sum_{k=1}^{M}\sum_{i=0}^{N} U_k^n(\epsilon, i)\right) \xrightarrow{\mathrm{d}} \left(\tilde{V}, \sum_{k=1}^{M}\sum_{i=0}^{N} U_k(\epsilon, i)\right)\ . \tag{4.25}$$

The convergence $(V^n, U^n(\epsilon)) \xrightarrow{\mathrm{d}} (\tilde{V}, U(\epsilon))$ then follows from (4.22)–(4.25) and Theorem 4.2 in [1].

Now by the definition of $\tau_k^n(\epsilon, i)$ and $\zeta_k^n(\epsilon, i)$,

$$\left|\int_0^T f(V_t^{n,1}) \cdot 1(\delta \leq V_t^n \leq K)dt - \sum_{k=1}^{\infty}\sum_{i=0}^{N} \int_0^T f(V_t^{n,1}) \cdot 1(t \in [\tau_k^n(\epsilon, i),\ \zeta_k^n(\epsilon, i)))dt\right|$$

$$\leq \|f\| \int_0^T [1(\delta - \epsilon \leq V_t^n \leq \delta) + 1(K \leq V_t^n \leq K + \epsilon)]dt$$

so by (4.15), we obtain from (4.20)

$$\begin{aligned} U^n(\epsilon) - \|f\| &\int_0^T [1(\delta - \epsilon \leq V_t^n \leq \delta) + 1(K \leq V_t^n \leq K + \epsilon)]dt \\ &\leq \int_0^T f(V_t^{n,1}) \cdot 1(\delta \leq V_t^n \leq K)dt \\ &\leq V^n(\epsilon) + \|f\| \int_0^T [1(\delta - \epsilon \leq V_t^n \leq \delta) + 1(K \leq V_t^n \leq K + \epsilon)]dt\ . \end{aligned} \tag{4.26}$$

Therefore, if we prove that as $\epsilon \to 0$

$$\begin{aligned} U(\epsilon) &\xrightarrow{\mathrm{d}} \int_0^T \left(\int_0^1 f(u\tilde{V}_t)du\right) 1(\delta \leq \tilde{V}_t \leq K)dt, \\ V(\epsilon) &\xrightarrow{\mathrm{d}} \int_0^T \left(\int_0^1 f(u\tilde{V}_t)du\right) 1(\delta \leq \tilde{V}_t \leq K)dt, \end{aligned} \tag{4.27}$$

then by applying Lemma 2.3 in [3] to (4.27) and taking into account (4.21), (4.2), we will then obtain (4.1). As before, we prove only the first of the results in (4.27); the proof of the second is similar.

In fact, we prove convergence with probability 1. The argument is almost identical to that in [3], but we give it here since it is used once again below. Since $a_i(\epsilon) > \delta$, we have from (4.17) and (4.22)

$$\left| U(\epsilon) - \sum_{k=1}^{\infty}\sum_{i=0}^{N}\int_0^1 f(u(a_i(\epsilon)-\epsilon))du[\zeta_k(\epsilon,i)\wedge T - \tau_k(\epsilon,i)\wedge T]\right| \le \frac{2\epsilon}{\delta}\|f\|T\ .$$

This tends to 0 as $\epsilon \to 0$, so we prove that

$$\begin{aligned}\lim_{\epsilon\to 0}\sum_{k=1}^{\infty}\sum_{i=0}^{N}[\zeta_k(\epsilon,i)\wedge T - \tau_k(\epsilon,i)\wedge T]\int_0^1 f(u(a_i(\epsilon)-\epsilon))du \\ = \int_0^T\left(\int_0^1 f(u\tilde{V}_t)du\right)1(\delta\le\tilde{V}_t\le K)dt\ .\end{aligned} \tag{4.28}$$

We can write

$$\begin{aligned}&\sum_{k=1}^{\infty}\sum_{i=0}^{N}[\zeta_k(\epsilon,i)\wedge T - \tau_k(\epsilon,i)\wedge T]\int_0^1 f(u(a_i(\epsilon)-\epsilon))du \\ &= \sum_{k=1}^{\infty}\sum_{i=0}^{N}\int_0^T\left(\int_0^1 f(u(a_i(\epsilon)-\epsilon))du\right)\cdot 1(\tau_k(\epsilon,i)\le t<\zeta_k(\epsilon,i))dt \\ &\equiv C_\epsilon\ .\end{aligned} \tag{4.29}$$

Note that if $x, y > \delta/2$, $|x-y| < 2\epsilon$, then

$$\begin{aligned}\left|\int_0^1 f(ux)du - \int_0^1 f(uy)du\right| &= \left|\frac{1}{x}\int_0^x f(u)du - \frac{1}{y}\int_0^y f(u)du\right| \\ &\le \left|\frac{1}{x}-\frac{1}{y}\right|\int_0^x f(u)du + \frac{1}{y}\left|\int_x^y f(u)du\right| \le \frac{8\epsilon}{\delta}\|f\|\ .\end{aligned}$$

Since $\tilde{V}_t \in [a_i(\epsilon)-\epsilon,\ a_i(\epsilon)+\epsilon]$, $t\in[\tau_k(\epsilon,i),\ \zeta_k(\epsilon,i))$, we then have

$$\begin{aligned}&\left|\int_0^1 f(u(a_i(\epsilon)-\epsilon)du\cdot 1(t\in[\tau_k(\epsilon,i),\ \zeta_k(\epsilon,i))) - \int_0^1 f(u\tilde{V}_t)du\cdot 1(t\in[\tau_k(\epsilon,i),\ \zeta_k(\epsilon,i)))\right| \\ &\le \frac{8\epsilon}{\delta}\|f\|\cdot 1(t\in[\tau_k(\epsilon,i),\ \zeta_k(\epsilon,i)))\ ,\end{aligned}$$

so by (4.29)

$$\left|C_\epsilon - \sum_{k=1}^{\infty}\sum_{i=0}^{N}\int_0^T\left(\int_0^1 f(u\tilde{V}_t)du\right)\cdot 1(t\in[\tau_k(\epsilon,i),\ \zeta_k(\epsilon,i)))dt\right| \le \frac{8\epsilon}{\delta}\|f\|T\ ,$$

whence

$$\left| C_\epsilon - \int_0^T \left(\int_0^1 f(u\tilde{V}_t)du \right) \cdot 1(\delta \le \tilde{V}_t \le K)dt \right|$$
$$\le \|f\| \int_0^T [1(\delta \le \tilde{V}_t \le \delta - \epsilon) + 1(K \le \tilde{V}_t \le K + \epsilon)]dt + \frac{8\epsilon}{\delta}\|f\|T \ .$$

Since the right-hand side of this inequality tends to 0 as $\epsilon \to 0$, we have proved (4.28). This completes the proof of the first assertion of Theorem 4.1 for bounded nonnegative $f(x)$. The general case is handled via a localization argument as in the proof of Theorem 2.1 in [3].

5 Tightness Results

The main purpose of this section is to prove several results on the tightness of some processes closely related to V^n. We start, however, with a number of technical results.

Introduce

$$B_t^{n,\ell} = \frac{S_{nt}^{n,\ell} - \rho_\ell^n nt}{\sqrt{n}}, \quad \ell = 1, 2\ , \ B_t^n = \frac{S_{nt}^n - nt}{\sqrt{n}}\ , \tag{5.1}$$

and let $B^{n,\ell} = (B_t^{n,\ell}, t \ge 0),\ \ell = 1, 2,\ B^n = (B_t^n, t \ge 0)$.

Lemma 5.1 *As $n \to \infty$,*

$$\frac{A_{nt}^{n,\ell}}{n} \xrightarrow{P} \lambda_\ell t,\ \frac{1}{\sqrt{n}} \sum_{i=1}^{\lfloor \sqrt{n}t \rfloor} s_i^{n,\ell} \xrightarrow{P} d_\ell\, t,\ \frac{1}{n} \sum_{i=1}^{\lfloor nt \rfloor} s_i^{n,\ell} \xrightarrow{P} d_\ell\, t, \quad \ell = 1, 2,$$
$$B^{n,\ell} \xrightarrow{d} \lambda_\ell^{1/2} \sigma_\ell W^\ell, \quad \ell = 1, 2, \qquad B^n \xrightarrow{d} (\sigma W_t + ct, t \ge 0),$$

where $W^\ell,\ \ell = 1, 2,$ and $W = (W_t, t \ge 0)$ are standard Brownian motions.

Proof. We sketch proofs of the second and third convergence results in the first line. The other statements follow from the assumptions in a standard manner. We have, for $\epsilon > 0$,

$$\left| \frac{1}{\sqrt{n}} \sum_{i=1}^{\lfloor \sqrt{n}t \rfloor} s_i^{n,\ell} - d_\ell\, t \right| \le \left| \frac{1}{\sqrt{n}} \sum_{i=1}^{\lfloor \sqrt{n}t \rfloor} \left(s_i^{n,\ell} 1(s_i^{n,\ell} \le \epsilon\sqrt{n}) - E s_i^{n,\ell} 1(s_i^{n,\ell} \le \epsilon\sqrt{n}) \right) \right|$$
$$+ \frac{1}{\sqrt{n}} \sum_{i=1}^{\lfloor \sqrt{n}t \rfloor} s_i^{n,\ell} 1(s_i^{n,\ell} > \epsilon\sqrt{n}) + tE s_1^{n,\ell} 1(s_1^{n,\ell} > \epsilon\sqrt{n}) + \left| \frac{\lfloor \sqrt{n}t \rfloor}{\sqrt{n}} E s_1^{n,\ell} - d_\ell t \right|.$$

The last term tends to 0 by (2.2). The second and third terms on the right tend to 0 by (2.6). Since the $s_i^{n,\ell}, i \ge 1$, are independent and identically distributed, the second moment of the first term on the right is no greater than

$$\frac{t}{\sqrt{n}} E(s_1^{n,\ell})^2 1(|s_1^{n,\ell}| \le \epsilon\sqrt{n}) \le t\epsilon E s_1^{n,\ell}$$

and goes to 0 as $n \to \infty$ and $\epsilon \to 0$. Hence, in view of Chebyshev's inequality, the first term tends in probability to 0 as $n \to \infty$ and $\epsilon \to 0$. This concludes the proof of the second convergence in the statement of the lemma. The third one is proved similarly. ■

Let

$$\beta_t^n = 1 - \alpha_{nt}^n.$$

Then definitions (2.9), (5.1) and equation (2.8) imply that V_t^n satisfies

$$V_t^n = V_0^n + B_t^n + \sqrt{n}\int_0^t 1(V_s^n = 0)ds + \sqrt{n}\int_0^t 1(V_s^n > 0)\beta_s^n ds \ . \tag{5.2}$$

We study properties of the processes on the right. For $\epsilon > 0$, we define the processes $K^{n,\epsilon} = (K_t^{n,\epsilon}, t \geq 0)$ by

$$K_t^{n,\epsilon} \quad = \quad \sqrt{n}\int_0^t 1(V_s^n > \epsilon)\beta_s^n ds. \tag{5.3}$$

Though the $K^{n,\epsilon}$ have continuous paths, we still consider them as random elements of $D[0,\infty)$.

Recall that a sequence of processes $\{X^n, n \geq 1\}$ in $D[0,\infty)$ is called C-tight if it is tight and all weak limit points of the sequence of their laws are laws of continuous processes [6, VI.3.25].

Lemma 5.2 *The sequence $\{K^{n,\epsilon}, n \geq 1\}$, where $K^{n,\epsilon} = (K_t^{n,\epsilon}, t \geq 0)$, is C-tight.*

Proof. Denote by $[u_i^{n,1}, v_i^{n,1}]$ and $[u_i^{n,2}, v_i^{n,2}]$, $i \geq 1$, the respective successive switchover periods (i.e., times during which the server is switching) from the first queue to the second and from the second queue back to the first. Let $\vartheta_t^{n,\ell}$ be the number of switchovers from queue ℓ started in $[0, nt]$. By (5.3),

$$K_t^{n,\epsilon} \quad = \quad \frac{1}{\sqrt{n}}\sum_{\ell=1}^{2}\sum_{i=1}^{\vartheta_t^{n,\ell}}\int_{u_i^{n,\ell}}^{v_i^{n,\ell}\wedge nt} 1(U_s^n > \epsilon\sqrt{n})ds. \tag{5.4}$$

So, if we define

$$\breve{K}_t^{n,\epsilon} \quad = \quad \frac{1}{\sqrt{n}}\sum_{\ell=1}^{2}\sum_{i=1}^{\vartheta_t^{n,\ell}}\int_{u_i^{n,\ell}}^{v_i^{n,\ell}} 1(U_s^n > \epsilon\sqrt{n})ds, \tag{5.5}$$

then, obviously,

$$\sup_{s\leq t}|K_s^{n,\epsilon} - \breve{K}_s^{n,\epsilon}| \leq \frac{1}{\sqrt{n}}\max_{l=1,2}\max_{1\leq i\leq\vartheta_t^{n,\ell}}(v_i^{n,\ell} - u_i^{n,\ell}). \tag{5.6}$$

Note that

$$P\left(\frac{\vartheta_t^{n,\ell}}{n} > \frac{t+1}{d_\ell}\right) \leq P\left(\sum_{i=1}^{\lfloor n(t+1)/d_\ell\rfloor} s_i^{n,\ell} \leq nt\right) \ , \tag{5.7}$$

and that the latter tends to 0 as $n \to \infty$ by Lemma 5.1. Recalling also that

$$v_i^{n,\ell} - u_i^{n,\ell} = s_i^{n,\ell}, \tag{5.8}$$

we get by Lemma 3.1 and (2.6) that the left-hand side of (5.6) tends in probability to 0 as $n \to \infty$. Thus the C-tightness of $\{K^{n,\epsilon}, n \geq 1\}$ will follow from the C-tightness of $\{\breve{K}^{n,\epsilon}, n \geq 1\}$.

To prove the latter, we will use repeatedly the concept of *strong majorization* [6, VI.3.34]: an increasing process $X = (X_t, t \geq 0)$ is said to strongly majo[illegible]

$$+ \frac{1}{\sqrt{n}} \sum_{i=1}^{\vartheta_t^{n,1}} s_i^{n,1} \cdot 1(S^{n,2}_{v_i^{n,1}} - S^{n,2}_{u_i^{n,1}} > \epsilon\sqrt{n}/4) ,$$

so by (5.8), for $\delta > 0$, since $u^{n,1}_{\vartheta_t^{n,1}} \le nt$,

$$\begin{aligned} P(\hat{K}_t^{n,\epsilon,1} > 0) \quad \le \quad & P\left(\frac{\vartheta_t^{n,1}}{n} > \frac{t+1}{d_1}\right) \\ & + P\left(\frac{1}{\sqrt{n}} \sup_{1 \le i \le \lfloor n\frac{t+1}{d_1}\rfloor} s_i^{n,1} > \delta\right) \\ & + P\left(\sup_{\substack{u \le nt \\ 0 \le v-u \le \delta\sqrt{n}}} |S_v^{n,1} - S_u^{n,1}| > \epsilon\sqrt{n}/4\right) \\ & + P\left(\sup_{\substack{u \le nt \\ 0 \le v-u \le \delta\sqrt{n}}} |S_v^{n,2} - S_u^{n,2}| > \epsilon\sqrt{n}/4\right) . \end{aligned} \tag{5.12}$$

The first term on the right of (5.12) goes to 0 as $n \to \infty$ by (5.7) and Lemma 5.1. The second term tends to 0 as $n \to \infty$ by Lemma 3.1 and (2.6).

Next,

$$\begin{aligned} & P\left(\sup_{\substack{u \le nt \\ 0 \le v-u \le \delta\sqrt{n}}} |S_v^{n,1} - S_u^{n,1}| > \epsilon\sqrt{n}/4\right) \\ & = P\left(\sup_{\substack{u \le t \\ 0 \le \sqrt{n}(v-u) \le \delta}} \left|\frac{S_{nv}^{n,1} - \rho_1^n nv}{\sqrt{n}} - \frac{S_{nu}^{n,1} - \rho_1^n nu}{\sqrt{n}} + \rho_1^n\sqrt{n}(v-u)\right| > \frac{\epsilon}{4}\right) \\ & \le P\left(\sup_{u \le t, |u-v| \le \gamma} \left|\frac{S_{nv}^{n,1} - \rho_1^n nv}{\sqrt{n}} - \frac{S_{nu}^{n,1} - \rho_1^n nu}{\sqrt{n}}\right| > \frac{\epsilon}{4} - \rho_1^n\delta\right) \\ & = P\left(\sup_{u \le t, |u-v| \le \gamma} |B_v^{n,1} - B_u^{n,1}| > \frac{\epsilon}{4} - \rho_1^n\delta\right) , \end{aligned} \tag{5.13}$$

where $\gamma > 0$ is arbitrary and n is large enough. Since by Lemma 5.1, $B^{n,1}$ converges in distribution to $\lambda_1^{1/2}\sigma_1 W^1$, and since the functional $X \to \sup_{s \le t} |X_s|$, $X \in D[0,\infty)$, is continuous almost everywhere with respect to the Wiener measure [10], we conclude from (5.13) that

$$\begin{aligned} & \overline{\lim_{n\to\infty}} P\left(\sup_{\substack{u \le nt \\ 0 \le v-u \le \delta\sqrt{n}}} |S_v^{n,1} - S_u^{n,1}| > \epsilon\sqrt{n}/4\right) \\ & \qquad \le P\left(\sup_{u \le t, |u-v| \le \gamma} |W_v^1 - W_u^1| \ge \lambda_1^{-1/2}\sigma_1^{-1}\left(\frac{\epsilon}{4} - \rho_1\delta\right)\right) , \end{aligned}$$

which goes to 0 as $\gamma \to 0$, if $\delta < \epsilon/4\rho_1$, by the continuity of Brownian motion. We have thus proved that the third term on the right of (5.12) tends to 0 as $n \to \infty$ if δ is small enough.

A similar argument applies to the last term on the right-hand side of (5.12). Therefore, since $\hat{K}^{n,\varepsilon,1}$ is increasing,

$$\lim_{n\to\infty} P(\sup_{s\le t} \hat{K}_s^{n,\epsilon,1} > 0) = 0 \ ,$$

as required.

We now prove that $\{\overline{K}^{n,\epsilon,1}, n \ge 1\}$ is C–tight. Call a switchover from queue 1 to queue 2 *sound* if at the time when it starts, the total unfinished work (which at that moment is the unfinished work at queue 2) is greater than $\epsilon\sqrt{n}/2$. Let $\overline{\vartheta}_t^{n,1}$ be the number of sound switchovers started in $[0, nt]$. By (5.10),

$$\overline{K}_t^{n,\epsilon,1} = \frac{1}{\sqrt{n}} \sum_{i=1}^{\overline{\vartheta}_t^{n,1}} \overline{s}_i^{n,1}, \tag{5.14}$$

where $\overline{s}_i^{n,1}$ is the duration of the ith sound switchover. Note that the soundness of a switchover is determined at its beginning, so the $\overline{s}_i^{n,1}$, $i \ge 1$, are i.i.d. and distributed as the $s_i^{n,1}$, $i \ge 1$.

We have by (5.14), for $t > 0, \delta > 0, \eta > 0, \gamma > 0$ and $\Lambda > 0$,

$$\begin{aligned} P(&\sup_{u,v\le t, |u-v|\le\delta} |\overline{K}_v^{n,\varepsilon,1} - \overline{K}_u^{n,\varepsilon,1}| > \eta) \\ &\le P(\overline{\vartheta}_t^{n,1} > \Lambda\sqrt{n}) + P(\sup_{v-\delta\le u\le v\le t} |\overline{\vartheta}_v^{n,1} - \overline{\vartheta}_u^{n,1}| > \gamma\sqrt{n}) \\ &\quad + P\left(\sup_{v-\gamma\le u\le v\le\Lambda} \left| \frac{1}{\sqrt{n}} \sum_{i=\lfloor u\sqrt{n}\rfloor+1}^{\lfloor v\sqrt{n}\rfloor} \overline{s}_i^{n,1} \right| > \eta \right). \end{aligned} \tag{5.15}$$

Now, if $\overline{\vartheta}_v^{n,1} - \overline{\vartheta}_u^{n,1} = m$, then the amount of work executed by the second server in the interval $[nu, nv]$ is no less than $(m-1)\varepsilon\sqrt{n}/2$ which takes time $(m-1)\varepsilon\sqrt{n}/2$. Hence $(m-1)\varepsilon\sqrt{n}/2 \le n(v-u)$ which leads to the estimate $\overline{\vartheta}_v^{n,1} - \overline{\vartheta}_u^{n,1} \le \frac{2\sqrt{n}}{\varepsilon}(v-u) + 1$, so that, for all n large enough,

$$\sup_{v-\delta\le u\le v\le t} |\overline{\vartheta}_v^{n,1} - \overline{\vartheta}_u^{n,1}| \le \frac{3\sqrt{n}}{\varepsilon}\delta; \;\; \overline{\vartheta}_t^{n,1} \le \frac{3\sqrt{n}}{\varepsilon} t.$$

Taking in (5.15) $\Lambda = \frac{3}{\varepsilon} t$ and $\gamma = \frac{3}{\varepsilon}\delta$, we get

$$\begin{aligned} \overline{\lim_{n\to\infty}} \; P(&\sup_{u,v\le t, |u-v|\le\delta} |\overline{K}_v^{n,\varepsilon,1} - \overline{K}_u^{n,\varepsilon,1}| > \eta) \\ &\le \overline{\lim_{n\to\infty}} \; P\left(\sup_{v-3\delta/\varepsilon\le u\le v\le 3t/\varepsilon} \left| \frac{1}{\sqrt{n}} \sum_{i=\lfloor u\sqrt{n}\rfloor+1}^{\lfloor v\sqrt{n}\rfloor} \overline{s}_i^{n,1} \right| > \eta \right), \end{aligned}$$

where the latter limit, by Lemma 5.1, is zero if $\frac{3\delta}{\varepsilon} d_1 \le \eta$. Therefore

$$\lim_{\delta\to 0} \overline{\lim_{n\to\infty}} \; P(\sup_{u,v\le t, |u-v|\le\delta} |\overline{K}_v^{n,\varepsilon,1} - \overline{K}_u^{n,\varepsilon,1}| > \eta) = 0,$$

which, since $\overline{K}_0^{n,\varepsilon,1} = 0$, proves the C-tightness of $\{\overline{K}_v^{n,\varepsilon,1}, n \geq 1\}$. The lemma is proved. ■

We next prove that the two rightmost processes in (5.2) are asymptotically bounded in probability.

Lemma 5.3 *We have*

$$\lim_{A\to\infty} \overline{\lim_{n\to\infty}} P\left(\sqrt{n}\int_0^t 1(V_s^n = 0)ds > A\right) = 0 , \tag{5.16}$$

$$\lim_{A\to\infty} \overline{\lim_{n\to\infty}} P\left(\sqrt{n}\int_0^t 1(V_s^n > 0)\beta_s^n ds > A\right) = 0 . \tag{5.17}$$

Proof. Let

$$\varphi_t^n = V_t^n - V_0^n - B_t^n - K_t^{n,1} . \tag{5.18}$$

By (5.3), (5.2) and the inequality $0 \leq \beta_t^n \leq 1$, we have for $0 < s < t$,

$$\varphi_t^n - \varphi_s^n = \sqrt{n}\int_s^t 1(V_u^n = 0)du + \sqrt{n}\int_s^t 1(0 < V_u^n \leq 1)\beta_u^n du \leq \sqrt{n}\int_s^t 1(V_u^n \leq 1)du ,$$

so, since $V_u^n \geq \phi_u^n + B_u^n$ by (5.18),

$$\phi_t^n - \phi_s^n \leq \sqrt{n}\int_s^t 1(\phi_u^n \leq 1 - B_u^n)du.$$

Therefore, by Lemma 1 in [2]

$$\phi_t^n \leq \sup_{s\leq t}(1 - B_s^n) \vee 0 .$$

Since the sequence $\{B^n, n \geq 1\}$ is C-tight by Lemma 5.1, we conclude that

$$\lim_{A\to\infty} \overline{\lim_{n\to\infty}} P(\varphi_t^n > A) = 0 .$$

Since by (5.2) and (5.18)

$$\varphi_t^n + K_t^{n,1} = \sqrt{n}\int_0^t 1(V_s^n = 0)\, ds + \sqrt{n}\int_0^t 1(V_s^n > 0)\beta_s^n\, ds,$$

an application of Lemma 5.2 completes the proof. ■

We are in need of two more technical lemmas. Introduce the processes

$$M_t^{n,\ell} = \frac{1}{\sqrt{n}} \sum_{i=1}^{A_{nt}^{n,\ell}+1} (\eta_i^{n,\ell} - \rho_\ell^n \xi_i^{n,\ell}), \quad \ell = 1, 2 , \quad M_t^n = M_t^{n,1} + M_t^{n,2}, \tag{5.19}$$

and recall that $\tau_i^{n,\ell}, i = 0, 1, \ldots$ denote the arrival times for $A^{n,\ell}$, $\ell = 1, 2$.

Lemma 5.4 *Define the filtration* $\mathbb{F}^n = (F_t^n, t \geq 0)$ *by* $F_t^n = F_t^{n,1} \vee F_t^{n,2} \vee \sigma(V_0^n) \vee \mathcal{N}$, *where* $F_t^{n,\ell} = G^{n,\ell}_{A^{n,\ell}_{nt}+1}$, $G_i^{n,\ell} = \sigma\{\eta_j^{n,\ell}, \xi_j^{n,\ell}, 1 \leq j \leq i\}$, $\ell = 1, 2$, *and* $\mathcal{N}$ *is the family* P*-null sets. Then* $\mathbb{F}^n$ *is well defined, the* $\tau_i^{n,\ell}/n, n = 0, 1, \ldots, \ell = 1, 2$ *are* $\mathbb{F}^n$*-stopping times and* $M^n = (M_t^n, t \geq 0)$ *is an* $\mathbb{F}^n$*-locally square integrable martingale with the predictable quadratic variation process*

$$\langle M^n \rangle_t = \frac{1}{n}[(\sigma_1^n)^2 A_{nt}^{n,1} + (\sigma_2^n)^2 A_{nt}^{n,2}] \; .$$

Proof. The proof is almost the same as the proof of Lemma 2 in [2]. In particular, the martingale property of M^n and the formula for its predictable quadratic variation process is deduced from the fact that the processes $\left(\sum_{i=1}^k (\eta_i^{n,\ell} - \rho_\ell^n \xi_i^{n,\ell}), k \geq 0\right)$, $\ell = 1, 2$, are locally square integrable martingales which have predictable quadratic variation processes $((\sigma_\ell^n)^2 k, k \geq 0)$ relative to the respective flows $(G_k^{n,\ell}, k \geq 0)$. ■

Note that the processes B^n, $(\beta_t^n, t \geq 0)$ and V^n are $\mathbf{F}^n$-adapted.

Introduce

$$\epsilon_t^{n,\ell} = \frac{\rho_\ell^n}{\sqrt{n}} \left[\sum_{i=1}^{A_{nt}^{n,\ell}+1} \xi_i^{n,\ell} - nt \right] - \frac{1}{\sqrt{n}} \eta^{n,\ell}_{A^{n,\ell}_{nt}+1}, \quad \ell = 1, 2 \; , \quad \epsilon_t^n = \epsilon_t^{n,1} + \epsilon_t^{n,2} \; . \tag{5.20}$$

Let ΔM_s^n denote the jump of M^n at s.

Lemma 5.5 *Under the hypotheses of Theorem 2.3, for* $t > 0$,

$$\langle M^n \rangle_t \xrightarrow{P} \sigma^2 t, \; \sum_{s \leq t} (\Delta M_s^n)^2 \xrightarrow{P} \sigma^2 t \; , \; \sup_{s \leq t} |\epsilon_s^n| \xrightarrow{\mathrm{P}} 0,$$

as $n \to \infty$.

Proof. The first convergence follows by the expression for $\langle M^n \rangle_t$ in Lemma 5.4, Lemma 5.1 and (2.1). For the second, note that since M^n is a process of locally bounded variation by (5.19), it is a purely discontinuous local martingale [6, I.4.14], [10, I.7], so its quadratic variation process $([M^n, M^n]_t, t \geq 0)$ is the sum of the squares of jumps: $[M^n, M^n]_t = \sum_{s \leq t} (\Delta M_s^n)^2$ [6, I.4.52], [10, I.8]. By Lemma 5.5.5 in [10], (2.4), (2.5) and Lemma 5.1 imply that the convergences $[M^n, M^n]_t \xrightarrow{P} \sigma^2 t$ and $\langle M^n \rangle_t \xrightarrow{P} \sigma^2 t$ are equivalent, so the second convergence of the lemma is a consequence of the first. The third convergence results from the inequalities

$$0 \leq \sum_{i=1}^{A_{nt}^{n,\ell}+1} \xi_i^{n,\ell} - nt \leq \xi^{n,\ell}_{A^{n,\ell}_{nt}+1},$$

condition (2.4) and Lemmas 3.1, 5.1. ■.

Let

$$\overline{V}_t^n = V_t^n - \epsilon_t^n \ . \tag{5.21}$$

Since V^n is $\mathbf{F}^n$-adapted and ε_t^n is F_t^n-measurable by (5.20), $\overline{V}^n = (\overline{V}_t^n, t \geq 0)$ is $\mathbf{F}^n$-adapted. By (5.1), (5.2), (5.19) and (5.20), we get the representation

$$\overline{V}_t^n = \overline{V}_0^n + \sqrt{n}(\rho^n - 1)t + \sqrt{n}\int_0^t 1(V_s^n = 0)ds + \sqrt{n}\int_0^t 1(V_s^n > 0)\beta_s^n ds + M_t^n \ . \tag{5.22}$$

Now squaring in (5.22), we have by Ito's formula (Theorem 2.3.1 [10]) that

$$\begin{aligned}(\overline{V}_t^n)^2 \quad &= \quad (\overline{V}_0^n)^2 + 2\sqrt{n}(\rho^n - 1)\int_0^t \overline{V}_s^n ds + 2\sqrt{n}\int_0^t \overline{V}_s^n 1(V_s^n = 0)ds \\ &\quad + 2\sqrt{n}\int_0^t \overline{V}_s^n 1(V_s^n > 0)\beta_s^n ds + 2\int_0^t \overline{V}_{s-}^n dM_s^n \\ &\quad + \sum_{s\leq t}(\Delta M_s^n)^2 \ ,\end{aligned} \tag{5.23}$$

where $\overline{V}_{s-}^n$ denotes the left-hand limit of $\overline{V}^n$ at s.

Lemma 5.6 *The sequences $\{\overline{V}^n, n \geq 1\}$ and $\{V^n, n \geq 1\}$ are C-tight.*

Proof. By (5.16), (5.17), the C-tightness of $\{B^n, n \geq 1\}$, and the convergence $V_0^n \xrightarrow{d} V_0$, the right-hand side of (5.2) is asymptotically bounded in probability, i.e.,

$$\lim_{A\to\infty} \overline{\lim_{n\to\infty}} P\left(\sup_{s\leq t} |V_s^n| > A\right) = 0, \quad t > 0 \ . \tag{5.24}$$

Then (5.21) and Lemma 5.5 yield

$$\lim_{A\to\infty} \overline{\lim_{n\to\infty}} P\left(\sup_{s\leq t} |\overline{V}_s^n| > A\right) = 0, \quad t > 0 \ . \tag{5.25}$$

We now check that, for any $T > 0$ and $\eta > 0$, we have that

$$\lim_{\delta\to 0} \overline{\lim_{n\to\infty}} \sup_{\tau\in S_T(\mathbb{F}^n)} P\left(\sup_{t\leq\delta} \left|(\overline{V}_{t+\tau}^n)^2 - (\overline{V}_\tau^n)^2\right| > \eta\right) = 0 \ , \tag{5.26}$$

where $S_T(\mathbb{F}^n)$ is the set of all $\mathbb{F}^n$-stopping times τ not greater than T.

Since the processes $(\int_0^t \overline{V}_{s-}^n dM_s^n, t \geq 0)$ and $(\sum_{s\leq t}(\Delta M_s^n)^2 - \langle M^n\rangle_t,\ t \geq 0)$ are $\mathbb{F}^n$-local martingales [10, Ch. 1, §8, Ch. 2, §2], the Lenglart-Rebolledo inequality [10, Theorem 1.9.3] yields, in view of (5.23), for $\epsilon > 0$,

$$\begin{aligned}P\left(\sup_{t\leq\delta}\left|(\overline{V}_{t+\tau}^n)^2 - (\overline{V}_\tau^n)^2\right| > \eta\right) &\leq \frac{\epsilon}{\eta} \\ &+ P\Big(2\sqrt{n}|\rho^n - 1|\int_\tau^{\tau+\delta} |\overline{V}_u^n|du + 2\sqrt{n}\int_\tau^{\tau+\delta} |\overline{V}_u^n|\, 1(V_u^n = 0)du \\ &+ 2\sqrt{n}\int_\tau^{\tau+\delta} |\overline{V}_u^n|\, 1(V_u^n > 0)\beta_u^n du + \langle M^n\rangle_{\tau+\delta} - \langle M^n\rangle_\tau > \epsilon\Big) \ .\end{aligned} \tag{5.27}$$

By (5.25) and the assumed limit $\sqrt{n}(\rho^n - 1) \to c$, we have

$$\lim_{\delta\to 0} \overline{\lim_{n\to\infty}} P\left(2\sqrt{n}|\rho^n - 1| \sup_{\substack{|s-t|\le\delta \\ s\le T}} \int_s^t |\overline{V}^n_u| du > \frac{\epsilon}{4}\right) = 0 \ . \tag{5.28}$$

By (5.21), we see that $|\overline{V}^n_s| 1(V^n_s = 0) = |\epsilon^n_s| 1(V^n_s = 0)$, so (5.16) and Lemma 5.5 yield

$$2\sqrt{n}\int_0^t |\overline{V}^n_s|\, 1(V^n_s = 0) ds \xrightarrow{\mathrm{P}} 0 \quad (n\to\infty), \quad t > 0 \ . \tag{5.29}$$

Next, for $\epsilon' > 0$, $0 < s < t$, we again use (5.21) and obtain

$$\begin{aligned} 2\sqrt{n}\int_s^t |\overline{V}^n_u|\, 1(V^n_u > 0)\beta^n_u du \ &\le\ 2\sqrt{n}\sup_{u\le t}|\epsilon^n_u| \int_s^t 1(V^n_u > 0)\beta^n_u du \\ &+\ 2\sqrt{n}\sup_{u\le t}|V^n_u| \int_s^t 1(V^n_u > \epsilon')\beta^n_u du \\ &+\ 2\sqrt{n}\epsilon' \int_s^t 1(V^n_u > 0)\beta^n_u du \ . \end{aligned} \tag{5.30}$$

The first term on the right tends to 0 in probability as $n\to\infty$ by Lemma 5.5 and (5.17). The third term tends in probability to zero as $n\to\infty$ and then $\epsilon' \to 0$ by (5.17). Finally, by (5.3), Lemma 5.2 and (5.24), we have for $\gamma > 0$,

$$\lim_{\delta\to 0} \overline{\lim_{n\to\infty}} P\left(2\sqrt{n}\sup_{u\le T}|V^n_u| \sup_{\substack{|s-t|<\delta \\ s\le T}} \int_s^t 1(V^n_u > \epsilon')\beta^n_u du > \gamma\right) = 0 \ .$$

Thus, by (5.30),

$$\lim_{\delta\to 0} \overline{\lim_{n\to\infty}} P\left(2\sqrt{n} \sup_{\substack{|s-t|\le\delta \\ s\le T}} \int_s^t |\overline{V}^n_u| \cdot 1(V^n_u > 0)\beta^n_u du > \frac{\epsilon}{4}\right) = 0 \ . \tag{5.31}$$

Lemma 5.5 easily implies that

$$\lim_{\delta\to 0} \overline{\lim_{n\to\infty}} P\left(\sup_{\substack{|s-t|<\delta \\ s\le T}} |\langle M^n\rangle_t - \langle M^n\rangle_s| > \frac{\epsilon}{4}\right) = 0 \ . \tag{5.32}$$

Applying (5.28), (5.29), (5.31) and (5.32) to (5.27) shows that (5.26) holds.

Now, by Aldous' condition (see, e.g. [10, Theorem 6.3]), (5.25) and (5.26) show that the sequence $\{(\overline{V}^n)^2, n\ge 1\}$ and hence $\{\overline{V}^n, n \ge 1\}$ is tight for the Skorohod topology . By [6, Proposition VI.3.26], it remains to prove that

$$\sup_{t\le T}|\Delta\overline{V}^n_t| \xrightarrow{\mathrm{P}} 0, \quad T > 0 \ .$$

But

$$\sup_{t\leq T}|\Delta\overline{V}_t^n| = \sup_{t\leq T}|\Delta M_t^n| \leq \max_{\ell=1,2}\left[\frac{1}{\sqrt{n}}\max_{1\leq i\leq A_{nT}^{n,\ell}+1}\eta_i^{n,\ell} + \frac{\rho_\ell^n}{\sqrt{n}}\max_{1\leq i\leq A_{nT}^{n,\ell}+1}\xi_i^{n,\ell}\right]$$

which tends to 0 in probability as $n \to \infty$ by Lemmas 3.1, 5.1 and (2.4), (2.5). This proves that $\{\overline{V}^n, n \geq 1\}$ is C-tight. The sequence $\{V^n, n \geq 1\}$ is then C-tight by Lemma 5.5 and (5.21). ■

By Prohorov's theorem, there exists a subsequence $\{V^{n'}, n' \geq 1\}$ and a continuous process $\tilde{V}$ such that $V^{n'} \xrightarrow{\mathrm{d}} \tilde{V}$. The next two lemmas deal with implications of this fact.

Lemma 5.7 *We have, for $\eta > 0$,*

$$\lim_{\epsilon\to 0}\overline{\lim_{n\to\infty}} P(\int_0^t 1(V_s^n < \epsilon)\,ds > \eta) = 0.$$

In particular, if the law of a process $\tilde{V} = (\tilde{V}_t, t \geq 0)$ is an accumulation point of the laws of $\{V^n, n \geq 1\}$, then

$$\int_0^t 1(\tilde{V}_s = 0)\,ds = 0 \; P-\,a.s.$$

Proof. Since $V^{n'} \xrightarrow{\mathrm{d}} \tilde{V}$ for a subsequence (n'), we have, for $\epsilon > 0$ and $\eta > 0$,

$$\underline{\lim}_{n'\to\infty} P\left(\int_0^t 1(V_s^{n'} < \epsilon)\,ds > \eta\right) \geq P\left(\int_0^t 1(\tilde{V}_s < \epsilon)\,ds > \eta\right),$$

and the second assertion of the lemma is a consequence of the first. To prove that, introduce the processes $Z^n = (Z_t^n, t \geq 0)$ by

$$Z_t^n = V_0^n + B_t^n + \sqrt{n}\int_0^t 1(V_s^n > 0)\beta_s^n\,ds,\; t \geq 0.$$

so that (5.2) is equivalent to

$$V_t^n = Z_t^n + \sqrt{n}\int_0^t 1(V_s^n = 0)\,ds,$$

which implies, since V_t^n is nonnegative, that $V^n = R(Z^n)$, where $R : D[0,\infty) \to D[0,\infty)$ is Skorohod's reflection map defined by

$$R(X)_t = X_t - \inf_{s\leq t} X_s \wedge 0,\; t \geq 0. \tag{5.33}$$

Now, if we define

$$\check{Z}_t^n = V_0^n + B_t^n,\; t \geq 0,\; \check{Z}^n = (\check{Z}_t^n, t \geq 0), \tag{5.34}$$

and introduce

$$\check{V}^n = R(\check{Z}^n), \tag{5.35}$$

then the process $Z^n - \check{Z}^n$ is increasing and (5.33) implies that $V_t^n \geq \check{V}_t^n, t \geq 0$. Hence, for $\epsilon > 0$,

$$P\left(\int_0^t 1(V_s^n < \epsilon)\, ds > \eta\right) \leq P\left(\int_0^t 1(\check{V}_s^n < \epsilon)\, ds > \eta\right). \tag{5.36}$$

Now, by (5.34) and Lemma 5.1, $\{\check{Z}^n, n \geq 1\}$ converges in distribution to the process $\check{Z} = (V_0 + \sigma W_t + ct, t \geq 0)$, where V_0 and $(W_t, t \geq 0)$ are independent. By the continuity of the reflection at continuous functions, we then deduce that $\check{V}^n \xrightarrow{d} R(\check{Z})$, i.e., by (5.36),

$$\begin{aligned}\overline{\lim_{\epsilon\to 0}}\,\overline{\lim_{n\to\infty}}\, P\left(\int_0^t 1(V_s^n < \epsilon)\, ds > \eta\right) &\leq \overline{\lim_{\epsilon\to 0}}\,\overline{\lim_{n\to\infty}} P\left(\int_0^t 1(\check{V}_s^n < \epsilon) ds > \eta\right)\\ \leq \overline{\lim_{\epsilon\to 0}} P\left(\int_0^t 1(R(\check{Z})_s \leq \epsilon)\, ds \geq \eta\right) &= P\left(\int_0^t 1(R(\check{Z})_s = 0)\, ds \geq \eta\right).\end{aligned}$$

Since $R(\check{Z})$ is a reflected Brownian motion, the latter probability is 0. ■

Lemma 5.8 *Under the conditions of Theorem 2.1, for $T > 0$,*

$$\lim_{n\to\infty} \sqrt{n} \int_0^T \beta_t^n V_t^n dt = dT\ .$$

Proof. We rely heavily on the argument in the proof of Theorem 4.1. The notation in that proof is used here. Again let a continuous process $\tilde{V} = (\tilde{V}_t, t \geq 0)$ be an accumulation point of $\{V^n, n \geq 1\}$. By Lemma 5.7, the conditions of Theorem 4.1 hold, so the results developed in the proof of Theorem 4.1 apply.

As in the proof of Theorem 4.1, it is enough to prove

$$\sqrt{n} \int_0^T V_t^n \beta_t^n \cdot 1(\delta \leq V_t^n \leq K) ds \xrightarrow{\mathrm{d}} d \int_0^T 1(\delta \leq \tilde{V}_t \leq K) dt \tag{5.37}$$

for any δ and K, $0 < \delta < K$, which satisfy (4.2). To check this, note first that, for $\eta > 0$,

$$\lim_{\delta\to 0} \overline{\lim_{n\to\infty}} P\left(\sqrt{n} \int_0^T V_t^n \beta_t^n \cdot 1(V_t^n \leq \delta)\, dt > \eta\right) = 0, \tag{5.38}$$

$$\lim_{K\to\infty} \overline{\lim_{n\to\infty}} P\left(\sqrt{n} \int_0^T V_t^n \beta_t^n \cdot 1(V_t^n \geq K)\, dt > \eta\right) = 0. \tag{5.39}$$

Limit (5.38) follows from (5.17). For (5.39), write

$$P\left(\sqrt{n} \int_0^T V_t^n \beta_t^n \cdot 1(V_t^n \geq K)\, dt > \eta\right) \leq P(\sup_{t\leq T} V_t^n \geq K)$$

and observe that the latter goes to 0 as $n \to \infty$ and $K \to \infty$ by the tightness of $\{V^n, n \geq 1\}$. Theorem 4.2 in [1] then implies the desired result by (5.38), (5.39) and the fact that, by Lemma 5.7 and the continuity of $\tilde{V}$,

$$\lim_{\delta\to 0} \int_0^T 1(\tilde{V}_t \leq \delta) dt = \int_0^T 1(\tilde{V}_t = 0) dt = 0,\ \lim_{K\to\infty} \int_0^T 1(\tilde{V}_t > K) dt = 0,\ P - a.s.\ .$$

So, we prove (5.37) next assuming (4.2).

Let $\omega^{n,1}$ (which is $\vartheta_T^{n,1}$ in the notation of Lemma 5.2) and $\omega^{n,2}$ (which is $\vartheta_T^{n,2}$ in Lemma 5.2) denote the number of respective switchovers from queue 1 to queue 2 and from queue 2 to queue 1 started in $[0, nT]$. By the definitions above (recall in particular that $\alpha_t^n = 0$ if the server is switching over at t), for $k \geq 1,\ 0 \leq i \leq N$,

$$\sum_{j=0}^{\vartheta_{i,k}^{n,1}} s_j^{n,1}(i,k) + \sum_{j=0}^{\vartheta_{i,k}^{n,2}} s_j^{n,2}(i,k)$$
$$\leq \int_{n(\tau_k^n(\epsilon,i)\wedge T)}^{n(\zeta_k^n(\epsilon,i)\wedge T)} (1-\alpha_t^n)dt$$
$$\leq \left(\max_{1\leq j\leq \omega^{n,1}} s_j^{n,1} + \max_{1\leq j\leq \omega^{n,2}} s_j^{n,2}\right)\cdot 1(\tau_k^n(\epsilon,i) < T) + \sum_{j=0}^{\vartheta_{i,k}^{n,1}+1} s_j^{n,1}(i,k) + \sum_{j=0}^{\vartheta_{i,k}^{n,2}+1} s_j^{n,2}(i,k)\ ,$$

so, since $V_t^n \in [a_i(\epsilon)-\epsilon, a_i(\epsilon)+\epsilon]$ on $[\tau_k^n(\epsilon,i), \zeta_k^n(\epsilon,i))$, we get

$$\begin{aligned}
&\frac{1}{\sqrt{n}}(a_i(\epsilon)-\epsilon)\left(\sum_{j=0}^{\vartheta_{i,k}^{n,1}} s_j^{n,1}(i,k) + \sum_{j=0}^{\vartheta_{i,k}^{n,2}} s_j^{n,2}(i,k)\right)\\
&\leq \sqrt{n}\int_{\tau_k^n(\epsilon,i)\wedge T}^{\zeta_k^n(\epsilon,i)\wedge T} \beta_t^n V_t^n dt\\
&\leq (a_i(\epsilon)+\epsilon)\cdot 1(\tau_k^n(\epsilon,i)<T)\frac{1}{\sqrt{n}}\left(\max_{1\leq j\leq\omega^{n,1}} s_j^{n,1} + \max_{1\leq j\leq\omega^{n,2}} s_j^{n,2}\right)\\
&+\frac{1}{\sqrt{n}}(a_i(\epsilon)+\epsilon)\left(\sum_{j=0}^{\vartheta_{i,k}^{n,1}+1} s_j^{n,1}(i,k) + \sum_{j=0}^{\vartheta_{i,k}^{n,2}+1} s_j^{n,2}(i,k)\right)\ .
\end{aligned}\tag{5.40}$$

Introduce

$$\tilde{A}_k^n(\epsilon,i) = \frac{1}{\sqrt{n}}(a_i(\epsilon)-\epsilon)\left(\sum_{j=0}^{\hat{\vartheta}_{i,k}^{n,1}} s_j^{n,1}(i,k) + \sum_{j=0}^{\hat{\vartheta}_{i,k}^{n,2}} s_j^{n,2}(i,k)\right), \tag{5.41}$$

$$\begin{aligned}
\hat{A}_k^n(\epsilon,i) &= \frac{1}{\sqrt{n}}(a_i(\epsilon)+\epsilon)\left(\sum_{j=0}^{\tilde{\vartheta}_{i,k}^{n,1}+1} s_j^{n,1}(i,k) + \sum_{j=0}^{\tilde{\vartheta}_{i,k}^{n,2}+1} s_j^{n,2}(i,k)\right)\\
&+(a_i(\epsilon)+\epsilon)\cdot 1(\tau_k^n(\epsilon,i)<T)\tilde{\omega}^n,
\end{aligned}\tag{5.42}$$

where

$$\tilde{\omega}^n = \frac{1}{\sqrt{n}}\left(\max_{1\leq j\leq\omega^{n,1}} s_j^{n,1} + \max_{1\leq j\leq\omega^{n,2}} s_j^{n,2}\right).$$

It was shown in the proof of Lemma 5.2 (see (5.7)) that $P\left(\frac{\omega^{n,\ell}}{n} > \frac{T+1}{d_\ell}\right) \to 0$, $\ell = 1,2$, as $n\to\infty$. Then, using (2.6) and Lemma 3.1, we get

$$\tilde{\omega}^n \xrightarrow{\mathrm{P}} 0\ (n\to\infty)\ . \tag{5.43}$$

Also, inequalities (5.40) and (4.14) (the latter holds obviously for the second queue too) yield

$$\tilde{A}_k^n(\epsilon, i) \leq \sqrt{n} \int_{\tau_k^n(\epsilon,i)\wedge T}^{\zeta_k^n(\epsilon,i)\wedge T} \beta_t^n V_t^n dt \leq \hat{A}_k^n(\epsilon, i) \; . \tag{5.44}$$

Next, (4.10), (4.12), and (4.13), and their analogues for the second queue imply, in view of (4.4) and Lemma 2.2 in [3], that

$$\left(\frac{1}{\sqrt{n}}\tilde{\vartheta}_{i,k}^{n,\ell}\right)_{\substack{0\leq i\leq N,\\ k\geq 1,\ell=1,2}} \xrightarrow{\text{d}} \left(\frac{\zeta_k(\epsilon,i)\wedge T - \tau_k(\epsilon,i)\wedge T}{a_i(\epsilon)-\epsilon}\left(\frac{1}{\rho_\ell}+\frac{1}{1-\rho_\ell}\right)^{-1}\right)_{\substack{0\leq i\leq N,\\ k\geq 1,\ell=1,2}} \tag{5.45}$$

$$\left(\frac{1}{\sqrt{n}}\hat{\vartheta}_{i,k}^{n,\ell}\right)_{\substack{0\leq i\leq N,\\ k\geq 1,\ell=1,2}} \xrightarrow{\text{d}} \left(\frac{\zeta_k(\epsilon,i)\wedge T - \tau_k(\epsilon,i)\wedge T}{a_i(\epsilon)+\epsilon}\left(\frac{1}{\rho_\ell}+\frac{1}{1-\rho_\ell}\right)^{-1}\right)_{\substack{0\leq i\leq N,\\ k\geq 1,\ell=1,2}} \tag{5.46}$$

Hence, by (5.41) - (5.43), Lemma 5.1, the equality $\rho_1 + \rho_2 = 1$ and again Lemma 2.2 in [3],

$$(\tilde{A}_k^n(\epsilon, i))_{\substack{0\leq i\leq N,\\ k\geq 1}} \xrightarrow{\text{d}} (\tilde{A}_k(\epsilon, i))_{\substack{0\leq i\leq N,\\ k\geq 1}} \; , \tag{5.47}$$

and

$$(\hat{A}_k^n(\epsilon, i))_{\substack{0\leq i\leq N,\\ k\geq 1}} \xrightarrow{\text{d}} (\hat{A}_k(\epsilon, i))_{\substack{0\leq i\leq N,\\ k\geq 1}} \; , \tag{5.48}$$

where

$$\tilde{A}_k(\epsilon, i) = \frac{a_i(\epsilon)-\epsilon}{a_i(\epsilon)+\epsilon}(\zeta_k(\epsilon,i)\wedge T - \tau_k(\epsilon,i)\wedge T)\rho_1\rho_2(d_1+d_2) \; , \tag{5.49}$$

$$\hat{A}_k(\epsilon, i) = \frac{a_i(\epsilon)+\epsilon}{a_i(\epsilon)-\epsilon}(\zeta_k(\epsilon,i)\wedge T - \tau_k(\epsilon,i)\wedge T)\rho_1\rho_2(d_1+d_2) \; . \tag{5.50}$$

Introduce

$$\tilde{A}^n(\epsilon) = \sum_{k=1}^{\infty}\sum_{i=1}^{N-1} \tilde{A}_k^n(\epsilon, i), \; \hat{A}^n(\epsilon) = \sum_{k=1}^{\infty}\sum_{i=0}^{N} \hat{A}_k^n(\epsilon, i) \; , \tag{5.51}$$

$$\tilde{A}(\epsilon) = \sum_{k=1}^{\infty}\sum_{i=1}^{N-1} \tilde{A}_k(\epsilon, i), \; \hat{A}(\epsilon) = \sum_{k=1}^{\infty}\sum_{i=0}^{N} \hat{A}_k(\epsilon, i) \; . \tag{5.52}$$

Note that the sums are P–a.s. finite (use (4.3) for the second line), so that the variables above are well defined.

Relations (5.47) and (5.48) imply, by the continuous mapping theorem, that for every $M = 1, 2, \ldots$,

$$\sum_{k=1}^{M}\sum_{i=1}^{N-1} \hat{A}_k^n(\epsilon, i) \xrightarrow{\text{d}} \sum_{k=1}^{M}\sum_{i=1}^{N-1} \hat{A}_k(\epsilon, i) \; ,$$

$$\sum_{k=1}^{M}\sum_{i=0}^{N} \tilde{A}_k^n(\epsilon, i) \xrightarrow{\text{d}} \sum_{k=1}^{M}\sum_{i=0}^{N} \tilde{A}_k(\epsilon, i)$$

as $n \to \infty$.

On the other hand, in a manner similar to (4.23) and (4.24),

$$\lim_{M\to\infty} \overline{\lim_{n\to\infty}} P\left(\left| \sum_{k=1}^{M} \sum_{i=1}^{N-1} \hat{A}_k^n(\epsilon, i) - \hat{A}^n(\epsilon) \right| > 0 \right) = 0 \ ,$$

$$\lim_{M\to\infty} P\left(\left| \sum_{k=1}^{M} \sum_{i=1}^{N-1} \hat{A}_k(\epsilon, i) - \hat{A}(\epsilon) \right| > 0 \right) = 0 \ ,$$

and Theorem 4.2 in [1] yields

$$\hat{A}^n(\epsilon) \xrightarrow{\mathrm{d}} \hat{A}(\epsilon) \quad (n \to \infty) \ . \tag{5.53}$$

Similarly,

$$\tilde{A}^n(\epsilon) \xrightarrow{\mathrm{d}} \tilde{A}(\epsilon) \quad (n \to \infty) \ . \tag{5.54}$$

Next, in analogy with the proof of (4.27), we get in view of (4.2) that, as $\epsilon \to 0$,

$$\hat{A}(\epsilon) \quad \xrightarrow{\mathrm{P}} \quad \rho_1\rho_2(d_1+d_2)\int_0^T 1(\delta \le \tilde{V}(t) \le K)dt \ , \tag{5.55}$$

$$\tilde{A}(\epsilon) \quad \xrightarrow{\mathrm{P}} \quad \rho_1\rho_2(d_1+d_2)\int_0^T 1(\delta \le \tilde{V}(t) \le K)dt \ . \tag{5.56}$$

Also, it is not difficult to see, using the definitions of $\tau_k^n(\epsilon, i$

(

By (5.16), (5.17), Lemma 5.5 and Lemma 5.8,

$$\sup_{s\leq t}|\delta^n_s| \xrightarrow{\mathrm{P}} 0 \quad (n\to\infty), \quad t>0\ . \tag{6.2}$$

We denote the left-hand side of (6.1) by X^n_t. Let $X^n = (X^n_t, t\geq 0)$. We next prove that X^n converges in distribution to X as $n\to\infty$. The process X^n is an $\mathbb{F}^n$-locally square integrable semimartingale [6, Ch. II, §2b)]. The process X as defined by (2.7) is a locally square integrable semimartingale as well with respect to the filtration generated by it [6, III.2.12]. We prove the convergence by applying Theorem IX.3.48 in [6] which gives conditions for convergence in distribution of a sequence of semimartingales to a semimartingale in terms of their predictable triplets [6, Ch. II, §2]. Therefore our first step is to identify the characteristics. Let $B'^n = (B'^n_t, t\geq 0)$ denote the first characteristic without truncation of X^n, let $\tilde{C}'^n = (\tilde{C}'^n_t, t\geq 0)$ denote its modified second characteristic without truncation and let $\nu^n = (\nu^n(ds,dx))$ denote its predictable measure of jumps [6, II.2.29, IX.3.25]. Then by (6.1)

$$B'^n_t = 2\sqrt{n}(\rho^n-1)\int_0^t \overline{V}^n_s ds + (2d+\sigma^2)t\ , \tag{6.3}$$

$$\tilde{C}'^n_t = 4\int_0^t (\overline{V}^n_s)^2 d\langle M^n\rangle_s\ . \tag{6.4}$$

This specifies the triplet of X^n.

Define, next, for $\alpha = (\alpha_t, t\geq 0)$, an element of the Skorohod space $D[0,\infty)$,

$$B_t(\alpha) = 2c\int_0^t (\alpha_s\vee 0)^{1/2}ds + (2d+\sigma^2)t, \quad t\geq 0, \tag{6.5}$$

$$C_t(\alpha) = 4\sigma^2\int_0^t (\alpha_s\vee 0)ds, \quad t\geq 0\ , \tag{6.6}$$

$$\nu([0,t],\Gamma)(\alpha) = 0,\ t\geq 0\ , \quad \Gamma \text{ is a Borel subset of } R\ ,$$

and let $B(\alpha) = (B_t(\alpha), t\geq 0), C(\alpha) = (C_t(\alpha), t\geq 0)$ and $\nu(\alpha) = (\nu(dt,dx)(\alpha))$. According to the definition of X in (2.7), its triplet of predictable characteristics is $(B(X), C(X), \nu(X))$. Since X is continuous, this triplet does not depend on a truncation function; in particular the triplet without truncation is the same.

Stated in another way, the distribution of X is the unique solution to the martingale problem associated with $(\mathcal{H}, X)$ and $(\mathcal{L}(X_0); B, C, \nu)$, in the sense of definition III.2.4 of [6], where $\mathcal{H}$ denotes the σ-field generated by X_0 and $\mathcal{L}(X_0)$ denotes the distribution of X_0.

Define next, as in IX.3.38 [6], for $a\geq 0$,

$$S_a(\alpha) = \inf(t: |\alpha_t|\geq a \text{ or } |\alpha_{t-}|\geq a), \tag{6.7}$$

$$S^n_a = \inf(t: |X^n_t|\geq a \text{ or } |X^n_{t-}|\geq a), \tag{6.8}$$

where α_{t-} denotes the left-hand limit at t. Let also $(\text{Var } B)_t(\alpha)$ denote the variation of $B(\alpha)$ on $[0,t]$ and $\mathbb{C}_1(R)$ denote the set of continuous bounded functions $g : R \to R$ which are equal to 0 in a neighborhood of 0.

By Theorem IX.3.48 [6], in order to prove that the X^n converge in distribution to X, we may check the following conditions (note that since X has no jumps, in the notation of the theorem, $B' = B$ and $\tilde{C}' = C$):

(i) *The local strong majorization hypothesis*: for all $a \geq 0$, there is an increasing continuous and deterministic function $F(a) = (F_t(a), t \geq 0)$ such that the stopped processes $\Big((\text{Var } B)_{t\wedge S_a(\alpha)}(\alpha), t \geq 0\Big)$, $\Big(C_{t\wedge S_a(\alpha)}(\alpha), t \geq 0\Big)$ and $\Big(\int_0^{t\wedge S_a(\alpha)} \int_R |x|^2 \nu(ds,dx)(\alpha), t \geq 0\Big)$ are strongly majorized by $F(a)$ for all $\alpha \in D[0,\infty)$.

(ii) *The local condition on big jumps*: for all $a \geq 0, t > 0$,

$$\lim_{b\to\infty} \sup_{\alpha\in D[0,\infty)} \int_0^{t\wedge S_a(\alpha)} \int_R |x|^2 1(|x| > b)\nu(ds,dx)(\alpha) = 0.$$

(iii) *Local uniqueness for the martingale problem associated with* $(\mathcal{H}, X)$ *and* $(\mathcal{L}(X_0); B, C, \nu)$ (see [6, III.2]).

(iv) *The continuity condition*: the maps $\alpha \to B_t(\alpha)$, $\alpha \to C_t(\alpha)$ and $\alpha \to \int_0^t \int_R g(x)\nu(ds,dx)$ are continuous for the Skorohod topology on $D[0,\infty)$ for all $t > 0$ and $g \in \mathbb{C}_1(R)$.

(v) $X_0^n \stackrel{\text{d}}{\to} X_0$.

(vi) $[\delta_{\text{loc}} - R_+]$ $\displaystyle\int_0^{t\wedge S_a^n} \int_R g(x)\nu^n(ds,dx) - \int_0^{t\wedge S_a^n} \int_R g(x)\nu(ds,dx)(X^n) \stackrel{\text{P}}{\to} 0$ for all $t > 0$, $a > 0$ and $g \in \mathbb{C}_1(R)$;

$[\sup -\beta'_{\text{loc}}]$ $\sup_{s\leq t} |B'^n_{s\wedge S_a^n} - B_{s\wedge S_a^n}(X^n) \stackrel{\text{P}}{\to} 0$ for all $t > 0$, $a > 0$;

$[\gamma'_{\text{loc}} - R_+]$ $\tilde{C}'^n_{t\wedge S_a^n} - C_{t\wedge S_a^n}(X^n) \stackrel{\text{P}}{\to} 0$ for all $t > 0$, $a > 0$;

(6.8a) $\displaystyle\lim_{b\to\infty} \overline{\lim_{n\to\infty}} P\left(\int_0^{t\wedge S_a^n} \int_R |x|^2 1(|x| > b)\nu^n(ds,dx) > \epsilon\right) = 0$ for all $t > 0$, $a > 0$ and $\epsilon > 0$.

This last condition is Equation (3.49) in [6].

We now check these 9 conditions in order. We have, by (6.5)–(6.7), (6.7), for $s < t$,

$$(\text{Var } B)_{t\wedge S_a(\alpha)}(\alpha) - (\text{Var } B)_{s\wedge S_a(\alpha)}(\alpha) \leq 2|c|a^{1/2}(t-s) + (2d + \sigma^2)(t-s)\ ,$$

$$C_{t\wedge S_a(\alpha)}(\alpha) - C_{s\wedge S_a(\alpha)}(\alpha) \leq 4\sigma^2 a(t-s),$$

$$\int_0^{t\wedge S_a(\alpha)} \int_R |x|^2 \nu(du,dx)(\alpha) - \int_0^{s\wedge S_a(\alpha)} \int_R |x|^2 \nu(du,dx)(\alpha) = 0\ .$$

This verifies condition (i) with $F_t(a) = K(a)t$, where $K(a)$ is large enough.

Condition (ii) holds since $\nu = 0$. Condition (iii) (local uniqueness) holds by Theorem III.2.40 [6] since the equation

$$Y_t = 2c \int_0^t (Y_s \vee 0)^{1/2} ds + (2d + \sigma^2)t + 2\sigma \int_0^t (Y_s \vee 0)^{1/2} dW_s + x \ ,$$

where $W = (W_t, t \geq 0)$ is a Wiener process, has a unique (weak) solution [5, Chapter IV, Theorems 1.1, 2.4, and 3.2, Example 8.2] for any $x \in R$, and since one can set, in the conditions of Theorem III.2.40 [6], $p_t B = B$, $p_t C = C$, $p_t \nu = \nu = 0$.

Condition (iv) is immediate from (6.5)–(6.7) since Skorohod convergence implies convergence at continuity points of the limit (e.g., use the argument of the proof of Theorem 6.2.2 in [10]). Condition (v) holds by the assumption $V_0^n \xrightarrow{d} V_0$, (5.21), and Lemma 5.5.

Consider condition $[\delta_{loc} - R_+]$ under (vi). Since $\nu = 0$, it is enough to prove that

$$\int_0^t \int_R |g(x)| \nu^n(ds, dx) \xrightarrow{\mathrm{P}} 0 \ .$$

Since, by the definition of $\mathbb{C}_1(R)$, for some $\epsilon > 0$, $g(x) = 0$ if $|x| < \epsilon$, and $g(x)$ is bounded, the latter integral converges in probability to zero as $n \to \infty$ if $\nu^n([0,t], \{|x| > \epsilon\}) \xrightarrow{\mathrm{P}} 0$. By Lemma 1, p. 424 in [10], this is implied by

$$\sup_{s \leq t} |\Delta X_s^n| \xrightarrow{\mathrm{P}} 0 \quad (n \to \infty), \quad t > 0 \ . \tag{6.9}$$

By the definition of X^n, $|\Delta X_s^n| \leq \Delta |\overline{V}_s^n|^2 + |\Delta \delta_s^n|$, and (6.9) holds by (6.2) and the C-tightness of $\overline{V}^n$.

Next, we check condition [Sup-β'_{loc}] under (vi). By definition of X^n, (6.3) and (6.5),

$$\begin{aligned}
\sup_{s \leq t} |B'^n_{s \wedge S_a^n} - B_{s \wedge S_a^n}(X^n)| \ &\leq \ 2|\sqrt{n}(\rho^n - 1) - c| \int_0^t |\overline{V}_s^n| ds \\
&\quad + 2|c| \int_0^t |\overline{V}_s^n - (((\overline{V}_s^n)^2 + \delta_s^n) \vee 0)^{1/2}| ds \\
&\leq \ 2|\sqrt{n}(\rho^n - 1) - c| t \sup_{s \leq t} |\overline{V}_s^n| + 2|c| t \sup_{s \leq t} |\delta_s^n|^{1/2},
\end{aligned}$$

and the latter converges in probability to 0 as $n \to \infty$, since $\sqrt{n}(\rho^n - 1) \to c$, $\{\overline{V}^n, n \geq 1\}$ is tight and $\sup_{s \leq t} |\delta_s^n| \xrightarrow{\mathrm{P}} 0$ by (6.2).

Now consider condition $[\gamma'_{loc} - R_+]$ under (vi). By (6.4), (6.6) and the definition of X^n

$$\left| \tilde{C}'^n_{t \wedge S_a^n} - C_{t \wedge S_a^n}(X^n) \right| \leq 4 \sup_{s \leq t} \left| \int_0^s (\overline{V}_u^n)^2 d\langle M^n \rangle_u - \sigma^2 \int_0^s (\overline{V}_u^n)^2 du \right| + 4\sigma^2 t \sup_{s \leq t} |\delta_s^n| \ . \tag{6.10}$$

The last term converges in probability to 0 by (6.2). Since $\{(\overline{V}^n)^2, n \geq 1\}$ is C-tight, we have, for $\eta > 0$,

$$\lim_{\delta \to 0} \overline{\lim_{n\to\infty}} P\left(\sup_{|u-v|<\delta, u,v \leq t} \left| (\overline{V}_u^n)^2 - (\overline{V}_v^n)^2 \right| > \eta \right) = 0,$$

which, in view of the second assertion of Lemma 5.5, is seen to yield (for a proof use [1, Problem 8, §2]).

$$\sup_{s\leq t} \left| \int_0^s (\overline{V}_u^n)^2 d\langle M^n\rangle_u - \sigma^2 \int_0^s (\overline{V}_u^n)^2 du \right| \xrightarrow{\text{P}} 0.$$

In view of (6.10), this concludes the verification of $[\gamma'_{loc} - D]$.

Finally, consider condition (6.8a) under (vi). Define

$$\tilde{L}_t^n = \int_0^t \int_R |x|^2 \cdot 1(|x| > b) \nu^n(ds, dx), \; t \geq 0.$$

Since $\nu^n(ds, dx)$ is the predictable measure of jumps of X^n, by (6.1) the process $\tilde{L}^n = (\tilde{L}_t^n, t \geq 0)$ is the $\mathbb{F}^n$-compensator of the process $L^n = (L_t^n, t \geq 0)$ defined by

$$L_t^n = 4 \sum_{s\leq t} (\overline{V}_{s-}^n)^2 (\Delta M_s^n)^2 \cdot 1(2|\overline{V}_{s-}^n||\Delta M_s^n| > b) \; . \tag{6.11}$$

We have for $\epsilon > 0$, $A > 0$,

$$P(\tilde{L}_t^n \geq \epsilon) \leq P\left(\sup_{s\leq t} |\overline{V}_s^n| > A \right) + P\left(\int_0^t 1(\overline{V}_{s-}^n \leq A) d\tilde{L}_s^n > \epsilon \right) \; . \tag{6.12}$$

The first term on the right goes in probability to 0 as $n \to \infty$ and $A \to \infty$ by the tightness of $\overline{V}^n$.

Next, letting

$$\gamma_t^{n,1} = \frac{1}{n} \tau_{\lfloor n(\lambda_1 t + 1) \rfloor}^{n,1}, \quad \gamma_t^{n,2} = \frac{1}{n} \tau_{\lfloor n(\lambda_2 t + 1) \rfloor}^{n,2} \; ,$$

we have for the second term on the right of (6.12), since $\gamma_t^{n,\ell} < t$ implies $A_{nt}^{n,\ell} \geq \lfloor n(\lambda_\ell t + 1) \rfloor$, $\ell = 1, 2$,

$$P\left(\int_0^t 1(\overline{V}_{s-}^n \leq A) d\tilde{L}_s^n > \epsilon \right) \leq P\left(\frac{1}{n} A_{nt}^{n,1} > \lambda_1 t + 1 - \frac{1}{n} \right)$$
$$+ P\left(\frac{1}{n} A_{nt}^{n,2} > \lambda_2 t + 1 - \frac{1}{n} \right) + P\left(\int_0^{t\wedge\gamma_t^{n,1}\wedge\gamma_t^{n,2}} 1(\overline{V}_{s-}^n \leq A) d\tilde{L}_s^n > \epsilon \right) \; .$$

Again the first two terms on the right tend to 0 as $n \to \infty$ by Lemma 5.1. It is thus left to prove that the last term on the right tends in probability to 0 as $n \to \infty$.

The process $(\int_0^t 1(\overline{V}_{s-}^n \leq A) d\tilde{L}_s^n, t \geq 0)$ is the $\mathbb{F}^n$-compensator of the process $(\int_0^t 1(\overline{V}_{s-}^n \leq A) dL_s^n, t \geq 0)$. Therefore, by (6.11) and, since, by Lemma 5.4, the $\frac{1}{n}\tau_i^{n,\ell}$, $\ell = 1, 2$, $i \geq 1$, are

$\mathbb{F}^n$-stopping times, the Lenglart-Rebolledo inequality (see, e.g., Theorem 1.9.3 [10]) implies that, for $\eta > 0$,

$$P\left(\int_0^{t\wedge\gamma_t^{n,1}\wedge\gamma_t^{n,2}} 1(\overline{V}_{s-}^n \le A)d\tilde{L}_s^n > \epsilon\right)$$
$$\le \frac{1}{\epsilon}\left(\eta + E\sup_{s\le t\wedge\gamma_t^{n,1}\wedge\gamma_t^{n,2}} 4A^2|\Delta M_s^n|^2\cdot 1\left(|\Delta M_s^n| > \frac{b}{2A}\right)\right)$$
$$+ P\left(4A^2\sum_{s\le t\wedge\gamma_t^{n,1}\wedge\gamma_t^{n,2}}(\Delta M_s^n)^2\cdot 1\left(|\Delta M_s^n| > \frac{b}{2A}\right)\ge\eta\right). \quad (6.13)$$

By the definitions of M^n (see (5.19)) and $\gamma_t^{n,1}$, $\gamma_t^{n,2}$,

$$E\sup_{s\le t\wedge\gamma_t^{n,1}\wedge\gamma_t^{n,2}}|\Delta M_s^n|^2 1\left(|\Delta M_s^n| > \frac{b}{2A}\right)$$
$$\le \frac{6}{n}E\sup_{i\le\lfloor n(\lambda_1 t+1)\rfloor}(\eta_i^{n,1}-\rho_1^n\xi_i^{n,1})^2\cdot 1\left(|\xi_i^{n,1}-\rho_1^n\eta_i^{n,1}| > \frac{b}{4A}\sqrt{n}\right)$$
$$+\frac{6}{n}E\sup_{i\le\lfloor n(\lambda_2(t+1))\rfloor}(\eta_i^{n,2}-\rho_2^n\xi_i^{n,2})^2\cdot 1\left(|\xi_i^{n,2}-\rho_2^n\eta_i^{n,2}| > \frac{b}{4A}\sqrt{n}\right)$$
$$\le 6(\lambda_1 t+1)E(\eta_i^{n,1}-\rho_1^n\xi_i^{n,1})^2\cdot 1\left(|\xi_1^{n,1}-\rho_1^n\eta_1^{n,1}| > \frac{b}{4A}\sqrt{n}\right)$$
$$+ 6(\lambda_2 t+1)E(\eta_i^{n,2}-\rho_2^n\xi_i^{n,2})^2\cdot 1\left(|\xi_1^{n,2}-\rho_2^n\eta_1^{n,2}| > \frac{b}{4A}\sqrt{n}\right),$$

which tends to 0 as $n\to\infty$ by (2.4) and (2.5). The third term on the right of (6.13) is not greater than

$$P\left(4A^2\left[\frac{6}{n}\sum_{i=1}^{\lfloor n(\lambda_1 t+1)\rfloor}(\xi_i^{n,1}-\rho_1^n\eta_i^{n,1})^2\cdot 1\left(|\eta_i^{n,1}-\rho_1^n\xi_i^{n,1}| > \frac{b}{4A}\sqrt{n}\right)\right.\right.$$
$$\left.\left.+\frac{6}{n}\sum_{i=1}^{\lfloor n(\lambda_2 t+1)\rfloor}(\eta_i^{n,2}-\rho_2^n\xi_i^{n,2})^2\cdot 1\left(|\xi_i^{n,2}-\rho_2^n\eta_i^{n,2}| > \frac{b}{4A}\sqrt{n}\right)\right]\ge\eta\right),$$

which again tends to 0 by (2.4) and (2.5). Therefore, by (6.13),

$$\overline{\lim_{n\to\infty}} P\left(\int_0^{t\wedge\gamma_t^{n,1}\wedge\gamma_t^{n,2}} 1(\overline{V}_{s-}^n \le A)d\tilde{L}_s^n > \epsilon\right)\le\frac{\eta}{\epsilon},$$

which completes the check of (6.8a) since η is arbitrary.

Thus, all the conditions of Theorem IX.3.48 [6] hold and, by the theorem, $\{X^n, n\ge 1\}$ converges in distribution to X, which by Lemma 2.1 is distributed as V^2. Since (5.21), Lemma 5.5, and (6.2) imply

$$P\left(\sup_{s\le t}|(V_s^n)^2 - X_s^n| > \delta\right)\to 0\ (n\to\infty),\quad t>0,\quad \delta>0,$$

we have that $(V^n)^2 \xrightarrow{\text{d}} V^2$, and hence that $V^n \xrightarrow{\text{d}} V$ since all the processes are nonnegative. ■

Proof of Theorem 2.2. The theorem follows by Theorem 2.1, Theorem 4.1 and Lemma 5.7. ■

Proof of Theorem 2.3 A basic equation for $\check{V}^n$ differs from (5.2) in that we have to substitute $\overline{r}^n - r^n(\sqrt{n}\check{V}^n_s)$ for β^n_s, i.e.,

$$\check{V}^n_t = V^n_0 + B^n_t + \sqrt{n}\int_0^t 1(\check{V}^n_s = 0)ds + \sqrt{n}\int_0^t 1(\check{V}^n_s > 0)(\overline{r}^n - r^n(\sqrt{n}\check{V}^n_s))ds\ .$$

After that the proof goes exactly the way it does for $\{V^n, n \geq 1\}$. We first prove the C-tightness of $\{\check{V}^n, n \geq 1\}$ by following the same steps as in Section 4. The only difference is that the $K^{n,\epsilon}$ are defined this time by

$$K^{n,\epsilon}_t = \sqrt{n}\int_0^t 1(\check{V}^n_s > \epsilon)(\overline{r}^n - r^n(\sqrt{n}\check{V}^n_s))\,ds,$$

and Lemma 5.2 is trivial because of $(r2)$.

As can be seen from the proof of Theorem 2.1 in Section 6, after the C-tightness has been proved, the only additional fact that is needed is the convergence

$$\sqrt{n}\int_0^t V^n_s\beta^n_s\,ds \xrightarrow{\text{P}} dt.$$

In the case of $\check{V}^n$, this amounts to proving that

$$\sqrt{n}\int_0^t \check{V}^n_s(\overline{r}^n - r^n(\sqrt{n}\check{V}^n_s))\,ds \xrightarrow{\text{P}} dt.$$

This turns out to be a simple consequence of $(r1)$, $(r2)$ and an analogue of Lemma 5.7 for $\check{V}^n$ which is proved in the same way. According to the latter result, for $\eta > 0$,

$$\lim_{\epsilon\to 0}\overline{\lim_{n\to\infty}}\,P\left(\int_0^t 1(\check{V}^n_s < \epsilon)\,ds > \eta\right) = 0. \tag{6.14}$$

Then

$$\begin{aligned}
P\left(\left|\sqrt{n}\int_0^t \check{V}^n_s(\overline{r}^n - r^n(\sqrt{n}\check{V}^n_s))\,ds - dt\right| > \eta\right)& \\
\leq\ P\left(\int_0^t \left|\sqrt{n}\check{V}^n_s(\overline{r}^n - r^n(\sqrt{n}\check{V}^n_s)) - d\right|\cdot 1(\check{V}^n_s < \epsilon)ds > \frac{\eta}{2}\right)& \\
+\ P\left(\int_0^t \left|\sqrt{n}\check{V}^n_s(\overline{r}^n - r^n(\sqrt{n}\check{V}^n_s)) - d\right|\cdot 1(\check{V}^n_s \geq \epsilon)\,ds > \frac{\eta}{2}\right)& \\
\leq\ P\left(\left(\sup_{x,n} x(\overline{r}^n - r^n(x)) + d\right)\int_0^t 1(\check{V}^n_s < \epsilon)ds > \frac{\eta}{2}\right)& \\
+\ 1\left(t\sup_{x\geq\sqrt{n}\epsilon}|x(\overline{r}^n - r^n(x)) - d| > \frac{\eta}{2}\right)&.
\end{aligned}$$

The probability on the right-hand side tends to 0 as $n \to \infty$ by $(r2)$ and (6.14), and the indicator goes to 0 as $n \to \infty$ by $(r1)$. ■.

References

[1] P. Billingsley, *Convergence of Probability Measures*, Wiley, New York, 1968.

[2] E. G. Coffman, Jr., A. A. Puhalskii, and M. I. Reiman, "Storage Limited Queues in Heavy Traffic," *Prob. Eng. Inf. Sci.*, 5 (1991), pp. 499–522.

[3] E. G. Coffman, Jr., A. A. Puhalskii, and M. I. Reiman, "Polling Systems with Zero Switchover Times: A Heavy Traffic Averaging Principle," *Ann. Appl. Prob.* (to appear).

[4] D. L. Iglehart and W. Whitt, "Multiple Channel Queues in Heavy Traffic, I and II," *Adv. Appl. Prob.*, 2 (1970), pp. 150–177 and 355–364.

[5] N. Ikeda and S. Watanabe, *Stochastic differential equations and diffusion processes*, North Holland, Amsterdam/Oxford/New York, 1989.

[6] J. Jacod and A. N. Shiryaev, *Limit Theorems for Stochastic Processes*, Springer, Berlin/Heidelberg/New York/London/Paris/Tokyo, 1987.

[7] J. B. Kruskal, "Work Scheduling Algorithms: A Nonprobabilistic Queueing Study (with Possible Application to No. 1 ESS)," *Bell Sys. Tech. J.*, 48 (1969), pp. 2963–2974.

[8] K. K. Leung, "Cyclic–Service Systems with Probabilistically–Limited Service," *IEEE J. Sel. Areas Comm.*, 9 (1991), pp. 185–193.

[9] H. Levy and M. Sidi, "Polling Systems: Applications, Modeling, and Optimization," *IEEE Trans. Comm.*, 38 (1990), pp. 1750–1760.

[10] R. Sh. Liptser and A. N. Shiryaev, *The Theory of Martingales*, Kluwer, Dordrecht/Boston/London, 1989.

[11] M. I. Reiman, "A Multiclass Feedback Queue in Heavy Traffic," *Adv. Appl. Prob.*, 2 (1988), pp. 179–207.

[12] M. I. Reiman and L. M. Wein, "Dynamic Scheduling of a Two–Class Queue with Setups," AT&T Bell Laboratories, Murray Hill, NJ 07974 (to appear).

[13] H. Takagi, "Queueing Analysis of Polling Models: An Update," in *Stochastic Analysis of Computer and Communication Systems*, H. Takagi (ed.) North Holland, Amsterdam, 1990.

[14] K. Yamada, "Diffusion Approximations for Storage Processes with General Release Rules", *Math. Oper. Res.*, 9 (1984), pp. 459–470.