MATHEMATICS OF OPERATIONS RESEARCH
Vol. 23, No. 2, May 1998
Printed in U.S.A.

POLLING SYSTEMS IN HEAVY TRAFFIC: A BESSEL PROCESS LIMIT

E. G. COFFMAN, JR., A. A. PUHALSKII, AND M. I. REIMAN

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.

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 performed by Leung (1991) is a modern example and a sequel to work of Kruskal (1969) on earlier switching systems. Introductions to a massive literature addressing many different applications can be found in Takagi (1986, 1990) and Levy and Sidi (1990).

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 Coffman, Puhalskii, and Reiman (1995). In a recent application of this principle, Reiman and Wein (1998) study set-up scheduling problems in two-class single-server queues. Olsen (1995) provides a heuristic refinement of the averaging principle that improves the quality of the resulting approximation for waiting time distributions in moderate loading.

A limitation of the results in Coffman et al. (1995) 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 Coffman et al. (1995) 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 Coffman et al. (1995) is analyzed in §3. Results for this

Received October 24, 1995; revised: December 18, 1996; October 9, 1997.
AMS 1991 subject classification. Primary: 60K25, 90B22.
OR/MS Index 1978 subject classification. Primary: Queues/Polling systems.
Key words. Polling systems, heavy traffic limit, averaging principles, Bessel process, semi-martingales.

0364-765X/98/2302/0257/$05.00

queueing system supply bounds for the polling system which lead to the averaging principle, as shown in §4. Further preliminaries are taken up in §5, where the tightness of a number of basic processes is proved. The development of §§3–5 culminates in the proofs in §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 from Jacod and Shiryaev (1987) and Liptser and Shiryaev (1989).

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 $d_i = 0$, $1 \le i \le M$, adapting the presentation of Coffman et al. (1995), which was for queue lengths, to the context of unfinished work. 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 (1970) as follows (see also Reiman (1988)). 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

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

(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 infinitesimal drift c and 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 Coffman et al. (1995) deals with queue lengths; converted to unfinished work, the principle states that, for any continuous function $f: R_+ \to R$ and any $T > 0$, we have

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

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

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

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

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. The purpose of the remainder of the paper is to formulate the argument precisely and to prove rigorously that it is correct.

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

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

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

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

For simplicity, let $M = 2$ and start the cycle at the moment the server switches to queue 1. On average, it takes time $\sqrt{n}x/(1 - \rho_1)$ to empty queue 1, d_1 to switch to queue 2, $\sqrt{n}x/(1 - \rho_2)$ 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

$$r^n(x) = \frac{x}{x + d/\sqrt{n}}, \tag{1.3}$$

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 $2d/\sigma^2 < 1$, 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

$$\pi(x) = \frac{a(ax)^{\beta} e^{-ax}}{\Gamma(\beta + 1)}, \qquad x \geq 0, \tag{1.4}$$

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 (1.4), 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 nth system, denote by $\tau_i^{n,l} = \xi_1^{n,l} + \cdots + \xi_i^{n,l}$, $i \geq 1$, $l = 1, 2$, the time of the ith arrival to the lth queue in terms of interarrival times $\xi_i^{n,l}$,

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

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

We assume that $\xi_i^{n,l}$, $i \geq 1$, $\eta_i^{n,l}$, $i \geq 1$, $s_i^{n,l}$, $i \geq 1$, $l = 1, 2$, are independent i.i.d. sequences, and that $\xi_i^{n,l} > 0$, $i \geq 1$. As in the previous section, we introduce, for $n = 1, 2, \ldots$ and $l = 1, 2$,

$$\lambda_l^n = (E\xi_1^{n,l})^{-1}, \qquad \mu_l^n = (E\eta_1^{n,l})^{-1}, \qquad d_l^n = Es_1^{n,l},$$

$$\rho_l^n = \frac{\lambda_l^n}{\mu_l^n}, \qquad \rho^n = \rho_1^n + \rho_2^n.$$

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

$$(\sigma_l^n)^2 = E(\eta_1^{n,l} - \rho_l^n \xi_1^{n,l})^2, \qquad l = 1, 2.$$

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

$$\lambda_l^n \to \lambda_l, \qquad \mu_l^n \to \mu_l > 0, \qquad \sigma_l^n \to \sigma_l > 0, \tag{2.1}$$

$$d_l^n \to d_l, \tag{2.2}$$

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 $l = 1, 2$, $\epsilon > 0$,

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

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

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

Also let

$$\sigma^2 = \lambda_1\sigma_1^2 + \lambda_2\sigma_2^2 > 0, \qquad \rho_l = \lambda_l/\mu_l,\ l = 1, 2. \tag{2.7}$$

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

$$V_t^n = \frac{1}{\sqrt{n}} U_{nt}^n, \quad t \ge 0, \qquad V^n = (V_t^n, t \ge 0). \tag{2.8}$$

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

$$dX_t = [2(d + c(X_t \vee 0)^{1/2}) + \sigma^2]dt + 2\sigma(X_t \vee 0)^{1/2}dW_t, \qquad X_0 \ge 0, \tag{2.9}$$

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

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

where the domain of L is

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

$C_K^2([0, \infty))$ 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 (Ikeda and Watanabe (1989), 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 assumed to have right-continuous with left-hand limits sample paths and considered as random elements of the Skorohod space $D[0, \infty)$ (see, e.g., Jacod and Shiryaev (1987), Liptser and Shiryaev (1989)), 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. Also, $\xrightarrow{P}$ denotes convergence in probability.

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*

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

Theorem 2.1 allows us to get the averaging principle for unfinished work. Let $U_t^{n,l}$, $t \geq 0$, denote the unfinished work at time t at queue $l = 1, 2$, and define $V_t^{n,l} = U_{nt}^{n,l}/\sqrt{n}$, $V^{n,l} = (V_t^{n,l}, t \geq 0)$.

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

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

To conclude this section, it is instructive to compare the result of Theorem 2.1 with a related Bessel process limit obtained by Yamada (1984, Theorem 1). (An example of this type for point processes is considered by Yamada (1986). Rosenkrantz (1984) considers an alternative approach to the problem studied in Yamada (1984).) Note that the process of total unfinished work satisfies the equation

$$U_t^n = U_0^n + S_t^n - \int_0^t 1(U_s^n > 0)\alpha_s^n\,ds, \tag{2.10}$$

where

$$S_t^n = S_t^{n,1} + S_t^{n,2}, \qquad S_t^{n,l} = \sum_{i=1}^{A_t^{n,l}} \eta_i^{n,l}, \quad l = 1, 2, \tag{2.11}$$

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

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

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 (2.10), if $U_s^n > 0$, then the instantaneous rate at which work leaves the system is α_s^n. The heuristic argument of §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

$$\breve{U}_t^n = U_0^n + S_t^n - \int_0^t 1(\breve{U}_s^n > 0)r^n(\breve{U}_s^n)\,ds \tag{2.12}$$

as an approximation for U^n. Equation (2.12) 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 (2.10). Moreover, it is plausible to conjecture that one can weaken the much more restrictive conditions of Yamada's result. Indeed, the techniques developed in the proof of Theorem 2.1 can be applied to prove the following generalization of Yamada's result. In this generalization, we assume that $\check{U}^n = (\check{U}_t^n, t \geq 0)$ is a nonnegative process satisfying (2.12), where $r^n(x)$, $x \geq 0$, is a nonnegative bounded function, not necessarily from (1.3). We further let $\check{V}_t^n = \check{U}_{nt}^n/\sqrt{n}$, $t \geq 0$, $\check{V}^n = (\check{V}_t^n, t \geq 0)$ and $\bar{r}^n = \sup_{x \geq 0} r^n(x)$. The previous notation is preserved.

THEOREM 2.3. *Assume that $r^n(x)$ satisfies the following conditions*:

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

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

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

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) and that we do not assume the condition $\rho^n \leq \bar{r}^n$. In addition, $r^n(x)$ does not have to be nondecreasing, the initial condition U_0^n does not need a second moment and the increments of S_t^n do not need fourth moments.

3. A Threshold Queue. In this section we prove an averaging principle for a single-server queue, called the threshold queue, which is central to our analysis. The threshold queue is basically the standard FIFO single-server queue described in Coffman et al. (1995) except that the threshold operates on the unfinished work, not the queue length. For a given parameter $h \geq 0$, busy periods of the threshold queue begin only when the unfinished work first exceeds h; busy periods terminate in the normal way, when no unfinished work remains. We say that the server switches on when the busy periods begin and switches off when the busy periods end. Those periods during which the server is switched off are called *accumulation* periods; such a period includes the usual idle period plus a period during which arrivals are accumulating in the queue. An accumulation period and its following busy period make up a cycle.

Threshold queues correspond in the obvious way to the queues in our two-queue polling system; for example, the accumulation periods of the threshold queue representing queue 1 correspond to the busy periods of queue 2. In our general approach to the proof of the averaging principle (cf. Theorem 2.2), the time interval $[0, T]$ is divided into subintervals sufficiently small that the total unfinished work in the system remains approximately constant during each. Then, during a subinterval, the behavior of the unfinished work at each queue is approximated by that of a threshold queue. The main result of this section (Theorem 3.1) shows that a threshold queue also obeys an averaging principle; the averaging principle for the polling system is derived as a consequence of the averaging principles for the threshold queues defined for the subintervals.

We use the notation of Coffman et al. (1995). 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

$$(3.1)\qquad \sup_n E(\xi^n)^2 < \infty, \qquad \sup_n E(\eta^n)^2 < \infty,$$

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

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

As in Coffman et al. (1995), within each busy period, at most one of the interarrival periods is allowed to be *exceptional*, i.e., to 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 ith cycle. If there are at least χ_i^n arrivals in the ith busy period, then the (χ_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

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

and that the joint distribution of $\zeta_i^n(r)$, the normal interarrival times, and the service times in the ith 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 $\bar{\xi}_1^n$, may have a distribution different from that of the generic interarrival time, and that

(3.4) $$\frac{\bar{\xi}_1^n}{\sqrt{n}} \xrightarrow{P} 0.$$

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 $X^n(0) = 0$.

The following result is well known and will be used several times in the remainder of the paper (see Iglehart and Whitt (1970) 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 $T > 0$*

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

PROOF. We proceed as in the proof of Theorem 3.1 in Coffman et al. (1995). Define the times

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

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

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

and

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

which immediately give the assertion of the theorem.

For $i\ge 2$, 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 $\{\xi_{i,k}^{n,1}, k\ge 1\}$ and $\{\xi_{i,k}^{n,2}, k\ge 1\}$ 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 $\{\eta_{i,k}^{n,l}, k\ge 1\}$, $i\ge 1$, $l=1,2$, be the 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 $\{\zeta_i^n(r), \xi_{i,k}^{n,l}, \eta_{i,k}^{n,l}, l=1,2, k\ge 1\}$ does not depend on $i=1,2,\dots$.

In a sense, the $\overline{\xi}_i^n$, $i\ge 2$, also represent exceptional interarrival times. By Lemma 3.1 in Coffman et al. (1995), we know that they satisfy conditions similar to those imposed on $\tilde{\xi}_i^n$, i.e.,

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

where

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

Moreover, as $n\to\infty$,

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

Define for $i\ge 1$

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

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

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

For homogeneity of notation, we further set $\overline{\zeta}_1^n(r) = \overline{\xi}_1^n$. As in Coffman et al. (1995), by (3.3) and (3.8), it is enough to prove (3.6) and (3.7) on the events

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

Define the interval lengths

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

so that by (3.5) and (3.10)–(3.12)

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

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)

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

and define as in (3.13) and (3.14)

$$
(3.16)\qquad
\begin{aligned}
\bar{u}_i^n &= \inf\left\{t > 0: \frac{1}{\sqrt{n}} S_i^{n,1}(\bar{A}_i^{n,1}(nt)) > a^n\right\},\\
\underline{u}_i^n &= \inf\left\{t > 0: \frac{1}{\sqrt{n}} S_i^{n,1}(\underline{A}_i^{n,1}(nt)) > a^n\right\},\\
\bar{v}_i^n &= \inf\left\{t > 0: \frac{1}{\sqrt{n}} [nt - S_i^{n,2}(\bar{A}_i^{n,2}(nt))] > \frac{1}{\sqrt{n}} S_i^{n,1}(\bar{A}_i^{n,1}(n\bar{u}_i^n))\right\},\\
\underline{v}_i^n &= \inf\left\{t > 0: \frac{1}{\sqrt{n}} [nt - S_i^{n,2}(\underline{A}_i^{n,2}(nt))] > \frac{1}{\sqrt{n}} S_i^{n,1}(\underline{A}_i^{n,1}(n\underline{u}_i^n))\right\},
\end{aligned}
$$

and

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

Note that since u_i^n, $\bar{u}_i^n$ and $\underline{u}_i^n$ are defined in terms of the same process $S_i^{n,1}$, we actually have $A_i^{n,1}(nu_i^n) = \bar{A}_i^{n,1}(n\bar{u}_i^n) = \underline{A}_i^{n,1}(n\underline{u}_i^n)$ so that

$$
(3.18)\qquad S_i^{n,1}(A_i^{n,1}(nu_i^n)) = S_i^{n,1}(\bar{A}_i^{n,1}(n\bar{u}_i^n)) = S_i^{n,1}(\underline{A}_i^{n,1}(n\underline{u}_i^n)).
$$

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

$$
(3.19)\qquad
\begin{aligned}
\bar{A}_i^{n,1}(t) &\le A_i^{n,1}(t) \le \underline{A}_i^{n,1}(t),\\
\underline{A}_i^{n,2}(t) &\le A_i^{n,2}(t) \le \bar{A}_i^{n,2}(t),
\end{aligned}
$$

on $\Gamma^n(r)$, and hence by (3.13), (3.16) and (3.18), for $1 \le i \le t\sqrt{n}$,

$$
(3.20)\qquad \underline{u}_i^n \le u_i^n \le \bar{u}_i^n, \ \underline{v}_i^n \le v_i^n \le \bar{v}_i^n,
$$

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

$$
(3.21)\qquad \underline{\gamma}_i^n - \underline{\gamma}_{i-1}^n \le \gamma_i^n - \gamma_{i-1}^n \le \bar{\gamma}_i^n - \bar{\gamma}_{i-1}^n,
$$

on $\Gamma^n(r)$.

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

First, note that by (3.15),

$$
\bar{A}_i^{n,1}(t) = \inf\left(k \ge 0 : \bar{\zeta}_i^n(r) + \sum_{j=1}^{k} \xi_{i,j}^{n,1} > t\right),
$$

$$
\bar{A}_i^{n,2}(t) = \inf\left(k \ge 0 : \sum_{j=1}^{k+1} \xi_{i,j}^{n,2} > t\right) + 1.
$$

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

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

and hence, by (3.3), (3.9), (3.12), and Lemma 2.1 in Coffman et al. (1995),

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

By Lemma 2.1 in Coffman et al. (1995) and (3.16), (3.2)

$$\sqrt{n}\bar{u}_i^n\xrightarrow{P}\frac{a\mu}{\lambda}. \tag{3.23}$$

Hence, by (3.22), $S_i^{n,1}(\bar{A}_i^{n,1}(n\bar{u}_i^n))/\sqrt{n}\xrightarrow{P}a$ which gives us by Lemma 2.1 in Coffman et al. (1995), (3.16), and (3.22)

$$\sqrt{n}\bar{v}_i^n\xrightarrow{P}\frac{a\mu}{\mu-\lambda},\qquad i\ge 1. \tag{3.24}$$

Then, by (3.17) and (3.23),

$$\sqrt{n}(\bar{\gamma}_i^n-\bar{\gamma}_{i-1}^n)\xrightarrow{P}a\mu\left(\frac{1}{\lambda}+\frac{1}{\mu-\lambda}\right),\qquad i\ge 1. \tag{3.25}$$

Since

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

and since $\bar{\gamma}_i^n-\bar{\gamma}_{i-1}^n$, $i\ge 1$, are identically distributed by construction, we would have, in view of Lemma 2.4 in Coffman et al. (1995),

$$\bar{\gamma}_{\lfloor\sqrt{n}t\rfloor}^n\xrightarrow{P}a\mu\left(\frac{1}{\lambda}+\frac{1}{\mu-\lambda}\right)t \tag{3.26}$$

provided

$$\lim_{k\to\infty}\overline{\lim_{n\to\infty}}\sqrt{n}P(\sqrt{n}(\bar{\gamma}_1^n-\bar{\gamma}_0^n)>k)=0. \tag{3.27}$$

By (3.17), this would follow from

$$\begin{aligned}&\lim_{k\to\infty}\overline{\lim_{n\to\infty}}\sqrt{n}P(\sqrt{n}\bar{u}_1^n>k)=0,\\ &\lim_{k\to\infty}\overline{\lim_{n\to\infty}}\sqrt{n}P(\sqrt{n}\bar{v}_1^n>k)=0.\end{aligned} \tag{3.28}$$

Consider the first limit. By (3.16)

$$P(\sqrt{n}\bar{u}_1^n > k) = P(S_1^{n,1}(\bar{A}_1^{n,1}(\sqrt{n}k)) \le \sqrt{n}a^n)$$

(3.29)

$$\le P(\bar{A}_1^{n,1}(\sqrt{n}k) \le \tfrac{1}{2}\lambda^n\sqrt{n}k) + P(S_1^{n,1}(\tfrac{1}{2}\lambda^n\sqrt{n}k) \le \sqrt{n}a^n).$$

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

$$\lim_{k\to\infty}\overline{\lim_{n\to\infty}}\sqrt{n}P\left(\left|S_1^{n,l}(k\sqrt{n}) - \frac{k}{\mu^n}\sqrt{n}\right| \ge k\sqrt{n}\epsilon\right) = 0, \qquad l = 1, 2. \tag{3.30}$$

By an analogue of (3.30) for interarrival times, and by (3.15), in analogy with the proof of (3.20) in Coffman et al. (1995),

$$\lim_{k\to\infty}\overline{\lim_{n\to\infty}}\sqrt{n}P(|\bar{A}_1^{n,l}(\sqrt{n}k) - \lambda^n k\sqrt{n}| \ge k\sqrt{n}\epsilon) = 0, \qquad l = 1, 2. \tag{3.31}$$

Relations (3.29)–(3.31) prove the first convergence in (3.28).

For the second convergence, we first prove that

$$\lim_{k\to\infty}\overline{\lim_{n\to\infty}}\sqrt{n}P\left(w_1^n > \frac{1}{3}\left(1 - \frac{\lambda^n}{\mu^n}\right)k\right) = 0, \tag{3.32}$$

where

$$w_1^n = \frac{1}{\sqrt{n}}S_1^{n,1}(\bar{A}_1^{n,1}(n\bar{u}_1^n)) - a^n. \tag{3.33}$$

Note that w_1^n is nonnegative. By the inequality

$$w_1^n \le \frac{1}{\sqrt{n}}\sup_{1\le j\le \bar{A}_1^{n,1}(n\bar{u}_1^n)}\eta_{1,j}^{n,1},$$

we have

$$\sqrt{n}P\left(w_1^n > \frac{1}{3}\left(1 - \frac{\lambda^n}{\mu^n}\right)k\right) \le \sqrt{n}P(\sqrt{n}\bar{u}_1^n > k)$$
$$+ \sqrt{n}P(\bar{A}_1^{n,1}(\sqrt{n}k) > \mu^n\sqrt{n}k) + \mu^n nkP\left(\eta_{1,1}^{n,1} > \frac{1}{3}\left(1 - \frac{\lambda^n}{\mu^n}\right)k\sqrt{n}\right).$$

We have proved that $\lim_{k\to\infty}\overline{\lim_{n\to\infty}}\sqrt{n}P(\sqrt{n}\bar{u}_1^n > k) = 0$; by (3.31), since $\lambda < \mu$, we have $\lim_{k\to\infty}\overline{\lim_{n\to\infty}}\sqrt{n}P(\bar{A}_1^{n,1}(\sqrt{n}k) > \mu^n\sqrt{n}k) = 0$. Finally, by Chebyshev's inequality,

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

and applying (3.1) and (3.2), we arrive at (3.32).

Going back to the proof of the second inequality in (3.28), write, by (3.16) and (3.33), in analogy with (3.29),

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

Putting together (3.30), (3.31), (3.32) and the inequality $\lambda < \mu$ yields the second convergence in (3.28). Thus, (3.26) has been proved. The proof of (3.6) is done.

To prove (3.7) on $\Gamma^n(r)$, we apply Lemma 2.4 in Coffman et al. (1995), i.e., we prove that

$$\begin{aligned}
\lim_{n\to\infty}\frac{1}{\sqrt{n}}\sum_{i=1}^{\lfloor\sqrt{n}t\rfloor} P\bigg(\bigg\{\bigg|\sqrt{n}\int_{\gamma_{i-1}^n}^{\gamma_i^n} f(X^n(s))\,ds\\
- \mu\Big(\frac{1}{\lambda} + \frac{1}{\mu-\lambda}\Big)\int_0^a f(u)\,du\bigg| > \epsilon\bigg\} \cap \Gamma^n(r)\bigg) = 0, \qquad \epsilon > 0,
\end{aligned} \tag{3.34}$$

and

$$\lim_{k\to\infty}\overline{\lim_{n\to\infty}}\sum_{i=1}^{\lfloor\sqrt{n}t\rfloor} P\bigg(\bigg\{\sqrt{n}\bigg|\int_{\gamma_{i-1}^n}^{\gamma_i^n} f(X^n(s))\,ds\bigg| > k\bigg\} \cap \Gamma^n(r)\bigg) = 0. \tag{3.35}$$

Note that (3.35) is easy. For, by the right inequality in (3.21) and the boundedness of f, we have, letting $\|\cdot\|$ denote the sup norm,

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

which tends to 0 as $k\to\infty$ by (3.27) and by the fact that the $(\bar{\gamma}_i^n - \bar{\gamma}_{i-1}^n)$, $i \ge 1$, are identically distributed.

By (3.14), (3.34) would follow if

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

and

$$\lim_{n\to\infty} \frac{1}{\sqrt{n}} \sum_{i=1}^{\lfloor \sqrt{n}t \rfloor} P\left(\left\{ \left| \sqrt{n} \int_0^{v_i^n} f(X^n(s+\beta_i^n))ds - \frac{\mu}{\mu-\lambda} \int_0^a f(u)du \right| > \epsilon \right\} \cap \Gamma^n(r) \right) = 0. \tag{3.36}$$

These limits have similar proofs; we prove only (3.36), which is more difficult.

First, by the second set of inequalities in (3.20) and the fact that $\{\overline{v}_i^n, i \geq 1\}$ and $\{\underline{v}_i^n, i \geq 1\}$ each consist of identically distributed random variables, for $\delta > 0$, $1 \leq i \leq t\sqrt{n}$,

$$P\left(\left\{ \left| \sqrt{n} v_i^n - \frac{a\mu}{\mu-\lambda} \right| > \delta \right\} \cap \Gamma^n(r) \right)$$

$$\leq P\left(\left| \sqrt{n}\overline{v}_1^n - \frac{a\mu}{\mu-\lambda} \right| > \delta \right) + P\left(\left| \sqrt{n}\underline{v}_1^n - \frac{a\mu}{\mu-\lambda} \right| > \delta \right),$$

and hence

$$\begin{aligned} &\overline{\lim_{n\to\infty}} \frac{1}{\sqrt{n}} \sum_{i=1}^{\lfloor \sqrt{n}t \rfloor} P\left(\left\{ \left| \sqrt{n} v_i^n - \frac{a\mu}{\mu-\lambda} \right| > \delta \right\} \cap \Gamma^n(r) \right) \\ &\leq t \overline{\lim_{n\to\infty}} P\left(\left| \sqrt{n}\overline{v}_1^n - \frac{a\mu}{\mu-\lambda} \right| > \delta \right) + t \overline{\lim_{n\to\infty}} P\left(\left| \sqrt{n}\underline{v}_1^n - \frac{a\mu}{\mu-\lambda} \right| > \delta \right) = 0, \end{aligned} \tag{3.37}$$

where the last equality follows by (3.24) and its counterpart for $\underline{v}_1^n$. Next,

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

$$\begin{aligned} &\leq P\left(\left\{ \int_0^{\sqrt{n}v_i^n \wedge a\mu/(\mu-\lambda)} \left| f\left(X^n\left(\frac{u}{\sqrt{n}} + \beta_i^n \right) \right) \right. \right. \right. \\ &\qquad \left. \left. \left. - f\left(a - \left(1 - \frac{\lambda}{\mu} \right) u \right) \right| du > \frac{\epsilon}{2} \right\} \cap \Gamma^n(r) \right) \\ &\quad + P\left(\left\{ \|f\| \cdot \left| \sqrt{n} v_i^n - \frac{a\mu}{\mu-\lambda} \right| > \frac{\epsilon}{2} \right\} \cap \Gamma^n(r) \right). \end{aligned}$$

Sum the second term on the right over $i = 1, \ldots, \lfloor t\sqrt{n} \rfloor$ and divide by $\sqrt{n}$. By (3.37), the result tends to 0 in probability as $n \to \infty$, so the proof of (3.36) will be finished by proving

$$\lim_{n\to\infty}\frac{1}{\sqrt{n}}\sum_{i=1}^{\lfloor\sqrt{n}t\rfloor}P\biggl(\biggl\{\int_0^{\sqrt{n}v_i^n\wedge a\mu/(\mu-\lambda)}\biggl|f\biggl(X^n\biggl(\frac{u}{\sqrt{n}}+\beta_i^n\biggr)\biggr) - f\biggl(a-\biggl(1-\frac{\lambda}{\mu}\biggr)u\biggr)\biggr|du>\frac{\epsilon}{2}\biggr\}\cap\Gamma^n(r)\biggr)=0. \tag{3.38}$$

We prove first that, for $\eta > 0$,

$$\overline{\lim_{n}}\frac{1}{\sqrt{n}}\sum_{i=1}^{\lfloor\sqrt{n}t\rfloor}P\biggl(\biggl\{\sup_{u\le\sqrt{n}v_i^n\wedge a\mu/(\mu-\lambda)}\biggl|X^n\biggl(\frac{u}{\sqrt{n}}+\beta_i^n\biggr) - \biggl(a-\biggl(1-\frac{\lambda}{\mu}\biggr)u\biggr)\biggr|>\eta\biggr\}\cap\Gamma^n(r)\biggr)=0. \tag{3.39}$$

By construction,

$$X^n(u+\beta_i^n)=a^n+w_i^n+\frac{S_i^{n,2}(A_i^{n,2}(nu))-nu}{\sqrt{n}},\qquad u\in[0,v_i^n],$$

where w_i^n is defined in analogy with (3.33), so,

$$\begin{aligned}
&P\biggl(\biggl\{\sup_{u\le\sqrt{n}v_i^n\wedge a\mu/(\mu-\lambda)}\biggl|X^n\biggl(\frac{u}{\sqrt{n}}+\beta_i^n\biggr)-\biggl(a-\biggl(1-\frac{\lambda}{\mu}\biggr)u\biggr)\biggr|>\eta\biggr\}\cap\Gamma^n(r)\biggr)\\
&\quad\le P\biggl(\biggl\{\sup_{u\le a\mu/(\mu-\lambda)}\biggl|\frac{S_i^{n,2}(A_i^{n,2}(\sqrt{n}u))}{\sqrt{n}}-\frac{\lambda}{\mu}u\biggr|>\frac{\eta}{3}\biggr\}\cap\Gamma^n(r)\biggr)\\
&\qquad+P\biggl(w_i^n>\frac{\eta}{3}\biggr)+1\biggl(|a^n-a|>\frac{\eta}{3}\biggr).
\end{aligned}$$

Since the distributions of $(\underline{A}_i^{n,2}(t),\,t\ge0)$, $(\bar{A}_i^{n,2}(t),\,t\ge0)$ and $(S_i^{n,2}(t),\,t\ge0)$ do not depend on i, we conclude from (3.32), (3.2) and (3.19) that the left-hand side of (3.39) is not greater than

$$\begin{aligned}
&t\,\overline{\lim_{n\to\infty}}\,P\biggl(\sup_{u\le a\mu/(\mu-\lambda)}\biggl|\frac{1}{\sqrt{n}}S_1^{n,2}(\bar{A}_1^{n,2}(\sqrt{n}u))-\frac{\lambda}{\mu}u\biggr|>\frac{\eta}{3}\biggr)\\
&\qquad+t\,\overline{\lim_{n\to\infty}}\,P\biggl(\sup_{u\le a\mu/(\mu-\lambda)}\biggl|\frac{1}{\sqrt{n}}S_1^{n,2}(\underline{A}_1^{n,2}(\sqrt{n}u))-\frac{\lambda}{\mu}u\biggr|>\frac{\eta}{3}\biggr)
\end{aligned}$$

which is zero by (3.22) and an analogous relation for $S_1^{n,2}(\underline{A}_1^{n,2}(\sqrt{n}u))/\sqrt{n}$; (3.39) is proved.

Now on the event

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

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

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

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,

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

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

so for η small enough

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

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

and so (3.38) follows from (3.39). Thus (3.36), (3.34) and (3.7) are proved. This completes the proof of the theorem. □

4. An averaging principle for the 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 $V^n \xrightarrow{d} \tilde{V}$ for 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 §4 in Coffman et al. (1995). The main result is the following.

THEOREM 4.1. *Assume that, in addition to the conditions of Theorem* 2.1, $V^n \xrightarrow{d} \tilde{V}$, *where* $\tilde{V} = (\tilde{V}_t,\ t \ge 0)$ *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_+,

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

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

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

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 Coffman et al. (1995). So, we prove (4.1) assuming (4.2). The idea of the proof is the same as in Coffman et al. (1995). Note that if considered in isolation an individual queue in our polling system passes through alternating periods of accumulating and serving requests, thus its behavior resembles the behavior of the threshold queue above, the distinction being that here the threshold is a random process: the queue starts being served when the unfinished work at this queue becomes equal to the total amount of the unfinished work in the system. According to the assumptions of the theorem, the (properly normalized and time-scaled) process of the total unfinished work is a continuous process in the limit. Therefore, we can divide the time axis into (random) intervals small enough for the total unfinished work during an interval to be close to a constant. Then during such an interval an individual-queue unfinished work is well approximated by the unfinished work in a threshold queue with the associated constant as a threshold. The proof of the theorem implements this program.

As in Coffman et al. (1995), choose $\epsilon \in (0, \delta/2)$ such that $N = (K - \delta)/\epsilon$ is an integer and, given $r(\epsilon) < \epsilon/2$, let, for $0 \le i \le N$,

$$a_i(\epsilon) = \delta + i\epsilon,$$

$$B_{r(\epsilon)}(\epsilon, i) = (a_i(\epsilon) - r(\epsilon), a_i(\epsilon) + r(\epsilon)),$$

$$C_{r(\epsilon)}(\epsilon, i) = (0, a_i(\epsilon) - \epsilon + r(\epsilon)) \cup (a_i(\epsilon) + \epsilon - r(\epsilon), \infty),$$

$$\zeta_0^n(\epsilon, i) = 0,$$

$$\tau_k^n(\epsilon, i) = \inf(t > \zeta_{k-1}^n(\epsilon, i) : V_t^n \in B_{r(\epsilon)}(\epsilon, i)), \qquad k \ge 1,$$

$$\zeta_k^n(\epsilon, i) = \inf(t > \tau_k^n(\epsilon, i) : V_t^n \in C_{r(\epsilon)}(\epsilon, i)), \qquad k \ge 1,$$

$$\zeta_0(\epsilon, i) = 0,$$

$$\tau_k(\epsilon, i) = \inf(t > \zeta_{k-1}(\epsilon, i) : \tilde{V}_t \in B_{r(\epsilon)}(\epsilon, i)), \qquad k \ge 1,$$

$$\zeta_k(\epsilon, i) = \inf(t > \tau_k(\epsilon, i) : \tilde{V}_t \in C_{r(\epsilon)}(\epsilon, i)), \qquad k \ge 1.$$

Thus, $[\tau_k^n(\epsilon, i), \zeta_k^n(\epsilon, i)]$ are intervals during which V_t^n ‘‘does not vary too much.’’

For the sequel, we note that, since $\tilde{V}$ is continuous, the argument of the proof of Lemma 4.1 in Coffman et al. (1995) applies to V^n and $\tilde{V}$ to give that

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

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

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

$$(4.4)\quad \begin{aligned}&(V^n, (\tau_k^n(\epsilon, i) \wedge T, \zeta_k^n(\epsilon, i) \wedge T)_{k\ge 1, 0\le i\le N})\\ &\qquad \xrightarrow{d} (\tilde{V}, (\tau_k(\epsilon, i) \wedge T, \zeta_k(\epsilon, i) \wedge T)_{k\ge 1, 0\le i\le N}),\end{aligned}$$

where convergence in distribution is in $D[0, \infty) \times R^\infty$ (Billingsley (1968)).

Let $n\kappa_j^{n,1}(i, k)$, $j \ge 0$, denote the successive times after $n\tau_k^n(\epsilon, i)$, when the unfinished work at queue 1 becomes equal to 0. These are times when switchovers from queue 1 to queue 2 start. We also let $\vartheta_{i,k}^{n,1}$ denote the number of accumulation-service cycles for queue 1 in $[\tau_k^n(\epsilon, i) \wedge T, \zeta_k^n(\epsilon, i) \wedge T]$, i.e.,

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

We now define two threshold queues approximating queue 1 on $[\tau_k^n(\epsilon, i) \wedge T, \zeta_k^n(\epsilon, i) \wedge T]$ whose unfinished work processes bound the unfinished work process of queue 1 from below and from above, respectively. We begin by introducing a threshold queue associated with queue 1 on $[\tau_k^n(\epsilon, i) \wedge T, \zeta_k^n(\epsilon, i) \wedge T]$.

Fixing i and k, we denote $\kappa_j^n = \kappa_j^{n,1}(i, k)$ and let $n\theta_j^n$, $j \ge 1$, denote the successive times after $n\kappa_0^n$ when the server starts serving queue 1; obviously, $\kappa_{j-1}^n < \theta_j^n < \kappa_j^n$, $j \ge 1$. Let the arrivals to queue 1 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 $(l-1)$th and lth of these arrivals. Obviously, $\{\tilde{\xi}_l^n, l \ge 2\}$ is a sequence of i.i.d. random variables with the distribution of the generic interarrival time for queue 1. Introduce independent replicas $\{\xi_{j,l}^{n,m}, l \ge 1\}$, $j \ge 1$, $m = 1, 2$, of the interarrival time sequence at queue 1 and independent replicas $\{\eta_{j,l}^{n,m}, l \ge 1\}$, $j \ge 1$, $m = 1, 2$, of the service time sequence at queue 1.

Given $h > 0$, let

$$\varphi_j^n = \inf(t > \kappa_{j-1}^n : V_t^{n,1} > h) \wedge \theta_j^n, \qquad j \ge 1,$$

$$\psi_j^n = \inf(t > \theta_j^n : V_t^{n,1} \le V_{\varphi_j^n}^{n,1}), \qquad j \ge 1.$$

Note that if $V_{\theta_j^n}^{n,1} \le h$, then $\varphi_j^n = \psi_j^n = \theta_j^n$. Let χ_j^n, $j \ge 1$, index the last arrival to queue 1 occurring no later than at $n\psi_j^n$ and let $\overline{\upsilon}_j^n$, $j \ge 1$, denote the time between $n\psi_j^n$ and the $(\chi_j^n + 1)$th arrival. By definition, $\overline{\upsilon}_j^n \le \overline{\xi}_{\chi_j^n+1}^n$.

Construct as follows a threshold queue with the threshold $h^n = \sqrt{n}h$. In the first cycle the interarrival times in the accumulation period are taken from the sequence $\{\tilde{\xi}_1^n, \tilde{\xi}_2^n, \dots, \tilde{\xi}_{\chi_1^n}^n, \tilde{\xi}_{\chi_1^n+1}^n, \xi_{1,1}^{n,1}, \xi_{1,2}^{n,1}, \cdots\}$. The associated service times for arrivals 1 through χ_1^n are those of corresponding arrivals to queue 1 and the subsequent service times are $\eta_{1,1}^{n,1}$, $\eta_{1,2}^{n,1}, \dots$.

Denoting the threshold queue normalized and time-scaled unfinished work at t by $\overline{V}_t^{n,1}$, define

$$\beta_1^n = \inf(t > 0 : \overline{V}_t^{n,1} > h).$$

Then $n\beta_1^n$ ends the first accumulation period. Obviously, $\beta_1^n = \varphi_1^n - \kappa_0^n$ if $V_{\varphi_1^n}^{n,1} > h$ which happens if $V_{\theta_1^n}^{n,1} > h$, in this case $\tilde{\xi}_{\chi_1^n+1}^n, \xi_{1,1}^{n,1}, \xi_{1,2}^{n,1}, \cdots$ are not actually used as interarrival times and $\eta_{1,1}^{n,1}, \eta_{1,2}^{n,1}, \cdots$ are not used as service times. At $n\beta_1^n$ service switches on. If $V_{\varphi_1^n}^{n,1} \le h$, which happens if $V_{\theta_1^n}^{n,1} \le h$, then interarrival times after $n\beta_1^n$ are taken from

$\{\xi^{n,2}_{1,l},\ l \geq 1\}$ and service times after $n\beta^n_1$ are taken from $\{\eta^{n,2}_{1,l},\ l \geq 1\}$ until time $n\phi^n_1$, where

$$\phi^n_1 = \inf(t > \beta^n_1 : \bar{V}^{n,1}_t \leq V^{n,1}_{\varphi^n_1}).$$

We then take $\xi^{n,2}_{1,\bar{\chi}^n_1}$ to be the last random variable in the sequence $\{\xi^{n,2}_{1,1}, \xi^{n,2}_{1,2}, \cdots\}$ that is actually realized as an interarrival time in $[n\beta^n_1, n\phi^n_1]$. In the case that $V^{n,1}_{\varphi^n_1} > h$, we define $\phi^n_1 = \beta^n_1$ and set $\bar{\chi}^n_1 = \xi^{n,2}_{1,\bar{\chi}^n_1} = 0$. In both cases, the first arrival after $n\phi^n_1$ is made to occur at time $n\phi^n_1 + \bar{v}^n_1$ and bring the same service time as the $(\chi^n_j + 1)$th arrival in queue 1 so that its interarrival time v^n_1 satisfies the inequality $v^n_1 \leq \bar{v}^n_1 + \xi^{n,2}_{1,\bar{\chi}^n_1+1} \leq \tilde{\xi}^n_{\chi^n_1+1} + \xi^{n,2}_{1,\bar{\chi}^n_1+1}$. The subsequent interarrival times are $\tilde{\xi}^n_{\chi^n_1+2}, \ldots, \tilde{\xi}^n_{\chi^n_2}$, and the service times are the same as for arrivals $\chi^n_1 + 2, \ldots, \chi^n_2$ in queue 1, where χ^n_2 is the index of the last arrival in queue 1 occurring no later than at $n\psi^n_2$. Arrivals $\chi^n_2 + 1, \chi^n_2 + 2, \cdots$ have the interarrival times $\tilde{\xi}^n_{\chi^n_2+1}, \xi^{n,1}_{2,1}, \xi^{n,1}_{2,2}, \cdots$ and the service times $\eta^{n,1}_{2,1}, \eta^{n,1}_{2,2}, \cdots$ until (and unless, i.e., these times are not used if $V^{n,1}_{\varphi^n_2} > h$) the threshold has been exceeded. After this has happened at $n\beta^n_2$, where

$$\beta^n_2 = \inf(t > \gamma^n_1 : \bar{V}^{n,1}_t > h),$$

and

$$\gamma^n_1 = \inf(t > \beta^n_1 : \bar{V}^{n,1}_t = 0),$$

and until $n\phi^n_2$, where

$$\phi^n_2 = \inf(t > \beta^n_2 : \bar{V}^{n,1}_t \leq V^{n,1}_{\varphi^n_2}),$$

the interarrival times are $\xi^{n,2}_{2,1}, \xi^{n,2}_{2,2}, \cdots$ and the service times are $\eta^{n,2}_{2,1}, \eta^{n,2}_{2,2}, \cdots$ (as above these are not used if $V^{n,1}_{\varphi^n_2} > h$ and hence $\phi^n_2 = \beta^n_2$). After $n\phi^n_2$, the next arrival brings the same service time as the arrival in the original queue terminating the interarrival time $\tilde{\xi}^n_{\chi^n_2+1}$ and occurs at time $n\phi^n_2 + \bar{v}^n_2$ (in both cases $V^{n,1}_{\varphi^n_2} > h$ and $V^{n,1}_{\varphi^n_2} \leq h$) so that its interarrival time satisfies $v^n_2 \leq \bar{v}^n_2 + \xi^{n,2}_{2,\bar{\chi}^n_2+1} \leq \tilde{\xi}^n_{\chi^n_2+1} + \xi^{n,2}_{2,\bar{\chi}^n_2+1}$, where $\xi^{n,2}_{2,\bar{\chi}^n_2}$ denotes the last random variable from $\{\xi^{n,2}_{2,1}, \xi^{n,2}_{2,2}, \cdots\}$ that is realized as an interarrival time in $[n\beta^n_2, n\phi^n_2]$ (again $\bar{\chi}^n_2 = \xi^{n,2}_{2,\bar{\chi}^n_2} = 0$ if $V^{n,1}_{\varphi^n_2} > h$). The subsequent interarrival times are $\tilde{\xi}^n_{\chi^n_2+2}, \cdots$ and the service times replicate those of queue 1 until the unfinished work hits 0 after which the cycle resumes.

That this is indeed a threshold queue with generic interarrival and service times distributed as in the original queue follows by Lemma 4.2 in Coffman et al. (1995). The exceptional arrivals, if any, are the ones occurring at $n\phi^n_j + \bar{v}^n_j$, $j \geq 1$. Note also that if $V^{n,1}_{\theta^n_j} \leq h$, then $\tilde{\xi}^n_{\chi^n_j+1}$ is used for constructing interarrival sequences in both accumulation and busy periods so that these sequences, generally, are dependent. This explains why we emphasised this assumption in Theorem 3.1.

We now check that $\bar{V}^{n,1}$ satisfies the conditions of Theorem 3.1. We need focus only on the part related to exceptional interarrival times and the time of the first arrival. Define

$$\zeta^n_j(r) = \max_{1\leq k\leq \lfloor \sqrt{n}r \rfloor} \tilde{\xi}^n_{\tilde{\chi}^n_j+k} + \max_{1\leq k\leq \lfloor \sqrt{n}r \rfloor} \xi^{n,2}_{j,k},$$

where $\tilde{\chi}^n_j$ indexes the first arrival in the original queue after κ^n_{j-1}. Noting that $\{\tilde{\xi}^n_{\chi^n_j+k}, k \geq 1\}$ is distributed as $\{\tilde{\xi}^n_k, k \geq 1\}$ and that $v^n_j \leq \tilde{\xi}^n_{\chi^n_j+1} + \xi^{n,2}_{j,\bar{\chi}^n_j+1}$, one can prove in analogy with Lemma 3.1 in Coffman et al. (1995) that the sequence $\{v^n_j, \zeta^n_j(r), j \geq 1\}$ satisfies the conditions of Theorem 3.1. Informally, this follows by the fact that since the number of arrivals in an accumulation-service cycle is of order $\sqrt{n}$, ''with high probability'' the exceptional arrival is among the first $\sqrt{n}r$ arrivals when r is ''big,'' similarly, $\xi^{n,2}_{j,\chi^n_j+1}$ belongs to the set $\{\xi^{n,2}_{j,1},\dots,\xi^{n,2}_{j,\sqrt{n}r}\}$ ''with high probability'' for r ''big.'' By a similar argument, $\tilde{\xi}^n_1/\sqrt{n}\xrightarrow{P}0$. Thus, the conditions of Theorem 3.1 hold for the associated threshold queue.

We now define a process $\tilde{V}^{n,1}$ as the process $\bar{V}^{n,1}$ corresponding to the threshold $h = a_i(\epsilon)-\epsilon$ and a process $\hat{V}^{n,1}$ as the process $\bar{V}^{n,1}$ corresponding to the threshold $h = a_i(\epsilon)+\epsilon$, and check that they represent a lower and upper bound, respectively, for $V^{n,1}$.

Since on the interval $[\tau^n_k(\epsilon,i),\zeta^n_k(\epsilon,i))$, the process V^n stays in the strip $(a_i(\epsilon)-\epsilon+r(\epsilon), a_i(\epsilon)+\epsilon-r(\epsilon))$, the process $V^{n,1}$ up-crosses the level $a_i(\epsilon)-\epsilon$ in every interval $[\kappa^n_{j-1},\kappa^n_j]$ belonging to $[\tau^n_k(\epsilon,i)\wedge T,\zeta^n_k(\epsilon,i)\wedge T]$, so the construction above yields

$$\tilde{V}^{n,1}_t=\begin{cases} V^{n,1}_{t-\tilde{\gamma}^n_{j-1}+\kappa^n_{j-1}}, & t\in[\tilde{\gamma}^n_{j-1},\tilde{\beta}^n_j),\quad 1\le j\le \vartheta^{n,1}_{i,k},\\ V^{n,1}_{t-\tilde{\beta}^n_j+\psi^n_j}, & t\in[\tilde{\beta}^n_j,\tilde{\gamma}^n_j),\quad 1\le j\le \vartheta^{n,1}_{i,k},\end{cases}$$

where, in analogy with the proof of Theorem 3.1, $n\tilde{\beta}^n_j$ denotes the jth time a busy period for $\tilde{V}^{n,1}$ starts and $n\tilde{\gamma}^n_j$ denotes the jth time when the queue empties. By the fact that f is nonnegative, we then get

$$\int_0^{\tilde{\vartheta}^n} f(\tilde{V}^{n,1}_t)\,dt\le\int_{\tau^n_k(\epsilon,i)\wedge T}^{\zeta^n_k(\epsilon,i)\wedge T} f(V^{n,1}_t)\,dt, \tag{4.5}$$

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

Similarly, since V^n does not exceed $a_i(\epsilon)+\epsilon$ on $[\tau^n_k(\epsilon,i)\wedge T,\zeta^n_k(\epsilon,i)\wedge T]$, the construction above implies that

$$V^{n,1}_t=\begin{cases} \hat{V}^{n,1}_{t-\kappa^n_{j-1}+\hat{\gamma}^n_{j-1}}, & t\in[\kappa^n_{j-1},\varphi^n_j),\quad 1\le j\le \vartheta^{n,1}_{i,k},\\ \hat{V}^{n,1}_{t-\varphi^n_j+\hat{\phi}^n_j}, & t\in[\varphi^n_j,\kappa^n_j),\quad 1\le j\le \vartheta^{n,1}_{i,k},\end{cases}$$

where $\hat{\phi}^n_j$ is defined as ϕ^n_j above corresponding to $h=a_i(\epsilon)+\epsilon$ and $n\hat{\gamma}^n_j$ is the jth time when the queue empties. Again, since f is nonnegative,

$$\begin{aligned}\int_{\tau^n_k(\epsilon,i)\wedge T}^{\zeta^n_k(\epsilon,i)\wedge T} f(V^{n,1}_t)\,dt &\le \int_{\tau^n_k(\epsilon,i)\wedge T}^{\kappa^n_0\wedge T} f(V^{n,1}_t)\,dt\\ &\quad+\int_{\kappa^n_{\vartheta^{n,1}_{i,k}}\wedge T}^{\zeta^n_k(\epsilon,i)\wedge T} f(V^{n,1}_t)\,dt+\int_0^{\hat{\vartheta}^n} f(\hat{V}^{n,1}_t)\,dt,\end{aligned}\tag{4.6}$$

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

Thus, recalling that $\|f\|=\sup_x f(x)$, by (4.5) and (4.6) we have the bounds

$$\begin{aligned}\int_0^{\tilde{\vartheta}^n} f(\tilde{V}^{n,1}_t)\,dt &\le \int_{\tau^n_k(\epsilon,i)\wedge T}^{\zeta^n_k(\epsilon,i)\wedge T} f(V^{n,1}_t)\,dt\\ &\le\int_0^{\hat{\vartheta}^n} f(\hat{V}^{n,1}_t)\,dt+[\kappa^n_0\wedge T-\tau^n_k(\epsilon,i)\wedge T]\|f\|\\ &\quad+[\zeta^n_k(\epsilon,i)\wedge T-\kappa^n_{\vartheta^{n,1}_{i,k}}\wedge T]\|f\|.\end{aligned}\tag{4.7}$$

We next apply Theorem 3.1 to $\tilde{V}^{n,1}$ and $\hat{V}^{n,1}$ to get the asymptotics of the bounds on the right and on the left.

Define

$$\tilde{\vartheta}^{n,1}_{i,k} = \min(j : \tilde{\gamma}^n_{j+1} > \zeta^n_k(\epsilon, i) \wedge T - \tau^n_k(\epsilon, i) \wedge T), \tag{4.8}$$

$$\hat{\vartheta}^{n,1}_{i,k} = \min(j : \hat{\gamma}^n_{j+1} > \zeta^n_k(\epsilon, i) \wedge T - \tau^n_k(\epsilon, i) \wedge T). \tag{4.9}$$

Obviously,

$$\hat{\vartheta}^{n,1}_{i,k} \le \vartheta^{n,1}_{i,k} \le \tilde{\vartheta}^{n,1}_{i,k}. \tag{4.10}$$

Let $U^n_k(\epsilon, i)$ and $V^n_k(\epsilon, i)$ denote respectively the lower bound in (4.7) with $\tilde{\vartheta}^n$ $(=\tilde{\gamma}^n_{\vartheta^{n,1}_{i,k}})$ changed to $\tilde{\omega}^n = \tilde{\gamma}^n_{\tilde{\vartheta}^{n,1}_{i,k}}$, and the upper bound in (4.7) with $\hat{\vartheta}^n$ $(=\hat{\gamma}^n_{\vartheta^{n,1}_{i,k}})$ changed to $\hat{\omega}^n = \hat{\gamma}^n_{\hat{\vartheta}^{n,1}_{i,k}}$. By (4.10),

$$U^n_k(\epsilon, i) \le \int_{\tau^n_k(\epsilon,i)\wedge T}^{\zeta^n_k(\epsilon,i)\wedge T} f(V^{n,1}_t)\,dt \le V^n_k(\epsilon, i). \tag{4.11}$$

We now show that, as $n \to \infty$,

$$U^n_k(\epsilon, i) \xrightarrow{d} U_k(\epsilon, i), \quad V^n_k(\epsilon, i) \xrightarrow{d} V_k(\epsilon, i), \quad k \ge 1,\ 0 \le i \le N, \tag{4.12}$$

where

$$\begin{aligned} U_k(\epsilon, i) &= \frac{a_i(\epsilon) - \epsilon}{a_i(\epsilon) + \epsilon}(\zeta_k(\epsilon, i) \wedge T - \tau_k(\epsilon, i) \wedge T)\int_0^1 f(u(a_i(\epsilon) - \epsilon))\,du, \\ V_k(\epsilon, i) &= \frac{a_i(\epsilon) + \epsilon}{a_i(\epsilon) - \epsilon}(\zeta_k(\epsilon, i) \wedge T - \tau_k(\epsilon, i) \wedge T)\int_0^1 f(u(a_i(\epsilon) + \epsilon))\,du, \end{aligned} \tag{4.13}$$

Let

$$\tilde{\vartheta}^{n,1}_{i,k}(t) = \min(j \ge 0 : \tilde{\gamma}^n_{j+1} > t) \tag{4.14}$$

and

$$\hat{\vartheta}^{n,1}_{i,k}(t) = \min(j \ge 0 : \hat{\gamma}^n_{j+1} > t). \tag{4.15}$$

Note that

$$\tilde{\vartheta}^{n,1}_{i,k} = \tilde{\vartheta}^{n,1}_{i,k}(\zeta^n_k(\epsilon, i) \wedge T - \tau^n_k(\epsilon, i) \wedge T), \tag{4.16}$$

$$\hat{\vartheta}^{n,1}_{i,k} = \hat{\vartheta}^{n,1}_{i,k}(\epsilon, i) \wedge T - \tau^n_k(\epsilon, i) \wedge T). \tag{4.17}$$

In the course of proving Theorem 3.1 we established (3.6). Since $\tilde{V}^{n,1}$ and $\hat{V}^{n,1}$ meet the conditions of Theorem 3.1 and the $\tilde{\gamma}^n_j$ and $\hat{\gamma}^n_j$ are analogues of the γ^n_j from (3.5), we can write for these processes, in analogy with (3.6),

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

By Lemma 2.1 in Coffman et al. (1995) and (4.14), (4.15),

$$\begin{aligned}\frac{\tilde{\vartheta}^{n,1}_{i,k}(t)}{\sqrt{n}} &\xrightarrow{P} \frac{t}{\mu_1(a_i(\epsilon)-\epsilon)}\left(\frac{1}{\lambda_1}+\frac{1}{\mu_1-\lambda_1}\right)^{-1},\\ \frac{\hat{\vartheta}^{n,1}_{i,k}(t)}{\sqrt{n}} &\xrightarrow{P} \frac{t}{\mu_1(a_i(\epsilon)+\epsilon)}\left(\frac{1}{\lambda_1}+\frac{1}{\mu_1-\lambda_1}\right)^{-1}.\end{aligned} \tag{4.18}$$

Lemma 2.2 in Coffman et al. (1995) then yields

$$\tilde{\gamma}^n_{\tilde{\vartheta}^{n,1}_{i,k}(t)} \xrightarrow{P} \frac{a_i(\epsilon)-\epsilon}{a_i(\epsilon)+\epsilon}\,t, \qquad \hat{\gamma}^n_{\hat{\vartheta}^{n,1}_{i,k}(t)} \xrightarrow{P} \frac{a_i(\epsilon)+\epsilon}{a_i(\epsilon)-\epsilon}\,t\,. \tag{4.19}$$

By Theorem 3.1 applied to $\tilde{V}^{n,1}$ and $\hat{V}^{n,1}$ and (4.19),

$$\begin{aligned}\int_0^{\tilde{\gamma}^n_{\tilde{\vartheta}^{n,1}_{i,k}(t)}} f(\tilde{V}^{n,1}_s)\,ds &\xrightarrow{P} \frac{a_i(\epsilon)-\epsilon}{a_i(\epsilon)+\epsilon}\,t\int_0^1 f(u(a_i(\epsilon)-\epsilon))\,du,\\ \int_0^{\hat{\gamma}^n_{\hat{\vartheta}^{n,1}_{i,k}(t)}} f(\hat{V}^{n,1}_s)\,ds &\xrightarrow{P} \frac{a_i(\epsilon)+\epsilon}{a_i(\epsilon)-\epsilon}\,t\int_0^1 f(u(a_i(\epsilon)+\epsilon))\,du.\end{aligned} \tag{4.20}$$

In view of (4.16), Lemma 2.2 in Coffman et al. (1995) shows that (4.4) and (4.20) imply

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

$$\int_0^{\hat{\omega}^n} f(\hat{V}^{n,1}_t)\,dt \xrightarrow{d} \frac{a_i(\epsilon)+\epsilon}{a_i(\epsilon)-\epsilon}\,(\zeta_k(\epsilon,i)\wedge T-\tau_k(\epsilon,i)\wedge T)\int_0^1 f(u(a_i(\epsilon)+\epsilon))\,du.$$

Since $|\kappa^n_0\wedge T-\tau^n_k(\epsilon,i)\wedge T| \xrightarrow{P} 0$ and $|\zeta^n_k(\epsilon,i)\wedge T-\kappa^n_{\vartheta^{n,1}_{i,k}}\wedge T| \xrightarrow{P} 0$ obviously hold, (4.12) is proved. Moreover, the same argument shows that

$$\begin{aligned}(V^n,(U^n_k(\epsilon,i))_{k\ge1,0\le i\le N}) &\xrightarrow{d} (\tilde{V},(U_k(\epsilon,i))_{k\ge1,0\le i\le N}),\\ (V^n,(V^n_k(\epsilon,i))_{k\ge1,0\le i\le N}) &\xrightarrow{d} (\tilde{V},(V_k(\epsilon,i))_{k\ge1,0\le i\le N}).\end{aligned} \tag{4.21}$$

Next, defining

$$U^n(\epsilon)=\sum_{k=1}^{\infty}\sum_{i=0}^{N} U^n_k(\epsilon,i), \qquad V^n(\epsilon)=\sum_{k=1}^{\infty}\sum_{i=0}^{N} V^n_k(\epsilon,i), \tag{4.22}$$

we need to prove that

$$(V^n,U^n(\epsilon)) \xrightarrow{d} (\tilde{V},U(\epsilon)), \qquad (V^n,V^n(\epsilon)) \xrightarrow{d} (\tilde{V},V(\epsilon)), \tag{4.23}$$

where

$$U(\epsilon) = \sum_{k=1}^{\infty} \sum_{i=0}^{N} U_k(\epsilon, i), \qquad V(\epsilon) = \sum_{k=1}^{\infty} \sum_{i=0}^{N} V_k(\epsilon, i). \tag{4.24}$$

We prove the first convergence result in (4.23); 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),

$$\begin{aligned} &\overline{\lim_{M\to\infty}}\ \overline{\lim_{n\to\infty}}\ P\left(\left| \sum_{k=1}^{M} \sum_{i=0}^{N} U_k^n(\epsilon, i) - U^n(\epsilon) \right| > 0 \right) \\ &\leq \overline{\lim_{M\to\infty}}\ \overline{\lim_{n\to\infty}}\ P\big(\min_{0\le i\le N} \zeta_M^n(\epsilon, i) \wedge (T+1) < T \big) \\ &\leq \overline{\lim_{M\to\infty}}\ P\big(\min_{0\le i\le N} \zeta_M(\epsilon, i) \wedge (T+1) \leq T \big) = 0. \end{aligned} \tag{4.25}$$

Analogously,

$$\sum_{k=1}^{M} \sum_{i=0}^{N} U_k(\epsilon, i) \xrightarrow{P} U(\epsilon) \qquad (M \to \infty). \tag{4.26}$$

Next, by (4.21) and the continuous mapping theorem, we have

$$\left(V^n, \sum_{k=1}^{M} \sum_{i=0}^{N} U_k^n(\epsilon, i) \right) \xrightarrow{d} \left(\tilde{V}, \sum_{k=1}^{M} \sum_{i=0}^{N} U_k(\epsilon, i) \right). \tag{4.27}$$

The convergence $(V^n, U^n(\epsilon)) \xrightarrow{d} (\tilde{V}, U(\epsilon))$ then follows from (4.24)–(4.27) and Theorem 4.2 in (Billingsley (1968)).

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

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

so by (4.11), we obtain from (4.22)

$$\begin{aligned} &U^n(\epsilon) - \|f\| \int_0^T [1(\delta - \epsilon \leq V_t^n \leq \delta) + 1(K \leq V_t^n \leq K + \epsilon)]\,dt \\ &\leq \int_0^T f(V_t^{n,1}) \cdot 1(\delta \leq V_t^n \leq K)\,dt \\ &\leq V^n(\epsilon) + \|f\| \int_0^T [1(\delta - \epsilon \leq V_t^n \leq \delta) + 1(K \leq V_t^n \leq K + \epsilon)]\,dt. \end{aligned} \tag{4.28}$$

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

$$U(\epsilon) \xrightarrow{d} \int_0^T \left(\int_0^1 f(u\tilde{V}_t)\,du \right) 1(\delta \le \tilde{V}_t \le K)\,dt,$$

(4.29)

$$V(\epsilon) \xrightarrow{d} \int_0^T \left(\int_0^1 f(u\tilde{V}_t)\,du \right) 1(\delta \le \tilde{V}_t \le K)\,dt,$$

then by applying Lemma 2.3 in Coffman et al. (1995) to (4.29) and taking into account (4.23), (4.2), we will then obtain (4.1). As before, we prove only the first of the results in (4.29); the proof to the second is similar.

In fact, we prove convergence with probability 1. The argument is almost identical to that in Coffman et al. (1995), but we give it here since it is used once again below. Since $a_i(\epsilon) > \delta$, we have from (4.13) and (4.24)

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

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

$$\lim_{\epsilon \to 0} \sum_{k=1}^{\infty} \sum_{i=0}^{N} [\zeta_k(\epsilon, i) \wedge T - \tau_k(\epsilon, i) \wedge T] \int_0^1 f(u(a_i(\epsilon) - \epsilon))\,du$$

(4.30)

$$= \int_0^T \left(\int_0^1 f(u\tilde{V}_t)\,du \right) 1(\delta \le \tilde{V}_t \le K)\,dt.$$

We can write

$$\sum_{k=1}^{\infty} \sum_{i=0}^{N} [\zeta_k(\epsilon, i) \wedge T - \tau_k(\epsilon, i) \wedge T] \int_0^1 f(u(a_i(\epsilon) - \epsilon))\,du$$

(4.31)

$$= \sum_{k=1}^{\infty} \sum_{i=0}^{N} \int_0^T \left(\int_0^1 f(u(a_i(\epsilon) - \epsilon))\,du \right) \cdot 1(\tau_k(\epsilon, i) \le t < \zeta_k(\epsilon, i))\,dt$$

$$\equiv C_\epsilon.$$

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

$$\left| \int_0^1 f(ux)\,du - \int_0^1 f(uy)\,du \right| = \left| \frac{1}{x} \int_0^x f(u)\,du - \frac{1}{y} \int_0^y f(u)\,du \right|$$

$$\le \left| \frac{1}{x} - \frac{1}{y} \right| \int_0^x f(u)\,du + \frac{1}{y} \left| \int_x^y f(u)\,du \right| \le \frac{8\epsilon}{\delta} \|f\|.$$

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

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

so by (4.31)

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

whence

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

Since by (4.2) the right-hand side of this inequality tends to 0 as $\epsilon \to 0$, we have proved (4.30). This completes the proof of the first assertion of Theorem 4.1 for bounded non-negative $f(x)$. The general case is handled via a localization argument as in the proof of Theorem 2.1 in Coffman et al. (1995).

5. Tightness results. The purpose of this section is to prove several results on the tightness of some processes closely related to the normalized and time-scaled unfinished-work process V^n defined in (2.8). We start, however, with a simple fact.

Introduce

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

where $S_t^{n,l}$ and S_t^n were defined in (2.11), and let $B^{n,l} = (B_t^{n,l}, t \geq 0)$, $l = 1, 2$, $B^n = (B_t^n, t \geq 0)$.

LEMMA 5.1. *As $n \to \infty$,*

$$\frac{1}{\sqrt{n}} \sum_{i=1}^{\lfloor \sqrt{n}t \rfloor} s_i^{n,l} \xrightarrow{P} d_l t, \quad \frac{1}{n} \sum_{i=1}^{\lfloor nt \rfloor} s_i^{n,l} \xrightarrow{P} d_l t, \quad \frac{A_{nt}^{n,l}}{n} \xrightarrow{P} \lambda_l t, \quad l = 1, 2,$$

$$B^{n,l} \xrightarrow{d} \lambda_l^{1/2} \sigma_l W^l, \quad l = 1, 2, \quad B^n \xrightarrow{d} (\sigma W_t + ct, t \geq 0),$$

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

PROOF. We give proofs of the first and second convergence results in the first line. Since, for $\epsilon > 0$, Var $s_1^{n,l} \le E(s_1^{n,l})^2 1(s_1^{n,l} > \epsilon\sqrt{n}) + \epsilon\sqrt{n} d_l^n$, it follows by (2.2) and (2.6) that Var $s_1^{n,l}/\sqrt{n} \to 0$ as $n \to \infty$. Hence,

$$\frac{1}{\sqrt{n}} \sum_{i=1}^{\lfloor \sqrt{n} t \rfloor} (s_i^{n,l} - d_l^n) \xrightarrow{P} 0, \qquad \frac{1}{n} \sum_{i=1}^{\lfloor nt \rfloor} (s_i^{n,l} - d_l^n) \xrightarrow{P} 0,$$

and both convergences follow by (2.2).

The third convergence follows from the convergence $1/n \sum_{i=1}^{\lfloor nt \rfloor} \xi_i^{n,l} \xrightarrow{P} \lambda_l^{-1} t$, which is proved similarly, and properties of the first-passage-time map (Whitt 1980; see also Coffman et al. 1995, Lemma 2.1). The claimed convergences in distribution hold by the Donsker-Prohorov invariance principle and (2.3). □

Let β_t^n denote the indicator of the event that the server is switching over at nt, i.e., $\beta_t^n = 1 - \alpha_{nt}^n$. Then definitions (2.8), (2.11), (5.1) and Equation (2.10) 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 $K^{n,\epsilon} = (K_t^{n,\epsilon}, t \ge 0)$ by

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

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 $\{X^n, n \ge 1\}$ 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 (Jacod and Shiryaev (1987), VI.3.25). Below, we repeatedly use the fact that $\{X^n, n \ge 1\}$ is C-tight if and only if, for all $T > 0$ and $\eta > 0$,

$$\lim_{H \to \infty} \overline{\lim_{n \to \infty}} P(|X_0^n| > H) = 0,$$

$$\lim_{\delta \to 0} \overline{\lim_{n \to \infty}} P\Big(\sup_{s,t \le T, |s-t| \le \delta} |X_t^n - X_s^n| > \eta\Big) = 0$$

(this follows, e.g., from Jacod and Shiryaev (1987), VI.3.26).

Another technical tool used below is the concept of strong majorization (Jacod and Shiryaev (1987), VI.3.34): Say that a process $X = (X_t, t \ge 0)$ strongly majorizes a process $Y = (Y_t, t \ge 0)$ if the process $X - Y = (X_t - Y_t, t \ge 0)$ is nondecreasing. If a sequence $\{X^n, n \ge 1\}$ of processes is C-tight and each X^n strongly majorizes a process Y^n, where X^n and Y^n are both nondecreasing and start at 0, then the sequence $\{Y^n, n \ge 1\}$ is C-tight (Jacod and Shiryaev (1987), VI.3.35).

LEMMA 5.2. *The sequence $\{K^{n,\epsilon}, n \ge 1\}$, where $K^{n,\epsilon} = (K_t^{n,\epsilon}, t \ge 0)$, is C-tight for every $\epsilon > 0$.*

PROOF. Denote by $[u_i^{n,1}, v_i^{n,1}]$ and $[u_i^{n,2}, v_i^{n,2}]$, $i \ge 1$, the respective successive switch-over 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 $\vartheta_t^{n,l}$ be the number of switchovers from queue l started in $[0, nt]$. By (5.3),

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

So, if we define

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

then, obviously,

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

Note that since

$$P\left(\frac{\vartheta_t^{n,l}}{n} > \frac{t+1}{d_l}\right) \le P\left(\sum_{i=1}^{\lfloor n(t+1)/d_l \rfloor} s_i^{n,l} \le nt\right),$$

Lemma 5.1 implies that

$$\lim_{n\to\infty} P\left(\frac{\vartheta_t^{n,l}}{n} > \frac{t+1}{d_l}\right) = 0. \tag{5.7}$$

Recalling also that

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

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

Define

$$\tilde{K}_t^{n,\epsilon,l} = \frac{1}{\sqrt{n}} \sum_{i=1}^{\vartheta_t^{n,l}} s_i^{n,l} \cdot 1(U_{v_i^{n,l}}^n > \epsilon\sqrt{n}), \qquad l = 1, 2. \tag{5.9}$$

Since U_s^n cannot increase on $[u_i^{n,l}, v_i^{n,l}]$, we have from (5.5) that $\breve{K}^{n,\epsilon}$ is strongly majorized by $\tilde{K}^{n,\epsilon} = \tilde{K}^{n,\epsilon,1} + \tilde{K}^{n,\epsilon,2}$, where $\tilde{K}^{n,\epsilon,l} = (\tilde{K}_t^{n,\epsilon,l}, t \ge 0)$, $l = 1, 2$. Therefore, it is enough to prove that $\{\tilde{K}^{n,\epsilon,1}, n \ge 1\}$ and $\{\tilde{K}^{n,\epsilon,2}, n \ge 1\}$ are each C-tight. By symmetry, we need only prove that $\{\tilde{K}^{n,\epsilon,1}, n \ge 1\}$ is C-tight.

Let

$$\bar{K}_t^{n,\epsilon,1} = \frac{1}{\sqrt{n}} \sum_{i=1}^{\vartheta_t^{n,1}} s_i^{n,1} \cdot 1(U_{u_i^{n,1}}^n > \epsilon\sqrt{n}/2), \tag{5.10}$$

$$\hat{K}_t^{n,\epsilon,1} = \frac{1}{\sqrt{n}} \sum_{i=1}^{\vartheta_t^{n,1}} s_i^{n,1} \cdot 1(U_{v_i^{n,1}}^n > \epsilon\sqrt{n}, U_{u_i^{n,1}}^n \le \epsilon\sqrt{n}/2). \tag{5.11}$$

By (5.9), $\tilde{K}^{n,\epsilon,1}$ is strongly majorized by $\bar{K}^{n,\epsilon,1} + \hat{K}^{n,\epsilon,1}$, where $\bar{K}^{n,\epsilon,1} = (\bar{K}_t^{n,\epsilon,1}, t \ge 0)$ and $\hat{K}^{n,\epsilon,1} = (\hat{K}_t^{n,\epsilon,1}, t \ge 0)$. We prove that $\{\bar{K}^{n,\epsilon,1}, n \ge 1\}$ is C-tight and that $\hat{K}_t^{n,\epsilon,1}$ tends in probability to 0 uniformly over finite intervals as $n \to \infty$. This will conclude the proof of the lemma.

We begin with the property of $\hat{K}^{n,\epsilon,1}$. Since there is no service on $[u_i^{n,1}, v_i^{n,1}]$, we have by (5.11) and (2.11),

$$\begin{aligned}\hat{K}_t^{n,\epsilon,1} &\le \frac{1}{\sqrt{n}}\sum_{i=1}^{\vartheta_t^{n,1}} s_i^{n,1}\cdot 1(S^n_{v_i^{n,1}} - S^n_{u_i^{n,1}} > \epsilon\sqrt{n}/2)\\ &\le \frac{1}{\sqrt{n}}\sum_{i=1}^{\vartheta_t^{n,1}} s_i^{n,1}\cdot 1(S^{n,1}_{v_i^{n,1}} - S^{n,1}_{u_i^{n,1}} > \epsilon\sqrt{n}/4)\\ &\quad + \frac{1}{\sqrt{n}}\sum_{i=1}^{\vartheta_t^{n,1}} s_i^{n,1}\cdot 1(S^{n,2}_{v_i^{n,1}} - S^{n,2}_{u_i^{n,1}} > \epsilon\sqrt{n}/4),\end{aligned}$$

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

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

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

Next,

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

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 $X \to \sup_{u\le t, |u-v|\le\gamma} |X_v - X_u|$, $X\in D[0,\infty)$, is continuous almost everywhere with respect to the Wiener measure (Liptser and Shiryaev (1989)), we conclude from (5.13) and (2.3) that

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

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 $\hat{K}^{n,\epsilon,1}$ is nondecreasing,

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

as required.

We now prove that $\{\bar{K}^{n,\epsilon,1}, n \ge 1\}$ 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 $\bar{\vartheta}_t^{n,1}$ be the number of sound switchovers started in $[0, nt]$. By (5.10),

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

where $\bar{s}_i^{n,1}$ is the duration of the ith sound switchover. Note that the soundness of a switchover is determined at its beginning, so the $\bar{s}_i^{n,1}$, $i \ge 1$, 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 $\Lambda > 0$,

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

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

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

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

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

where the latter limit, by Lemma 5.1, is zero if $(3\delta/\epsilon)d_1 < \eta$. Therefore,

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

which, since $\bar{K}_0^{n,\epsilon,1} = 0$, proves the C-tightness of $\{\bar{K}_v^{n,\epsilon,1}, n \geq 1\}$. 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*

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

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

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 $V_u^n \geq \varphi_u^n + B_u^n$ by (5.18),

$$\varphi_t^n - \varphi_s^n \leq \sqrt{n}\int_s^t 1(\varphi_u^n \leq 1 - B_u^n)\,du.$$

Therefore, by Lemma 1 in Coffman, Puhalskii, and Reiman (1991)

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

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

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

and recall that $\tau_i^{n,l}$, $i = 1, 2, \cdots$ denote arrival times for $A^{n,l}$, $l = 1, 2$.

LEMMA 5.4. *Define the filtration* $\mathbb{F}^n = (F^n_t,\, t \ge 0)$ *by* $F^n_t = F^{n,1}_t \vee F^{n,2}_t \vee \sigma(s^{n,l}_i,\, i = 1, 2, \dots, l = 1, 2) \vee \sigma(V^n_0) \vee \mathcal{N}$, *where* $F^{n,l}_t = G^{n,l}_{A^{n,l}_{nt}+1}$, $G^{n,l}_i = \sigma(\eta^{n,l}_j,\, \xi^{n,l}_j,\, 1 \le j \le i)$, $l = 1, 2$, *and* $\mathcal{N}$ *is the family of P-null sets. Then* $\mathbb{F}^n$ *is well defined, the* $\tau^{n,l}_i/n$, $i = 1, 2, \dots$, $l = 1, 2$, *are* $\mathbb{F}^n$*-stopping times, the processes* $(A^{n,l}_{nt},\, t \ge 0)$, $l = 1, 2$, *are* $\mathbb{F}^n$*-predictable and* $M^n = (M^n_t,\, t \ge 0)$ *is an* $\mathbb{F}^n$*-locally square-integrable martingale with the predictable quadratic-variation process*

$$\langle M^n \rangle_t = \frac{1}{n}\left[(\sigma^n_1)^2 A^{n,1}_{nt} + (\sigma^n_2)^2 A^{n,2}_{nt}\right].$$

PROOF. The proof is almost the same as the proof of Lemma 2 in Coffman et al. (1991). 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 $(\sum_{i=1}^k (\eta^{n,l}_i - \rho^n_l \xi^{n,l}_i),\, k \ge 0)$, $l = 1, 2$, are locally square-integrable martingales which have the predictable quadratic-variation processes $((\sigma^n_l)^2 k,\, k \ge 0)$ relative to the respective flows $(G^{n,l}_k,\, k \ge 0)$. □

Note that the processes B^n, $(\beta^n_t,\, t \ge 0)$ and V^n are $\mathbb{F}^n$-adapted.

Introduce

$$\epsilon^{n,l}_t = \frac{\rho^n_l}{\sqrt{n}}\left[\sum_{i=1}^{A^{n,l}_{nt}+1} \xi^{n,l}_i - nt\right] - \frac{1}{\sqrt{n}}\eta^{n,l}_{A^{n,l}_{nt}+1}, \qquad l = 1, 2, \qquad \epsilon^n_t = \epsilon^{n,1}_t + \epsilon^{n,2}_t. \tag{5.20}$$

Let ΔM^n_s denote the jump of M^n at s.

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

$$\langle M^n \rangle_t \xrightarrow{P} \sigma^2 t, \qquad \sum_{0<s\le t} (\Delta M^n_s)^2 \xrightarrow{P} \sigma^2 t, \qquad \sup_{s \le t} |\epsilon^n_s| \xrightarrow{P} 0,$$

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, (2.1) and (2.7). 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 (Jacod and Shiryaev (1987), I.4.14; Liptser and Shiryaev (1989), I.7), so its quadratic-variation process $([M^n, M^n]_t,\, t \ge 0)$ is the sum of the squares of jumps: $[M^n, M^n]_t = \sum_{s \le t} (\Delta M^n_s)^2$ (Jacod and Shiryaev (1987), I.4.52; Liptser and Shiryaev (1989), I.8). By Lemma 5.5.5 in Liptser and Shiryaev (1989), (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

$$0 \le \sum_{i=1}^{A^{n,l}_{nt}+1} \xi^{n,l}_i - nt \le \xi^{n,l}_{A^{n,l}_{nt}+1},$$

conditions (2.4) and (2.5) and Lemmas 3.1 and 5.1. □

Let

$$\bar{V}^n_t = V^n_t - \epsilon^n_t. \tag{5.21}$$

Since V^n is $\mathbb{F}^n$-adapted and ϵ^n_t is F^n_t-measurable by (5.20), $\bar{V}^n = (\bar{V}^n_t,\, t \ge 0)$ is $\mathbb{F}^n$-adapted. By (5.1), (5.2), (5.19) and (5.20), we get the representation

$$(5.22)\qquad \begin{aligned}\bar V_t^n = \bar V_0^n + \sqrt{n}(\rho^n - 1)t + \sqrt{n}\int_0^t 1\,(V_s^n = 0)ds \\ + \sqrt{n}\int_0^t 1(V_s^n > 0)\beta_s^n ds + (M_t^n - M_0^n).\end{aligned}$$

Now squaring in (5.22), we have by Ito's formula (Theorem 2.3.1 in Liptser and Shiryaev (1989)) that

$$(5.23)\qquad \begin{aligned}(\bar V_t^n)^2 = (\bar V_0^n)^2 + 2\sqrt{n}(\rho^n - 1)\int_0^t \bar V_s^n ds + 2\sqrt{n}\int_0^t \bar V_s^n 1\,(V_s^n = 0)ds \\ + 2\sqrt{n}\int_0^t \bar V_s^n 1(V_s^n > 0)\beta_s^n ds + 2\int_0^t \bar V_{s-}^n dM_s^n + \sum_{0<s\le t}(\Delta M_s^n)^2,\end{aligned}$$

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

LEMMA 5.6. *The sequences* $\{\bar V^n, n \ge 1\}$ *and* $\{V^n, n \ge 1\}$ *are C-tight.*

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

$$(5.24)\qquad \lim_{A\to\infty}\overline{\lim_{n\to\infty}} P(\sup_{s\le t}|V_s^n| > A) = 0, \qquad t > 0.$$

Then (5.21) and Lemma 5.5 yield

$$(5.25)\qquad \lim_{A\to\infty}\overline{\lim_{n\to\infty}} P(\sup_{s\le t}|\bar V_s^n| > A) = 0, \qquad t > 0.$$

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

$$(5.26)\qquad \lim_{\delta\to 0}\overline{\lim_{n\to\infty}}\sup_{\tau\in S_T(\mathbb{F}^n)} P(\sup_{t\le\delta}|(\bar V_{t+\tau}^n)^2 - (\bar V_\tau^n)^2| > \eta) = 0,$$

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

Since the processes

$$\left(\int_0^t \bar V_{s-}^n dM_s^n, t \ge 0\right) \quad \text{and} \quad \left(\sum_{0<s\le t}(\Delta M_s^n)^2 - \langle M^n\rangle_t, \qquad t \ge 0\right)$$

are $\mathbb{F}^n$-local martingales (Liptser and Shiryaev (1989), Ch. 1, §8, Ch. 2, §2), the Lenglart-Rebolledo inequality (Liptser and Shiryaev (1989), Theorem 1.9.3) yields, in view of (5.23), for $\epsilon > 0$,

$$P(\sup_{t\le\delta}|(\bar V^n_{t+\tau})^2 - (\bar V^n_\tau)^2| > \eta) \le \frac{\epsilon}{\eta}$$

$$(5.27)\qquad + P\biggl(2\sqrt{n}\,|\rho^n - 1|\int_\tau^{\tau+\delta} |\bar V^n_u|\,du + 2\sqrt{n}\int_\tau^{\tau+\delta} |\bar V^n_u|\,1\,(V^n_u = 0)du$$

$$+ 2\sqrt{n}\int_\tau^{\tau+\delta} |\bar V^n_u|\,1(V^n_u > 0)\beta^n_u du + \langle M^n\rangle_{\tau+\delta} - \langle M^n\rangle_\tau > \epsilon\biggr).$$

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

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

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

$$(5.29)\qquad 2\sqrt{n}\int_0^t |\bar V^n_s|\,1\,(V^n_s = 0)ds \overset{P}{\to} 0(n\to\infty), \qquad t > 0.$$

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

$$2\sqrt{n}\int_s^t |\bar V^n_u|\,1(V^n_u > 0)\beta^n_u du$$

$$(5.30)\qquad \le 2\sqrt{n}\sup_{u\le t}|\epsilon^n_u| \int_s^t 1(V^n_u > 0)\beta^n_u du$$

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

The first term on the right tends in probability to 0 as $n\to\infty$ by Lemma 5.5 and (5.17). The third term tends in probability to 0 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 $\gamma > 0$,

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

Thus, by (5.30),

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

Lemma 5.5 easily implies that

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

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., Liptser and Shiryaev (1989), Theorem 6.3.1), (5.25) and (5.26) imply that the sequence $\{(\bar{V}^n)^2, n \ge 1\}$ and hence $\{\bar{V}^n, n \ge 1\}$ is tight for the Skorohod topology. By Proposition VI.3.26 in Jacod and Shiryaev (1987), it remains to prove that

$$\sup_{t\le T} |\Delta \bar{V}^n_t| \stackrel{P}{\to} 0, \qquad T > 0.$$

In view of (5.19),

$$\sup_{t\le T} |\Delta \bar{V}^n_t| = \sup_{t\le T} |\Delta M^n_t| \le \max_{l=1,2}\biggl[\frac{1}{\sqrt{n}} \max_{1\le i\le A^{n,l}_{nT}+1} \eta^{n,l}_i + \frac{\rho^n_l}{\sqrt{n}} \max_{1\le i\le A^{n,l}_{nT}+1} \xi^{n,l}_i\biggr]$$

which tends to 0 in probability as $n\to\infty$ by Lemmas 3.1 and 5.1 and (2.3), (2.4) and (2.5). This proves that $\{\bar{V}^n, n\ge 1\}$ is C-tight. The sequence $\{V^n, n \ge 1\}$ is then C-tight by Lemma 5.5 and (5.21). □

By Prohorov's theorem, Lemma 5.6 makes it certain that there exists a subsequence $\{V^{n'}, n' \ge 1\}$ and a continuous process $\tilde{V}$ such that $V^{n'} \stackrel{d}{\to} \tilde{V}$. The next two lemmas deal with implications of this fact.

LEMMA 5.7. *We have, for $\eta > 0$,*

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

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

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

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

$$\underline{\lim}_{n\to\infty} P\biggl(\int_0^t 1(V^{n'}_s < \epsilon)\,ds > \eta\biggr) \ge P\biggl(\int_0^t 1(\tilde{V}_s < \epsilon)\,ds > \eta\biggr),$$

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

$$Z^n_t = V^n_0 + B^n_t + \sqrt{n}\int_0^t 1(V^n_s > 0)\beta^n_s\,ds, \qquad t \ge 0,$$

so that (5.2) is equivalent to

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

Since V_t^n is nonnegative and $\sqrt{n}\int_0^t 1\,(V_s^n = 0)\,ds$ increases only when V_t^n equals 0, we conclude that $V^n = R(Z^n)$, where $R : D[0,\infty) \to D[0,\infty)$ is Skorohod's reflection map. In the one-dimensional case it is well known to be equivalently defined by (for $C[0,\infty)$, the result is in Ikeda and Watanabe (1989), the result for $D[0,\infty)$ is a special case of a more general result in Chen and Mandelbaum (1991))

$$R(X)_t = X_t - \inf_{s \le t} X_s \wedge 0, \qquad t \ge 0. \tag{5.33}$$

Now, if we define

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

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 $V_t^n \ge \check{V}_t^n$, $t \ge 0$. Hence, for $\epsilon > 0$,

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

Now, by (5.34) and Lemma 5.1, $\{\check{Z}^n, n \ge 1\}$ converges in distribution to the process $\check{Z} = (V_0 + \sigma W_t + ct,\, t \ge 0)$, where V_0 and $(W_t, t \ge 0)$ are independent. By the continuity of the reflection (Whitt (1980), Theorem 6.4), we then deduce that $\check{V}^n \xrightarrow{d} R(\check{Z})$, i.e., by (5.36),

$$\overline{\lim_{e\to 0}}\ \overline{\lim_{n\to\infty}} P\left(\int_0^t 1(V_s^n < \epsilon)\,ds > \eta\right) \le \overline{\lim_{\epsilon\to 0}}\ \overline{\lim_{n\to\infty}} P\left(\int_0^t 1(\check{V}_s^n < \epsilon)\,ds > \eta\right)$$

$$\le \overline{\lim_{\epsilon\to 0}}\, P\left(\int_0^t 1(R(\check{Z})_s \le \epsilon)\,ds \ge \eta\right) = P\left(\int_0^t 1(R(\check{Z})_s = 0)\,ds \ge \eta\right).$$

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

The next lemma shows that, "on average," $\sqrt{n}\,\beta_t^n$ behaves as d/V_t^n substantiating the heuristic argument of §1.

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 $\tilde{V} = (\tilde{V}_t,\, t \ge 0)$ be an accumulation point of $\{V^n, n \ge 1\}$. 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

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

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

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

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

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

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

and observe that the latter goes to 0 as $n \to \infty$ and $K \to \infty$ by the tightness of $\{V^n, n \ge 1\}$. Theorem 4.2 in Billingsley (1968) then implies the desired result by (5.38), (5.39) and the fact that, by Lemma 5.7 and the continuity of $\tilde{V}$,

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

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

Recall that $\kappa_j^{n,1}(i,k)$, $j \ge 0$, are the successive times after $n\tau_k^n(\epsilon, i)$, when the unfinished work at queue 1 becomes equal to 0, and $\vartheta_{i,k}^{n,1}$ is the number of cycles accumulation-service for queue 1 in $[\tau_k^n(\epsilon, i)\wedge T, \zeta_k^n(\epsilon, i)\wedge T]$. We denote the switchover times starting after $n\tau_k^n(\epsilon, i)$ by $s_0^{n,1}(i,k)$, $s_1^{n,1}(i,k)$, ... Obviously, they are independent and distributed as $s_1^{n,1}$. We introduce a similar notation for queue 2: $n\kappa_j^{n,2}(i,k)$, $j \ge 0$, denote the successive times after $n\tau_k^n(\epsilon, i)$, when the unfinished work at queue 2 becomes equal to 0, $\vartheta_{i,k}^{n,2}$ denotes the number of cycles accumulation-service for queue 2 in $[\tau_k^n(\epsilon, i)\wedge T, \zeta_k^n(\epsilon, i)\wedge T]$, $s_0^{n,2}(i,k)$, $s_1^{n,2}(i,k)$, $\cdots$ denote the switchover times from queue 2 to queue 1 that start at $n\kappa_0^{n,2}(i,k)$, $n\kappa_1^{n,2}(i,k)$, ... We note again that $s_0^{n,2}(i,k)$, $s_1^{n,2}(i,k)$, ... are independent and distributed as $s_1^{n,2}$.

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 $k \ge 1$, $0 \le i \le N$,

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

$$\cdot 1(\tau_k^n(\epsilon, i) < T) + \sum_{j=0}^{\vartheta_{i,k}^{n,1}+1} s_j^{n,1}(i,k) + \sum_{j=0}^{\vartheta_{i,k}^{n,2}+1} s_j^{n,2}(i,k),$$

so, since $V_t^n \in [a_i(\epsilon)-\epsilon, a_i(\epsilon)+\epsilon]$ on $[\tau_k^n(\epsilon, i), \zeta_k^n(\epsilon, i))$ and $\beta_t^n = 1-\alpha_{nt}^n$, we get

$$\frac{1}{\sqrt{n}}(a_i(\epsilon)-\epsilon)\left(\sum_{j=0}^{\vartheta_{i,k}^{n,1}} s_j^{n,1}(i,k)+\sum_{j=0}^{\vartheta_{i,k}^{n,2}} s_j^{n,2}(i,k)\right)$$

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

Introduce

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

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

where $\tilde{\vartheta}_{i,k}^{n,1}$ and $\hat{\vartheta}_{i,k}^{n,1}$ are defined by (4.8) and (4.9) respectively, $\tilde{\vartheta}_{i,k}^{n,2}$ and $\hat{\vartheta}_{i,k}^{n,2}$ denote their counterparts for queue 2 and

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

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

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

Also, inequalities (5.40) and (4.10) (an analogue of the latter holds obviously for queue 2 as well) yield

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

Next, (4.18), (4.16), and (4.17), and their analogues for queue 2 imply, in view of (4.4) and Lemma 2.2 in Coffman et al. (1995), that

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

$$(5.46)\qquad \left(\frac{1}{\sqrt{n}}\hat{\vartheta}^{n,l}_{i,k}\right)_{\substack{0\le i\le N,\\ k\ge 1, l=1,2}} \xrightarrow{d} \left(\frac{\zeta_k(\epsilon,i)\wedge T-\tau_k(\epsilon,i)\wedge T}{a_i(\epsilon)+\epsilon}\left(\frac{1}{\rho_l}+\frac{1}{1-\rho_l}\right)^{-1}\right)_{\substack{0\le i\le N,\\ k\ge 1, l=1,2}}.$$

Hence, by (5.41)–(5.43), Lemma 5.1, the equality $\rho_1+\rho_2=1$ and again Lemma 2.2 in Coffman et al. (1995),

$$(5.47)\qquad (\tilde{A}^n_k(\epsilon,i))_{\substack{0\le i\le N,\\ k\ge 1}} \xrightarrow{d} (\tilde{A}_k(\epsilon,i))_{\substack{0\le i\le N,\\ k\ge 1}},$$

and

$$(5.48)\qquad (\hat{A}^n_k(\epsilon,i))_{\substack{0\le i\le N,\\ k\ge 1}} \xrightarrow{d} (\hat{A}_k(\epsilon,i))_{\substack{0\le i\le N,\\ k\ge 1}},$$

where

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

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

Introduce

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

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

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 $M=1,2,\dots,$

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

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

as $n\to\infty$.

On the other hand, in a manner similar to (4.25),

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

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

and Theorem 4.2 in Billingsley (1968) yields

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

Similarly,

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

Next, in analogy with the proof of (4.29), we get in view of (4.2) that, as $\epsilon \to 0$,

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

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

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

$$\begin{aligned} \sum_{k=1}^{\infty} \sum_{i=1}^{N-1} \sqrt{n} \int_{\tau_k^n(\epsilon,i)\wedge T}^{\zeta_k^n(\epsilon,i)\wedge T} \beta_t^n V_t^n\,dt &\le \sqrt{n} \int_0^T \beta_t^n V_t^n \cdot 1(\delta \le V_t^n \le K)\,dt \\ &\le \sum_{k=1}^{\infty} \sum_{i=0}^{N} \sqrt{n} \int_{\tau_k^n(\epsilon,i)\wedge T}^{\zeta_k^n(\epsilon,i)\wedge T} \beta_t^n V_t^n\,dt, \end{aligned}$$

so by (5.44) and (5.51)

$$\tilde{A}^n(\epsilon) \le \sqrt{n} \int_0^T \beta_t^n V_t^n \cdot 1(\delta \le V_t^n \le K)\,dt \le \hat{A}^n(\epsilon).$$

The latter and (5.53)–(5.56) imply (5.37) by Lemma 2.3 in Coffman, Puhalskii, and Reiman (1995). □

6. Proofs of theorems.

PROOF OF THEOREM 2.1. By (5.23) and (5.21),

$$(\bar{V}_t^n)^2 + \delta_t^n = (\bar{V}_0^n)^2 + 2\sqrt{n}(\rho^n - 1) \int_0^t \bar{V}_s^n\,ds + (2d + \sigma^2)t + 2 \int_0^t \bar{V}_{s-}^n\,dM_s^n, \tag{6.1}$$

where

$$\delta_t^n = 2\sqrt{n}\int_0^t \epsilon_s^n \cdot 1\,(V_s^n = 0)\,ds + 2\sqrt{n}\int_0^t \epsilon_s^n \beta_s^n \cdot 1(V_s^n > 0)\,ds$$

$$+\, 2dt - 2\sqrt{n}\int_0^t \beta_s^n V_s^n\,ds + \sigma^2 t - \sum_{s\le t} (\Delta M_s^n)^2.$$

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

$$\sup_{s\le t} |\delta_s^n| \xrightarrow{P} 0(n \to \infty), \qquad t > 0. \tag{6.2}$$

We denote the left-hand side of (6.1) by X_t^n. Let $X^n = (X_t^n, t \ge 0)$. We next prove that X^n converges in distribution to X as $n \to \infty$. By (6.1), the process X^n is an $\mathbb{F}^n$-locally square-integrable semimartingale (Jacod and Shiryaev (1987), Ch. II, §2b)). The process X as defined by (2.9) is a locally square-integrable semimartingale as well with respect to the filtration generated by it (Jacod and Shiryaev (1987), III.2.12). We prove the convergence by applying Theorem IX.3.48 in Jacod and Shiryaev (1987) which gives conditions for convergence in distribution of a sequence of semimartingales to a semimartingale in terms of their predictable characteristics (Jacod and Shiryaev (1987), Ch. II, §2). Therefore our first step is to identify the characteristics. Let $B'^n = (B_t'^n, t \ge 0)$ denote the first characteristic without truncation of X^n, let $\tilde{C}'^n = (\tilde{C}_t'^n, t \ge 0)$ denote its modified second characteristic without truncation and let $\nu^n = (\nu^n(ds, dx))$ denote its predictable measure of jumps (Jacod and Shiryaev (1987), II.2.29, IX.3.25). Then by (6.1)

$$B_t'^n = 2\sqrt{n}(\rho^n - 1)\int_0^t \bar{V}_s^n\,ds + (2d + \sigma^2)t, \tag{6.3}$$

$$\tilde{C}_t'^n = 4\int_0^t (\bar{V}_s^n)^2\, d\langle M^n\rangle_s. \tag{6.4}$$

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

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

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

$$\nu([0, t], \Gamma)(\alpha) = 0, \qquad t \ge 0, \qquad \Gamma \text{ is a Borel subset of } R, \tag{6.7}$$

and let $B(\alpha) = (B_t(\alpha), t \ge 0)$, $C(\alpha) = (C_t(\alpha), t \ge 0)$ and $\nu(\alpha) = (\nu(dt, dx)(\alpha))$. According to the definition of X in (2.9), 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 $(\mathscr{H}, X)$ and $(\mathscr{L}(X_0); B, C, \nu)$, in the sense of definition III.2.4 of Jacod and Shiryaev (1987), where $\mathscr{H}$ denotes the σ-field generated by X_0 and $\mathscr{L}(X_0)$ denotes the distribution of X_0.

Define next, as in IX.3.38 of Jacod and Shiryaev (1987), for $a \geq 0$,

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

(6.9) $$S_a^n = \inf(t : |X_t^n| \geq a \text{ or } |X_{t-}^n| \geq a),$$

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

By Theorem IX.3.48 of Jacod and Shiryaev (1987), 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 $((\text{Var } B)_{t\wedge S_a(\alpha)}(\alpha), t \geq 0)$, $(C_{t\wedge S_a(\alpha)}(\alpha), t \geq 0)$ and $(\int_0^{t\wedge S_a(\alpha)} \int_R |x|^2 \nu(ds, dx)(\alpha),$ $t \geq 0)$ 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* $(\mathscr{H}, X)$ *and* $(\mathscr{L}(X_0);$ $B, C, \nu)$ (see Jacod and Shiryaev (1987), 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) $X_0^n \xrightarrow{d} X_0$.

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

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

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

(6.9a) $\lim_{b\to\infty} \overline{\lim}_{n\to\infty} P(\int_0^{t\wedge S_a^n} \int_R |x|^2 1(|x| > b)\nu^n(ds, dx) > \epsilon) = 0$ for all $t > 0$, $a > 0$ and $\epsilon > 0$.

This last condition is Equation (3.49) in Jacod and Shiryaev (1987).

We now check these 9 conditions in order. We have, by (6.5)–(6.8), 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 of Jacod and Shiryaev (1987) 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 (Ikeda and Watanabe (1989), 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 of Jacod and Shiryaev (1987), $p_t B = B$, $p_t C = C$, $p_t \nu = \nu = 0$.

Condition (iv) follows from (6.5)–(6.7) by the argument of the proof of Theorem 6.2.2 in Liptser and Shiryaev (1989) since Skorohod convergence implies convergence at continuity points of the limit.

Condition (v) holds by the assumption $V_0^n \xrightarrow{d} V_0$, (5.21), Lemma 5.5 and (6.2).

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

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

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 $\nu^n([0, t], \{|x| > \epsilon\}) \xrightarrow{P} 0$. By Lemma 5.5.1 in Liptser and Shiryaev (1989), this is implied by

$$\sup_{s \leq t} |\Delta X_s^n| \xrightarrow{P} 0 (n \to \infty), \qquad t > 0. \tag{6.10}$$

By the definition of X^n, $|\Delta X_s^n| \leq \Delta |\bar{V}_s^n|^2 + |\Delta \delta_s^n|$, and (6.10) holds by (6.2) and the C-tightness of $\bar{V}^n$ (use Proposition VI.3.26 in Jacod and Shiryaev (1987)).

Next, we check condition $[\sup - \beta'_{\text{loc}}]$ under (vi). By the definition of X^n, (6.3) and (6.5),

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

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

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

$$\begin{aligned}&|\tilde{C}'^n_{t \wedge S_a^n} - C_{t \wedge S_a^n}(X^n)| \\ &\quad \leq 4 \sup_{s \leq t} \left| \int_0^s (\bar{V}_u^n)^2 d\langle M^n \rangle_u - \sigma^2 \int_0^s (\bar{V}_u^n)^2 du \right| + 4\sigma^2 t \sup_{s \leq t} |\delta_s^n|.\end{aligned} \tag{6.11}$$

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

$$\lim_{\delta\to 0}\overline{\lim_{n\to\infty}} P\Big(\sup_{|u-v|<\delta, u,v\le t} |(\bar V_u^n)^2 - (\bar V_v^n)^2| > \eta\Big) = 0,$$

which, in view of the second assertion of Lemma 5.5, is seen to yield (the idea of the proof is as in (Billingsley (1968), Problem 8, §2))

$$\sup_{s\le t}\left|\int_0^s (\bar V_u^n)^2\, d\langle M^n\rangle_u - \sigma^2\int_0^s (\bar V_u^n)^2\, du\right| \xrightarrow{P} 0.$$

In view of (6.11), this concludes the verification of $[\gamma'_{\text{loc}} - R_+]$.

Finally, consider condition (6.9a) under (vi). Define

$$\tilde L_t^n = \int_0^t\int_R |x|^2\cdot 1(|x| > b)\,\nu^n(ds, dx), \qquad t\ge 0.$$

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

$$P(\tilde L_t^n \ge \epsilon) \le P(\sup_{s\le t}|\bar V_s^n| > A) + P\left(\int_0^t 1(\bar V_{s-}^n \le A)\,d\tilde L_s^n > \epsilon\right). \tag{6.12}$$

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

Next, letting

$$\gamma_t^{n,1} = \tfrac{1}{n}\tau^{n,1}_{\lfloor n(\lambda_1 t+1)\rfloor}, \qquad \gamma_t^{n,2} = \tfrac{1}{n}\tau^{n,2}_{\lfloor n(\lambda_2 t+1)\rfloor},$$

we have for the second term on the right of (6.12), since $A_{nt}^{n,l} \ge \lfloor n(\lambda_l t + 1)\rfloor$ when $\gamma_t^{n,l} < t$, $l = 1, 2$,

$$\begin{aligned} &P\left(\int_0^t 1(\bar V_{s-}^n \le A)\,d\tilde L_s^n > \epsilon\right)\\ &\quad\le P\left(\frac{1}{n}A_{nt}^{n,1} > \lambda_1 t + 1 - \frac{1}{n}\right) + P\left(\frac{1}{n}A_{nt}^{n,2} > \lambda_2 t + 1 - \frac{1}{n}\right)\\ &\quad\quad + P\left(\int_0^{t\wedge\gamma_t^{n,1}\wedge\gamma_t^{n,2}} 1(\bar V_{s-}^n \le A)\,d\tilde L_s^n > \epsilon\right). \end{aligned}$$

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 $n\to\infty$.

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

$$L_t^n = 4\sum_{0<s\le t} (\bar V_{s-}^n)^2(\Delta M_s^n)^2\cdot 1(2|\bar V_{s-}^n||\Delta M_s^n| > b). \tag{6.13}$$

Accordingly, the process $(\int_0^t 1(\bar V_{s-}^n \le A)\,d\tilde L_s^n, t\ge 0)$ is the $\mathbb{F}^n$-compensator of the process $(\int_0^t 1(\bar V_{s-}^n \le A)\,dL_s^n, t\ge 0)$. Therefore, by (6.13) and, since, by Lemma 5.4, $\gamma_t^{n,1}$ and $\gamma_t^{n,2}$ are $\mathbb{F}^n$-stopping times, the Lenglart-Rebolledo inequality implies that, for $\eta > 0$,

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

$$\le \frac{1}{\epsilon}\left(\eta + E\sup_{s\le t\wedge\gamma_t^{n,1}\wedge\gamma_t^{n,2}} 4A^2|\Delta M_s^n|^2\cdot 1\left(|\Delta M_s^n| > \frac{b}{2A}\right)\right) \tag{6.14}$$

$$+ P\left(4A^2\sum_{0<s\le t\wedge\gamma_t^{n,1}\wedge\gamma_t^{n,2}} (\Delta M_s^n)^2\cdot 1\left(|\Delta M_s^n| > \frac{b}{2A}\right) \ge \eta\right).$$

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 (6.14) 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 (6.14),

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

which completes the check of (6.9a) since η is arbitrary.

Thus, all the conditions of Theorem IX.3.48 in Jacod and Shiryaev (1987) 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

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

we have that $(V^n)^2 \xrightarrow{d} V^2$, and hence that $V^n \xrightarrow{d} V$ 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. By (2.12) and the definition of $\breve{V}^n$, a basic equation for $\breve{V}^n$ has the form

$$\breve{V}_t^n = V_0^n + \breve{B}_t^n + \sqrt{n}\bar{r}^n \int_0^t 1(\breve{V}_s^n = 0)\,ds + \sqrt{n}\int_0^t 1(\breve{V}_s^n > 0)(\bar{r}^n - r^n(\sqrt{n}\breve{V}_s^n))\,ds,$$

where

$$\breve{B}_t^n = \frac{S_{nt}^n - \bar{r}^n nt}{\sqrt{n}}.$$

The equation is similar to Equation (5.2) and the proof goes exactly the way it does for $\{V^n, n \geq 1\}$. We first prove the C-tightness of $\{\breve{V}^n, n \geq 1\}$ by following the same steps as in §5. The only difference is that the $K^{n,\epsilon}$ are defined this time by

$$K_t^{n,\epsilon} = \sqrt{n}\int_0^t 1(\breve{V}_s^n > \epsilon)(\bar{r}^n - r^n(\sqrt{n}\breve{V}_s^n))\,ds,$$

and Lemma 5.2 is trivial because of $(r1)$.

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

$$\sqrt{n}\int_0^t V_s^n \beta_s^n\,ds \xrightarrow{P} dt.$$

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

$$\sqrt{n}\int_0^t \breve{V}_s^n(\bar{r}^n - r^n(\sqrt{n}\breve{V}_s^n))\,ds \xrightarrow{P} dt.$$

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

$$\lim_{\epsilon\to 0}\overline{\lim_{n\to\infty}} P\left(\int_0^t 1(\breve{V}_s^n < \epsilon)\,ds > \eta\right) = 0. \tag{6.15}$$

Then, for $\epsilon > 0$,

$$P\left(\left|\sqrt{n}\int_0^t \check{V}_s^n(\bar{r}^n - r^n(\sqrt{n}\check{V}_s^n))ds - dt\right| > \eta\right)$$

$$\le P\left(\int_0^t |\sqrt{n}\check{V}_s^n(\bar{r}^n - r^n(\sqrt{n}\check{V}_s^n)) - d| \cdot 1(\check{V}_s^n < \epsilon)ds > \frac{\eta}{2}\right)$$

$$+ P\left(\int_0^t |\sqrt{n}\check{V}_s^n(\bar{r}^n - r^n(\sqrt{n}\check{V}_s^n)) - d| \cdot 1(\check{V}_s^n \ge \epsilon)ds > \frac{\eta}{2}\right)$$

$$\le P\left((\sup_{x,n} x(\bar{r}^n - r^n(x)) + d)\int_0^t 1(\check{V}_s^n < \epsilon)ds > \frac{\eta}{2}\right)$$

$$+ 1(t \sup_{x \ge \sqrt{n}\epsilon} |x(\bar{r}^n - r^n(x)) - d| > \tfrac{\eta}{2}).$$

The probability on the right-hand side tends to 0 as $n \to \infty$ and $\epsilon \to 0$ by $(r2)$ and (6.15), and the indicator goes to 0 as $n \to \infty$ by $(r1)$. □

References

Billingsley, P. (1968). *Convergence of Probability Measures*. Wiley, New York.

Chen, H., A. Mandelbaum (1991). Leontief systems, RBV's and RBM's. M. H. A. Davis, R. J. Elliott, eds., *Proceedings of the Imperial College Workshop on Applied Stochastic Processes*. Gordon and Breach Science Publishers, London.

Coffman, Jr., E. G., A. A. Puhalskii, M. I. Reiman (1991). Storage limited queues in heavy traffic. *Prob. Eng. Inf. Sci.* **5** 499–522.

———, ———, ——— (1995). Polling systems with zero switchover times: A heavy traffic averaging principle. *Ann. Appl. Prob.* **5** 681–719.

Iglehart, D. L., W. Whitt (1970). Multiple channel queues in heavy traffic, I and II. *Adv. Appl. Prob.* **2** 150–177 and 355–364.

Ikeda, N., S. Watanabe (1989). *Stochastic Differential Equations and Diffusion Processes*. North Holland, Amsterdam/Oxford/New York.

Jacod, J., A. N. Shiryaev (1987). *Limit Theorems for Stochastic Processes*. Springer, Berlin/Heidelberg/New York/London/Paris/Tokyo.

Kruskal, J. B. (1969). Work scheduling algorithms: A nonprobabilistic queueing study (with possible application to No. 1 ESS). *Bell Sys. Tech. J.* **48** 2963–2974.

Leung, K. K. (1991). Cyclic-service systems with probabilistically-limited service. *IEEE J. Sel. Areas Comm.* **9** 185–193.

Levy, H., M. Sidi (1990). Polling systems: Applications, modeling, and optimization. *IEEE Trans. Comm.* **38** 1750–1760.

Liptser, R. Sh., A. N. Shiryaev (1989). *Theory of Martingales*. Kluwer, Dordrecht/Boston/London.

Olsen, T. (1998). Approximations for the waiting time in polling models with and without state-dependent set ups. TR 98-1. Industrial and Operations Engineering Dept., University of Michigan, Ann Arbor, MI.

Reiman, M. I. (1988). A multiclass feedback queue in heavy traffic. *Adv. Appl. Prob.* **2** 179–207.

———, L. M. Wein (1995). Dynamic scheduling of a two-class queue with setups. *Oper. Res.* to appear.

Rosenkrantz, W. A. (1984). Weak convergence of a sequence of queueing and storage processes to a singular diffusion. F. Baccelli, G. Fayolle, eds., *Lect. Notes in Control Inform. Sci.*, vol. 60. Springer-Verlag, Berlin-Heidelberg-New York-Tokyo, 257–272.

Takagi, H. (1986). *Analysis of Polling Systems*. The MIT Press, Cambridge, MA.

——— (1990). Queueing analysis of polling models: An update. H. Takagi, ed., *Stochastic Analysis of Computer and Communication Systems*. North Holland, Amsterdam.

Whitt, W. (1980). Some useful functions for functional limit theorems. *Math. Oper. Res.* **5** 67–85.

Yamada, K. (1984). Diffusion approximations for storage processes with general release rules. *Math. Oper. Res.* **9** 459–470.

——— (1986). Multi-dimensional Bessel processes as heavy traffic limits of certain tandem queues. *Stoch. Proc. Appl.* **23** 35–56.

E. G. Coffman, Jr.: Bell Laboratories, Lucent Technologies, Murray Hill, New Jersey 07974

A. A. Puhalskii: Institute for Problems in Information Transmission, Moscow, Russia 101447

M. I. Reiman: Bell Laboratories, Lucent Technologies, Murray Hill, New Jersey 07974