Queueing Systems 28 (1998) 157–190

A critically loaded multirate link with trunk reservation

A.A. Puhalskii [a,*] and M.I. Reiman [b]

[a] *Institute for Problems in Information Transmission, 101447 Moscow, Russia*
[b] *Bell Laboratories, Lucent Technologies, Murray Hill, 07974, USA*

We consider a loss system model of interest in telecommunications. There is a single service facility with N servers and no waiting room. There are K types of customers, with type i customers requiring A_i servers simultaneously. Arrival processes are Poisson and service times are exponential. An arriving type i customer is accepted only if there are $R_i\,(\geqslant A_i)$ idle servers. We examine the asymptotic behavior of the above system in the regime known as critical loading where both N and the offered load are large and almost equal. We also assume that $R_1, \ldots, R_{K-1}$ remain bounded, while $R_K^N \to \infty$ and $R_K^N/\sqrt{N} \to 0$ as $N \to \infty$. Our main result is that the K dimensional "queue length" process converges, under the appropriate normalization, to a particular K dimensional diffusion. We show that a related system with preemption has the same limit process. For the associated optimization problem where accepted customers pay, we show that our trunk reservation policy is asymptotically optimal when the parameters satisfy a certain relation.

Keywords: stochastic service system, multirate link, trunk reservation

1. Introduction and summary

In this paper we consider a stochastic service system model of interest for telecommunication systems. The system we model consists of several traffic streams with possibly different bandwidths sharing a link, with a trunk reservation mechanism used for call admission control. The model is of interest for both multirate circuit switching and broadband packet switching systems analyzed using the notion of effective bandwidth (cf. [20]). The link is assumed to consist of N circuits, which we model as a single service facility with N servers and no additional waiting room. There are K types of calls (corresponding to the customers in our model), with type i calls arriving in a Poisson process of rate λ_i. Type i calls require A_i servers (A_i is an integer) for the entire duration of the call, which has an exponential distribution with mean μ_i^{-1}. An arriving type i call is accepted if there are at least $R_i \geqslant A_i$ idle circuits at the moment of its arrival. The integer quantities R_i are referred to as trunk reservation parameters. Any arriving call that is not accepted is blocked from the system and never returns (this is a loss model).

The above model gives rise to a finite state, K-dimensional Markov process. Solving for either transient or steady state performance characteristics of this system

* The research of this author was supported in part by Bell Laboratories.

becomes prohibitive as N becomes large. Systems with N large are of practical interest, motivating the examination of asymptotics as $N \to \infty$. Suppose we let $N \to \infty$, with $\lambda_i^N \to \infty$ as $N \to \infty$, $1 \leqslant i \leqslant K$. We hold μ_i and A_i fixed, $1 \leqslant i \leqslant K$. Let

$$\rho(N) = N^{-1} \sum_{i=1}^{K} A_i \lambda_i^N / \mu_i$$

denote the normalized offered load on the link. Suppose that $\rho(N) \to \rho$ as $N \to \infty$. There are three regimes to consider: $\rho < 1$, known as underloaded; $\rho > 1$, known as overloaded (also called heavy traffic); and $\rho = 1$, known as critically loaded. Roughly speaking, the blocking probabilities in underloaded, critically loaded, and overloaded systems respectively are exponentially small in N, $O(N^{-1/2})$ and $O(1)$. From a practical point of view, the blocking probabilities in underloaded systems are typically too small (there is too much idleness, corresponding to wasted resources), while the blocking probabilities in overloaded systems are too large (unacceptable service quality). Our focus in this paper is on the critically loaded regime. Results for a multirate link with trunk reservation in the overloaded regime are available in [2,3,8].

In the "homogeneous" case, where $A_i = A_j$ and $\mu_i = \mu_j$, for $1 \leqslant i, j \leqslant K$, the analysis of this model can be carried out using a one-dimensional birth-death process (and we can, without loss of generality, assume that $A_i = 1$, $1 \leqslant i \leqslant K$, achieving this effect by taking the number of servers to be $\lfloor N/A_i \rfloor$). This is because once a customer has been admitted into the system, its type becomes irrelevant to the further evolution of the system. It is known that a policy of the above form with $A_1 = R_1 \leqslant R_2 \leqslant \cdots \leqslant R_K$ (known as a trunk reservation policy) is optimal for the control of this system if the objective is to maximize the long run average reward earned, where accepted type i customers pay r_i, with $r_1 \geqslant r_2 \geqslant \cdots \geqslant r_K$ [16]. With $A_i \neq A_j$, there are counterexamples showing that there sometimes is no trunk reservation policy that is optimal [21].

The behavior of the optimal trunk reservation level for the homogeneous case with $K = 2$, $A_i = 1$ and $r_1 > r_2$ under critical loading was examined in [19]. There it was assumed that $\mu = 1$ (which, in this case, is without loss of generality) and

$$\lambda_i^N = \alpha_i N + \beta_i \sqrt{N}, \quad i = 1, 2,$$

where critical loading is equivalent to $\alpha_1 + \alpha_2 = 1$. If $0 < \alpha_1, \alpha_2 < 1$, it was shown in [19] that $R_1^*(N) = 1$ and $R_2^*(N)/\log N \to -1/(2\log\alpha_1)$, where $R_i^*(N)$ are optimal trunk reservation parameters in the Nth system. Let $b_i(N)$ denote the blocking probability of type i customers in the Nth system. With no trunk reservation ($R_1 = R_2 = 1$) $b_1(N) = b_2(N) \equiv b(N)$, and it is known [10] that $\sqrt{N} b(N) \to h(\beta)$, where h is the hazard rate of a standard normal distribution and $\beta = \beta_1 + \beta_2$. Straightforward asymptotics on the birth–death process stationary distribution can be used to show that, if $R_1^N = 1$, $R_2^N \to \infty$, and $R_2^N/\sqrt{N} \to 0$ as $N \to \infty$, then

$$\sqrt{N} b_1(N) \to 0 \quad \text{and} \quad \sqrt{N} b_2(N) \to \alpha_2^{-1} h(\beta).$$

This corresponds to a transfer of all type 1 blocking to type 2 with no increase in average blocking (to order $N^{-1/2}$). On this scale, with error o($N^{-1/2}$), this is the same as would be achieved by the following infeasible scheme that gives preemptive priority to type 1 calls. Admit all calls if there is an idle server. When a type 1 call arrives to find no idle server, it preempts a type 2 call. We interpret b_2 here as the fraction of type 2 calls that are either blocked on arrival or preempted while in service.

The purpose of this paper is to generalize, to a certain extent, the above results for the homogeneous case to the heterogeneous case, where $A_i \neq A_j$ and/or $\mu_i \neq \mu_j$. Although trunk reservation may not be optimal, we restrict our attention to trunk reservation controls because trunk reservation is simple to implement and analyze. Our main result is that an appropriately normalized K-dimensional queue length process for the system with trunk reservation converges in distribution to a reflected diffusion process on a K-dimensional half-space when $R_K^N \to \infty$ and $R_K^N/\sqrt{N} \to 0$ as $N \to \infty$. We show that the same process arises as the limit in a system that gives preemptive priority to all but type K calls. Thus, just as in the homogeneous case, the limit behavior of the systems with trunk reservation and preemption are identical. When the parameters satisfy a certain relation (equation (2.11)) we prove that trunk reservation is asymptotically optimal.

The results of this paper indicate a substantial robustness in the system behavior to variation in trunk reservation parameters. All that is needed to obtain the limit diffusion process is $R_K^N \to \infty$ and $R_K^N/\sqrt{N} \to 0$. Thus the choice of R_K^N does not need to depend on the values of $\beta_1, \dots, \beta_K$ or $\alpha_1, \dots, \alpha_K$ as long as $\alpha_1 + \dots + \alpha_K = 1$.

In addition to the references indicated above, there are a few papers related to loss systems and/or trunk reservation that deserve mention. A survey of some work on loss networks is contained in [13]. Asymptotics for product form loss networks are provided in [7,11], with the former focusing on the critically loaded case. The model of the present paper without trunk reservation ($R_i = A_i,\ 1 \leqslant i \leqslant K$) is a product form loss system; asymptotics under critical load are provided in [18]. The need for trunk reservation in telephone networks is described in [1]. Some mathematical models and results related to trunk reservation are given in [12,14].

The rest of the paper is organized as follows. In section 2 we provide a construction of the stochastic processes of interest to us in a manner that facilitates our proofs. We also state our main theorems. In section 3 we prove some properties of the reflection mapping that we use. Some preliminary results are proved in section 4. Sections 5–8 contain, respectively, the proofs of theorems 1–4.

2. Statement of main results

Consider an N-server queue with K types of customers and no additional waiting room. The customers of type i, $i = 1, \dots, K$, arrive in a Poisson process of rate λ_i^N, have exponentially distributed service times with mean μ_i^{-1}, require A_i servers simultaneously for their entire service, and are accepted if the number of idle servers at their arrival time is not less than $R_i^N \geqslant A_i$. (Preemption of customers in service is not allowed.) For type i customers, we denote by $B_i^N(t)$ the number of arrivals by t; $D_i^N(t)$, the number of service completions by t; and $Q_i^N(t)$, the number in the system at t. We assume that all the processes under consideration are right-continuous with left limits.

The queue-length processes satisfy the equations

$$Q_i^N(t) = Q_i^N(0) + \int_0^t \mathbf{1}\left(\sum_{j=1}^K A_j Q_j^N(s-) \leqslant N - R_i^N\right) \mathrm{d}B_i^N(s) - D_i^N(t),$$
$$1 \leqslant i \leqslant K, \tag{2.1}$$

where $Q_j^N(s-)$ denotes the left-hand limit. By hypothesis, $B_i^N = (B_i^N(t),\ t \geqslant 0)$, $i = 1, 2, \dots, K$, are Poisson processes with rate λ_i^N. For the D_i^N we use the representation

$$D_i^N(t) = \sum_{j=1}^{\lfloor N/A_i \rfloor} \int_0^t \mathbf{1}\big(Q_i^N(s-) \geqslant j\big)\, \mathrm{d}S_{i,j}(s), \quad i = 1, 2, \dots, K, \tag{2.2}$$

where $S_{i,j} = (S_{i,j}(t),\ t \geqslant 0)$, $j = 1, 2, \dots, \lfloor N/A_i \rfloor$, $i = 1, 2, \dots, K$, are independent Poisson processes with respective rates μ_i, $i = 1, 2, \dots, K$.

We assume that $Q^N(0) = (Q_1^N(0), \dots, Q_K^N(0))$, B_i^N, $S_{i,j}$, $i = 1, 2, \dots, K$, $j = 1, \dots, \lfloor N/A_i \rfloor$, are mutually independent. Also we assume that

$$\frac{A_i \lambda_i^N}{\mu_i} = \alpha_i N + \beta_i \sqrt{N}, \quad i = 1, 2, \dots, K, \tag{2.3}$$

with some α_i and β_i such that

$$\sum_{i=1}^K \alpha_i = 1, \quad \alpha_i > 0, \quad i = 1, \dots, K, \tag{2.4}$$

and $-\infty < \beta_i < \infty$, $1 \leqslant i \leqslant K$, and denote

$$\beta = \sum_{i=1}^K \beta_i. \tag{2.5}$$

With an infinite number of servers, the expected number of type i calls in the system in equilibrium is λ_i^N/μ_i. With these quantities used for centering, we next introduce

$$X_i^N(t) = \frac{A_i}{\sqrt{N}}\left(Q_i^N(t) - \frac{\lambda_i^N}{\mu_i}\right), \quad i = 1, 2, \dots, K,$$
$$X^N(t) = \big(X_1^N(t), \dots, X_K^N(t)\big), \qquad X^N = \big(X^N(t),\ t \geqslant 0\big). \tag{2.6}$$

In the theorems below convergence in distribution for processes is understood as weak convergence of their laws in an appropriate Skorohod space. A sequence of processes is called C-tight if the sequence of their laws is tight and any limit point is a law of a continuous process [9]. The Skorohod space of right-continuous functions with left limits taking values in $\mathbb{R}^K$ is denoted by $D([0,\infty),\mathbb{R}^K)$. Let $(W_i(t),\ t \geqslant 0)$, $i = 1,\dots,K$, be independent standard Brownian motions.

Using (2.3)–(2.6) a bit of algebra yields that

$$\sum_{i=1}^{K} A_i Q_i^N(t) \leqslant N \quad \text{is equivalent to} \quad \sum_{i=1}^{K} X_i^N(t) + \beta \leqslant 0.$$

Let

$$\Theta = \left\{ x = (x_1,\dots,x_K) \in \mathbb{R}^K\colon\ \sum_{i=1}^{K} x_i + \beta \leqslant 0 \right\},$$

and let $\partial\Theta$ denote the boundary of Θ.

Theorem 1. Suppose that $R_i^N/\sqrt{N} \to 0$ as $N \to \infty$, $1 \leqslant i \leqslant K$. Assume that the random vectors $X^N(0)$ converge in distribution to random vector $X_0 = (X_{1,0}, \dots, X_{K,0}) \in \Theta$. Then the sequence $\{X^N,\ N \geqslant 1\}$ is C-tight in $D([0,\infty),\mathbb{R}^K)$, and if $X = ((X_1(t),\dots,X_K(t)),\ t \geqslant 0)$ is a limit point in distribution of $\{X^N,\ N \geqslant 1\}$, then P-a.s.

$$\begin{aligned}
&X_i(t) = X_{i,0} - \mu_i \int_0^t X_i(s)\,\mathrm{d}s + \sqrt{2A_i\alpha_i\mu_i}\,W_i(t) - \phi_i(t), \quad i = 1,\dots,K,\\
&X(t) = \bigl(X_1(t),\dots,X_K(t)\bigr) \in \Theta,\\
&\phi_i(t), \quad i = 1,\dots,K, \text{ are nondecreasing and continuous}, \quad \phi_i(0) = 0,\\
&\phi_i(t) = \int_0^t \mathbf{1}\bigl(X(s) \in \partial\Theta\bigr)\,\mathrm{d}\phi_i(s), \quad i = 1,\dots,K.
\end{aligned}$$

The above equations do not completely specify the process X. In particular, the boundary "reflection" terms $\phi_i(\cdot)$ are not uniquely determined. Loosely speaking, any inward pointing "direction of reflection" is allowed above (even one that is time and state dependent). The next theorem introduces additional conditions that define X uniquely.

Theorem 2. Suppose that

$$\sup_N R_i^N < \infty, \quad i = 1,\dots,K-1, \quad \lim_{N\to\infty} R_K^N = \infty \quad \text{and} \quad \lim_{N\to\infty} R_K^N/\sqrt{N} = 0.$$

Assume that the random vectors $X^N(0)$ converge in distribution to the random vector $X_0 = (X_{1,0},\dots,X_{K,0}) \in \Theta$. Then the sequence $\{X^N,\ N \geqslant 1\}$ converges in distribution in $D([0,\infty),\mathbb{R}^K)$ as $N \to \infty$, to the process $X = ((X_1(t),\dots,X_K(t)),\ t \geqslant 0)$ defined by

$$X_i(t) = X_{i,0} - \mu_i \int_0^t X_i(s)\,\mathrm{d}s + \sqrt{2A_i\alpha_i\mu_i}\,W_i(t), \quad 1 \leqslant i \leqslant K-1, \tag{2.7}$$

$$X_K(t) = X_{K,0} - \mu_K \int_0^t X_K(s)\,\mathrm{d}s + \sqrt{2A_K\alpha_K\mu_K}\,W_K(t) - \phi_K(t), \tag{2.8}$$
$$X(t) = \big(X_1(t), \ldots, X_K(t)\big) \in \Theta,$$
$$\phi_K(t) \text{ is nondecreasing and continuous}, \quad \phi_K(0) = 0,$$
$$\phi_K(t) = \int_0^t \mathbf{1}\big(X(s) \in \partial\Theta\big)\,\mathrm{d}\phi_K(s).$$

The limit process X can also be described as a (multidimensional) diffusion process with instantaneous reflection on the half-space Θ characterized by a state-dependent drift vector whose ith component, $d_i(\mathbf{x})$, $1 \leqslant i \leqslant K$, is given by $d_i(\mathbf{x}) = -\mu_i x_i$, a diagonal infinitesimal covariance matrix whose ith component is $2A_i\alpha_i\mu_i$, and a constant reflection direction $-e_K$ (where $e_K \in \mathbb{R}^K$ is a unit vector in direction K). As shown in the proof of theorem 2, the reflection direction that arises here is actually due to a "boundary-layer" effect as opposed to simply the behavior of the process on the boundary. The latter is more typical in queueing applications. The boundary layer consists of a narrow strip (whose width converges to zero) with a large drift (that grows unboundedly).

Let

$$Y^N(t) = -\sum_{i=1}^{K} X_i^N(t) - \beta, \tag{2.9}$$

and $Y^N = (Y^N(t),\ t \geqslant 0)$. With $\mu_i = \mu$, $1 \leqslant i \leqslant K$, the above result simplifies, with the limit process being one-dimensional.

Corollary 1. Under the assumptions of theorem 2, and with the additional condition that $\mu_i = \mu$, $1 \leqslant i \leqslant K$, the sequence $\{Y^N,\ N \geqslant 1\}$ converges in distribution in $D([0,\infty),\mathbb{R})$ as $N \to \infty$ to the process $Y = (Y(t),\ t \geqslant 0)$ defined by $Y_0 = -\sum_{i=1}^K X_{i,0} - \beta$ and

$$Y(t) = Y_0 - \mu \int_0^t Y(s)\,\mathrm{d}s - \mu\beta t - \sum_{i=1}^{K} \sqrt{2A_i\alpha_i\mu_i}\,W_i(t) + \phi_K(t), \quad Y(t) \geqslant 0,$$
$$\phi_K(t) \text{ is nondecreasing and continuous}, \quad \phi_K(0) = 0,$$
$$\phi_K(t) = \int_0^t \mathbf{1}\big(Y(s) = 0\big)\,\mathrm{d}\phi_K(s).$$

The limit process Y arising here is a (one-dimensional) diffusion process with instantaneous reflection on the half-line $[0,\infty)$ characterized by a state-dependent drift $\mathrm{d}(x) = -\mu x$ and an infinitesimal variance $\sum_{i=1}^K 2A_i\alpha_i\mu_i$. This is in fact a reflected Ornstein–Uhlenbeck process. The stationary distribution of this process is straightforward to obtain [17].

We next consider the preemptive-priority-to-type-i-calls scheme. All calls are admitted if there are enough idle servers. When a type i, $i = 1,\ldots,K-1$, call arrives to find an insufficient number of idle servers, it preempts as many type K calls as is required to have at least A_i idle servers. We assume that type K calls are not preempted for an arriving type i call if the number of idle servers plus the number serving type K calls is less than A_i.

For this system, we use the same assumptions as above for arrival and service processes. Let $\overline{Q}_i^N(t)$, $1 \leqslant i \leqslant K$, denote the number of type i calls in the system at t, and let

$$\overline{X}_i^N(t) = \frac{A_i}{\sqrt{N}}\left(\overline{Q}_i^N(t) - \frac{\lambda_i^N}{\mu_i}\right), \quad i = 1, \dots, K,$$
$$\overline{X}^N(t) = \left(\overline{X}_1^N(t), \dots, \overline{X}_K^N(t)\right), \quad \overline{X}^N = \left(\overline{X}^N(t),\ t \geqslant 0\right).$$

Theorem 3. Assume that the random vectors $\overline{X}^N(0)$ converge in distribution to the random vector $X_0 = (X_{1,0}, \dots, X_{K,0}) \in \Theta$. Then the sequence $\{\overline{X}^N,\ N \geqslant 1\}$ converges in distribution in $D([0,\infty), R^K)$ as $N \to \infty$ to the process X of theorem 2.

For a certain subset of parameter values we prove that the trunk reservation policy considered in theorem 2 is asymptotically optimal. In order to discuss the issue of optimality we need to first introduce costs into our model. Although our result is only for a subset of cases, we develop the cost formulation for the general case. Suppose that each time a type i call is blocked we incur a cost of c_i. Fix a sequence of trunk reservation policies that satisfy the hypotheses of theorem 2, and let $C^N(t)$ denote the total cost incurred up to time t by the Nth policy. We compare this to $C^{\pi_N,N}(t)$, the total cost incurred up to time t by the policy $\pi_N \in \Pi(N)$. (The set of policies we consider, $\Pi(N)$, is defined in section 8.) Let $\widehat{C}^N(t) = N^{-1/2}C^N(t)$ and $\widehat{C}^{\pi_N,N}(t) = N^{-1/2}C^{\pi_N,N}(t)$. We assume that the costs $c_1, \dots, c_K$ are such that

$$\frac{c_K\mu_K}{A_K} \leqslant \frac{c_i\mu_i}{A_i}, \quad 1 \leqslant i < K. \tag{2.10}$$

Intuitively, this condition makes type K the least expensive to block. We prove asymptotic optimality under a more restrictive condition, namely,

$$\frac{c_K}{A_K} \leqslant \frac{c_i}{A_i}, \quad 1 \leqslant i \leqslant K, \tag{2.11a}$$

and

$$\mu_K \leqslant \mu_i, \quad 1 \leqslant i \leqslant K. \tag{2.11b}$$

Note that if $\mu_i = \mu$, $1 \leqslant i \leqslant K$, then (2.11) holds when (2.10) holds. The following result is proved in section 8.

Theorem 4. If (2.11) holds, and $X^N(0)$ converges in distribution to X_0, then for every nondecreasing, bounded and continuous function $f:\mathbb{R}_+\to\mathbb{R}_+$, $\lim_{N\to\infty} Ef(\widehat{C}^N(t))$ exists, and

$$\varliminf_{N\to\infty} Ef\big(\widehat{C}^{\pi_N,N}(t)\big) \geqslant \lim_{N\to\infty} Ef\big(\widehat{C}^N(t)\big).$$

Although we prove theorem 4 under the conditions (2.11), we conjecture that our sequence of trunk reservation policies is asymptotically optimal (possibly in a weaker sense than that of theorem 4) whenever (2.10) holds.

3. Properties of the reflection mapping

The following lemma is proved in the same manner as in [6, lemma 1].

Lemma 3.1. For given functions $z(t)$ and $a(t)$, where $0 \leqslant a(t) < \infty$, suppose that a continuous function $\phi(t)$ with $\phi(0) \leqslant z(0)\vee 0$, satisfies for all $0 \leqslant s \leqslant t$, the inequality

$$\phi(t)-\phi(s) \leqslant \int_s^t a(u)\mathbf{1}\big(\phi(u)\leqslant z(u)\big)\,\mathrm{d}u.$$

Then for $0 \leqslant t < \infty$,

$$\phi(t) \leqslant \sup_{s\leqslant t} z(s) \vee 0.$$

Lemma 3.2. For given functions $x^n(t)$, $f^n(t)$, and $\varepsilon^n(t)$ let the functions $y^n(t)$ satisfy the equation

$$y^n(t) = x^n(t) + \phi^n(t), \tag{3.1}$$

with

$$\phi^n(t) = n\int_0^t f^n(s)\mathbf{1}\big(y^n(s) < \varepsilon^n(s)\big)\,\mathrm{d}s.$$

If for every $T > 0$,

$$\varliminf_{n\to\infty} \inf_{t\leqslant T} f^n(t) > 0, \qquad \lim_{n\to\infty} \sup_{t\leqslant T} \big|\varepsilon^n(t)\big| = 0, \tag{3.2}$$

$$\lim_{n\to\infty} \sup_{t\leqslant T} \big|x^n(t)-x(t)\big| = 0, \tag{3.3}$$

where $x(t)$ is a continuous function with $x(0) \geqslant 0$, then, for every $T > 0$,

$$\lim_{n\to\infty} \sup_{t\leqslant T} \big|y^n(t)-y(t)\big| = 0,$$
$$\lim_{n\to\infty} \sup_{t\leqslant T} \big|\phi^n(t)-\phi(t)\big| = 0,$$

where $(y(t), \phi(t),\ t \geqslant 0)$ is the solution of the Skorohod problem for $(x(t),\ t \geqslant 0)$ with reflection at 0, i.e.,

$$\begin{aligned}&y(t) = x(t) + \phi(t), \quad y(t) \geqslant 0,\\&\phi(t) \text{ is nondecreasing, continuous}, \quad \phi(0) = 0,\\&\phi(t) = \int_0^t \mathbf{1}\bigl(y(s) = 0\bigr)\,\mathrm{d}\phi(s).\end{aligned}$$

Proof. We proceed along the lines of the proof of theorem 2 in [6].

We first prove that for $0 < T < 1$

$$\varliminf_{n\to\infty} \inf_{t\leqslant T} y^n(t) \geqslant 0. \tag{3.4}$$

Denote

$$\bar\phi^n(t) = n\int_0^t f^n(s)\mathbf{1}\bigl(y^n(s) \geqslant \varepsilon^n(s)\bigr)\,\mathrm{d}s.$$

Then, since

$$y^n(t) = x^n(t) + n\int_0^t f^n(s)\,\mathrm{d}s - \bar\phi^n(t), \tag{3.5}$$

we have that

$$\bar\phi^n(t) = n\int_0^t f^n(s)\mathbf{1}\left(\bar\phi^n(s) \leqslant x^n(s) + n\int_0^s f^n(u)\,\mathrm{d}u - \varepsilon^n(s)\right)\mathrm{d}s,$$

hence, by lemma 3.1, for n large enough,

$$\bar\phi^n(t) \leqslant \sup_{s\leqslant t}\left(x^n(s) + n\int_0^s f^n(u)\,\mathrm{d}u - \varepsilon^n(s)\right)\vee 0, \quad t \leqslant T,$$

so, again with the use of (3.5),

$$\begin{aligned}y^n(t) \geqslant{}& x^n(t) + n\int_0^t f^n(u)\,\mathrm{d}u - \sup_{s\leqslant t}\left(x^n(s) + n\int_0^s f^n(u)\,\mathrm{d}u\right)\vee 0\\&- \sup_{s\leqslant t}|\varepsilon^n(s)|, \quad t \leqslant T.\end{aligned} \tag{3.6}$$

Since $x^n(t)$ is bounded in n on $[0, T]$ and $\inf_{t\leqslant T} f^n(t) > 0$ for n large, we have that, for any $\delta > 0$ and n large,

$$x^n(t) + n\int_0^t f^n(u)\,\mathrm{d}u \geqslant \sup_{s\leqslant (t-\delta)\vee 0}\left(x^n(s) + n\int_0^s f^n(u)\,\mathrm{d}u\right)\vee 0, \quad t \leqslant T.$$

Hence, by (3.6), for n large enough,

$$y^n(t) \geqslant \inf_{(t-\delta)\vee 0\leqslant s\leqslant t}\left(x^n(t) - x^n(s) + n\int_s^t f^n(u)\,\mathrm{d}u\right)\wedge 0 - \sup_{s\leqslant t}\bigl|\varepsilon^n(t)\bigr|, \quad t \leqslant T.$$

Therefore,

$$\inf_{t\leqslant T} y^n(t) \geqslant -\sup_{\substack{|s-t|\leqslant\delta\\0\leqslant s,t\leqslant T}} \bigl|x^n(t)-x^n(s)\bigr| - \sup_{t\leqslant T}\bigl|\varepsilon^n(t)\bigr|.$$

The right-hand side goes to 0 as $n\to\infty$ and $\delta\to 0$ by (3.3) and (3.2). Inequality (3.4) is proved.

We now check that the sequence $\{\phi^n(\cdot),\ n\geqslant 1\}$ is compact for the local uniform topology, i.e., check the conditions of Arzelà–Ascoli's theorem:

$$\lim_{\delta\to 0}\overline{\lim_{n\to\infty}}\sup_{\substack{|s-t|<\delta\\ s,t\leqslant T}}\bigl|\phi^n(t)-\phi^n(s)\bigr| = 0. \tag{3.7}$$

Since for $t\geqslant s\geqslant t_0$,

$$\begin{aligned}&\bigl(\phi^n(t)-\phi^n(t_0)\bigr)-\bigl(\phi^n(s)-\phi^n(t_0)\bigr)\\ &\quad = n\int_s^t f^n(u)\mathbf{1}\bigl(\phi^n(u)-\phi^n(t_0)<\varepsilon^n(u)-x^n(u)-\phi^n(t_0)\bigr)\,\mathrm{d}u,\end{aligned}$$

lemma 3.1 yields

$$\begin{aligned}\phi^n(t)-\phi^n(t_0) &\leqslant \sup_{t_0\leqslant s\leqslant t}\bigl(\varepsilon^n(s)-x^n(s)-\phi^n(t_0)\bigr)\vee 0\\ &\leqslant \bigl(-x^n(t_0)-\phi^n(t_0)\bigr)\vee 0+\sup_{s\leqslant t}\bigl|\varepsilon^n(s)\bigr|+\sup_{t_0\leqslant s\leqslant t}\bigl|x^n(s)-x^n(t_0)\bigr|.\end{aligned}$$

Since $-x^n(t_0)-\phi^n(t_0)=-y^n(t_0)$, we obtain that

$$\sup_{\substack{|s-t|\leqslant\delta\\ s,t\leqslant T}}\bigl|\phi^n(t)-\phi^n(s)\bigr|\leqslant -\inf_{t\leqslant T}\bigl(y^n(t)\wedge 0\bigr)+\sup_{t\leqslant T}\bigl|\varepsilon^n(t)\bigr|+\sup_{\substack{|s-t|\leqslant\delta\\ s,t\leqslant T}}\bigl|x^n(t)-x^n(s)\bigr|.$$

The right-hand side goes to 0 as $n\to\infty$ and $\delta\to 0$ by (3.4) and the hypotheses. Convergence (3.7) is proved.

Let (n') be a subsequence such that $\phi^{n'}(\cdot)\to\tilde{\phi}(\cdot)$ uniformly on bounded intervals for some continuous increasing function $\tilde{\phi}=(\tilde{\phi}(t),\ t\geqslant 0)$, $\tilde{\phi}(0)=0$. Equation (3.1) and limit (3.3) imply that

$$\lim_{n'\to\infty}\sup_{t\leqslant T}\bigl|y^{n'}(t)-\tilde{y}(t)\bigr|=0,\quad T>0,$$

where $\tilde{y}(t)$, $t\geqslant 0$, is a continuous function, equal to $x(0)$ at 0, and

$$\tilde{y}(t)=x(t)+\tilde{\phi}(t).$$

It is left to show that $\tilde{y}(\cdot)$ is the reflection of $x(\cdot)$ at 0. The fact that $\tilde{y}(\cdot)$ is nonnegative follows by (3.4). So the proof is completed by checking that $\tilde{\phi}$ increases only when $\tilde{y}(t)=0$, i.e.,

$$\tilde{\phi}(t)=\int_0^t\mathbf{1}\bigl(\tilde{y}(s)=0\bigr)\,\mathrm{d}\tilde{\phi}(s). \tag{3.8}$$

This is done as in [6]. Indeed, by the hypotheses,

$$\phi^n(t) = \int_0^t \mathbf{1}\bigl(y^n(s) \leqslant \varepsilon^n(s)\bigr)\, \mathrm{d}\phi^n(s),$$

hence since $\varepsilon^n(\cdot) \to 0$ uniformly on bounded intervals,

$$\phi^n(t) \leqslant \int_0^t \mathbf{1}\bigl(y^n(s) \leqslant \varepsilon\bigr)\, \mathrm{d}\phi^n(s), \quad t \leqslant T,$$

for any $\varepsilon > 0$ and all n large enough. Furthermore, since $y^{n'}(\cdot) \to \tilde{y}(\cdot)$ uniformly on bounded intervals,

$$\varlimsup_{n' \to \infty} \phi^{n'}(t) \leqslant \varlimsup_{n' \to \infty} \int_0^t \mathbf{1}\bigl(\tilde{y}(s) \leqslant 2\varepsilon\bigr)\, \mathrm{d}\phi^{n'}(s), \quad t \leqslant T,$$

and the latter is not greater than $\int_0^t \mathbf{1}(\tilde{y}(s) \leqslant 2\varepsilon)\, \mathrm{d}\tilde{\phi}(s)$ since $\phi^{n'}(\cdot) \to \tilde{\phi}(\cdot)$ and the set $\{s \leqslant T\colon\ \tilde{y}(s) \leqslant 2\varepsilon\}$ is closed. Hence

$$\tilde{\phi}(t) \leqslant \int_0^t \mathbf{1}\bigl(\tilde{y}(s) \leqslant 2\varepsilon\bigr)\, \mathrm{d}\tilde{\phi}(s),$$

which implies by the arbitrariness of ε that

$$\tilde{\phi}(t) \leqslant \int_0^t \mathbf{1}\bigl(\tilde{y}(s) = 0\bigr)\, \mathrm{d}\tilde{\phi}(s),$$

proving (3.8). The lemma is proved. □

Lemma 3.3. 1. Suppose that $(x(t),\ t \geqslant 0)$ is real-valued, nonnegative, nondecreasing, and right continuous. If a function $(y(t),\ t \geqslant 0)$ satisfies the equation

$$y(t) = x(t) - a \int_0^t y(s)\, \mathrm{d}s,$$

where $a \geqslant 0$, then $y(t) \geqslant 0,\ t \geqslant 0$.

2. Suppose that the real-valued functions $(x(t),\ t \geqslant 0)$ and $(x'(t),\ t \geqslant 0)$ are right continuous with left hand limits, and $(\phi(t),\ t \geqslant 0)$, $(\phi'(t),\ t \geqslant 0)$ are nondecreasing and right continuous with $\phi(0) = \phi'(0) = 0$. Suppose in addition that $(y(t),\ t \geqslant 0)$ and $(y'(t),\ t \geqslant 0)$ are nonnegative and satisfy the equations

$$y(t) = x(t) - a \int_0^t y(s)\, \mathrm{d}s + \phi(t), \qquad y'(t) = x'(t) - a \int_0^t y'(s)\, \mathrm{d}s + \phi'(t),$$

where $a \geqslant 0$. If the function $(x(t) - x'(t),\ t \geqslant 0)$ is nonnegative and nondecreasing, and

$$\phi(t) = \int_0^t \mathbf{1}\bigl(y(s) = 0\bigr)\, \mathrm{d}\phi(s), \quad t \geqslant 0,$$

then $\phi'(t) \geqslant \phi(t)$, $t \geqslant 0$.

Proof. Part 1 follows by the fact that

$$y(t) = \mathrm{e}^{-at} x(t) + a\mathrm{e}^{-at} \int_0^t \big(x(t) - x(s)\big)\, \mathrm{e}^{as}\, \mathrm{d}s.$$

For part 2, introduce $\hat{y}(t),\ t \geqslant 0$, by

$$\hat{y}(t) = x(t) - a \int_0^t \hat{y}(s)\, \mathrm{d}s + \phi'(t). \tag{3.9}$$

We prove that

$$\hat{y}(t) \geqslant y(t), \quad t \geqslant 0, \tag{3.10}$$

which implies the claimed relationship because we can write

$$\phi'(t) - \phi(t) = \hat{y}(t) - y(t) + a \int_0^t \big[\hat{y}(s) - y(s)\big]\mathrm{d}s.$$

Suppose the contrary, i.e., that $\hat{y}(t_0) < y(t_0)$ for some $t_0 > 0$ and let

$$t_1 = \inf\big\{t < t_0\colon\ \hat{y}(s) < y(s) \text{ for all } s \in [t, t_0]\big\}.$$

Since

$$\phi(t) = -\left[\inf_{s \leqslant t} \left(x(s) - a \int_0^s y(u)\, \mathrm{d}u\right) \wedge 0\right], \quad t \geqslant 0,$$

the function $\phi(t)$ does not jump when $x(t)$ jumps upwards which easily implies that positive jumps of $y(t)$ are not greater than the respective jumps of $\hat{y}(t)$, hence $t_1 < t_0$. Next, since $y'(t) \leqslant \hat{y}(t)$ by (3.9) and part 1 of the lemma, the latter function is nonnegative, so since $y(t) > \hat{y}(t)$ on $(t_1, t_0]$, we conclude that $y(t)$ is positive for these t, therefore, by the definition of $\phi(t)$, it follows that $\phi(t_1) = \phi(t_0)$. Also, it can be seen as follows that $\hat{y}(t_1) = y(t_1)$. When $t_1 = 0$, $\hat{y}(t_1) = y(t_1)$ by the hypotheses. If $t_1 > 0$ and $\hat{y}(t_1) < y(t_1)$, then $y(t_1) > 0$, so $\Delta y(t_1) = \Delta x(t) \leqslant \Delta\hat{y}(t)$, which contradicts the definition of t_1. Putting all these facts together, we can write

$$\begin{aligned}\hat{y}(t_0) &= \hat{y}(t_1) + \big(x(t_0) - x(t_1)\big) - a \int_{t_1}^{t_0} \hat{y}(s)\, \mathrm{d}s + \big(\phi'(t_0) - \phi'(t_1)\big) \\ &> y(t_1) + \big(x(t_0) - x(t_1)\big) - a \int_{t_1}^{t_0} y(s)\, \mathrm{d}s + \big(\phi(t_0) - \phi(t_1)\big) = y(t_0).\end{aligned}$$

The contradiction proves (3.10) and, hence, the lemma. □

4. Preliminary results

Let $\mathcal{F}^N(t)$ be the σ-field generated by $Q_i^N(0)$, $B_i^N(s)$, $S_{i,j}(s)$, $i = 1, \ldots, K$, $j = 1, \ldots, \lfloor N/A_i \rfloor$, $s \leqslant t$, and the family of P-null sets, and let $\mathbb{F}^N = (\mathcal{F}^N(t),\ t \geqslant 0)$

be the corresponding filtration (note that it is right continuous since B_i^N and $S_{i,j}$ are piecewise constant [5]). For $i = 1, \dots, K$, let

$$M_{i,B}^N(t) = \frac{A_i}{\sqrt{N}}\bigl(B_i^N(t) - \lambda_i^N t\bigr), \tag{4.1}$$

$$M_{i,D}^N(t) = \frac{A_i}{\sqrt{N}}\left(D_i^N(t) - \mu_i \int_0^t Q_i^N(s)\,\mathrm{d}s\right), \tag{4.2}$$

$$M_{i,A}^N(t) = \int_0^t \mathbf{1}\left(\sum_{j=1}^K X_j^N(s-) \leqslant -\beta - \frac{R_i^N}{\sqrt{N}}\right)\mathrm{d}M_{i,B}^N(s). \tag{4.3}$$

With this notation and (2.3)–(2.6), the equations (2.1) take the form

$$X_i^N(t) = X_i^N(0) - \mu_i \int_0^t X_i^N(s)\,\mathrm{d}s + M_{i,A}^N(t) - M_{i,D}^N(t) - \sqrt{N}\gamma_i^N \int_0^t \mathbf{1}\left(\sum_{j=1}^K X_j^N(s) > -\beta - \frac{R_i^N}{\sqrt{N}}\right)\mathrm{d}s, \tag{4.4}$$

where $\gamma_i^N = A_i\lambda_i^N/N$. Note that by our assumptions

$$\gamma_i^N \to \alpha_i\mu_i > 0, \quad i = 1, \dots, K. \tag{4.5}$$

For the results from martingale theory used in the rest of the paper, we refer the reader to [9,15].

Lemma 4.1. The processes $M_{i,A}^N = (M_{i,A}^N(t),\ t \geqslant 0)$ and $M_{i,D}^N = (M_{i,D}^N(t),\ t \geqslant 0)$, $i = 1, \dots, K$, are $\mathbb{F}^N$-locally square integrable martingales that are pairwise orthogonal and have respective predictable quadratic variation processes

$$\langle M_{i,A}^N\rangle(t) = A_i^2 \frac{\lambda_i^N}{N} \int_0^t \mathbf{1}\left(\sum_{j=1}^K X_j^N(s) \leqslant -\beta - \frac{R_i^N}{\sqrt{N}}\right)\mathrm{d}s,$$

$$\langle M_{i,D}^N\rangle(t) = A_i^2 \mu_i \int_0^t \frac{Q_i^N(s)}{N}\,\mathrm{d}s.$$

Proof. The Poisson processes B_i^N and $S_{i,j}$, $i = 1, \dots, K$, $j = 1, \dots, \lfloor N/A_i \rfloor$, have the respective $\mathbb{F}^N$-compensators $(\lambda_i^N t,\ t \geqslant 0)$ and $(\mu_i t,\ t \geqslant 0)$. Since by (2.2) and (4.2)

$$M_{i,D}^N(t) = \frac{A_i}{\sqrt{N}} \sum_{j=1}^{\lfloor N/A_i \rfloor} \int_0^t \mathbf{1}\bigl(Q_i^N(s-) \geqslant j\bigr)\,\mathrm{d}\bigl(S_{i,j}(s) - \mu_i s\bigr),$$

and the processes B_i^N, $S_{i,j}$, $i = 1, \dots, K$, $j = 1, \dots, \lfloor N/A_i \rfloor$, are mutually independent, we derive by the property of quadratic variation processes [9,15] and using also (4.1) that $M_{i,D}^N$ and $M_{i,B}^N = (M_{i,B}^N(t),\ t \geqslant 0)$, $i = 1, \ldots, K$, are locally square integrable martingales with

$$\langle M_{i,D}^N\rangle(t) = A_i^2 \mu_i \int_0^t \frac{Q_i^N(s)}{N}\,\mathrm{d}s, \qquad \langle M_{i,B}^N\rangle(t) = A_i^2 \frac{\lambda_i^N}{N} t,$$

$$\langle M_{i,D}^N, M_{j,B}^N\rangle(t) = 0, \qquad \langle M_{i,D}^N, M_{j,D}^N\rangle(t) = 0, \qquad \langle M_{i,B}^N, M_{j,B}^N\rangle(t) = 0, \quad i \neq j.$$

The formula for the quadratic variation of $M_{i,A}^N$ and the pairwise orthogonality of $M_{i,A}^N$, $M_{i,D}^N$, $i = 1, \ldots, K$, follows by (4.3). $\square$

In the sequel, we repeatedly use the following version of the Lenglart–Rebolledo inequality [9, lemma I.3.30; 15, theorem I.9.3].

Lemma 4.2. Let $M = (M(t),\ t \geqslant 0)$ be a locally square integrable martingale defined on a stochastic basis $(\Omega, \mathcal{F}, \mathbb{F}, P)$ with $M(0) = 0$ and let $\langle M\rangle = (\langle M\rangle(t),\ t \geqslant 0)$ be its predictable quadratic variation process. Then, for any finite $\mathbb{F}$-stopping time τ and any $a > 0$, $b > 0$,

$$P\Big(\sup_{t \leqslant \tau} \big|M(t)\big| \geqslant a\Big) \leqslant \frac{b}{a^2} + P\big(\langle M\rangle(\tau) \geqslant b\big).$$

Lemma 4.3. The sequences $\{M_{i,A}^N,\ N \geqslant 1\}$, $\{M_{i,D}^N,\ N \geqslant 1\}$, $i = 1, \ldots, K$, are C-tight.

Proof. Tightness in D follows in a standard manner by Aldous's condition [9,15] if we apply the Lenglart–Rebolledo inequality (lemma 4.2), lemma 4.1, (2.3) and the fact that $Q_i^N(t) \leqslant N$, $i = 1, \ldots, K$. The C-tightness follows since the jumps of $M_{i,A}^N$, $M_{i,D}^N$ are of size $A_i/\sqrt{N}$, $i = 1, \ldots, K$ [9]. $\square$

By (2.9) and (4.4),

$$Y^N(t) = Y^N(0) + \sum_{i=1}^K \mu_i \int_0^t X_i^N(s)\,\mathrm{d}s + M_D^N(t) - M_A^N(t) + \phi^N(t), \tag{4.6}$$

where

$$M_D^N(t) = \sum_{i=1}^K M_{i,D}^N(t), \tag{4.7}$$

$$M_A^N(t) = \sum_{i=1}^K M_{i,A}^N(t), \tag{4.8}$$

and

$$\phi^N(t) = \sqrt{N} \sum_{i=1}^N \gamma_i^N \int_0^t \mathbf{1}\left(Y^N(s) < \frac{R_i^N}{\sqrt{N}}\right)\mathrm{d}s. \tag{4.9}$$

The Ito-type formula given by the next lemma plays a major part in the proof of theorem 2. It can be derived by using general results for semimartingales [9,15], but we find a direct approach to be more appropriate.

Lemma 4.4. The process Y^N is nonnegative and if $f(x)$, $x \geqslant 0$, is a real-valued Borel function, then

$$\begin{aligned}
&f\big(Y^N(t)\big) \\
&\quad = f\big(Y^N(0)\big) + \sqrt{N}\sum_{i=1}^K \frac{\mu_i}{A_i}\int_0^t \left[f\left(Y^N(s)+\frac{A_i}{\sqrt{N}}\right) - f\big(Y^N(s)\big)\right] X_i^N(s)\,\mathrm{d}s \\
&\qquad + \sum_{i=1}^K \lambda_i^N \int_0^t \left[f\left(Y^N(s)+\frac{A_i}{\sqrt{N}}\right) - f\big(Y^N(s)\big)\right]\mathbf{1}\left(Y^N(s) < \frac{R_i^N}{\sqrt{N}}\right)\mathrm{d}s \\
&\qquad + \sum_{i=1}^K \lambda_i^N \int_0^t \left[f\left(Y^N(s)+\frac{A_i}{\sqrt{N}}\right) + f\left(Y^N(s)-\frac{A_i}{\sqrt{N}}\right)\right. \\
&\qquad\qquad \left. -2f\big(Y^N(s)\big)\right]\mathbf{1}\left(Y^N(s) \geqslant \frac{R_i^N}{\sqrt{N}}\right)\mathrm{d}s \\
&\qquad + \sum_{i=1}^K \frac{\sqrt{N}}{A_i}\int_0^t \left[f\left(Y^N(s-)+\frac{A_i}{\sqrt{N}}\right) - f\big(Y^N(s-)\big)\right]\mathrm{d}M_{i,D}^N(s) \\
&\qquad + \sum_{i=1}^K \frac{\sqrt{N}}{A_i}\int_0^t \left[f\left(Y^N(s-)-\frac{A_i}{\sqrt{N}}\right) - f\big(Y^N(s-)\big)\right]\mathrm{d}M_{i,A}^N(s).
\end{aligned}$$

Proof. The fact that $Y^N(t)$ is nonnegative follows by (2.3)–(2.6), (2.9) and the inequality

$$\sum_{i=1}^N A_i Q_i^N(t) \leqslant N.$$

Next, since $Y^N(t)$ is piecewise constant and right continuous,

$$f\big(Y^N(t)\big) = f\big(Y^N(0)\big) + \sum_{0<s\leqslant t}\big(f\big(Y^N(s)\big) - f\big(Y^N(s-)\big)\big).$$

By (4.6)–(4.9) and (4.1)–(4.3), the jumps of Y^N are

$$\begin{aligned}
\Delta Y^N(t) &\equiv Y^N(t) - Y^N(t-) = \Delta M_D^N(t) - \Delta M_A^N(t) \\
&= \sum_{i=1}^K \frac{A_i}{\sqrt{N}}\Delta D_i^N(t) - \sum_{i=1}^K \frac{A_i}{\sqrt{N}}\mathbf{1}\left(Y^N(t-) \geqslant \frac{R_i^N}{\sqrt{N}}\right)\Delta B_i^N(t).
\end{aligned}$$

Since the jumps of D_i^N and B_i^N are of size 1 and are disjoint, we get

$$f\big(Y^N(t)\big) = f\big(Y^N(0)\big) + \sum_{i=1}^K \int_0^t \left[f\left(Y^N(s-) + \frac{A_i}{\sqrt{N}}\right) - f\big(Y^N(s-)\big)\right] \mathrm{d}D_i^N(s)$$
$$+ \sum_{i=1}^K \int_0^t \left[f\left(Y^N(s-) - \frac{A_i}{\sqrt{N}}\right) - f\big(Y^N(s-)\big)\right]$$
$$\times \mathbf{1}\left(Y^N(s-) \geqslant \frac{R_i^N}{\sqrt{N}}\right) \mathrm{d}B_i^N(s). \tag{4.10}$$

If we take into account that by (4.1) and (4.2),

$$B_i^N(t) = \lambda_i^N t + \frac{\sqrt{N}}{A_i} M_{i,B}^N(t), \quad i = 1, \ldots, K,$$
$$D_i^N(t) = \mu_i \int_0^t Q_i^N(s)\,\mathrm{d}s + \frac{\sqrt{N}}{A_i} M_{i,D}^N(t), \quad i = 1, \ldots, K$$

and by (2.6),

$$Q_i^N(t) = \frac{\lambda_i^N}{\mu_i} + \frac{\sqrt{N}}{A_i} X_i^N(t),$$

and use also (4.3) and (2.9), equation (4.10) assumes the form required by the lemma. The lemma is proved. □

5. Proof of theorem 1

We denote below by $\xrightarrow{\mathrm{d}}$ convergence in distribution, and by $\xrightarrow{\mathrm{P}}$ convergence in probability. We prove first that

$$\lim_{a\to\infty} \overline{\lim_{N\to\infty}} P\left(\sup_{t\leqslant T} \left|X_i^N(t)\right| > a\right) = 0, \quad T > 0, \quad i = 1, \ldots, K. \tag{5.1}$$

By (2.9) and (4.6)–(4.9), we have that, for $0 \leqslant s \leqslant t$,

$$\phi^N(t) - \phi^N(s) \leqslant \sqrt{N} \sum_{i=1}^N \gamma_i^N \int_s^t \mathbf{1}\left(Y^N(u) < \frac{\max_{1\leqslant i\leqslant K} R_i^N}{\sqrt{N}}\right) \mathrm{d}u$$
$$= \sqrt{N} \sum_{i=1}^N \gamma_i^N \int_s^t \mathbf{1}\left(\phi^N(u) < -Y^N(0) - \sum_{i=1}^K \mu_i \int_0^u X_i^N(v)\,\mathrm{d}v - M_D^N(u) + M_A^N(u) + \frac{\max_{1\leqslant i\leqslant K} R_i^N}{\sqrt{N}}\right) \mathrm{d}u.$$

By lemma 3.1,

$$\phi^N(t) \leqslant \sup_{s\leqslant t} \left[\frac{\max_{1\leqslant i\leqslant K} R_i^N}{\sqrt{N}} - Y^N(0) - \sum_{i=1}^K \mu_i \int_0^s X_i^N(u)\,\mathrm{d}u - M_D^N(s) + M_A^N(s)\right] \vee 0$$

$$\leqslant \frac{\max_{1\leqslant i\leqslant K} R_i^N}{\sqrt{N}} + \left|Y^N(0)\right| + \sum_{i=1}^K \mu_i \int_0^t \left|X_i^N(u)\right| \mathrm{d}u$$
$$+ \sup_{s\leqslant t} \left|M_D^N(s)\right| + \sup_{s\leqslant t} \left|M_A^N(s)\right|, \quad t \geqslant 0. \tag{5.2}$$

Since by (4.4), (2.9) and (4.9),

$$\left|X_i^N(t)\right| \leqslant \left|X_i^N(0)\right| + \mu_i \int_0^t \left|X_i^N(s)\right| \mathrm{d}s + \left|M_{i,A}^N(t)\right| + \left|M_{i,D}^N(t)\right| + \phi^N(t),$$
$$i = 1, \dots, K,$$

we then obtain by (5.2), using also (2.9), (4.7) and (4.8), that

$$\sum_{i=1}^K \left|X_i^N(t)\right| \leqslant 2 \sum_{i=1}^K \left|X_i^N(0)\right| + \frac{\max_{1\leqslant i\leqslant K} R_i^N}{\sqrt{N}} + |\beta| + 2 \sum_{i=1}^K \sup_{s\leqslant t} \left|M_{i,A}^N(s)\right|$$
$$+ 2 \sum_{i=1}^K \sup_{s\leqslant t} \left|M_{i,D}^N(s)\right| + 2 \max_{1\leqslant j\leqslant K} \mu_j \int_0^t \sum_{i=1}^K \left|X_i^N(s)\right| \mathrm{d}s.$$

Gronwall's inequality then yields

$$\sum_{i=1}^K \left|X_i^N(t)\right| \leqslant \left(2 \sum_{i=1}^K \left|X_i^N(0)\right| + \frac{\max_{1\leqslant i\leqslant K} R_i^N}{\sqrt{N}} + |\beta| + 2 \sum_{i=1}^K \sup_{s\leqslant t} \left|M_{i,A}^N(s)\right|\right.$$
$$\left. + 2 \sum_{i=1}^K \sup_{s\leqslant t} \left|M_{i,D}^N(s)\right|\right) \mathrm{e}^{2 \max_{1\leqslant i\leqslant K} \mu_i t},$$

implying (5.1) in view of lemma 4.3, the hypothesis that $R_i^N/\sqrt{N} \to 0$, and the convergence $X^N(0) \xrightarrow{\mathrm{d}} X_0$.

By (5.1), (2.9), (4.6) and lemma 4.3, we have that

$$\lim_{a\to\infty} \overline{\lim_{N\to\infty}} P\big(\phi_N(t) > a\big) = 0, \tag{5.3}$$

so that by (4.9) and (4.5)

$$\int_0^t \mathbf{1}\left(Y^N(s) < \frac{R_i^N}{\sqrt{N}}\right) \mathrm{d}s \xrightarrow{\mathrm{P}} 0, \quad i = 1, \dots, K. \tag{5.4}$$

The latter yields by lemma 4.1, (2.9) and (2.3)

$$\left\langle M_{i,A}^N\right\rangle(t) \xrightarrow{\mathrm{P}} A_i \alpha_i \mu_i t, \quad i = 1, \dots, K. \tag{5.5}$$

Also (5.1) implies, by (2.3) and (2.6), that $A_i Q_i^N(t)/N \xrightarrow{\mathrm{P}} \alpha_i$, hence by lemma 4.1 and Lebesgue's dominated convergence theorem

$$\left\langle M_{i,D}^N\right\rangle(t) \xrightarrow{\mathrm{P}} A_i \alpha_i \mu_i t, \quad i = 1, \dots, K. \tag{5.6}$$

Since by lemma 4.1 the locally square integrable martingales $M_{i,A}^N$, $M_{i,D}^N$, $i = 1, \dots, K$, are pairwise orthogonal and have jumps not greater than $\max_{1\leqslant i\leqslant K} A_i/\sqrt{N}$, we obtain from (5.5) and (5.6) (see, e.g., [15]) that

$$\bigl(M_{i,A}^N, M_{i,D}^N\bigr)_{1\leqslant i\leqslant K} \xrightarrow{\mathrm{d}} \bigl(\sqrt{A_i\alpha_i\mu_i}\, W_{i,A}, \sqrt{A_i\alpha_i\mu_i}\, W_{i,D}\bigr)_{1\leqslant i\leqslant K},$$

where $W_{i,A}$, $W_{i,D}$, $1 \leqslant i \leqslant K$, are independent standard Brownian motions. Therefore,

$$\bigl(M_{i,A}^N - M_{i,D}^N\bigr)_{1\leqslant i\leqslant K} \xrightarrow{\mathrm{d}} \bigl(\sqrt{2A_i\alpha_i\mu_i}\, W_i\bigr)_{1\leqslant i\leqslant K}, \tag{5.7}$$

and, in view of the notation (4.7) and (4.8),

$$M_D^N - M_A^N \xrightarrow{\mathrm{d}} \sqrt{2\textstyle\sum_{i=1}^K A_i\alpha_i\mu_i}\, W, \tag{5.8}$$

where W is a standard Brownian motion. Let

$$Z^N(t) = Y^N(0) + \sum_{i=1}^K \mu_i \int_0^t X_i^N(s)\,\mathrm{d}s + M_D^N(t) - M_A^N(t).$$

By (2.9), (5.1) and (5.8), the sequence $\{Z^N,\ N \geqslant 1\}$, with $Z^N = (Z^N(t),\ t \geqslant 0)$, is C-tight. By (4.6) we also have that

$$Y^N(t) = Z^N(t) + \phi^N(t).$$

Note that by (4.9)

$$\phi^N(t) = \sqrt{N}\int_0^t \Biggl[\sum_{\substack{i=1\\ i\neq i_0}}^K \gamma_i^N \mathbf{1}\biggl(Y^N(s) < \frac{R_i^N}{\sqrt{N}}\biggr) + \gamma_{i_0}^N\Biggr] \mathbf{1}\biggl(Y^N(s) < \frac{\max_{1\leqslant i\leqslant K} R_i^N}{\sqrt{N}}\biggr)\mathrm{d}s,$$

where $R_{i_0}^N = \max_{1\leqslant i\leqslant K} R_i^N$. The term in brackets is not less than $\min_{1\leqslant i\leqslant K}\gamma_i^N$, and by (4.5)

$$\varliminf_N \min_{1\leqslant i\leqslant K} \gamma_i^N > 0.$$

So we can apply lemma 3.2 to Y^N, Z^N, ϕ^N to get by Prohorov's theorem and Skorohod's embedding, in view of the C-tightness of $\{Z^N,\ N \geqslant 1\}$, that the sequence $\{(Y^N,\ Z^N,\ \phi^N),\ n \geqslant 1\}$ is weakly relatively sequentially compact with continuous limits. Furthermore, if (Z, Y, ϕ) with $Z = (Z(t),\ t \geqslant 0)$, $Y = (Y(t),\ t \geqslant 0)$ and $\phi = (\phi(t),\ t \geqslant 0)$ is a limit point in distribution of $\{(Z^N, Y^N, \phi^N),\ N \geqslant 1\}$, then P-a.s.

$$\begin{aligned}
&Y(t) = Z(t) + \phi(t), \quad Y(t) \geqslant 0,\\
&\phi(t) \text{ is nondecreasing and continuous}, \quad \phi(0) = 0,\\
&\phi(t) = \int_0^t \mathbf{1}\bigl(Y(s) = 0\bigr)\,\mathrm{d}\phi(s).
\end{aligned} \tag{5.9}$$

The weak relative sequential compactness with continuous limits of $\{\phi^N,\ N \geqslant 1\}$ implies, by Prohorov's theorem and (4.9), that the sequence $\{(\phi_1^N, \dots, \phi_K^N),\ N \geqslant 1\}$ is C-tight, where

$$\phi_i^N(t) = \sqrt{N}\gamma_i^N \int_0^t \mathbf{1}\left(Y^N(s) < \frac{R_i^N}{\sqrt{N}}\right) \mathrm{d}s. \tag{5.10}$$

Then by (4.4), (5.1), (5.7) and the convergence $X^N(0) \xrightarrow{\mathrm{d}} X_0$, the sequence $\{(X_1^N, \dots, X_K^N, \phi_1^N, \dots, \phi_K^N),\ N \geqslant 1\}$ is C-tight and any limit point in distribution $(X_1, \dots, X_K, \phi_1, \dots, \phi_K)$ satisfies P-a.s.

$$X_i(t) = X_{i,0} - \mu_i \int_0^t X_i(s)\,\mathrm{d}s + \sqrt{2A_i\alpha_i\mu_i}\, W_i(t) - \phi_i(t), \quad i = 1, \dots, K,$$

where $\phi_i(t)$ are nondecreasing and continuous, $\phi_i(0) = 0$. Since by (2.9), $-X_1 - X_2 - \cdots - X_K - \beta$ is a limit point in distribution of $\{Y^N,\ N \geqslant 1\}$ and as we have seen, $Y(t) \geqslant 0$ for any such point, we conclude that $(X_1, \dots, X_K) \in \Theta$.

Finally, since $\phi_1 + \cdots + \phi_K$ is a limit point in distribution of ϕ^N by (4.9) and (5.10), we derive from (5.9) that

$$\begin{aligned}&\int_0^t \mathbf{1}\big(\big(X_1(s), \dots, X_K(s)\big) \in \Theta \setminus \partial\Theta\big)\, \mathrm{d}\phi_i(s) \\ &\quad\leqslant \int_0^t \mathbf{1}\big(Y(s) > 0\big)\, \mathrm{d}(\phi_1 + \cdots + \phi_K)(s) = 0.\end{aligned}$$

The theorem is proved. □

6. Proof of theorem 2

We first prove that, in the notation of theorem 1,

$$\phi_i^N(t) \xrightarrow{\mathrm{P}} 0 \quad \text{as } N \to \infty,\ t > 0,\ i = 1, \dots, K-1, \tag{6.1}$$

after which the proof is straightforward. The proof of (6.1) is carried out with the use of the Ito formula of lemma 4.4. Let $R^N = R_K^N - 1 - \max_{1 \leqslant i \leqslant K} A_i$, and $p > 0$ to be chosen later. We assume that N is large enough so that

$$R^N > 2 \max_{1 \leqslant i \leqslant K} A_i \quad \text{and} \quad R^N > \max_{1 \leqslant i \leqslant K-1} R_i^N.$$

Also, let $k_N(x)$, $x \geqslant 0$, be a continuous function with values in $[0, 1]$, such that $k_N(0) = 1$, and $k_N(x) = 0$, for $x \geqslant 1/\sqrt{N}$. Define the function h_N by

$$\begin{aligned}h_N(x) &= \mathrm{e}^{-p\sqrt{N}x}, \quad 0 \leqslant x \leqslant \frac{R^N}{\sqrt{N}}, \\ h_N(x) &= \mathrm{e}^{-pR^N} - p\sqrt{N}\,\mathrm{e}^{-pR^N}\left(x - \frac{R^N}{\sqrt{N}}\right) \\ &\quad + p^2 N\,\mathrm{e}^{-pR^N} \int_{R^N/\sqrt{N}}^x \int_{R^N/\sqrt{N}}^y k_N\left(z - \frac{R^N}{\sqrt{N}}\right) \mathrm{d}z\,\mathrm{d}y, \quad x > \frac{R^N}{\sqrt{N}}.\end{aligned}$$

It is easy to see that h_N is twice differentiable and

$$\begin{aligned}|h_N'(x)| &\leqslant (p^2+p)\sqrt{N}\,\mathrm{e}^{-p(R^N-2\max_{1\leqslant i\leqslant K}A_i)},\\ &x\geqslant \frac{R^N-2\max_{1\leqslant i\leqslant K}A_i}{\sqrt{N}},\end{aligned}\tag{6.2}$$

$$|h_N''(x)|\leqslant p^2N\,\mathrm{e}^{-pR^N},\quad x\geqslant\frac{R^N}{\sqrt{N}},\tag{6.3}$$

$$h_N''(x)=0,\quad x\geqslant\frac{R^N+1}{\sqrt{N}}.\tag{6.4}$$

We now estimate the terms in the Ito formula of lemma 4.4 with $f=h_N$. Since Y^N is tight by theorem 1 and (2.9), and h_N grows at most linearly,

$$\frac{h_N(Y^N(t))}{\sqrt{N}}\xrightarrow{\mathrm{P}}0,\qquad\frac{h_N(Y^N(0))}{\sqrt{N}}\xrightarrow{\mathrm{P}}0.\tag{6.5}$$

Next, denote the terms on the right of the Ito formula starting from the second by I_1, I_2, I_3, I_4 and I_5, respectively.

Since $R^N\to\infty$ and the X_i^N are tight by theorem 1, estimate (6.2) implies that

$$\begin{aligned}&\int_0^t\left[h_N\left(Y^N(s)+\frac{A_i}{\sqrt{N}}\right)-h_N\big(Y^N(s)\big)\right]X_i^N(s)\mathbf{1}\left(Y^N(s)>\frac{R^N-A_i}{\sqrt{N}}\right)\mathrm{d}s\xrightarrow{\mathrm{P}}0,\\ &\quad i=1,\dots,K.\end{aligned}\tag{6.6}$$

Next, since h_N is bounded in N on $[0,R^N/\sqrt{N}]$, the X_i^N are tight and by (5.4),

$$\int_0^t\mathbf{1}\left(Y^N(s)<\frac{R_K^N}{\sqrt{N}}\right)\mathrm{d}s\xrightarrow{\mathrm{P}}0,$$

we have that

$$\begin{aligned}&\int_0^t\left[h_N\left(Y^N(s)+\frac{A_i}{\sqrt{N}}\right)-h_N\big(Y^N(s)\big)\right]X_i^N(s)\mathbf{1}\left(Y^N(s)\leqslant\frac{R^N-A_i}{\sqrt{N}}\right)\mathrm{d}s\xrightarrow{\mathrm{P}}0,\\ &\quad i=1,\dots,K,\end{aligned}$$

which, by (6.6), yields

$$\frac{I_1}{\sqrt{N}}\xrightarrow{\mathrm{P}}0.\tag{6.7}$$

We now prove that

$$\frac{I_4}{\sqrt{N}}\xrightarrow{\mathrm{P}}0,\qquad\frac{I_5}{\sqrt{N}}\xrightarrow{\mathrm{P}}0.\tag{6.8}$$

By the Lenglart–Rebolledo inequality (lemma 4.2) and since the predictable quadratic variation process of $M_{i,D}^N$ is continuous by lemma 4.1, for $\eta > 0$, $\varepsilon > 0$,

$$P\biggl(\biggl|\int_0^t \biggl[h_N\biggl(Y^N(s-) + \frac{A_i}{\sqrt{N}}\biggr) - h_N\bigl(Y^N(s-)\bigr)\biggr]\,\mathrm{d}M_{i,D}^N(s)\biggr| > \eta\biggr)$$
$$\leqslant \frac{\varepsilon}{\eta^2} + P\biggl(\int_0^t \biggl[h_N\biggl(Y^N(s) + \frac{A_i}{\sqrt{N}}\biggr) - h_N\bigl(Y^N(s)\bigr)\biggr]^2 \mathrm{d}\bigl\langle M_{i,D}^N\bigr\rangle(s) > \varepsilon\biggr),$$
$$i = 1,\dots,K. \tag{6.9}$$

The form of $\langle M_{i,D}^N\rangle$ in lemma 4.1 and the fact that $A_i Q_i^N(s)/N \leqslant 1$, yield, by the argument used when proving (6.7),

$$\int_0^t \biggl[h_N\biggl(Y^N(s) + \frac{A_i}{\sqrt{N}}\biggr) - h_N\bigl(Y^N(s)\bigr)\biggr]^2 \mathrm{d}\bigl\langle M_{i,D}^N\bigr\rangle(s) \xrightarrow{\mathrm{P}} 0, \quad i = 1,\dots,K,$$

which by (6.9) implies the first convergence in (6.8). The second one is proved similarly, with the use of the form of $\langle M_{i,A}^N\rangle$ in lemma 4.1.

Convergences (6.5), (6.7) and (6.8) imply that

$$\frac{I_2 + I_3}{\sqrt{N}} \xrightarrow{\mathrm{P}} 0. \tag{6.10}$$

We represent the latter sum as

$$I_2 + I_3 = I_1' + I_2' + I_3' + I_4', \tag{6.11}$$

where

$$I_1' = \sum_{i=1}^{K-1} \lambda_i^N \int_0^t \biggl[h_N\biggl(Y^N(s) + \frac{A_i}{\sqrt{N}}\biggr) - h_N\bigl(Y^N(s)\bigr)\biggr]\mathbf{1}\biggl(Y^N(s) < \frac{R_i^N}{\sqrt{N}}\biggr)\mathrm{d}s$$
$$+ \lambda_K^N \int_0^t \biggl[h_N\biggl(Y^N(s) + \frac{A_K}{\sqrt{N}}\biggr) - h_N\bigl(Y^N(s)\bigr)\biggr]$$
$$\times \mathbf{1}\biggl(Y^N(s) < \frac{\min_{1\leqslant i\leqslant K-1} R_i^N}{\sqrt{N}}\biggr)\mathrm{d}s,$$

$$I_2' = \int_0^t \Biggl(\sum_{i=1}^{K-1} \lambda_i^N \biggl[h_N\biggl(Y^N(s) + \frac{A_i}{\sqrt{N}}\biggr) + h_N\biggl(Y^N(s) - \frac{A_i}{\sqrt{N}}\biggr) - 2h_N\bigl(Y^N(s)\bigr)\biggr]$$
$$\times \mathbf{1}\biggl(Y^N(s) \geqslant \frac{R_i^N}{\sqrt{N}}\biggr) + \lambda_K^N\biggl[h_N\biggl(Y^N(s) + \frac{A_K}{\sqrt{N}}\biggr) - h_N\bigl(Y^N(s)\bigr)\biggr]$$
$$\times \mathbf{1}\biggl(Y^N(s) \geqslant \frac{\min_{1\leqslant i\leqslant K-1} R_i^N}{\sqrt{N}}\biggr)\Biggr)\mathbf{1}\biggl(Y^N(s) < \frac{R^N - \max_{1\leqslant i\leqslant K} A_i}{\sqrt{N}}\biggr)\mathrm{d}s,$$

$$I_3' = \int_0^t \left(\sum_{i=1}^{K-1} \lambda_i^N \left[h_N\left(Y^N(s) + \frac{A_i}{\sqrt{N}}\right) + h_N\left(Y^N(s) - \frac{A_i}{\sqrt{N}}\right) - 2h_N\big(Y^N(s)\big) \right] \right.$$
$$\left. + \lambda_K^N \left[h_N\left(Y^N(s) + \frac{A_K}{\sqrt{N}}\right) - h_N\big(Y^N(s)\big) \right] \right)$$
$$\times \mathbf{1}\left(\frac{R^N - \max_{1 \leqslant i \leqslant K} A_i}{\sqrt{N}} \leqslant Y^N(s) < \frac{R_K^N}{\sqrt{N}} \right) \mathrm{d}s,$$
$$I_4' = \sum_{i=1}^{K} \lambda_i^N \int_0^t \left[h_N\left(Y^N(s) + \frac{A_i}{\sqrt{N}}\right) + h_N\left(Y^N(s) - \frac{A_i}{\sqrt{N}}\right) - 2h_N\big(Y^N(s)\big) \right]$$
$$\times \mathbf{1}\left(Y^N(s) \geqslant \frac{R_K^N}{\sqrt{N}} \right) \mathrm{d}s.$$

By (6.4) and since $R_K^N - A_i \geqslant R^N + 1$,

$$I_4' = 0. \tag{6.12}$$

Consider I_3'. By (6.2), we have that

$$\sqrt{N} \left| \int_0^t \left[h_N\left(Y^N(s) + \frac{A_K}{\sqrt{N}}\right) - h_N\big(Y^N(s)\big) \right] \right.$$
$$\left. \times \mathbf{1}\left(\frac{R^N - \max_{1 \leqslant i \leqslant K} A_i}{\sqrt{N}} \leqslant Y^N(s) < \frac{R_K^N}{\sqrt{N}} \right) \mathrm{d}s \right|$$
$$\leqslant \big(p^2 + p\big)\big(\gamma_K^N\big)^{-1} A_K \, \mathrm{e}^{-p(R^N - \max_{1 \leqslant i \leqslant K} A_i)} \phi^N(t),$$

where $\phi^N(t)$ is defined in (4.9).

By (4.5), (5.3) and the convergence $R^N \to \infty$ as $N \to \infty$, we conclude that

$$\sqrt{N} \int_0^t \left[h_N\left(Y^N(s) + \frac{A_K}{\sqrt{N}}\right) - h_N\big(Y^N(s)\big) \right]$$
$$\times \mathbf{1}\left(\frac{R^N - \max_{1 \leqslant i \leqslant K} A_i}{\sqrt{N}} \leqslant Y^N(s) < \frac{R_K^N}{\sqrt{N}} \right) \mathrm{d}s \xrightarrow{\mathrm{P}} 0.$$

A similar argument shows that

$$\sqrt{N} \int_0^t \left[h_N\left(Y^N(s) + \frac{A_i}{\sqrt{N}}\right) + h_N\left(Y^N(s) - \frac{A_i}{\sqrt{N}}\right) - 2h_N\big(Y^N(s)\big) \right]$$
$$\times \mathbf{1}\left(\frac{R^N - \max_{1 \leqslant i \leqslant K} A_i}{\sqrt{N}} \leqslant Y^N(s) < \frac{R_K^N}{\sqrt{N}} \right) \mathrm{d}s \xrightarrow{\mathrm{P}} 0,$$

so that recalling (2.3)

$$\frac{I_3'}{\sqrt{N}} \xrightarrow{\mathrm{P}} 0. \tag{6.13}$$

Turning to I_2', we have, by the definition of h_N, that

$$I_2' = N\int_0^t \mathrm{e}^{-p\sqrt{N}Y^N(s)}\left[\sum_{i=1}^{K-1}\frac{\lambda_i^N}{N}\left(\mathrm{e}^{-pA_i}+\mathrm{e}^{pA_i}-2\right)\mathbf{1}\left(Y^N(s)\geqslant\frac{R_i^N}{\sqrt{N}}\right)\right.$$
$$\left.+\frac{\lambda_K^N}{N}\left(\mathrm{e}^{-pA_K}-1\right)\right]\mathbf{1}\left(Y^N(s)\geqslant\frac{\min_{1\leqslant i\leqslant K-1}R_i^N}{\sqrt{N}}\right)$$
$$\times\mathbf{1}\left(Y^N(s)<\frac{R^N-\max_{1\leqslant i\leqslant K}A_i}{\sqrt{N}}\right)\mathrm{d}s. \tag{6.14}$$

By (2.3), as $N\to\infty$,

$$\sum_{i=1}^{K-1}\frac{\lambda_i^N}{N}\left(\mathrm{e}^{-pA_i}+\mathrm{e}^{pA_i}-2\right)+\frac{\lambda_K^N}{N}\left(\mathrm{e}^{-pA_K}-1\right)$$
$$\to\sum_{i=1}^{K-1}\frac{\alpha_i\mu_i}{A_i}\left(\mathrm{e}^{-pA_i}+\mathrm{e}^{pA_i}-2\right)+\frac{\alpha_K\mu_K}{A_K}\left(\mathrm{e}^{-pA_K}-1\right). \tag{6.15}$$

Since the sum over i on the right is of order p^2 for small p and $(\mathrm{e}^{-pA_K}-1)$ is negative and of order p, the limit in (6.15) is negative for all p small enough. Hence, for these p, and N large enough, the term in brackets in the integral in (6.14) is negative, so that $I_2'\leqslant 0$. This and (6.10)–(6.13) imply, for suitable $p>0$, that for all N large enough

$$\frac{I_1'}{\sqrt{N}}\geqslant\delta^N, \tag{6.16}$$

where $\delta^N\xrightarrow{\mathrm{P}}0$ as $N\to\infty$.

By the definition of h_N, for N large, all the integrals in the expression for I_1' are nonpositive, so (6.16) and the definition of h_N yield

$$\frac{\lambda_i^N}{\sqrt{N}}\int_0^t\mathrm{e}^{-p\sqrt{N}Y^N(s)}\left(1-\mathrm{e}^{-pA_i}\right)\mathbf{1}\left(Y^N(s)<\frac{R_i^N}{\sqrt{N}}\right)\mathrm{d}s\xrightarrow{\mathrm{P}}0,\quad i=1,2,\dots,K-1,$$

which is equivalent to (6.1) by (5.10), (2.3), (4.5) and boundedness of the R_i^N.

By (6.1) and theorem 1, any limit point in distribution of $\{X^N,\ N\geqslant 1\}$ satisfies

$$X_i(t)=X_{i,0}-\mu_i\int_0^t X_i(s)\,\mathrm{d}s+\sqrt{2A_i\alpha_i\mu_i}\,W_i(t),\quad i=1,\dots,K-1,$$
$$X_K(t)=X_{K,0}-\mu_K\int_0^t X_K(s)\,\mathrm{d}s+\sqrt{2A_K\alpha_K\mu_K}\,W_K(t)-\phi_K(t),$$
$$X(t)=\left(X_1(t),\dots,X_K(t)\right)\in\Theta,$$
$$\phi_K(t)\text{ is continuous, increasing,}\quad\phi_K(0)=0,$$
$$\phi_K(t)=\int_0^t\mathbf{1}\left(X(s)\in\partial\Theta\right)\mathrm{d}\phi_K(s).$$

The theorem would follow if this specified $X(t)$ uniquely. Note that

$$Y(t) = -\sum_{i=1}^{K} X_i(t) - \beta,$$

so that

$$Y(t) = -\sum_{i=1}^{K} X_{i,0} - \beta + \sum_{i=1}^{K-1} (\mu_i - \mu_K) \int_0^t X_i(s)\,\mathrm{d}s - \mu_K \int_0^t Y(s)\,\mathrm{d}s - \beta\mu_K t$$
$$- \sum_{i=1}^{K} \sqrt{2A_i\alpha_i\mu_i}\, W_i(t) + \phi_K(t), \tag{6.17}$$

$Y(t) \geqslant 0$, and

$$\phi_K(t) = \int_0^t \mathbf{1}\big(Y(s) = 0\big)\,\mathrm{d}\phi_K(s),$$

so that $(X_1(t), \ldots, X_{K-1}(t),\ Y(t))$ is a solution of a semimartingale problem of diffusion type with normal reflection in the domain $R^{K-1} \times R_+$ [15, chapter 10, section 2]. It has a unique solution by [15, theorem 10.2.2]. The theorem is proved. □

Proof of corollary 1. As a consequence of theorem 2, we have $Y^N \xrightarrow{\mathrm{d}} Y$, where Y is defined in (6.17). When $\mu_i = \mu$, $1 \leqslant i \leqslant K$, we can write (6.17) as

$$Y(t) = Y(0) - \mu \int_0^t Y(s)\,\mathrm{d}s - \mu\beta t - \sum_{i=1}^{K} \sqrt{2A_i\alpha_i\mu_i}\, W_i(t) + \phi_K(t).$$

7. Proof of theorem 3

The proof proceeds along the lines of the proof of theorem 1, so we omit some details. The equations for the queue-length processes are easily seen to be

$$\overline{Q}_i^N(t) = \overline{Q}_i^N(0) + \int_0^t \mathbf{1}\left(\sum_{j=1}^{K-1} A_j \overline{Q}_j^N(s-) \leqslant N - A_i \right) \mathrm{d}B_i^N(s) - \overline{D}_i^N(t),$$
$$i = 1, \ldots, K-1, \tag{7.1}$$

$$\overline{Q}_K^N(t) = \overline{Q}_K^N(0) + \int_0^t \mathbf{1}\left(\sum_{j=1}^{K} A_j \overline{Q}_j^N(s-) \leqslant N - A_K \right) \mathrm{d}B_K^N(s) - \widehat{D}_K^N(t), \tag{7.2}$$

where

$$\overline{D}_i^N(t) = \sum_{j=1}^{\lfloor N/A_i \rfloor} \int_0^t \mathbf{1}\big(\overline{Q}_i^N(s-) \geqslant j\big)\,\mathrm{d}S_{i,j}(s), \quad i = 1, \ldots, K,$$

$$\widehat{D}_K^N(t) = \overline{D}_K^N(t) + \sum_{j=1}^{K-1} \int_0^t \ell_j^N(s-)\mathbf{1}\left(\sum_{r=1}^{K} A_r \overline{Q}_r^N(s-) > N - A_j\right) \times \mathbf{1}\left(\sum_{r=1}^{K-1} A_r \overline{Q}_r^N(s-) \leqslant N - A_j\right) \mathrm{d}B_j^N(s),$$

$$\ell_j^N(s) = \left\lceil \frac{\sum_{r=1}^K A_r \overline{Q}_r^N(s) - (N - A_j)}{A_K} \right\rceil \vee 0, \tag{7.3}$$

and $\lceil x \rceil$ is the smallest integer that is not smaller than x. Introducing

$$\overline{M}_{i,A}^N(t) = \frac{A_i}{\sqrt{N}} \int_0^t \mathbf{1}\left(\sum_{j=1}^{K-1} A_j \overline{Q}_j^N(s-) \leqslant N - A_i\right) \left(\mathrm{d}B_i^N(s) - \lambda_i^N \,\mathrm{d}s\right), \quad i = 1, \ldots, K-1,$$

$$\overline{M}_{K,A}^N(t) = \frac{A_K}{\sqrt{N}} \int_0^t \mathbf{1}\left(\sum_{j=1}^{K} A_j \overline{Q}_j^N(s-) \leqslant N - A_K\right) \left(\mathrm{d}B_K^N(s) - \lambda_K^N \,\mathrm{d}s\right),$$

$$\overline{M}_{i,D}^N(t) = \frac{A_i}{\sqrt{N}} \left(\overline{D}_i^N(t) - \mu_i \int_0^t \overline{Q}_i^N(s)\,\mathrm{d}s\right), \quad i = 1, \ldots, K,$$

we reduce (7.1) and (7.2), in analogy with (4.4) to

$$\overline{X}_i^N(t) = \overline{X}_i^N(0) - \mu_i \int_0^t \overline{X}_i^N(s)\,\mathrm{d}s + \overline{M}_{i,A}^N(t) - \overline{M}_{i,D}^N(t) - \sqrt{N}\gamma_i^N \int_0^t \mathbf{1}\left(\sum_{j=1}^{K-1} \overline{X}_j^N(s) > \sqrt{N}\alpha_K - \sum_{j=1}^{K-1} \beta_j - \frac{A_i}{\sqrt{N}}\right)\mathrm{d}s, \quad i = 1, \ldots, K-1, \tag{7.4}$$

$$\overline{X}_K^N(t) = \overline{X}_K^N(0) - \mu_K \int_0^t \overline{X}_K^N(s)\,\mathrm{d}s + \overline{M}_{K,A}^N(t) - \overline{M}_{K,D}^N(t) - \sqrt{N}\gamma_K^N \int_0^t \mathbf{1}\left(\sum_{j=1}^{K} \overline{X}_j^N(s) > -\beta - \frac{A_K}{\sqrt{N}}\right)\mathrm{d}s - \sum_{j=1}^{K-1} \sqrt{N}\gamma_j^N \frac{A_K}{A_j} \int_0^t \ell_j^N(s)\mathbf{1}\left(\sum_{r=1}^{K} \overline{X}_r^N(s) > -\beta - \frac{A_j}{\sqrt{N}}\right) \times \mathbf{1}\left(\sum_{r=1}^{K-1} \overline{X}_r^N(s) \leqslant \sqrt{N}\alpha_K - \sum_{r=1}^{K-1} \beta_r - \frac{A_j}{\sqrt{N}}\right)\mathrm{d}s + \overline{M}_E^N(t), \tag{7.5}$$

where

$$\overline{M}_E^N(t) = \sum_{j=1}^{K-1} \frac{A_K}{\sqrt{N}} \int_0^t \ell_j^N(s-) \mathbf{1}\left(\sum_{r=1}^{K} \overline{X}_r^N(s-) > -\beta - \frac{A_j}{\sqrt{N}} \right) \times \mathbf{1}\left(\sum_{r=1}^{K-1} \overline{X}_r^N(s-) \leqslant \sqrt{N}\alpha_K - \sum_{r=1}^{K-1} \beta_r - \frac{A_j}{\sqrt{N}} \right) \big(\mathrm{d}B_j^N(s) - \lambda_j^N \, \mathrm{d}s \big). \quad (7.6)$$

Also, as in lemmas 4.1 and 4.3,

$$\overline{M}_{i,A}^N = \big(\overline{M}_{i,A}^N(t),\ t \geqslant 0\big), \qquad \overline{M}_{i,D}^N = \big(M_{i,D}^N(t),\ t \geqslant 0\big), \quad i = 1, \ldots, K,$$

are pairwise orthogonal, locally square integrable martingales and the sequences $\{\overline{M}_{i,A}^N,\ N \geqslant 1\}$, $\{\overline{M}_{i,D}^N,\ N \geqslant 1\}$ are C-tight.

We next show that

$$\lim_{N\to\infty} P\left(\int_0^t \mathbf{1}\left(\sum_{j=1}^{K-1} \overline{X}_j^N(s) > \sqrt{N}\alpha_K - \sum_{j=1}^{K-1} \beta_j - \frac{A_i}{\sqrt{N}} \right) \mathrm{d}s > 0 \right) = 0, \quad i = 1, \ldots, K-1. \quad (7.7)$$

Let $\widetilde{X}_i^N(t) = \overline{X}_i^N(t)/\sqrt{N},\ i = 1, \ldots, K-1.$ Since $\alpha_K > 0$, limit (7.7) would follow by

$$\sup_{t \leqslant T} \big|\widetilde{X}_i^N(t)\big| \xrightarrow{\mathrm{P}} 0, \quad T > 0,\ i = 1, \ldots, K-1. \quad (7.8)$$

By (7.4),

$$\widetilde{X}_i^N(t) = \widetilde{X}_i^N(0) - \mu_i \int_0^t \widetilde{X}_i^N(s) \, \mathrm{d}s + \frac{\overline{M}_{i,A}^N(t) - \overline{M}_{i,D}^N(t)}{\sqrt{N}} - \gamma_i^N \int_0^t \mathbf{1}\left(\sum_{j=1}^{K-1} \widetilde{X}_j^N(s) > \alpha_K - \frac{1}{\sqrt{N}} \sum_{j=1}^{K-1} \beta_j - \frac{A_i}{N} \right) \mathrm{d}s, \quad i = 1, \ldots, K-1. \quad (7.9)$$

Since $\{\overline{M}_{i,A}^N,\ N \geqslant 1\}$ and $\{\overline{M}_{i,D}^N,\ N \geqslant 1\}$ are tight,

$$\sup_{t \leqslant T} \frac{\big|\overline{M}_{i,A}^N(t) - \overline{M}_{i,D}^N(t)\big|}{\sqrt{N}} \xrightarrow{\mathrm{P}} 0, \quad i = 1, \ldots, K-1, \quad (7.10)$$

and since $\overline{X}_i^N(0) \xrightarrow{\mathrm{d}} X_{i,0}$, $i = 1, \ldots, K-1$, we have that

$$\widetilde{X}_i^N(0) \xrightarrow{\mathrm{P}} 0, \quad i = 1, \ldots, K-1. \quad (7.11)$$

Also since $\gamma_i^N \to \alpha_i\mu_i$ by (4.5), it is easy to see from (7.9), with the use of Gronwall's inequality, that the sequence $\{(\widetilde{X}_1^N,\dots,\widetilde{X}_{K-1}^N),\ N \geqslant 1\}$ is C-tight in $D([0,\infty),\mathbb{R}^{K-1})$. For $\eta > 0$, let a subsequence (N') be such that

$$\lim_{N'} P\Big(\sup_{t\leqslant T}\big|\widetilde{X}_i^{N'}(t)\big| \geqslant \eta\Big) = \overline{\lim_{N}}\, P\Big(\sup_{t\leqslant T}\big|\widetilde{X}_i^{N}(t)\big| \geqslant \eta\Big), \quad i = 1,\dots,K-1, \tag{7.12}$$

and

$$\big(\widetilde{X}_1^{N'},\dots,\widetilde{X}_{K-1}^{N'}\big) \xrightarrow{\mathrm{d}} \big(\widetilde{X}_1,\dots,\widetilde{X}_{K-1}\big), \tag{7.13}$$

where $\widetilde{X}_i = (\widetilde{X}_i(t),\ t \geqslant 0)$, $i = 1,\dots,K-1$, are continuous processes. Let $\varepsilon \in (0,\alpha_K/4T)$. Introduce, for $x(\cdot) \in D([0,\infty),\ R)$, the first passage times

$$\tau_\varepsilon\big(x(\cdot)\big) = \inf\big(t \geqslant 0\colon\ \big|x(t)\big| + \varepsilon t \geqslant \alpha_K/2\big) \wedge T. \tag{7.14}$$

Then τ_ε is continuous at continuous functions (a proof can be carried out in a manner similar to the proof in [15, theorem 6.2.3]). Denoting

$$\widetilde{X}^{N'} = \big(\widetilde{X}_1^{N'},\dots,\widetilde{X}_{K-1}^{N'}\big), \qquad \widetilde{X} = \big(\widetilde{X}_1,\dots,\widetilde{X}_{K-1}\big),$$
$$\widetilde{Y}^N = \sum_{i=1}^{K-1}\widetilde{X}_i^N, \qquad \widetilde{Y} = \sum_{i=1}^{K-1}\widetilde{X}_i,$$

we derive from (7.13) that

$$\big(\widetilde{X}^{N'},\tau_\varepsilon\big(\widetilde{Y}^{N'}\big)\big) \xrightarrow{\mathrm{d}} \big(\widetilde{X},\tau_\varepsilon\big(\widetilde{Y}\big)\big). \tag{7.15}$$

By (7.14), $\sup_{t<\tau_\varepsilon(x(\cdot))}|x(t)| < \alpha_K/2$ for $x(\cdot) \in D([0,\infty),\mathbb{R})$ (by convention $\sup_\emptyset = 0$), so

$$\sup_{t<\tau_\varepsilon(\widetilde{Y}^N)}\big|\widetilde{Y}^N(t)\big| < \alpha_K - \frac{1}{\sqrt{N}}\sum_{j=1}^{K-1}\beta_j - \frac{A_i}{N}, \quad i = 1,\dots,K-1,$$

for N large enough. Hence by (7.9), for these N,

$$\widetilde{X}_i^N\big(t\wedge\tau_\varepsilon\big(\widetilde{Y}^N\big)\big) = \widetilde{X}_i^N(0) - \mu_i\int_0^{t\wedge\tau_\varepsilon(\widetilde{Y}^N)}\widetilde{X}_i^N(s)\,\mathrm{d}s$$
$$+ \frac{\overline{M}_{i,A}^N(t\wedge\tau_\varepsilon(\widetilde{Y}^N)) - \overline{M}_{i,D}^N(t\wedge\tau_\varepsilon(\widetilde{Y}^N))}{\sqrt{N}}, \quad i = 1,\dots,K-1.$$

Taking the latter limit along the subsequence (N'), we obtain by (7.10), (7.11), (7.15) and the continuity of $\widetilde{X}$ and $\widetilde{Y}$, and the random time change theorem [4, section 17], that P-a.s.

$$\widetilde{X}_i\big(t\wedge\tau_\varepsilon\big(\widetilde{Y}\big)\big) = -\mu_i\int_0^{t\wedge\tau_\varepsilon(\widetilde{Y})}\widetilde{X}_i(s)\,\mathrm{d}s,$$

which obviously implies that

$$\widetilde{X}_i(t) = 0, \quad t \leqslant \tau_\varepsilon(\widetilde{Y}), \quad i = 1, \dots, K-1. \tag{7.16}$$

We now show that $\tau_\varepsilon(\widetilde{Y}) = T$. Assume the contrary. Then by (7.14), $|\widetilde{Y}(\tau_\varepsilon(\widetilde{Y}))| + \varepsilon\tau_\varepsilon(\widetilde{Y}) = \alpha_K/2$. Since by the choice of ε, $\varepsilon\tau_\varepsilon(\widetilde{Y}) \leqslant \varepsilon T \leqslant \alpha_K/4$, it follows that $|\widetilde{Y}(\tau_\varepsilon(\widetilde{Y}))| \geqslant \alpha_K/4 > 0$ which contradicts (7.16) and the fact that $\widetilde{Y} = \sum_{i=1}^{K-1} \widetilde{X}_i$. Thus $\tau_\varepsilon(\widetilde{Y}) = T$, and (7.12), (7.13), (7.16) imply that

$$\lim_N P\Big(\sup_{t\leqslant T} \big|\widetilde{X}_i^N(t)\big| \geqslant \eta\Big) = 0, \quad i = 1, \dots, K-1.$$

completing the proof of (7.8), and hence (7.7).

Introduce, in analogy with the proof of theorem 1,

$$\overline{Y}^N(t) = -\sum_{i=1}^{K} \overline{X}_i^N(t) - \beta. \tag{7.17}$$

Also denote

$$\begin{aligned} \varepsilon_i^N(t) &= \sqrt{N}\gamma_i^N \int_0^t \mathbf{1}\Bigg(\sum_{j=1}^{K-1} \overline{X}_j^N(s) > \sqrt{N}\alpha_K - \sum_{j=1}^{K-1}\beta_j - \frac{A_i}{\sqrt{N}}\Bigg)\mathrm{d}s, \\ & i = 1, \dots, K-1, \end{aligned} \tag{7.18}$$

$$\begin{aligned} \delta_i^N(t) &= \sqrt{N}\gamma_i^N \frac{A_K}{A_i}\int_0^t \ell_i^N(s)\mathbf{1}\Bigg(\sum_{j=1}^{K} \overline{X}_j^N(s) > -\beta - \frac{A_i}{\sqrt{N}}\Bigg) \\ &\quad \times \mathbf{1}\Bigg(\sum_{j=1}^{K-1} \overline{X}_j^N(s) > \sqrt{N}\alpha_K - \sum_{j=1}^{K-1}\beta_j - \frac{A_i}{\sqrt{N}}\Bigg)\mathrm{d}s, \\ & i = 1, \dots, K-1. \end{aligned} \tag{7.19}$$

Limit (7.7) obviously implies that

$$\sup_{t\leqslant T} \big|\varepsilon_i^N(t)\big| \xrightarrow{\mathrm{P}} 0, \qquad \sup_{t\leqslant T} \big|\delta_i^N(t)\big| \xrightarrow{\mathrm{P}} 0, \quad i = 1, \dots, K-1. \tag{7.20}$$

Also, by (7.4), (7.5), and (7.17),

$$\begin{aligned} \overline{Y}^N(t) &= \overline{Y}^N(0) + \sum_{i=1}^{K} \mu_i \int_0^t \overline{X}_i^N(s)\,\mathrm{d}s \\ &\quad + \sum_{i=1}^{K} \big(\overline{M}_{i,D}^N(t) - \overline{M}_{i,A}^N(t)\big) + \sum_{i=1}^{K-1} \varepsilon_i^N(t) - \sum_{i=1}^{K-1} \delta_i^N(t) - \overline{M}_E^N(t) \\ &\quad + \sqrt{N}\Bigg[\sum_{i=1}^{K-1} \gamma_i^N \frac{A_K}{A_i}\int_0^t \ell_i^N(s)\mathbf{1}\bigg(\overline{Y}^N(s) < \frac{A_i}{\sqrt{N}}\bigg)\mathrm{d}s \end{aligned}$$

$$+\gamma_K^N \int_0^t \mathbf{1}\left(\overline{Y}^N(s) < \frac{A_K}{\sqrt{N}}\right) \mathrm{d}s\Bigg]. \tag{7.21}$$

For the sequel, note that, as it follows by (7.3),

$$0 \leqslant \ell_i^N(s) \leqslant \left\lceil \frac{A_i}{A_K} \right\rceil, \quad i = 1, \ldots, K-1. \tag{7.22}$$

Next, since by (7.6), $\overline{M}_E^N = (\overline{M}_E^N(t),\ t \geqslant 0)$ is a locally square integrable martingale with the predictable quadratic variation process

$$\langle \overline{M}_E^N \rangle(t) = \sum_{i=1}^{K-1} \frac{A_K^2 \lambda_i^N}{N} \int_0^t \left(\ell_i^N(s)\right)^2 \mathbf{1}\left(\overline{Y}^N(s) < \frac{A_i}{\sqrt{N}}\right) \times \mathbf{1}\left(\sum_{j=1}^{K-1} \overline{X}_j^N(s) \leqslant \sqrt{N}\alpha_K - \sum_{j=1}^{K-1} \beta_j - \frac{A_i}{\sqrt{N}}\right) \mathrm{d}s, \tag{7.23}$$

it is easy to see, in view of (7.22), as in lemma 4.3, that the sequence $\{\overline{M}_E^N,\ N \geqslant 1\}$ is C-tight.

Limit (7.20), tightness of $\{\overline{M}_{i,A}^N,\ N \geqslant 1\}$ and $\{\overline{M}_{i,D}^N,\ N \geqslant 1\}$, and the convergence $\overline{X}^N(0) \xrightarrow{\mathrm{d}} X_0$ imply by (7.4) that the sequences $\{\overline{X}_i^N,\ N \geqslant 1\}$, $1 \leqslant i \leqslant K-1$, are C-tight. Moreover, applying the argument of the proof of (5.1) and (5.4) in theorem 1 to $\overline{X}_i^N$, $i = 1, \ldots, K$, and $\overline{Y}^N$, and using (7.5), (7.17)–(7.22) and C-tightness of the sequences $\{\overline{M}_{i,A}^N,\ N \geqslant 1\}$, $\{\overline{M}_{i,D}^N,\ N \geqslant 1\}$ and $\{\overline{M}_E^N,\ N \geqslant 1\}$, we can prove that

$$\lim_{a\to\infty} \varlimsup_{N} P\Big(\sup_{t\leqslant T} \left|\overline{X}_K^N(t)\right| > a\Big) = 0, \tag{7.24}$$

and

$$\int_0^t \mathbf{1}\left(\overline{Y}^N(s) < \frac{A_i}{\sqrt{N}}\right) \mathrm{d}s \xrightarrow{\mathrm{P}} 0, \quad i = 1, \ldots, K. \tag{7.25}$$

This, as in the proof of theorem 1, yields

$$\left(\overline{M}_{i,A}^N - \overline{M}_{i,D}^N\right)_{i=1,\ldots,K} \xrightarrow{\mathrm{d}} \left(\sqrt{2A_i\alpha_i\mu_i}\, W_i\right)_{i=1,\ldots,K}. \tag{7.26}$$

Also, it follows from (7.25), (7.22), (2.3) and (7.23), that $\langle \overline{M}_E^N \rangle(t) \xrightarrow{\mathrm{P}} 0$, and the Lenglart–Rebolledo inequality (lemma 4.2) implies that

$$\sup_{t\leqslant T} \left|\overline{M}_E^N(t)\right| \xrightarrow{\mathrm{P}} 0, \quad T > 0. \tag{7.27}$$

An application of lemma 3.2 to (7.21) shows as in the proof of theorem 1, in view of (7.20), (7.22), (7.24), (7.26) and (7.27), that the sequence $\{(\overline{X}_1^N, \ldots, \overline{X}_K^N),\ N \geqslant 1\}$ is C-tight and any limit point in distribution is as in the assertion of that theorem.

Further, the first convergence in (7.20) implies, by (7.18) and (7.4), that an analog of (6.1) holds so that, by the argument that completed the proof of theorem 2, the limit is the process $(X_1, \dots, X_K)$. The result follows.

8. Asymptotic optimality

We define a policy π in terms of the "controls" $\{(\psi_i^{\pi,N}(t),\ t \geqslant 0),\ 1 \leqslant i \leqslant K\}$ that it generates. Let the (stochastic) control process $(\psi_i^{\pi,N}(t),\ t \geqslant 0)$ be defined such that $\psi_i^{\pi,N}(t-)$ indicates if a type i arrival at time t is blocked: if $\psi_i^{\pi,N}(t-) = 0$, then the customer is accepted, and if $\psi_i^{\pi,N}(t-) = 1$, the customer is blocked. We then can write

$$C^{\pi,N}(t) = \sum_{i=1}^{K} c_i \int_0^t \psi_i^{\pi,N}(s-) \mathrm{d}B_i^N(s).$$

In analogy with (2.1) we can write

$$Q_i^{\pi,N}(t) = Q_i^N(0) + \int_0^t \bigl(1 - \psi_i^{\pi,N}(s-)\bigr) \mathrm{d}B_i^N(s) - D_i^N(t), \quad 1 \leqslant i \leqslant K.$$

Let

$$\phi_i^{\pi,N}(t) = \sqrt{N}\gamma_i^N \int_0^t \psi_i^{\pi,N}(s)\,\mathrm{d}s.$$

Then in analogy with (4.4) we have

$$X_i^{\pi,N}(t) = X_i^N(0) - \mu_i \int_0^t X_i^{\pi,N}(s)\,\mathrm{d}s + M_{i,A}^{\pi,N}(t) - M_{i,D}^{\pi,N}(t) - \phi_i^{\pi,N}(t), \tag{8.1}$$

where

$$M_{i,A}^{\pi,N}(t) = \int_0^t \bigl(1 - \psi_i^{\pi,N}(s-)\bigr)\,\mathrm{d}M_{i,B}^N(s),$$

and

$$M_{i,D}^{\pi,N}(t) = \frac{A_i}{\sqrt{N}} \left(D_i^N(t) - \mu_i \int_0^t Q_i^{\pi,N}(s)\,\mathrm{d}s \right).$$

We are now in a position to define $\Pi(N)$, the set of policies that we consider. A policy $\pi \in \Pi(N)$ if and only if

(i) $\psi_i^{\pi,N}$ is $\mathbb{F}^N$-adapted, $1 \leqslant i \leqslant K$, and

(ii) $\sum_{i=1}^K X_i^{\pi,N}(t) \leqslant -\beta,\ t \geqslant 0$.

These conditions are unrestrictive from a practical point of view. Condition (i) simply requires that the policy not use any information about the future. Condition (ii) requires that accepted customers must find a sufficient number of idle servers to serve them.

Proof of theorem 4. Since easily

$$\widehat{C}^N(t) = \sum_{i=1}^K \frac{c_i}{A_i} \int_0^t \mathbf{1}\left(Y^N(s-) < \frac{R_i^N}{\sqrt{N}}\right) \mathrm{d}M_{i,B}^N(s) + \sum_{i=1}^K \frac{c_i}{A_i} \phi_i^N(t),$$

it follows from the proof of theorem 2 that $\widehat{C}^N(t) \xrightarrow{\mathrm{d}} (c_K/A_k)\phi_K(t)$, so, in particular, for every continuity point x of the distribution of $\phi_K(t)$,

$$\lim_{N\to\infty} P\big(\widehat{C}^N(t) > x\big) = P\left(\frac{c_K}{A_K}\phi_K(t) > x\right). \tag{8.2}$$

We thus prove that, for every $x > 0$,

$$\varliminf_{N\to\infty} P\big(\widehat{C}^{\pi_N,N}(t) > x\big) \geqslant P\left(\frac{c_K}{A_K}\phi_K(t) > x\right). \tag{8.3}$$

Let (N') denote a subsequence that attains the lim inf in (8.3). If, for some $\varepsilon > 0$ and $i = 1, 2, \ldots, K$,

$$\varlimsup_{N'\to\infty} P\left(\left|\int_0^t \psi_i^{\pi_{N'},N'}(s)\,\mathrm{d}s\right| > \varepsilon\right) > 0,$$

an application of the Lenglart–Rebolledo inequality (lemma 4.2) easily implies that $\widehat{C}^{\pi_{N''},N''}(t) \xrightarrow{\mathrm{P}} \infty$ for a subsequence (N'') of (N') , so that (8.3) trivially holds in this case. We now assume that for every $i = 1, 2, \ldots, K$

$$\int_0^t \psi_i^{\pi_{N'},N'}(s)\,\mathrm{d}s \xrightarrow{\mathrm{P}} 0. \tag{8.4}$$

Again by the Lenglart–Rebolledo inequality (lemma 4.2) we then have that, for arbitrary $\varepsilon > 0$,

$$\begin{aligned}&\varliminf_{N\to\infty} P\big(\widehat{C}^{\pi_N,N}(t) > x\big)\\ &\quad\geqslant \varliminf_{N'\to\infty} P\left(\frac{1}{\sqrt{N'}}\sum_{i=1}^K c_i\lambda_i^{N'} \int_0^t \psi_i^{\pi_{N'},N'}(s)\,\mathrm{d}s > x + \varepsilon\right).\end{aligned} \tag{8.5}$$

We now estimate the sum in the right parentheses.

Introduce the processes $(\widehat{X}_i^{\pi_N,N}(t),\ t \geqslant 0),\ 1 \leqslant i \leqslant K-1$, defined by

$$\widehat{X}_i^{\pi_N,N}(t) = X_i^N(0) - \mu_i \int_0^t \widehat{X}_i^{\pi_N,N}(s)\,\mathrm{d}s + M_i^{\pi_N,N}(t), \quad 1 \leqslant i \leqslant K-1,$$

where $M_i^{\pi_N,N}(t) = M_{i,A}^{\pi_N,N}(t) - M_{i,D}^{\pi_N,N}(t)$. Since $\phi_i^{\pi_N,N}(t)$ is nondecreasing, by part 1 of lemma 3.3,

$$X_i^{\pi_N,N}(t) \leqslant \widehat{X}_i^{\pi_N,N}(t), \quad t \geqslant 0,\ 1 \leqslant i \leqslant K-1. \tag{8.6}$$

Let the process $(\widehat{Y}^{\pi_N,N}(t),\ t \geqslant 0)$ be nonnegative and satisfy the equation

$$\widehat{Y}^{\pi_N,N}(t) = Y^N(0) + \sum_{i=1}^{K-1}(\mu_i - \mu_K)\int_0^t \widehat{X}_i^{\pi_N,N}(s)\,\mathrm{d}s - \mu_K\beta t - \sum_{i=1}^{K} M_i^{\pi_N,N}(t) - \mu_K\int_0^t \widehat{Y}^{\pi_N,N}(s)\,\mathrm{d}s + \hat{\phi}_K^{\pi_N,N}(t), \quad (8.7)$$

where $(\hat{\phi}_K^{\pi_N,N}(t),\ t \geqslant 0)$ is nondecreasing, $\hat{\phi}_K^{\pi_N,N}(0) = 0$, and

$$\hat{\phi}_K^{\pi_N,N}(t) = \int_0^t \mathbf{1}\big(\widehat{Y}^{\pi_N,N}(s) = 0\big)\,\mathrm{d}\hat{\phi}_K^{\phi_N,N}(s), \quad t \geqslant 0. \quad (8.8)$$

Existence (and uniqueness) of $(\widehat{Y}^{\pi_N,N}(t),\ t \geqslant 0)$ follows by existence (and uniqueness) of the solution to the corresponding Skorohod problem [22].

Defining

$$Y^{\pi_N,N}(t) = -\sum_{i=1}^{K} X_i^{\pi_N,N}(t) - \beta,$$

we obviously have that $Y^{\pi_N,N}(t) \geqslant 0$ and

$$Y^{\pi_N,N}(t) = Y^N(0) + \sum_{i=1}^{K-1}(\mu_i - \mu_K)\int_0^t X_i^{\pi_N,N}(s)\,\mathrm{d}s - \mu_K\beta t - \sum_{i=1}^{K} M_i^{\pi_N,N}(t) - \mu_K\int_0^t Y^{\pi_N,N}(s)\,\mathrm{d}s + \sum_{i=1}^{K}\phi_i^{\pi_N,N}(t).$$

Comparing the latter with (8.7), recalling that $\widehat{Y}^{\pi_N,N}(t)$ is nonnegative and taking into account (8.8), the inequalities (8.6) and $\mu_i \geqslant \mu_K$, $1 \leqslant i \leqslant K-1$, we conclude, by part 2 of lemma 3.3, that

$$\hat{\phi}_K^{\pi_N,N}(t) \leqslant \sum_{i=1}^{K}\phi_i^{\pi_N,N}(t).$$

Therefore, since $c_i/A_i \geqslant c_K/A_K$, $1 \leqslant i \leqslant K-1$,

$$\sum_{i=1}^{K} c_i\lambda_i^N\int_0^t \psi_i^{\pi_N,N}(s)\,\mathrm{d}s = \sqrt{N}\sum_{i=1}^{K}\frac{c_i}{A_i}\phi_i^{\pi_N,N}(t) \geqslant \sqrt{N}\frac{c_K}{A_K}\hat{\phi}_K^{\pi_N,N}(t)$$

so that by (8.5)

$$\varliminf_{N\to\infty} P\big(\widehat{C}^{\pi_N,N}(t) > x\big) \geqslant \varliminf_{N'\to\infty} P\bigg(\frac{c_K}{A_K}\hat{\phi}_K^{\pi_{N'},N'}(t) > x + \varepsilon\bigg). \quad (8.9)$$

If we define

$$\widehat{X}_K^{\pi_N,N}(t) = X_K^N(0) - \mu_K \int_0^t \widehat{X}_K^{\pi_N,N}(s)\,\mathrm{d}s + M_K^{\pi_N,N}(t) - \hat{\phi}_K^{\pi_N,N}(t),$$

then

$$\widehat{Y}^{\pi_N,N}(t) = -\sum_{i=1}^{K} \widehat{X}_i^{\pi_N,N}(t) - \beta.$$

In analogy with the proof of theorem 2, using (8.4) we obtain

$$\big(M_i^{\pi_N,N}\big)_{1\leqslant i\leqslant K} \xrightarrow{\mathrm{d}} \big(\sqrt{2A_i\alpha_i\mu_i}\,W_i\big)_{1\leqslant i\leqslant K},$$

so $\widehat{X}_i^{\pi_N,N} \to \widehat{X}_i,\ 1 \leqslant i < K$, where $\widehat{X}_i$ solves

$$\widehat{X}_i(t) = X_{i,0} - \mu_i \int_0^t \widehat{X}_i(s)\,\mathrm{d}s + \sqrt{2A_i\alpha_i\mu_i}\,W_i(t).$$

It is not difficult to deduce then, by (6.17), (8.7) and (8.8), that $\widehat{Y}^{\pi_N,N} \xrightarrow{\mathrm{d}} Y$ and $\hat{\phi}_K^{\pi_N,N} \xrightarrow{\mathrm{d}} \phi_K$. The inequality (8.3) follows by (8.9), Fatou's lemma and the arbitrariness of ε.

As a consequence of (8.3), we have, for every nonnegative, nondecreasing, bounded and continuous function $f(x)$,

$$\varliminf_{N\to\infty} Ef\big(\widehat{C}^{\pi_N,N}(t)\big) \geqslant Ef\left(\frac{c_K}{A_K}\phi_K(t)\right),$$

where, in view of (8.2), equality is attained when π is the trunk reservation policy. The theorem is proved.

Acknowledgement

We would like to thank Phil Hunt for insightful discussions on this topic.

References

[1] J.M. Akinpelu, The overload performance of engineered networks with nonhierarchical and hierarchical routing, AT&T Tech. J. 63 (1984) 1261–1281.
[2] N.G. Bean, R.J. Gibbens and S. Zachary, The performance of single resource loss systems in multiservice networks, in: *The Fundamental Role of Teletraffic in the Evolution of Telecommunications Networks*, eds. J. Labetoulle and J.W. Roberts (Elsevier, Amsterdam, 1994).
[3] N.G. Bean, R.J. Gibbens and S. Zachary, Asymptotic analysis of single resource loss systems in heavy traffic, with applications to integrated networks, Adv. in Appl. Probab. 27 (1995) 273–292.
[4] P. Billingsley, *Convergence of Probability Measures* (Wiley, New York, 1968).
[5] P. Bremaud, *Point Processes and Queues: Martingale Dynamics* (Springer, Berlin, 1980).

[6] E.G. Coffman, Jr., A.A. Pukhalskii and M.I. Reiman, Storage limited queues in heavy traffic, Probab. Engrg. Inform. Sci. 5 (1991) 499–522.

[7] P.J. Hunt and F.P. Kelly, On critically loaded loss networks, Adv. in Appl. Probab. 21 (1989) 831–841.

[8] P.J. Hunt and C.N. Laws, Optimization via trunk reservation in single resource loss systems under heavy traffic, Ann. Appl. Probab. 7 (1997) 1058–1079.

[9] J. Jacod and A.N. Shiryaev, *Limit Theorems for Stochastic Processes* (Springer, Berlin, 1987).

[10] D.L. Jagerman, Some properties of the Erlang loss function, Bell System Tech. J. 53 (1974) 525–551.

[11] F.P. Kelly, Blocking probabilities in large circuit-switched networks, Adv. in Appl. Probab. 18 (1986) 473–505.

[12] F.P. Kelly, Routing and capacity allocation in networks with trunk reservation, Math. Oper. Res. 15 (1990) 771–793.

[13] F.P. Kelly, Loss networks, Ann. Appl. Probab. 1 (1991) 319–378.

[14] P. Key, Optimal control and trunk reservation in loss networks, Probab. Engin. Inform. Sci. 4 (1990) 203–242.

[15] R.Sh. Liptser and A.N. Shiryaev, *Theory of Martingales* (Kluwer, Dordrecht, 1989).

[16] B.L. Miller, A queueing reward system with several customer classes, Manag. Sci. 16 (1969) 234–245.

[17] M.I. Reiman, Asymptotically optimal trunk reservation for large trunk groups, in: *Proceedings of the 28th IEEE CDC* (1989) pp. 2536–2541.

[18] M.I. Reiman, A critically loaded multiclass Erlang loss system, Queueing Systems, 9 (1991) 65–82.

[19] M.I. Reiman, Optimal trunk reservation for a critically loaded link, in: *Teletraffic and Datatraffic in a Period of Change*, eds. A. Jensen and V.B. Iversen (North-Holland, Amsterdam, 1991).

[20] K.W. Ross, *Multiservice Loss Models for Broadband Telecommunication Networks* (Springer, Berlin, 1995).

[21] K.W. Ross and D. Tsang, Optimal circuit access policies in an ISDN environment: a Markov decision approach, IEEE Trans. Commun. (1989) 934–939.

[22] T.S. Zhang, On the strong solutions of one-dimensional stochastic differential equations with reflecting boundary, Stoch. Process. Appl. 50 (1994) 135–147.