The Annals of Applied Probability
2003, Vol. 13, No. 3, 1151–1206

CONTROL OF END-TO-END DELAY TAILS IN A MULTICLASS NETWORK: LWDF DISCIPLINE OPTIMALITY

BY ALEXANDER L. STOLYAR

Lucent Technologies, Bell Labs

We consider a multiclass queueing network with N customer classes, each having an arbitrary fixed route through the network. (Thus, the network is *not* necessarily feedforward.) We show that the *largest weighted delay first* (LWDF) discipline is an optimal scheduling discipline in the network in the following sense. Let w_i be the (random) instantaneous largest end-to-end delay of a class i customer in the network in stationary regime. For any set of positive constants $\alpha_1, \dots, \alpha_N$, the LWDF discipline associated with this set maximizes (among all disciplines) the quantity

$$\min_{i=1,\dots,N}\left[\alpha_i \lim_{n\to\infty}\frac{-1}{n}\log P(w_i > n)\right] = \lim_{n\to\infty}\frac{-1}{n}\log P(r > n), \tag{1}$$

where $r \doteq \max_i w_i/\alpha_i$ is the *maximal weighted delay* in the network. [This result is a generalization of the single-server result proved by A. L. Stolyar and K. Ramanan in *Ann. Appl. Probab.* **11** (2001) 1–48.]

As the key element of the proof, we establish the following *critical node property*: *In a LWDF network, there exists a most likely path to build large r, which is a most likely path to do so in one of the network nodes in isolation.* Such a most likely path has a very simple structure: its parameters [and the optimal value of (1)] can be computed by solving a finite-dimensional optimization problem for each network node.

1. Introduction.

Consider the following queueing network control problem. Find a scheduling (queueing) discipline such that

$$P\{w_i > T_i\} \le \delta_i, \qquad i = 1, \dots, N, \tag{1.1}$$

where N is the number of traffic flows, w_i is the steady state end-to-end delay for flow i, $T_i > 0$ is a predefined delay threshold (or deadline) and δ_i is the maximal acceptable deadline violation probability.

This problem appears in many applications, in particular in modern data communication networks where (1.1) is one of the typical *quality of service* (QoS) or *service level* requirements.

It has been shown recently that, in a single-server system, the largest weighted delay first (LWDF) scheduling discipline defined below [or a related discipline, such as generalized longest queue first (GLQF)] is an optimal discipline to satisfy

Received September 2000; revised October 2001.
AMS 2000 subject classifications. 60F10, 90B12, 60K25.
Key words and phrases. Queueing networks, scheduling, large deviations, fluid limit, end-to-end delay.

requirements (1.1) in the *large deviations* [31] and *heavy traffic* [11, 25, 32] asymptotic regimes.

The LWDF discipline, with positive parameters $\alpha_1, \dots, \alpha_N$, is defined roughly as follows. It always picks for service the customer (message)

$$c^* = \arg\max_c W(c)/\alpha_{i(c)}, \tag{1.2}$$

where the maximum is taken over the customers c available for service, $W(c)$ is the customer's current delay (i.e., *age* in the system) and $i(c)$ is its class (i.e., the index of the flow it belongs to).

The large deviations regime is when the system load is fixed, but $\delta_i \downarrow 0$ and $T_i \uparrow \infty$ in such a way that

$$-\log(\delta_i)/T_i \to \alpha_i^{-1} > 0. \tag{1.3}$$

More precisely, conditions (1.1) are replaced by the following asymptotic "tail" conditions:

$$\beta(w_i) \geq \alpha_i^{-1}, \qquad i = 1, \dots, N, \tag{1.4}$$

where we use the notation

$$\beta(X) \doteq \lim_{n\to\infty} -\frac{1}{n} \log P(X > n)$$

(assuming the limit exists).

Finding a discipline which would satisfy conditions (1.4) is equivalent to solving the following *optimization* problem:

$$\max_G \min_{i=1,\dots,N} \alpha_i \beta(w_i), \tag{1.5}$$

where the maximization is over scheduling disciplines G. Indeed, a discipline G satisfying (1.4) exists if and only if the maximum in (1.5) is 1 or greater.

It has been proved in [31] that, in a single-server system, the LWDF discipline with parameters α_i is an optimal solution of problem (1.5). In this paper we extend that result to a network setting. Namely, we prove the following.

MAIN RESULT. *The LWDF discipline* (*with parameters* α_i) *is an optimal solution of problem* (1.5) *in a queueing network of arbitrary topology.*

We emphasize that, with the *network* LWDF discipline, scheduling rule (1.2) is applied by each node to the set of customers in the node's queue (at the time of scheduling), and the delay (age) W of a customer is the time elapsed from its arrival at the network.

An important feature of the (network) LWDF discipline is that it is sufficiently "local" in that each node only needs to know the ages of the customers present in the node at the time of scheduling. In real applications, the LWDF discipline can be implemented by making each customer carry its age and making each node appropriately increment ages when forwarding customers to other nodes. With such an implementation, the scheduling is completely decentralized.

We prove our main result using large deviations techniques. As in [31], we note that if for a given discipline G all $\beta(w_i)$ are well defined, then

$$\min_{i=1,\dots,N} \alpha_i \beta(w_i) = \beta(r), \tag{1.6}$$

where $r \doteq \max_{i=1,\dots,N} w_i/\alpha_i$ is the *maximal weighted delay* (in stationary regime). It is well known in large deviations theory that the value of $\beta(r)$ is determined by a *most likely path* to build large r in the system. [Roughly, a most likely path is a most likely—lowest cost—trajectory of the input flows in the corresponding *fluid* system, such that the system starts from zero state and r reaches level 1; $\beta(r)$ is the cost of this path.]

Consider a fixed network node *in isolation*, which employs the LWDF discipline. (A node in isolation is a single-server system obtained from the network by removing all other nodes and removing all flows not passing through this node.) We know from the results of [31] that, for such a single-server system, there exists a most likely path which has a simple special structure. Parameters of this *simple* path (and its cost) can be computed by solving a finite dimensional optimization problem.

The key element of our analysis is the following (quite surprising) property.

CRITICAL NODE PROPERTY OF THE LWDF NETWORK. *In an LWDF network, there exists a most likely path to build large r, which is a most likely path to do so in one of the network nodes in isolation.*

More precisely, consider a simple *most likely path for each node in isolation, and pick a path which has the lowest cost among them. Then this path is a most likely path for the network.*

This means that a most likely path for the LWDF network [and the maximum in (1.5), attained on the LWDF discipline] can be found by solving a finite dimensional optimization problem for each network node.

To find a most likely path for the LWDF network, we use a novel approach to the crucial "path reduction" problem, that is, the problem of reducing an arbitrary path to a simpler and lower cost path such that the desired property still holds. (In our case the desired property is that r reaches value 1.) The standard "finite interval" approach is to find "special" time points such that the (fluid) input flow trajectories between those points can be replaced by linear ones, which "simplifies" the path without increasing its cost and without "compromising" the desired path property.

In contrast, our technique is based on an "infinitesimal interval" argument. Roughly speaking, we take an arbitrary (fluid) path such that r reaches 1, and for each $y \in [0, 1]$ we look at the infinitesimal cost $c(y)\,dy$ of raising r from y

to $y+dy$ when r crosses level y for the first time. Looking at the network evolution ("derivatives") at the time of such passage, we can construct a simple path which has the cost $c(y)$. This path is associated with one of the nodes j and is such that r reaches level 1 in both the "node j in isolation" system and in the entire network. On the other hand, the cost of the original path is lower bounded by $\int_0^1 c(y)\,dy$. Thus, loosely speaking, by picking y with the smallest $c(y)$, we obtain a simple path with cost not exceeding that of the original path, and still having the desired property. This means that a most likely path for the network is one of such simple paths. We believe this technique of finding a most likely path may be useful for other models too.

For a network of arbitrary topology (i.e., not even feedforward), optimality of the LWDF discipline and the critical node property are very surprising. One might expect that a most likely path to build large delays in the network would typically involve the interaction and "cooperation" between nodes. (E.g., it is shown in [26] that a most likely path to build a large *queue*, even in a simple feedforward network with first-in-first-out (FIFO) discipline, may involve such node interaction. Moreover, in a nonfeedforward network, even the issue of *stability* is nontrivial.) However, our results show that the LWDF discipline is very well behaved: the network is stable and, moreover, the optimal value of problem (1.5) is not "worse" (i.e., not less) than it would be for one of the nodes in isolation.

A prior result that suggests that the network LWDF should not "behave badly" is the stability of the longest-in-system (LIS) discipline [3], which is a special case of the LWDF with all α_i's equal. (This paper was under review when we became aware of [6], where the LIS stability result is generalized to show stability of a more general multiclass network with earliest-due-date-first-served (EDDFS) discipline. The main difference between LWDF and EDDFS is as follows: with LWDF, customer mutual priorities may change with time, if weights α_i of different flows are different; with EDDFS, mutual customer priorities do not change.) We also mention a recent network stability result for the GLQF discipline [18], which is related to LWDF but differs from it in that a flow with the largest weighted total *queue* in the *entire* network gets a priority along the flow's *entire* route.

The large deviations *queue length* asymptotics for the GLQF in a single-server system with two input flows were obtained in [4]. The issue of GLQF optimality was not considered there. More important, the technique of finding most likely paths in [4] would be hard to extend beyond the single-server–two-flows case. A different technique, used in [31], allows us to consider the case of arbitrary number of flows and prove optimality of both LWDF [in the sense of (1.5)] and GLQF (in the sense of an analogous problem involving queue length distributions); however, the technique is still confined to the single-server case. The technique of finding a most likely path used in this paper, based on the infinitesimal interval argument, allows us to extend the LWDF optimality results to a network.

We refer the reader to [4] or [31] for a more detailed review of work related to large deviations asymptotics of the delay and queue length distribution tails. A good review of large deviations methods in the analysis of communication systems, including earlier results on GLQF, can be found in [28].

Our main result, Theorem 3.2, consists of two statements. The first one is the upper bound on the maximal weighted delay distribution exponential decay rate $\beta(r)$ for the LWDF discipline. The second statement is the (same) lower bound on $\beta(r)$ for an arbitrary discipline. We prove the lower bound under very general assumptions on the input flow process. To prove the upper bound for the LWDF, we make a Markov assumption on the input flows. We need a Markov assumption to be able to use classical Wentzell–Freidlin constructions [13] in the proof. In [31], for a single-server system, both the lower and the upper bounds are proved without an additional Markov assumption, because there Loynes construction [17] was used (in a way analogous to [19]) to represent a stationary system state process via the input flow process. Loynes construction does *not* apply to a general topology network. However, it *does* apply to a *feedforward* network (see [19]), and hence for feedforward networks our results (both lower and upper bounds) hold without the Markov assumption.

Let us discuss our results in the context of the original network control problem (1.1). Existing approaches for providing end-to-end network delay guarantees of the form (1.1) include using the generalized processor sharing (GPS) and the earliest deadline first (EDF) disciplines. If the GPS is employed in a network, then setting of appropriate *GPS weights* in each node is required to achieve statistical multiplexing gains (see [12, 15]). The EDF discipline is known to be optimal for providing deterministic, worst case, delay guarantees in a network (see [14, 16]), but it also needs parameter tuning when used in conjunction with statistical multiplexing (see [2]). The obvious reason for that is the fact that the EDF scheduling rule does not take into account desired bounds δ_i on the violation probabilities in (1.1), at least not directly.

The LWDF discipline with parameters α_i set to

$$\alpha_i^{-1} = -\log(\delta_i)/T_i, \tag{1.7}$$

as suggested by (1.3), has the following advantages:

1. The LWDF asymptotic optimality shows that it typically has larger feasibility region than GPS, that is, is able to support more flows with the desired QoS.
2. Scheduling is completely decentralized, as long as messages can carry their ages.
3. The LWDF needs no parameter tuning since it directly takes into account both T_i and δ_i (unlike EDF).
4. There is no need to choose GPS weights (unlike GPS).

The technique used in this paper also allows us to prove optimality of the GLQF in a network. (Following [31], we call this discipline *largest weighted unfinished* *work first* (LWWF) and consider two different versions of it.) The optimality is in the sense of the problem

$$\max_G \min_{i=1,\dots,N} \alpha_i \beta(q_i), \tag{1.8}$$

where q_i is the amount of unfinished work (i.e., total queue length) of class i in the network in stationary regime. The proof of this result is omitted since it is just a simplified version of the proof of our main result.

The paper is organized as follows. In Section 2 we introduce basic notation, definitions and conventions used in the paper. We describe the network model, define the LWDF discipline and formulate our main result, Theorem 3.2, in Section 3. In Section 4 we, first, consider properties of the LWDF discipline in the network with discrete input flows, then extend the definition of the LWDF to fluid input flows and, finally, study properties of the sample paths of the fluid LWDF network. Section 5 is central to this paper: we find a most likely path to build large maximal weighted delay in the fluid LWDF network, and we prove the critical node property. We prove the lower bound on the $\beta(r)$ for any scheduling discipline in Section 6. In Section 7 we establish further properties of the family of sample paths of the fluid LWDF network. In Section 8 we consider large deviations properties of the sequence of scaled processes for the original (discrete) LWDF network. In Section 9 we consider the LWDF network with Markov input flows and, finally, in Section 10 prove the upper bound on $\beta(r)$ for such network. In Section 11 we present the LWWF (GLQF) optimality result [for problem (1.8)]. The Appendix contains the proof of one technical result.

2. Basic notation and definitions. We denote by $\mathbb{R}$ the set of real numbers, and by $\Lambda(\cdot)$ the Lebesgue measure on $\mathbb{R}$. For an integer $k \ge 1$, we define $\mathbb{R}^k_+ \doteq \{y = (y_1, \dots, y_k) \in \mathbb{R}^k : y_i \ge 0,\ \text{for } i = 1, \dots, k\}$ and $\mathbb{R}^k_- \doteq \{y \in \mathbb{R}^k : -y \in \mathbb{R}^k_+\}$.

We let $a \wedge b$ and $a \vee b$ denote respectively the minimum and the maximum of real numbers a and b. For a finite set S, $|S|$ denotes its cardinality. The infimum of a function over an empty set is interpreted as ∞.

The derivative, right derivative and partial right derivative of a function h on variable t (where they exist) are denoted by h', $(d^+/dt)h$ and $(\partial^+/\partial t)h$, respectively.

A pair of integer indices (i, k) is sometimes written as (ik), where it cannot cause confusion; similarly, in subscripts, we often write h_{ik} instead of $h_{i,k}$.

Let $\mathcal{D}$ (respectively $\mathcal{D}_{(b)}$, for a fixed $b \in \mathbb{R}$) be the space of RCLL functions (i.e., right continuous functions with left limits) on $(-\infty, \infty)$ (resp. $[b, \infty)$). Unless otherwise specified, we assume $\mathcal{D}$ and $\mathcal{D}_{(b)}$ are endowed with the topology of uniform convergence on compact sets (u.o.c.). We use $h(t-)$ to denote the left limit $\lim_{u\uparrow t} h(u)$ of the function h at the point t. As measurable spaces, we always assume that $\mathcal{D}$ and $\mathcal{D}_{(b)}$ are endowed with the σ-algebra generated by the cylinder sets.

We now define some subspaces of $\mathcal{D}$ and $\mathcal{D}_{(b)}$, related as follows:

$$\mathcal{A} \subset \mathcal{C} \subset \mathcal{I} \subset \mathcal{D}, \qquad \mathcal{S} \subset \mathcal{I},$$

$$\mathcal{A}_{(b)} \subset \mathcal{C}_{(b)} \subset \mathcal{I}_{(b)} \subset \mathcal{D}_{(b)}, \qquad \mathcal{S}_{(b)} \subset \mathcal{I}_{(b)}.$$

Let $\mathcal{I}$ be the subset of nondecreasing functions in $\mathcal{D}$, $\mathcal{C}$ be the subset of continuous functions in $\mathcal{I}$, let $\mathcal{A}$ be the subset of absolutely continuous functions in $\mathcal{C}$ and let $\mathcal{S}$ be the subset of functions in $\mathcal{I}$ which are nondecreasing, piecewise constant and have only a finite number of jumps on any finite time interval. Let $\mathcal{I}_{(b)}$, $\mathcal{C}_{(b)}$, $\mathcal{A}_{(b)}$ and $\mathcal{S}_{(b)}$ be similarly defined subsets of $\mathcal{D}_{(b)}$. Also, if S is one of the spaces defined above, then let S_+ denote the subset of nonnegative functions in S, for $S \subseteq \mathcal{D}_{(b)}$ let S_c denote the subset of functions h in S such that $h(b) = c$ and let $S_{+,c} = S_+ \cap S_c$. (E.g., $\mathcal{I}_{(b),+}$ is the subset of nondecreasing nonnegative functions in $\mathcal{D}_{(b)}$, and $\mathcal{A}_{(0),+,0}$ is the subset of nondecreasing absolutely continuous functions h in the interval $[0, \infty)$ such that $h(0) = 0$.)

Somewhat of an exception from our notational conventions will be the space $S_{+,0} \subseteq S_+$ (where S can be $\mathcal{I}, \mathcal{C}, \mathcal{A}$ or $\mathcal{S}$) which consists of the functions h such that $h(t) = 0$ for at least one t (which implies of course that $h(s) = 0,\ s \in (-\infty, t]$).

We assume that the subspaces inherit the topology and σ-algebra of the original space. Given any space S, we assume that the k times product space S^k has the product topology and product σ-algebra defined in the natural way.

For any $s \geq 0$ and $h = (h_1, \dots, h_k) \in \mathcal{D}^k$, we define the norm

$$\|h\|_s \doteq \max_{i=1,\dots,k} \sup_{-s \leq t \leq s} |h_i(s)|.$$

For $h \in \mathcal{D}^k_{(b)}$ the norm $\|h\|_s$ is defined similarly with $-s$ replaced by b in the above display. Thus the u.o.c. convergence in $\mathcal{D}^k$ or $\mathcal{D}^k_{(b)}$ is equivalent to convergence in the corresponding norm $\|\cdot\|_s$ for all $s > 0$.

We will also need to consider the following *weak convergence* of nondecreasing functions. Let the functions h and $h^{(n)}$, $n = 1, 2, \dots$, be elements of the space $\mathcal{I}_{(b)}$ (or $\mathcal{I}$). Then we say that the sequence $\{h^{(n)}\}$ converges weakly to h, and we denote this fact by

$$h^{(n)} \Rightarrow h \qquad \text{as } n \to \infty,$$

if $h^{(n)}(t) \to h(t)$ in every point t of the *open* interval (b, ∞) [resp. $(-\infty, \infty)$], where h is continuous. (We emphasize that in the case of the space $\mathcal{I}_{(b)}$, the convergence at the boundary point b is not required.) The convergence

$$(h_1^{(n)}, \dots, h_k^{(n)}) \Rightarrow (h_1, \dots, h_k)$$

in a product space is equivalent to $h_m^{(n)} \Rightarrow h_m$ for all m.

We define the scaling operator Γ^c, $c > 0$, for $h \in \mathcal{D}^k$ (or $\mathcal{D}^k_{(b)}$) as follows:

$$(\Gamma^c h)(t) \doteq \frac{1}{c} h(ct). \tag{2.1}$$

For a scalar a, $\Gamma^c a \doteq a/c$; and $\Gamma^c(Y_1, \dots, Y_k) \doteq (\Gamma^c Y_1, \dots, \Gamma^c Y_k)$, where each Y_m can be either a function or a scalar.

Given an operator $A : S \to S'$, where S and S' belong to the function spaces defined above, we say that A is scalable if, for every $h \in S$ and $c > 0$,

$$A(\Gamma^c h) = \Gamma^c(Ah). \tag{2.2}$$

For any scalar function $h = (h(t), t \in C)$, $C \subseteq \mathbb{R}$, we define the shift operator θ_a, $a \in \mathbb{R}$, and the truncation operators $\zeta_{a_1}^{a_2}$, ζ^{a_2}, ζ_{a_1}, $a_1, a_2 \in \mathbb{R}$, in the standard way:

$$\begin{aligned}
(\theta_a h)(t) &= h(a+t), && t \in \{s \in \mathbb{R} \mid a + s \in C\},\\
(\zeta_{a_1}^{a_2} h)(t) &= h(t), && t \in C \cap [a_1, a_2],\\
(\zeta^{a_2} h)(t) &= h(t), && t \in C \cap (-\infty, a_2],\\
(\zeta_{a_1} h)(t) &= h(t), && t \in C \cap [a_1, \infty).
\end{aligned}$$

For a set of functions (with possibly different domains), the shift and truncation operators are applied componentwise.

Let $\mathbf{\Omega} \doteq (\Omega, \mathcal{F}, P)$ be a probability space. We assume that $\mathbf{\Omega}$ is large enough to support all the independent random processes that we use in the paper. We denote by $P_*(B)$ the inner measure (with respect to the probability P) of an arbitrary subset $B \subseteq \Omega$. If $B \in \mathcal{F}$, then $P_*(B) = P(B)$. Given any subset B of a topological space, we use $\bar{B}$ and B° to denote its closure and interior respectively.

Typically, we follow the convention of using bold font for stochastic processes and lightface Roman font for deterministic processes.

We now give the definition of a large deviation principle (LDP) ([10], page 5).

DEFINITION 2.1 (LDP). Let $\mathcal{X}$ be a topological space and $\mathcal{B}$ a σ-algebra on $\mathcal{X}$ (which is not necessarily the Borel σ-algebra). A sequence of random variables $\{\mathbf{X}_n\}$ on Ω taking values in $\mathcal{X}$ is said to satisfy the LDP with good rate function I if, for all $B \in \mathcal{B}$,

$$\limsup_{n\to\infty} \frac{1}{n} \log P(\mathbf{X}_n \in B) \le -\inf_{x \in \bar{B}} I(x),$$

and

$$\liminf_{n\to\infty} \frac{1}{n} \log P(\mathbf{X}_n \in B) \ge -\inf_{x \in B^\circ} I(x),$$

where $I : \mathcal{X} \to \mathbb{R} \cup \{\infty\}$ is a function with compact level sets.

3. The model and main results.

3.1. *The model.* We consider a multiclass queueing network with a finite set of nodes J. Each network node j is a single-server queue with the fixed service rate $\mu_j > 0$. There is a finite number of exogenous input flows (customer classes) forming the set $N = \{1, 2, \dots, N\}$. (We use the same symbol N for both the set and its cardinality.) A class i customer has its prescribed route through the network,

$$\hat{j}(i, 1), \dots, \hat{j}(i, k), \dots, \hat{j}(i, K_i),$$

where K_i is the route length and $\hat{j}(i, k) \in J$ is the node on the kth step of the route. After completing service in node $\hat{j}(i, k)$ the customer enters node $\hat{j}(i, k + 1)$ or (if $k = K_i$) leaves the network. We assume that the route for any class i has no loops; that is, it does not go through any node more than once. (This assumption is not essential; see the discussion in Section 3.3.)

A class i customer in the kth node of its route will be called a *type* (i, k) *customer*, or an (i, k)-*customer*. The flow of (i, k)-customers [i.e., the flow of class i customers arriving in node $\hat{j}(i, k)$] we will be called the (i, k)-*flow*.

Denote by G the set of all customer types and by G_j the subset of G listing the types of customers to be served in node $j \in J$:

$$G = \{(i, k) \mid k = 1, 2, \dots, K_i; i \in N\}, \qquad G_j = \{(i, k) \in G \mid \hat{j}(i, k) = j\}.$$

We assume that the set $\mathbf{f} = (\mathbf{f}_1, \dots, \mathbf{f}_N)$ of exogenous input flow processes satisfies the following assumption.

ASSUMPTION 3.1. (i) Each flow $\mathbf{f}_i$ is a random process on $\mathbf{\Omega}$ that takes values in $\mathcal{S}_{(0),+}$.

(ii) The flows $\mathbf{f}_i$, $i = 1, \dots, N$, are mutually independent.

(iii) For each i, and every $a \ge 0$ and $T \ge 0$, the sequence of processes $\{\theta_a[\zeta_a^{a+T}\mathbf{f}_i^{(n)} - \mathbf{f}_i^{(n)}(a)], n = 1, 2, \dots\}$ satisfies a LDP with good rate function J_T^i given by

$$J_T^i(h) \doteq \begin{cases} \int_0^T L_i(h'(s))\, ds, & \text{if } h \in \zeta^T \mathcal{A}_{(0),+,0}, \\ \infty, & \text{otherwise}, \end{cases} \tag{3.1}$$

where $\mathbf{f}^{(n)} = \Gamma^n \mathbf{f}$ is the scaled version of $\mathbf{f}$ as defined in (2.1), L_i is a convex lower semicontinuous function taking values in $[0, \infty]$ and such that $L_i(\lambda_i) = 0$ for some $\lambda_i \in (0, \infty)$, $L_i(x) > 0$ for $x \neq \lambda_i$ and $\lim_{x \to \infty} L_i(x)/x = \infty$.

The process $\mathbf{f}_i(t)$ represents the cumulative amount of work of class i (in terms of the required amount of service) that has entered the system by time t. A jump in $\mathbf{f}_i(\cdot)$ at time t corresponds to a "customer" arrival, with the *service requirement* of that customer equal to the size of the jump $\mathbf{f}_i(t) - \mathbf{f}_i(t-)$. We assume that the service requirement of a customer is the same in each node on its route (also a nonessential assumption; see Section 3.3). Thus, the *service time* of a customer by node j is equal to the customer's service requirement divided by μ_j.

Assumption 3.1 implies, in particular, that the mean exogenous arrival rate for flow i is equal to λ_i. (If, in addition, $\mathbf{f}_i$ would have stationary increments, then λ_i would simply be the mean service requirement of the class i customers arriving within any unit time interval.)

We always assume that the nominal load of every node j is strictly less than its service rate, that is,

$$\sum_{(i,k)\in G_j} \lambda_i < \mu_j \qquad \forall\, j \in J. \tag{3.2}$$

The function $L_i(\cdot)$ is sometimes called the *local rate function* for flow i. We will call the functional $J_T^i(h)$ the *cost* (associated with flow i) of a function $h \in \zeta^T \mathcal{I}_{(0),+}$.

For fixed $T \geq T_1$, the functional

$$J_{T,T_1}(f) \doteq \sum_{i=1}^{N} J_{T-T_1}^i\big(\theta_{T_1}\big[\zeta_{T_1}^T f_i - f_i(T_1)\big]\big) \tag{3.3}$$

will be called the *cost* of a function $f \in \mathcal{I}_{(0)}^N$ in the interval $[T_1, T]$. Equivalently,

$$J_{T,T_1}(f) = \sum_{i\in N} \int_{T_1}^{T} L_i\big(f_i'(s)\big)\, ds$$

if all functions f_i are absolutely continuous in the interval $[T_1, T]$, and $J_{T,T_1}(f) = \infty$ otherwise.

We put $J_T(f) \doteq J_{T,0}(f)$ to simplify notation.

For some results in this paper we will need the following stronger (Markov) assumption on the input flows.

ASSUMPTION 3.2. The flows $\mathbf{f}_i, i = 1, \ldots, N$, are mutually independent. Each flow $\mathbf{f}_i$ can be described as a "Markov modulated Poisson flow." Namely, for each i, there exists an underlying (modulating) continuous time irreducible Markov chain with finite set of states indexed by $1, 2, \ldots, \nu_i$. When this chain is in state m, the customers arrive according to a Poisson process of the rate λ_{im}, and the service requirements of the customers are i.i.d. equal in distribution to a positive bounded random variable Y_{im}.

Assumption 3.2 is stronger than Assumption 3.1 in the sense that if input flows satisfy Assumption 3.2, then they also satisfy Assumption 3.1 for any fixed combination of initial states of the modulating Markov chains (and there is only a finite number of those combinations). Assumption 3.2 also implies the existence and uniqueness (in distribution) of the "stationary increments" versions of the input processes for which Assumption 3.1 holds.

3.2. *Main result.* Given a sample path of the random process describing the evolution of the network, let $\tau_{ik}(t)$ denote the arrival time into the network of the "oldest" class i customer present in the network at time t and not completely served by node $j = \hat{j}(i,k)$ by time t. By convention, $\tau_{ik}(t) = t$ if there are no such customers at time t; $\tau_{i0}(t) \equiv t$; and $\tau_{ik}(\cdot)$ is a right-continuous function. We refer to $\tau_i \doteq (\tau_i(t))$, with $\tau_i(t) \doteq \min_k \tau_{ik}(t) = \tau_{i,K_i}(t)$ as the class i *backlog* sample path (or backlog process, depending on the context). Suppose we are given the set of positive *weights*

$$\alpha_i > 0, \qquad i \in N.$$

Then we define the class i *delay* w_i, the *weighted delay* r_i and the *maximal weighted delay* r in terms of the backlog sample path as follows. For every t and $i = 1, \dots, N$,

$$w_i(t) \doteq t - \tau_i(t), \tag{3.4}$$

$$r_i(t) \doteq w_i(t)/\alpha_i, \tag{3.5}$$

$$r(t) \doteq \max_i r_i(t). \tag{3.6}$$

Our goal is to find a discipline that is optimal in the sense that it maximizes the exponential decay rate of the stationary distribution of the maximal weighted delay $\mathbf{r}(\cdot)$. It is shown in [31] that in a single node system, the largest weighted delay first discipline is optimal. The main result of this paper, formulated below in Theorem 3.2, is that the LWDF is also optimal in a network, if the LWDF is understood as the *network* LWDF discipline defined as follows.

DEFINITION 3.1 (The network LWDF discipline for discrete input flows). For any customer c present in the network, its (current) weighted delay is defined as the ratio $W(c)/\alpha_{i(c)}$, where $W(c)$ is the customer delay (i.e., the time elapsed since its arrival at the network) and $i(c)$ is the customer's class. The LWDF discipline is a nonpreemptive, work conserving discipline that always chooses for service in any node j the customer with the largest weighted delay (from the customers present in node j). In case of a tie, by convention the LWDF discipline chooses the customer with the highest class index i.

We will need some additional definitions.

For each j, consider an artificial single node system which we will call *node j in isolation*. Namely, consider the subset $N_j \subseteq N$ of flows which have node j on their route, and assume that node j is the only node in which those flows need to be served. It follows from the results of [31] that in such single node system with the LWDF discipline, the exponential decay rate of the maximal weighted delay $\mathbf{r}(0)$ in stationary regime is equal to

$$\lim \frac{1}{n} \log P\left(\frac{1}{n}\mathbf{r}(0) > 1\right) = -J_{*,j},$$

where $J_{*,j}$ is defined as the optimal value of the following optimization problem:

$$J_{*,j} = \inf_{\nu} \frac{1}{\gamma} \sum_{i \in \tilde{N}_j} (1 - \alpha_i \gamma) L_i(x_i), \tag{3.7}$$

where

$$\nu = \{\tilde{N}_j \subseteq N_j, (x_i, i \in \tilde{N}_j)\}, \tag{3.8}$$

$$\gamma = \frac{\sum_{m \in \tilde{N}_j} x_m - \mu_j}{\sum_{m \in \tilde{N}_j} \alpha_m x_m} \tag{3.9}$$

and minimization is subject to the constraints

$$x_i > 0, \qquad i \in \tilde{N}_j, \tag{3.10}$$

$$0 < \gamma < \min_{i \in \tilde{N}_j} \frac{1}{\alpha_i}, \tag{3.11}$$

$$\frac{1}{\alpha_i} \le \gamma \qquad \forall\, i \in N_j \setminus \tilde{N}_j. \tag{3.12}$$

Let us write

$$J_* \doteq \min_{j \in J} J_{*,j}. \tag{3.13}$$

(Note that the case $J_* = \infty$ is possible.)

Consider the class $\overline{\mathcal{G}}$ of *work-nonabandoning* queueing (or scheduling) disciplines G such that the following hold:

(a) customers are not allowed to skip any node on their routes;

(b) a customer can leave any node on its route only after it is completely served by that node.

Now we are in position to formulate our main result.

THEOREM 3.2. (i) *Consider the network with the LWDF scheduling discipline, and suppose that Assumption* 3.2 *holds. Then a stationary random process describing network evolution exists, is unique in distribution and is such that the (stationary) distribution of the maximal weighted delay* $\mathbf{r}(0)$ *satisfies the condition*

$$\lim_{n\to\infty} \frac{1}{n} \log P\left(\frac{1}{n}\mathbf{r}(0) > 1\right) = -J_*. \tag{3.14}$$

(ii) *Suppose that Assumption* 3.1 *holds. Then there exists* $T^0 \in (0, \infty)$ *such that, for any queueing discipline* $G \in \overline{\mathcal{G}}$, *and any* $t > T^0$, *we have the following lower bound*:

$$\liminf_{n\to\infty} \frac{1}{n} \log P_*\left(\frac{1}{n}\mathbf{r}(nt) > 1\right) \ge -J_*, \tag{3.15}$$

where $\mathbf{r}(\cdot)$ *is the maximal weighted delay associated with the discipline* G.

In the second statement of the theorem, it is only assumed that $\mathbf{r}(\cdot)$ is a well-defined (but not necessarily measurable) function on the probability space. [That is why we do not call $\mathbf{r}(\cdot)$ a process, and use the inner measure P_* in the formulation.] Also, no assumption is made about the existence or nonexistence of a stationary version of the process. For any stationary version, if it exists, (3.15) of course implies

$$\liminf_{n\to\infty} \frac{1}{n} \log P_*\left(\frac{1}{n}\mathbf{r}(0) > 1\right) \geq -J_*.$$

Theorem 3.2 shows that indeed the LWDF discipline in the network maximizes the exponential decay rate of the stationary distribution of the maximal weighted delay $\mathbf{r}(\cdot)$. Moreover, the optimal decay rate J_* is equal to the minimum of the decay rates taken for each node j in isolation. As we already discussed in the Introduction, this *critical node* property is quite surprising.

3.3. *Comments on the model assumptions.*

COMMENT 1. The assumptions that a customer service requirement remains the same in all nodes of its route and that each route has no loops are not essential. We make them to simplify the exposition. We could assume that class i input flow is characterized by the function describing the *number* of exogenous arrivals and the set of K_i (route length) functions describing the cumulative service requirement (as a function of the *number* of arrivals) for each *step* k on the route [but the node $\hat{j}(i,k)$ may be same for different steps k]. All we need to make our technique work is that *marginally*, for each *node* j on the class i route, Assumption 3.1 holds for the function of *time* describing the cumulative amount of work arrived in the network and intended for node j.

COMMENT 2. The specific Assumption 3.2 requiring that each input flow is a “Markov modulated Poisson flow” can be greatly relaxed (or changed). The essential purpose of a Markov assumption on the input flows (which is only used in Sections 9 and 10) is to make the process describing the LWDF network evolution Markov. This allows us to use Wentzell–Freidlin constructions in the proof of the large deviations upper bound [Theorem 3.2(i)] for the network of arbitrary topology, not necessarily feedforward. Beyond this main purpose, specifics of a Markov assumption are less important. The form of our Assumption 3.2, on the one hand, is general enough to cover many interesting applications and, on the other hand, makes the exposition of the upper bound proof (in Sections 9 and 10) relatively simple.

COMMENT 3. As already mentioned in the Introduction, if the network is feedforward, then the Markov Assumption 3.2 can be relaxed to Assumption 3.1 (plus the assumption of stationary increments) in Theorem 3.2(i). In this case, similarly to [31], the proof would use the form (derived in this paper) of an optimal network simple element f^0 and a Loynes type construction of the stationary state process (as in [19]).

COMMENT 4. In Assumption 3.1, the condition that J_T^i defined by (3.1) is a good rate function (i.e., has compact level sets) in fact follows from the conditions imposed on the local rate function L_i. To see this, let us fix function $h \in \mathcal{A}_{(0),+,0}$. Let h_m be the piecewise linear approximation of h with step T/m. Then we know that, as $m \to \infty$, $h_m \to h$ u.o.c., and (for the derivatives) $h'_m \to h'$ almost everywhere. The lower semicontinuity of L_i and Fatou's lemma imply that $\liminf J_T^i(h_m) \geq J_T^i(h)$. It is easy to see (using the convexity of L_i) that, for any sequence $\{h^{(n)},\ n = 1, 2, \dots\}$, such that $h^{(n)} \to h$ uniformly, and for any fixed m,

$$\liminf J_T^i(h^{(n)}) \geq J_T^i(h_m).$$

This implies $\liminf J_T^i(h^{(n)}) \geq J_T^i(h)$ and therefore J_T^i is lower semicontinuous. It remains to show that the level sets of J_T^i are precompact, which is easily established using the superlinearity of L_i, and we are done.

COMMENT 5. Throughout our analysis, we do *not* exclude the case $J_* = \infty$. This may seem redundant since, first, we will see that Assumption 3.2 implies $J_* < \infty$ and, second, the lower bound in Theorem 3.2(ii) holds trivially when $J_* = \infty$. We do this, however, because if Markov Assumption 3.2 is relaxed (see Comment 2) or is replaced by Assumption 3.1 (if the network is feedforward, see Comment 3), then the case $J_* = \infty$ is possible and the upper bound in Theorem 3.2(i) is nontrivial. Therefore, our analysis (with possibly adjusted Sections 9 and 10) covers such cases too.

4. The largest weighted delay first discipline in the network.

4.1. *The LWDF discipline for discrete input flows.* Let us recall that we denote by $\mathcal{S}_{+,0}$ the subset of functions $h \in \mathcal{S}_+$ satisfying the following additional condition:

$$\text{There exists finite } c, \text{ such that } h(t) = 0, t < c. \tag{4.1}$$

Consider $f \in \mathcal{S}_{+,0}^N$, which we will interpret as a realization of the set of input flows $f = (f_i,\ i \in N)$, with each flow f_i being defined only in the time interval $[b_i, \infty)$, where

$$b_i = b_i(f_i) \doteq 0 \wedge \sup\{t : f_i(t) = 0\}.$$

Clearly, $b_i \leq 0$. [Extending the domain of each f_i to the entire real axis by putting $f_i(t) = 0$ for $t < b_i$ is a notational convenience.]

Given $f \in \mathcal{S}_{+,0}^N$ (interpreted as described above), consider the evolution of the network starting time 0. According to the definitions introduced earlier,

$\tau_{ik}(t), t \geq 0$, is the arrival time at the network of the "oldest" class i customer which is not completely served by node $j = \hat{j}(i,k)$ by time t. [By convention, $\tau_{ik}(t) = t$ if there are no such customers at time t; and $\tau_{ik}(\cdot)$ is a right-continuous function.] Let us also denote by $\hat{f}_{ik}(t), t \geq 0$, the total amount of service received by time t from node $j = \hat{j}(i,k)$ by class i customers arriving in the network in the interval $[b_i, t]$; and by $f_{ik}(t), t \geq 0$, denote the total amount of work of class i which arrived in the network in $[b_i, t]$ and arrived in node $j = \hat{j}(i,k)$ by time t. According to our conventions,

$$\tau \doteq \big((\tau_{ik}(t), t \geq 0), (ik) \in G\big) \in \mathcal{S}_{(0)}^{|G|},$$
$$\bar{f} \doteq \big((f_{ik}(t), t \geq 0), (ik) \in G\big) \in \mathcal{S}_{(0),+}^{|G|},$$
$$\hat{f} \doteq \big((\hat{f}_{ik}(t), t \geq 0), (ik) \in G\big) \in \mathcal{C}_{(0),+}^{|G|};$$

moreover, each function $\hat{f}_{ik}(\cdot)$ is Lipschitz continuous with the Lipschitz constant μ_j, $j = \hat{j}(i,k)$.

Now, suppose queueing discipline in the network is LWDF. Suppose also that the following condition on $\tau(0)$ holds:

$$\tau_{i,K_i}(0) \equiv \tau_i(0) = b_i \qquad \forall i. \tag{4.2}$$

(This condition is nothing more than a convention that the input flow i is considered only from the time of arrival of the oldest flow i customer still present in the network at time 0.)

Then it is clear that evolution of the network in the time interval $[0, \infty)$, that is, $(\hat{f}, \tau, \bar{f})$, is uniquely determined by the 4-tuple $(f, \hat{f}(0), \tau(0), \bar{f}(0))$. Moreover, in this case, the following properties hold:

(i) For every class i,

$$\widehat{f}_{i,K_i}(t) \leq \widehat{f}_{i,K_i-1}(t) \leq \cdots \leq \widehat{f}_{i,1}(t) \leq f_i(t), \qquad t \in [0, \infty). \tag{4.3}$$

(ii) For every $(i,k) \in G$,

$$\tau_{ik}(t) = \sup\{s \leq \tau_{i,k-1}(t) \mid f_i(s) \leq \hat{f}_{ik}(t)\}, \qquad t \in [0, \infty),$$

and, for every $i \in N$,

$$\begin{aligned} b_i = \tau_i(0) \leq \tau_i(t) \equiv \tau_{i,K_i}(t) \leq \tau_{i,K_i-1}(t) \leq \cdots \leq \tau_{i,1}(t) \leq \tau_{i,0}(t) \\ \equiv t, \qquad t \in [0, \infty). \end{aligned} \tag{4.4}$$

(iii) For any $(i,k) \in G$,

$$f_{ik}(t) = f_i\big(\tau_{i,k-1}(t)-\big), \qquad t \in [0, \infty), \tag{4.5}$$

and, for every class i,

$$f_{i,K_i}(t) \leq f_{i,K_i-1}(t) \leq \cdots \leq f_{i,1}(t) \equiv f_i(t), \qquad t \in [0, \infty). \tag{4.6}$$

(iv) For every $(i,k) \in G$,

$$\widehat{f}_{ik}(t) \in \bigl[f_i\bigl(\tau_{i,k}(t)-\bigr), f_i\bigl(\tau_{i,k}(t)\bigr) \wedge f_{ik}(t)\bigr], \qquad t \in [0,\infty). \tag{4.7}$$

[Note that both cases $f_i(\tau_{i,k}(t)) < f_{ik}(t)$ and $f_i(\tau_{i,k}(t)) > f_{ik}(t)$ are possible: the former case when there is more than one (i,k)-customer in node $j = \hat{j}(i,k)$ and the latter case when there is no (i,k)-customer in node j, but there is at least one in the network "upstream" from node j.]

(v) At any time t, for any node j, there may be at most one flow $(ik) \in G_j$ such that

$$\widehat{f}_{ik}(t) \in \bigl(f_i\bigl(\tau_{i,k}(t)-\bigr), f_i\bigl(\tau_{i,k}(t)\bigr)\bigr).$$

(The latter condition is implied by the nonpreemptiveness of LWDF.)

Thus, the LWDF discipline defines a deterministic operator A_d which maps a 4-tuple $(f, \hat{f}(0), \tau(0), \bar{f}(0))$ satisfying additional conditions (i)–(v) (for $t = 0$) into a triple $(\hat{f}, \tau, \bar{f}) = A_d(f, \hat{f}(0), \tau(0), \bar{f}(0))$ which (along with f) describes the sample path of the system in the interval $[0, \infty)$.

A 4-tuple $(f, \hat{f}, \tau, \bar{f})$ such that $(f, \hat{f}(0), \tau(0), \bar{f}(0))$ satisfies additional conditions (i)–(v) (with $t = 0$) and $(\hat{f}, \tau, \bar{f}) = A_d(f, \hat{f}(0), \tau(0), \bar{f}(0))$, we will call a *sample path of LWDF network with discrete input flows*, or just a *discrete sample path* (DSP). Any discrete sample path $(f, \hat{f}, \tau, \bar{f})$ satisfies conditions (i)–(v) for all $t \geq 0$.

REMARK. Operator A_d can be viewed as *the* queueing discipline itself, LWDF in our case. It should be clear that the set of properties (i)–(v) is *not* a definition of operator A_d—those properties obviously do *not* characterize the LWDF discipline completely. All we will need to know is that, for any discrete sample path, properties (i)–(v) do hold.

4.2. *Extension of the LWDF discipline to fluid inputs.* In this section we define sample paths of a fluid system as limits of the discrete sample paths. The formal procedure we use is very similar to that in [30, 31].

A 4-tuple $(f, \hat{f}, \tau, \bar{f})$ with $f \in \mathcal{I}^N_{+,0}$, $\widehat{f} \in \mathcal{C}^{|G|}_{(0),+}$, $\tau \in \mathcal{I}^{|G|}_{(0)}$ and $\bar{f} \in \mathcal{I}^{|G|}_{(0),+}$, we will call a *sample path of LWDF network with fluid input flows*, or just a *fluid sample path* (FSP), if there exists a sequence of discrete sample paths $\{(f^{(n)}, \hat{f}^{(n)}, \tau^{(n)}, \bar{f}^{(n)}),\ n = 1, 2, \ldots\}$ such that

$$\bigl(f^{(n)}, \hat{f}^{(n)}, \tau^{(n)}, \bar{f}^{(n)}\bigr) \Rightarrow (f, \hat{f}, \tau, \bar{f}), \qquad n \to \infty. \tag{4.8}$$

(We remind that the weak convergence denoted by "$\Rightarrow$" means convergence in every interior point of continuity of the limit, for each component function.)

REMARK. It should be clear (and is important for our analysis) that some of the component functions of a fluid sample path may be *discontinuous*. Moreover, any discrete sample path is a fluid sample path.

We define operator A as follows. For every

$$(f, \hat{f}(0), \tau(0), \bar{f}(0)) \in \mathcal{I}_{+,0}^{N} \times \mathbb{R}_{+}^{|G|} \times \mathbb{R}^{|G|} \times \mathbb{R}_{+}^{|G|},$$

$A(f, \hat{f}(0), \tau(0), \bar{f}(0))$ is the *subset* of triples

$$(\hat{f}, \tau, \bar{f}) \in \mathcal{C}_{(0),+}^{|G|} \times \mathcal{I}_{(0)}^{|G|} \times \mathcal{I}_{(0),+}^{|G|}$$

having $(\hat{f}(0), \tau(0), \bar{f}(0))$ as the initial condition and such that $(f, \hat{f}, \tau, \bar{f})$ is a fluid sample path. Notice that $A(f, \hat{f}(0), \tau(0), \bar{f}(0))$ may be an empty set.

REMARK. We see that the multivalued operator A is an extention (via the above limiting procedure) of the operator A_d. Operator A can be viewed as a "definition" of the LWDF discipline for the networks with fluid input flows.

We will say that $(f, \hat{f}, \tau, \bar{f})$ is a fluid sample path with *zero initial condition*, or *starting from* 0, if

$$f_i(0) = 0 \qquad \forall\, i \in N$$

and

$$\hat{f}_{ik}(0) = 0, \qquad \tau_{ik}(0) = 0, \qquad \bar{f}_{ik}(0) = 0 \qquad \forall\, (ik) \in G.$$

Lemma 4.1 below describes properties of fluid sample paths. To formulate it, we need more notation.

For every $t \geq 0$ and every class $i \in N$, denote

$$\tau_i(t) \doteq \tau_{i,K_i}(t),$$
$$\tau_{i0}(t) \equiv t.$$

For every $t \geq 0$ and every flow $(ik) \in G$, let

$$\tau_{ik}^{*}(t) \doteq \sup\{s \leq t \mid f_i(s) \leq \widehat{f}_{ik}(t)\}$$

and let

$$w_{ik}(t) \doteq t - \tau_{ik}(t), \qquad r_{ik}(t) \doteq w_{ik}(t)/\alpha_i,$$
$$w_{ik}^{*}(t) \doteq t - \tau_{ik}^{*}(t), \qquad r_{ik}^{*}(t) \doteq w_{ik}^{*}(t)/\alpha_i.$$

It is easy to see that inequality

$$\tau_{ik}(t) \leq \tau_{ik}^{*}(t) \tag{4.9}$$

holds for all $t \geq 0$. However, a strict inequality in (4.9) is possible. To illustrate this, consider the case when $f_i(\cdot)$ has a constant value (does not increase) in an interval $[t_1, t_2]$, where $t_2 \le t$ and either $t_2 = t$ or t_2 is a point of increase of f_i. If at time t function $\widehat{f}_{ik}(t)$ reaches value $f_i(t_1)$, then, by the above definition, $\tau^*_{ik}(t)$ jumps to the value t_2, that is, to the end of the "flat region" of f_i, or to t, whichever is less. However, each DSP from a sequence which defines the FSP under consideration may be such that $f_i^{(n)}$ does increase in the interval $[t_1, t_2]$, the *set* of jump points of $f_i^{(n)}(\cdot)$ becomes asymptotically dense in $[t_1, t_2]$ as $n \to \infty$, but $f_i^{(n)}(t_2) - f_i^{(n)}(t_1) \to 0$. As a result, if service of flow (i, k) in a (nonzero length) interval $(t, t + \delta)$ is "held back" by service of another flow, it is possible that $\tau_{ik}(\cdot)$ does not jump at t, or jumps to a value which is strictly less than t_2. [Also, since τ_{ik} is right continuous with left limits (RCLL), the strict inequality in (4.9) will hold in some interval to the right of, and including, time t.] As a simple example of such "holding back" situation, suppose that there is another flow (i', k') served by the same node $j = \hat{j}(i', k') = \hat{j}(i, k)$, and such that $\alpha_{i'} = \alpha_i$ (this is just for simplicity) and $t_1 \le \tau^*_{i'k'}(t) < t_2$. In this case, $\tau_{ik}(t) \le \tau^*_{i'k'}(t)$.

Let us write the following: for every $t \ge 0$ and every node $j \in J$,

$$r_{(j)}(t) \doteq \max_{(ik)\in G_j} r_{ik}(t),$$

$$f_{(j)}(t) \doteq \sum_{(ik)\in G_j} f_{ik}(t),$$

$$\widehat{f}_{(j)}(t) \doteq \sum_{(ik)\in G_j} \widehat{f}_{ik}(t);$$

for any subset $B \subseteq G$,

$$f_{(B)}(t) \doteq \sum_{(ik)\in B} f_{ik}(t),$$

$$\widehat{f}_{(B)}(t) \doteq \sum_{(ik)\in B} \widehat{f}_{ik}(t),$$

$$r_{(B)}(t) \doteq \max_{(ik)\in B} r_{ik}(t);$$

and

$$r(t) \doteq \max_{j\in J} r_{(j)}(t) \equiv r_{(G)}(t), \qquad r^*(t) \doteq \max_{(ik)\in G} r^*_{ik}(t).$$

The following functions naturally have the meaning of queue lengths:

$$q_{ik}(t) \doteq f_{ik}(t) - \widehat{f}_{ik}(t), \qquad (ik) \in G,$$

$$q_{(j)}(t) \doteq \sum_{(ik)\in G_j} q_{ik}(t), \qquad j \in J,$$

$$q_{(B)}(t) \doteq \sum_{(ik)\in B} q_{ik}(t), \qquad B \subseteq G.$$

Let us also introduce the following functions:

$$\bar{r}_{ik}(t) \doteq \sup_{0\le s\le t} r_{ik}(s), \qquad (ik) \in G,$$

$$\bar{r}_{(j)}(t) \doteq \max_{(ik)\in G_j} \bar{r}_{ik}(t), \qquad j \in J,$$

$$\bar{r}_{(B)}(t) \doteq \max_{(ik)\in B} \bar{r}_{ik}(t), \qquad B \subseteq G,$$

$$\bar{r}(t) \doteq \max_{j\in J} \bar{r}_{(j)}(t) \equiv \bar{r}_{(G)}(t).$$

LEMMA 4.1. *For any fluid sample path* $(f, \hat{f}, \tau, \bar{f})$, *the following properties hold*:

(a) *Ordering properties*. *For every class* $i \in N$ *and all* $t \in [0, \infty)$,

$$\widehat{f}_{i,K_i}(t) \le \widehat{f}_{i,K_i-1}(t) \le \cdots \le \widehat{f}_{i,1}(t) \le f_i(t), \tag{4.10}$$

$$\tau_i(0) \le \tau_i(t) \equiv \tau_{i,K_i}(t) \le \tau_{i,K_i-1}(t) \le \cdots \le \tau_{i,1}(t) \le \tau_{i,0}(t) \equiv t, \tag{4.11}$$

$$f_{i,K_i}(t) \le f_{i,K_i-1}(t) \le \cdots \le f_{i,1}(t) = f_i(t). \tag{4.12}$$

(b) *Lipschitz properties*. *For every* $(i, k) \in G$, *the function* $(\widehat{f}_{ik}(t),\ t \ge 0)$ *is nondecreasing Lipschitz continuous*,

$$\widehat{f}_{ik}(t_2) - \widehat{f}_{ik}(t_1) \le \mu_j(t_2 - t_1), \qquad t_1 \le t_2, \tag{4.13}$$

where $j = \hat{j}(i, k)$.

For every $(i, k) \in G$, *the function* $(r_{ik}(t),\ t \ge 0)$ *is "Lipschitz above," that is*,

$$r_{ik}(t_2) - r_{ik}(t_1) \le \frac{1}{\alpha_i}(t_2 - t_1), \qquad t_1 \le t_2; \tag{4.14}$$

moreover, *for any* $B \subseteq G$, *the function* $(r_{(B)}(t),\ t \ge 0)$ *can be represented as* $r_{(B)}(t) = h_1(t) + h_2(t)$, *where* h_1 *is an absolutely continuous function with*

$$h_1'(t) \le d_{\max}$$

[*where* $d_{\max} \doteq 1/(\min_i \alpha_i)$], *and* h_2 *is nonincreasing purely singular* (*with respect to Lebesgue measure*).

For every $(i, k) \in G$, *the function* $(\bar{r}_{ik}(t),\ t \ge 0)$ *is nondecreasing Lipschitz continuous*,

$$\bar{r}_{ik}(t_2) - \bar{r}_{ik}(t_1) \le \frac{1}{\alpha_i}(t_2 - t_1), \qquad t_1 \le t_2; \tag{4.15}$$

and moreover, *for any* $B \subseteq G$, *the function* $(\bar{r}_{(B)}(t),\ t \ge 0)$ *is nondecreasing Lipschitz continuous*,

$$\bar{r}_{(B)}(t_2) - \bar{r}_{(B)}(t_1) \le d_{\max}(t_2 - t_1), \qquad t_1 \le t_2. \tag{4.16}$$

(c) *Work conservation laws. For every* $j \in J$, *any* $t \geq 0$ *and any* $a \geq 0$,

$$\begin{aligned}\widehat{f}_{(j)}(t+a) - \widehat{f}_{(j)}(t) \\ = \mu_j a + 0 \wedge \inf_{s \in [t,t+a]} \big[f_{(j)}(s) - \widehat{f}_{(j)}(t) - \mu_j (s-t)\big],\end{aligned} \tag{4.17}$$

and if in addition the infimum above is negative, then there exists $\xi \in [t, t+a]$ *such that*

$$q_{(j)}(\xi) = f_{(j)}(\xi) - \widehat{f}_{(j)}(\xi) = 0 \tag{4.18}$$

and

$$\tau_{ik}(\xi) = \tau_{i,k-1}(\xi) \qquad \forall\, (ik) \in G_j. \tag{4.19}$$

Suppose, for some $j \in J$, $t \geq 0$ *and* $a > 0$,

$$\sum_{(ik) \in G_j} f_i(t+a) - \widehat{f}_{(j)}(t) - \mu_j a < 0. \tag{4.20}$$

Then there exists $\xi \in [t, t+a]$ *such that* (4.18) *and* (4.19) *hold.*

PROOF. Consider a fixed FSP and a fixed sequence of DSP $\{(f^{(n)}, \hat{f}^{(n)}, \tau^{(n)}, \bar{f}^{(n)}),\ n = 1, 2, \ldots\}$ which defines it, that is, the convergence (4.8) holds. This sequence will be referred to as a *defining* sequence. The proofs of all statements make use of the fact that all component functions (of all DSP's of the defining sequence and the FSP) are RCLL and nondecreasing, and of the definition of weak convergence.

(a) The ordering properties follow trivially from the corresponding properties of each DSP.

(b) The Lipschitz properties are basically trivial. We only note that the decomposition $r_{(B)}(t) = h_1(t) + h_2(t)$ is obtained from the fact that $r_{(B)}$ is Lipschitz above and has bounded variation (which in turn itself follows from the Lipschitz-above property).

(c) The conservation law

$$\begin{aligned}\widehat{f}^{(n)}_{(j)}(t+a) - \widehat{f}^{(n)}_{(j)}(t) \\ = \mu_j a + 0 \wedge \inf_{s \in [t,t+a]} \big[f^{(n)}_{(j)}(s) - \widehat{f}^{(n)}_{(j)}(t) - \mu_j (s-t)\big]\end{aligned} \tag{4.21}$$

holds for each DSP of the defining sequence—this is just a form of the standard reflection mapping for the unfinished work. If we note that all $\hat{f}^{(n)}_{(j)}(\cdot)$ and $\hat{f}_{(j)}(\cdot)$ are continuous, and

$$f_{(j)}(t-) - \hat{f}_{(j)}(t) \geq 0, \qquad f^{(n)}_{(j)}(t-) - \hat{f}^{(n)}_{(j)}(t) \geq 0 \qquad \forall\, n,$$

then it can be easily shown that, as $n \to \infty$, the RHS and LHS of (4.21) converge to those of (4.17).

To prove (4.18), we use the definition of $q_{(j)}(\cdot)$ and (4.17) to obtain the following expression:

$$\begin{aligned} q_{(j)}(t+a) &= q_{(j)}(t) + f_{(j)}(t+a) - f_{(j)}(t) - \mu_j a \\ &\qquad - 0 \wedge \inf_{s\in[t,t+a]} [q_{(j)}(t) + f_{(j)}(s) - f_{(j)}(t) - \mu_j(s-t)]. \end{aligned} \tag{4.22}$$

Let us denote by $u(s)$ the value of the function under the infimum in point $s \in [t, t+a]$. Note that $u(s)$ is Lipschitz below [which means $-u(s)$ is Lipschitz above], it can decrease at most at the rate μ_j. Then there must exist $\xi \in [t, t+a]$ such that u attains its infimum over $[t, \xi]$ at point ξ, $u(\xi) < 0$, and the derivative $u'(\xi) < 0$. (See Lemma 4.3 below.) Putting $a = \xi - t$ in (4.22), we obtain (4.18). To prove (4.19) we observe the following. Since $u'(\xi) < 0$, then, for each DSP from the defining sequence, we can pick a point $\xi^{(n)} > \xi$ such that node j is empty at time $\xi^{(n)}$, and $\xi^{(n)} \downarrow \xi$ as $n \to \infty$. Using this and the fact that all functions $\tau_{ik}(\cdot)$ and $\tau_{ik}^{(n)}(\cdot)$ are nondecreasing right continuous, we obtain (4.19).

Property (4.20) follows immediately, since, for all t,

$$\sum_{(ik)\in G_j} f_i(t) \geq f_{(j)}(t).$$

□

For a given fluid sample path $(f, \hat{f}, \tau, \bar{f})$, a time point $t \in (0, \infty)$ will be called *regular* if the derivatives

$$\hat{f}'_{(B)}(t), \bar{r}'_{(B)}(t),\ r'_{(B)}(t) \qquad \forall\, B \subseteq G$$

exist and are finite. (Note that $t = 0$ is *not* a regular point.)

The Lipschitz properties described in Lemma 4.1 imply the following: *for any fluid sample path* $(f, \hat{f}, \tau, \bar{f})$, *almost all* (*with respect to Lebesgue measure*) *points* $t \in [0, \infty)$ *are regular.*

Note that in any regular point t the finite derivatives $\tau'_{ik}(t)$, $(ik) \in G$, also exist since $\tau_{ik}(t) = t - \alpha_i r_{ik}(t)$.

For a node j at time t, we will use the notation

$$\tilde{G}_j(t) \doteq \{(ik) \in G_j \mid r_{ik}(t) = r_{(j)}(t)\}$$

for the subset of flows on which $r_{(j)}(t)$ is attained.

It follows from the above definition that in any regular point t, for any $j \in J$,

$$\tau'_{ik}(t) = 1 - \alpha_i r'_{(j)}(t) \qquad \forall\, (ik) \in \tilde{G}_j(t).$$

Very often we will consider the following condition on a node $j \in J$ at time t:

$$\tau_{ik}(t) = \tau^*_{ik}(t) < \tau_{i,k-1}(t) \qquad \text{for at least one } (ik) \in \tilde{G}_j(t). \tag{4.23}$$

This condition in particular implies that $r_{(j)}(t) > 0$ and that [for any (ik) for which the condition holds] $\tau_{ik}(t)$ is a *right point of increase* of $f_i(\cdot)$, that is, $f_i(s) > f_i(\tau_{ik}(t))$ for any $s > \tau_{ik}(t)$.

LEMMA 4.2. *Consider a fluid sample path* $(f, \hat{f}, \tau, \bar{f})$ *such that for some fixed* T_1 *and* T_2, $0 \le T_1 \le T_2$, *for all flows* $i \in N$, *the function* f_i *is continuous in the interval* $[T_1, \infty)$, *and*

$$\widehat{f}_{i,K_i}(T_2) \ge f_i(T_1) \quad \text{and} \quad \tau_i(T_2) \ge T_1. \tag{4.24}$$

Then the following properties hold:

(a) *Flow conservation—For any* $(i, k) \in G$ *and* $t \ge T_2$,

$$f_{ik}(t) = \hat{f}_{i,k-1}(t) \tag{4.25}$$

and

$$\hat{f}_{i,k}(t) = f_i(\tau_{i,k}(t)). \tag{4.26}$$

(b) *If for some* $j \in J$ *and* $t \ge T_2$, *property* (4.23) *holds* [*which in particular means that* $r_{(j)}(t) > 0$], *then there exists* $\varepsilon_1 > 0$ *such that*, *for any* $\varepsilon \in [0, \varepsilon_1]$,

$$\widehat{f}_{(B)}(t+\varepsilon) - \widehat{f}_{(B)}(t) = \mu_j \varepsilon,$$

where $B = \tilde{G}_j(t)$.

(c) "*Instantaneous critical node*" *property*,

$$r(t) = r^*(t), \qquad t \ge T_2.$$

Suppose in addition that all functions f_i *are absolutely continuous in the interval* $[T_1, \infty)$ *and*, *moreover*, *for any* $T \ge T_1$ *the cost of* f *in* $[T_1, T]$ *is finite*, *that is*,

$$J_{T,T_1}(f) = \sum_{i \in N} \int_{T_1}^{T} L_i(f_i'(s))\, ds < \infty.$$

Then the following additional properties hold:

(d) *For each* $j \in J$ *consider the set* $H_j \subseteq [T_2, \infty)$ *of regular points* t *such that*

$$\exists\, (ik) \in G_j, \qquad \tau_{ik}'(t) = 0 \quad \text{and} \quad \hat{f}_{ik}'(t) > 0. \tag{4.27}$$

Then, H_j *has zero Lebesgue measure*

$$\Lambda(H_j) = 0, \qquad j \in J. \tag{4.28}$$

(e) *For any node* $j \in J$, *in any regular point* $t \in [T_2, \infty) \setminus H_j$ *such that the condition* (4.23) *holds*, *we have*

$$\widehat{f}'_{ik}(t) = f_i'(\tau_{ik}(t))\tau_{ik}'(t) \qquad \forall\, (ik) \in \tilde{G}_j(t), \tag{4.29}$$

and

$$\sum_{(ik) \in \tilde{G}_j(t), \tau_{ik}'(t) > 0, f_i'(\tau_{ik}(t)) > 0} f_i'(\tau_{ik}(t))\tau_{ik}'(t) = \sum_{(ik) \in \tilde{G}_j(t)} \widehat{f}'_{ik}(t) = \mu_j. \tag{4.30}$$

Before we prove Lemma 4.2, let us introduce notation used later in the paper, which explains why statement (c) of the lemma is called the instantaneous critical node property.

Let us write

$$\tilde{J}(t) \doteq \{j \in J \mid r_{(j)}(t) = r(t);\ \text{condition (4.23) holds}\}.$$

[Recall that (4.23) implies $r_{(j)}(t) > 0$.] We will call any node $j \in \tilde{J}(t)$ an *instantaneous critical node* of the network at time t. The crucial statement (c) of Lemma 4.2 implies the following property.

COROLLARY 4.1. *Within the conditions of Lemma* 4.2, *for any* $t \geq T_2$ *such that* $r(t) > 0$, *there exists at least one instantaneous critical node*; *that is*, $\tilde{J}(t)$ *is nonempty*.

Indeed, according to statement (c), for such t there exists at least one flow $(ik) \in G$ such that

$$r_{ik}(t) = r_{ik}^*(t) = r(t) = r^*(t).$$

Let l be the maximal integer from the set $\{1, \dots, k\}$, such that $\tau_{il}(t) < \tau_{i,l-1}(t)$. [Such l exists because $\tau_{i0}(t) = t$.] For the flow $(il) \in G$ we also must have

$$r_{il}(t) = r_{il}^*(t) = r(t) = r^*(t),$$

which means that, for $j = \hat{j}(i, l)$, we have

$$r_{(j)}(t) = r(t), \qquad (i, l) \in \tilde{G}_j(t) \quad \text{and} \quad \tau_{il}(t) = \tau_{il}^*(t) < \tau_{i,l-1}(t).$$

Thus, $j \in \tilde{J}(t)$.

When $\tilde{J}(t)$ is nonempty, we will write

$$\tilde{j}(t) \doteq \min\{j \mid j \in \tilde{J}(t)\}.$$

[The function $\tilde{j}(t)$ can be any mapping that picks a unique well-defined representative element of the subset $\tilde{J}(t)$.]

For every regular point t such that $r(t) > 0$, we define

$$\tilde{N}(t) \doteq \{i \mid \exists k, (ik) \in \tilde{G}_{\tilde{j}(t)}(t), \tau_{ik}'(t) > 0, f_i'(\tau_{ik}(t)) > 0\}. \tag{4.31}$$

PROOF OF LEMMA 4.2. As in the proof of Lemma 4.1, throughout this proof we consider a fixed FSP and a defining fixed sequence of DSP $\{(f^{(n)}, \hat{f}^{(n)}, \tau^{(n)}, \bar{f}^{(n)}), n = 1, 2, \dots\}$, that is, the sequence such that the convergence (4.8) holds. Again, we heavily employ the fact that all component functions of the FSP and of each DSP are nondecreasing RCLL.

(a) To prove (4.25) we note that, trivially, $\hat{f}_{i,k-1}(t) \geq f_{ik}(t)$, because the same inequality holds for every DSP of the defining sequence. Let us prove by contradiction that the equality must hold. Suppose not. Then we write

$$\varepsilon = \frac{\hat{f}_{i,k-1}(t) - f_{ik}(t)}{3} > 0. \tag{4.32}$$

We can always choose $\varepsilon_1 > 0$ arbitrarily small and such that both $\hat{f}_{i,k-1}(\cdot)$ and $f_{ik}(\cdot)$ are continuous in point $t + \varepsilon_1$, and therefore, for all sufficiently large n,

$$f_{ik}^{(n)}(t + \varepsilon_1) < f_{ik}(t) + \varepsilon \quad \text{and} \quad \hat{f}_{i,k-1}^{(n)}(t + \varepsilon_1) > f_{ik}(t) + 2\varepsilon.$$

The only way the latter inequalities may hold is if the node $\hat{j}(i, k-1)$ is serving an $(i, k-1)$-customer at time $t + \varepsilon_1$ and the amount of service the customer already received is greater than ε. This implies that, for all large n, the function $f_i^{(n)}(\cdot)$ is such that, for some point $s = s^{(n)} \leq t + \varepsilon_1$,

$$f_i^{(n)}(s) > \hat{f}_{i,k-1}(t) + 2\varepsilon \geq \hat{f}_{i,k-1}(T_2) + 2\varepsilon \geq f_i(T_1) + 2\varepsilon$$

and

$$f_i^{(n)}(s) - f_i^{(n)}(s-) > \varepsilon.$$

Since all functions $f_i^{(n)}(\cdot)$ and $f_i(\cdot)$ are nondecreasing and $f_i(\cdot)$ is continuous in $[T_1, \infty)$, the weak convergence $f_i^{(n)}(\cdot) \Rightarrow f_i(\cdot)$ cannot hold. This contradiction proves that (4.32) is impossible, and therefore (4.25) is proved.

To prove (4.26), we notice that the inequality

$$\hat{f}_{i,k}(t) \leq f_i(\tau_{i,k}(t)) \tag{4.33}$$

must hold for all $t \geq 0$ (not only $t \geq T_2$). Indeed, $\hat{f}_{i,k}^{(n)}(s) \leq f_i^{(n)}(\tau_{i,k}^{(n)}(s))$ trivially holds for all n and all $s \geq 0$. For any fixed $t \geq 0$, we can pick a point $s > t$, arbitrarily close t, such that both functions $\hat{f}_{i,k}(\cdot)$ and $\tau_{i,k}(\cdot)$ are continuous in point s. Then

$$\liminf_{n \to \infty} f_i^{(n)}(\tau_{i,k}^{(n)}(s)) \geq \hat{f}_{i,k}(s) \geq \hat{f}_{i,k}(t).$$

This easily implies

$$f_i(\tau_{i,k}(s)) \geq \hat{f}_{i,k}(t),$$

and since s can be arbitrarily close to t, we obtain (4.33).

We need to show that, actually, for $t \geq T_2$, the equality in (4.33) holds. Suppose not, that is,

$$\hat{f}_{i,k}(t) < f_i(\tau_{i,k}(t)). \tag{4.34}$$

If $\tau_{i,k}(t) = T_1$, then the contradiction is immediate:

$$\hat{f}_{i,k}(t) \geq \hat{f}_{i,k}(T_2) \geq f_i(T_1) = f_i(\tau_{i,k}(t)).$$

Suppose, $\tau_{i,k}(t) > T_1$. Then we can pick $\varepsilon > 0$, arbitrarily close to 0, such that $\hat{f}_{i,k}(\cdot)$ and $\tau_{i,k}(\cdot)$ are both continuous at $t+\varepsilon$. Since $\tau_{i,k}^{(n)}(t+\varepsilon) \to \tau_{i,k}(t+\varepsilon)$ and $f_i^{(n)}(\cdot)$ converges to $f_i(\cdot)$ uniformly in a small neighborhood of $t+\varepsilon$, it is easy to see that, for the defining DSP with index n, the residual service requirement of the head-of-the-line customer of type (i,k) (if any) in node $\hat{j}(i,k)$ at time $\tau_{i,k}^{(n)}(t+\varepsilon)$, converges to 0 as $n \to \infty$. This residual service requirement is exactly equal to $f_i^{(n)}(\tau_{i,k}^{(n)}(t+\varepsilon)) - \hat{f}_{i,k}^{(n)}(t+\varepsilon)$. Taking the limit on n of the above difference, we obtain $f_i(\tau_{i,k}(t+\varepsilon)) = \hat{f}_{i,k}(t+\varepsilon)$ and, since ε can be arbitrarily small,

$$f_i\bigl(\tau_{i,k}(t)\bigr) = \hat{f}_{i,k}(t).$$

This is a contradiction to (4.34) which finally proves (4.26).

(b) It is easy to see that if we choose $\varepsilon_1 > 0$ to be sufficiently small, then for any $\varepsilon_2 \in (0, \varepsilon_1)$, all DSP (from the defining sequence) with sufficiently large n are such that in the interval $[t+\varepsilon_2, t+\varepsilon_1]$, first, node j cannot be idle and, second, only customers from the subset $\tilde{G}_j(t)$ may be served. The desired property easily follows.

(c) The proof is by contradiction. Suppose $r(t) > r^*(t)$. Then there exists at least one $(ml) \in G$ such that

$$r_{ml}(t) = r(t) > r^*(t) \tag{4.35}$$

and

$$\tau_{ml}(t) < \tau_{m,l-1}(t). \tag{4.36}$$

[Indeed, pick any (mk) such that $r_{mk}(t) = r(t)$, and let l be the maximal element of the set $\{1, \dots, k\}$ such that (4.36) holds. Such l is well defined, because by our convention $\tau_{m0}(t) \equiv t$ for any m.]

Let us denote $\hat{r} = r^*(t)$. Let us fix $\varepsilon_1 > 0$ such that

$$\frac{t - \tau_{m,l-1}(t)}{\alpha_m} < \hat{r} + \varepsilon_1 < r_{ml}(t) = r(t),$$

and a small $\varepsilon_2 = \varepsilon_2(\varepsilon_1) > 0$ such that $\varepsilon_2/\alpha_{\min} < \varepsilon_1$, and therefore

$$\hat{r} + \frac{\varepsilon_2}{\alpha_{\min}} < \hat{r} + \varepsilon_1,$$

where $\alpha_{\min} \doteq \min_{i \in N} \alpha_i$.

If we recall the definition of the functions $\tau_{ik}^*(\cdot)$, we see that the inequality in (4.35) implies that, for any flow $i \in N$, $\tau_i(t) \le t - \alpha_i r(t) < t - \alpha_i \hat{r}$ and the function $f_i(\cdot)$ is constant in the interval $[\tau_i(t), t - \alpha_i \hat{r})$. This in turn implies [again, using the definition of $\tau_{ik}^*(\cdot)$] the following.

OBSERVATION 1. *For any $\varepsilon_3 > 0$ and for any flow $i \in N$, we have*

$$f_i^{(n)}(t - \alpha_i \hat{r} - \varepsilon_3) - \hat{f}_{i,K_i}^{(n)}(t) \to 0, \qquad n \to \infty.$$

In other words, the defining DSP sequence is such that the amount of class i work arrived in the network before time $t - \alpha_i \hat{r} - \varepsilon_3$ and still present (anywhere in the network) at time t, vanishes as $n \to \infty$.

Observation 1, the LWDF scheduling rule, and the work-conservation property of the LWDF (i.e., the fact that node j cannot be idle as long as there is some unfinished work already arrived in the node), allow us to verify the following.

OBSERVATION 2. *For all defining DSP with sufficiently large n, all class m customers that arrived in the network before time $t - \alpha_m(\hat{r} + \varepsilon_1)$ will be completely served by node $j = \hat{j}(m, l)$ before time $t + \varepsilon_2$.*

Indeed, for all sufficiently large n, all class m work that arrived in the network before time $t - \alpha_m(\hat{r} + \varepsilon_1)$ has arrived in node j by the time $t + (1/3)\varepsilon_2$ [because $\tau_{m,l-1}(t) > t - \alpha_m(\hat{r} + \varepsilon_1)$], and the residual amount of this work at time t converges to 0 as $n \to \infty$ (by Observation 1). The amount of work of all other classes $i \neq m$ which can possibly "compete" for server j in the interval $[t + (1/3)\varepsilon_2, t + (2/3)\varepsilon_2]$ [with class m work arrived in the network before $t - \alpha_m(\hat{r} + \varepsilon_1)$] also converges to 0 (again, by Observation 1). This implies Observation 2.

We see that, for all large n,

$$r_{ml}^{(n)}(t + \varepsilon_2) \le \hat{r} + \varepsilon_1 + \varepsilon_2/\alpha_m.$$

From the fact that each function $\tau_{ml}^{(n)}$ and τ_{ml} is nondecreasing RCLL, we obtain

$$r_{ml}(t + \varepsilon_2) \le \hat{r} + \varepsilon_1 + \varepsilon_2/\alpha_m.$$

Since ε_2 can be chosen arbitrarily small, we derive (again using the fact that τ_{ml} is nondecreasing RCLL)

$$r_{ml}(t) \le \hat{r} + \varepsilon_1,$$

a contradiction to the assumption (4.35).

(d) For any fixed j consider the function

$$h(t) = \sum_{(ik) \in G_j} \int_{T_1}^{\tau_{ik}(t)} L_i\big(f_i'(s)\big)\, ds.$$

This function is clearly RCLL and nondecreasing. Consider any point $t \in H_j$ and $(ik) \in \tilde{G}_j(t)$ such that

$$\tau_{ik}'(t) = 0 \quad \text{and} \quad \hat{f}_{ik}'(t) > 0. \tag{4.37}$$

Since a simple relation $\hat{f}_{ik}(s) = f_i(\tau_{ik}(s))$ holds for $s \geq t$, we see that condition (4.37) is possible only if $f_i(\cdot)$ has (positive) infinite derivative in point $\tau_{ik}(t)$. Using the fact that the local rate function $L_i(\cdot)$ is convex and $\lim_{x\to\infty} L_i(x)/x = \infty$, it is easy to verify that t is a point where $h(\cdot)$ has infinite right derivative:

$$\frac{d^+}{dt}h(t) = +\infty.$$

Lebesgue measure of the set of points where a nondecreasing function has infinite right derivative is 0.

(e) These properties follow from the property (b) and the definitions of the regular point and the set H_j. $\square$

The following elementary Lemma 4.3 is very useful in this paper. (We omit a rather straightforward proof.)

LEMMA 4.3. *Suppose a function $h \in \mathcal{D}_{(0)}$ is Lipschitz above; that is, there exists a constant a such that, for any t_1, t_2, $0 \leq t_1 \leq t_2$,*

$$h(t_2) - h(t_1) \leq a(t_2 - t_1).$$

Define the function $\bar{h} \in \mathcal{D}_{(0)}$ as

$$\bar{h}(t) = \sup_{0\leq s\leq t} h(s), \qquad t \geq 0.$$

Then the function $\bar{h}$ is nondecreasing Lipschitz continuous with Lipschitz constant equal to a, and, for any $t \geq 0$,

$$\bar{h}'(t) > 0 \quad \textit{implies} \quad h(t) = \bar{h}(t) \quad \textit{and} \quad h'(t) = \bar{h}'(t).$$

5. LWDF network: optimal fluid sample path to build maximal weighted delay. Critical node property. This section contains the results which are central to our analysis. Considering each LWDF network node in isolation, we construct a set f^0 of fluid input flows which has a simple special structure. Then we show that f^0 is an optimal (lowest cost) path to build maximal weighted delay r (starting from zero initial condition), in *both* the entire network and one of the nodes in isolation.

Consider a fixed node j. We remind the reader that $N_j \subseteq N$ denotes the subset of classes (flows) which have this node on their route. Consider a fixed subset $\tilde{N}_j \subseteq N_j$ and a set of numbers $x_i > 0$, $i \in \tilde{N}_j$, such that

$$0 < \gamma < \min_{i\in\tilde{N}_j} \frac{1}{\alpha_i}, \tag{5.1}$$

where

$$\gamma = \frac{\sum_{m\in\tilde{N}_j} x_m - \mu_j}{\sum_{m\in\tilde{N}_j} \alpha_m x_m}. \tag{5.2}$$

(The existence of such $\tilde{N}_j$ and a corresponding set of x_i is obvious: for arbitrary $i \in N_j$ we can choose $\tilde{N}_j = \{i\}$ and arbitrary $x_i > \mu_j$.)

Note that (5.2) can be rewritten as

$$\sum_{m \in \tilde{N}_j} x_m(1 - \alpha_m \gamma) = \mu_j. \tag{5.3}$$

Consider the following set of fluid input flows $f^0 = (f_i^0 = (f_i^0(t), t \in \mathbb{R}), i \in N)$, which we call a *network simple element* (associated with the node j and the parameters $\tilde{N}_j, x_i,\ i \in \tilde{N}_j$):

$$f_i^0(t) \doteq \begin{cases} \lambda_i t, & t \in [0, \infty),\ \text{for } i \in N \setminus \tilde{N}_j, \\ x_i t, & t \in [0, T_i^0],\ \text{for } i \in \tilde{N}_j, \\ x_i T_i^0 + \lambda_i (t - T_i^0), & t \in [T_i^0, \infty),\ \text{for } i \in \tilde{N}_j, \end{cases} \tag{5.4}$$

where

$$T^0 \doteq \frac{1}{\gamma} \quad \text{and} \quad T_i^0 \doteq (1 - \alpha_i \gamma) T^0 = T^0 - \alpha_i. \tag{5.5}$$

This definition of f^0 trivially implies that, for all i, $f_i^0(t) \equiv 0$ for $t \le 0$.

It is easy to verify directly (see also [31]) that f^0 has, in particular, the following properties:

$$\text{all functions } f_i^0(\cdot) \text{ are strictly increasing continuous in } [0, \infty), \tag{5.6}$$

$$J_{T^0}(f^0) = \frac{1}{\gamma} \sum_{i \in \tilde{N}_j} (1 - \gamma \alpha_i) L_i(x_i), \tag{5.7}$$

$$\sum_{i \in \tilde{N}_j} f_i^0(T_i^0) = \mu_j T^0. \tag{5.8}$$

LEMMA 5.1. *Consider any fluid sample path* $(f^0, \widehat{f}, \tau, \bar{f})$ *with* f^0 *being a network simple element defined by* (5.4). (*This FSP necessarily starts from zero.*) *Then the maximal weighted delay* $r^0(\cdot)$ *associated with this FSP is such that*

$$r^0(T^0) \ge 1. \tag{5.9}$$

PROOF. For this FSP, in addition to Lemma 4.1, all statements of Lemma 4.2 hold with $T_1 = T_2 = 0$. Given properties (5.6)–(5.8) and the relation $\widehat{f}_{ik}(t) = f_i(\tau_{ik}(t))$, if (5.9) would not hold, we would obtain

$$\begin{aligned} \widehat{f}_{(j)}(T^0) &= \sum_{(ik) \in G_j} \widehat{f}_{ik}(T^0) \\ &\ge \sum_{(ik) \in G_j, i \in \tilde{N}_j} f_i(\tau_{ik}(T^0)) > \sum_{(ik) \in G_j,\ i \in \tilde{N}_j} f_i(T_i^0) = \mu_j T^0. \end{aligned}$$

This contradicts the conservation law (4.17). $\square$

For a fixed node j consider the following optimization problem, which is the same as the problem (3.7)–(3.12), but with relaxed constraints:

$$J_{*,j} = \inf_{\nu} \frac{1}{\gamma} \sum_{i \in \tilde{N}_j} (1 - \alpha_i \gamma) L_i(x_i), \tag{5.10}$$

where

$$\nu = \{\tilde{N}_j \subseteq N_j, (x_i, i \in \tilde{N}_j)\}, \tag{5.11}$$

γ is a function of ν defined by (5.2) and minimization is subject to the constraints

$$x_i \geq 0, \qquad i \in \tilde{N}_j, \tag{5.12}$$

and (5.1).

We use the same notation $J_{*,j}$ for the infima in the problems (5.10) and (3.7) because the following lemma shows that they are indeed equal.

LEMMA 5.2. *Consider the optimization problem* (5.10)–(5.12), (5.2), (5.1). *The following properties hold*:

(a) $J_{*,j} < \infty$ *if and only if*

$$\exists\, \nu \text{ such that } \qquad \sum_{i \in \tilde{N}_j} x_i > \mu_j, L_i(x_i) < \infty \ \forall i \in \tilde{N}_j. \tag{5.13}$$

(b) *If condition* (5.13) *holds*, *then the infimum* $J_{*,j} < \infty$ *is attained. Any optimization variable* ν *on which the minimum is attained satisfies additional conditions*:

$$x_i \geq \lambda_i, \qquad i \in \tilde{N}_j, \tag{5.14}$$

and

$$\frac{1}{\alpha_i} \leq \gamma \qquad \forall i \in N_j \setminus \tilde{N}_j. \tag{5.15}$$

Moreover, *if for all* $i \in N_j$ *the local rate function* $L_i(\cdot)$ *has zero right derivative in point* λ_i, $(d^+/dx)L_i(\lambda_i) = 0$, *then*

$$x_i > \lambda_i, \qquad i \in \tilde{N}_j. \tag{5.16}$$

(c) *If condition* (5.13) *does not hold*, *then the infimum* $J_{*,j} = \infty$ *is attained on any* ν *satisfying the constraints of the problem.*

PROOF. Let us denote by

$$g(\nu) = \sum_{i \in \tilde{N}_j} \left(\frac{1}{\gamma} - \alpha_i\right) L_i(x_i) \tag{5.17}$$

the function being minimized, by $x_i(\nu)$ the x_i-component of ν and by $\gamma(\nu)$ the function of ν defined by (5.2).

(a) The "only if" claim is trivial. Let us prove that (5.13) implies $J_{*,j} < \infty$. Consider arbitrary ν satisfying (5.13). Then $\gamma(\nu) > 0$, but the second inequality in (5.1) may not hold. If the latter is the case, let us pick any $i \in \tilde{N}_j$ with the largest α_i, and remove this i from $\tilde{N}_j$. It is easy to see that ν modified this way is such that $\gamma(\nu) > 0$ still holds. If necessary, we repeat this "removal" procedure until the second inequality in (5.1) also holds. (At the latest, the procedure will stop when only one flow remains in the set $\tilde{N}_j$—in this case the desired inequality holds for sure.) The value of ν we obtain is such that $g(\nu) < \infty$, and we are done. (Note that we can also remove from $\tilde{N}_j$ all the flows i with $x_i = 0$.)

(b) Consider a sequence $\{\nu_l,\ l = 1, 2, \ldots\}$, such that each ν_l satisfies all the problem constraints, $g(\nu_l) \downarrow J_{*,j}$ and the subset $\tilde{N}_j$ (a component of ν_l) is constant. It is easy to see from (5.17) that $\gamma(\nu_l)$ must stay bounded away from 0, and it stays bounded away from infinity due to (5.1). We will choose a subsequence of $\{\nu_l\}$ (which we still denote by $\{\nu_l\}$) such that $\gamma(\nu_l)$ and each $x_i(\nu_l)$ converge to some limits which we will denote by γ and x_i, respectively. Obviously,

$$0 < \gamma \le \min_{i \in \tilde{N}_j} 1/\alpha_i = a, \tag{5.18}$$

all $x_i \ge 0$, but some of the x_i may be infinite.

Consider the case when some x_i are infinite. This is possible only if $\gamma = a$. Denote by $B \subseteq \tilde{N}_j$ the subset of those i with $1/\alpha_i = a = \gamma$. Note that $1 - \alpha_i \gamma(\nu_l) \to 0$ for all $i \in B$. Obviously, x_i may be infinite only for $i \in B$. Also, if $x_i = \infty$, then

$$x_i(\nu_l)\big(1 - \alpha_i \gamma(\nu_l)\big) \to 0, \tag{5.19}$$

because otherwise, for this i,

$$\limsup L_i\big(x_i(\nu_l)\big)\big(1 - \alpha_i \gamma(\nu_l)\big) = \infty$$

[recall that $L_i(y)/y \to \infty$ as $y \to \infty$]. This implies that (5.19) holds for all $i \in B$. [If (5.19) would not hold for some $i \in B$, then $x_i = \infty$ for that i, which is impossible.] This in turn implies that $\tilde{N}_j \setminus B$ is nonempty, because, by (5.3), condition (5.19) cannot hold for all $i \in \tilde{N}_j$. We can make the following observation: *If we were to exclude B a priori from the subset $\tilde{N}_j$, then, for the appropriately modified subsequence $\{\nu_l\}$ we would have the following*: *all the problem constraints are satisfied for each ν_l* (*except maybe a finite number of them*); $\gamma(\nu_l) \to \gamma$; $x_i(\nu_l) \to x_i$ *and x_i is finite for each $i \in \tilde{N}_j$*; *and* $\limsup g(\nu_l) \le J_{*,j}$, *implying* $\lim g(\nu_l) = J_{*,j}$. If we write $\nu_* = \lim \nu_l$, we see that $g(\nu_*) = J_{*,j}$.

Thus, we have proved the existence of ν_* such that $g(\nu_*) = J_{*,j}$, $x_i = x_i(\nu_*) < \infty$ for all $i \in \tilde{N}_j$, and ν_* satisfies all constraints of the problem except maybe the fact that the equality $\gamma = \gamma(\nu_*) = 1/\alpha_i$ holds for some $i \in \tilde{N}_j$. If the latter occurs, then, since (5.2) and (5.3) are equivalent, we can exclude such flows i

from the subset $\tilde{N}_j$, which does not change the value of γ and can only decrease the value of $g(\nu_*)$. For this modified ν_* all constraints are satisfied. We have proved that the infimum is attained.

Now, let us prove that any ν_* such that $g(\nu_*) = J_{*,j}$ must satisfy the additional conditions (5.14) and (5.15). If $x_i < \lambda_i$ were to hold for some $i \in \tilde{N}_j$, then we could do the following. Let us continuously increase this x_i from its "original" value to λ_i. Then γ will also increase continuously. (Condition $\gamma < 1/\alpha_i$ implies that the derivative of γ on x_i is positive.) If the increasing γ will "hit" $1/\alpha_m$ for some $m \in \tilde{N}_j$, then we will remove that m from $\tilde{N}_j$ (without changing γ, as already explained above), and then keep increasing x_i. Note that, in this procedure of changing x_i to λ_i, the value of γ cannot hit $1/\alpha_i$. Indeed, if this were to happen at some point, then the class i itself and all $m \in \tilde{N}_j$ with $1/\alpha_m \le 1/\alpha_i$ could be removed, and (since γ has strictly increased) we would get a contradiction to (5.3). Since γ has strictly increased and $L_i(\lambda_i) = 0$, we have strictly decreased the value of $g(\nu_*)$—a contradiction which proves (5.14).

The proof of (5.15) essentially repeats that of (5.14). Suppose (5.15) does not hold. Let us pick $i \in N_j \setminus \tilde{N}_j$ such that $\gamma < 1/\alpha_i$, include this i in $\tilde{N}_j$, and put initially $x_i = 0$. The new ν_* satisfies all the constraints of the problem. Then repeating the procedure described above, we could change x_i from 0 to λ_i (maybe excluding some m from the original $\tilde{N}_j$ in the process), and construct a modified ν_* with strictly smaller $g(\nu_*)$.

Finally, if we had $x_i = \lambda_i$ for some $i \in \tilde{N}_j$, and $(d^+/dx)L_i(\lambda_i) = 0$, then $g(\nu_*)$ could be improved. Indeed, $(\partial^+/\partial x_i)\gamma(\nu_*) > 0$, and therefore $(\partial^+/\partial x_i)g(\nu_*) < 0$. This contradiction proves (5.16).

(c) Statement (c) trivially follows from statement (a). □

Let us write

$$J_* \doteq \min_{j \in J} J_{*,j}. \tag{5.20}$$

REMARK. If Assumption 3.2 holds, then $L_i(x_i) < \infty$ for all $x_i \ge 0$ and all $i \in N$. Therefore, by Lemma 5.2(a), $J_{*,j} < \infty$ for all nodes j, which implies $J_* < \infty$.

Let us choose a node j for which the minimum in (5.20) is attained, and choose ν_*, that is, the set of optimization variables (5.11), which solves problem (5.10) [and problem (3.7)] for this j. We will call the network simple element f^0 associated with this node j and ν_* an *optimal network simple element*. By our construction, $J_* \doteq J_{T^0}(f^0)$. The name "optimal network simple element" for f^0 is justified by the following lemma. (In the rest of this paper, f^0 denotes some a priori chosen optimal network simple element associated with a fixed node j.)

LEMMA 5.3. *The following property holds*:

$$J_* = \inf J_s(f), \tag{5.21}$$

where the infimum is over

$$s \ge 0, \qquad \{f \in \mathcal{A}_{+,0} \mid \zeta_0 f \in \mathcal{A}_{(0),+,0}; \\ \exists (\hat{f}, \tau, \bar{f}) \in A(f, 0, 0, 0), r(s) \ge 1\}. \tag{5.22}$$

Moreover, *the infimum is attained on the optimal network simple element* f^0.

We will call the property described by this lemma the *critical node* property because it implies that f^0 is a lowest cost path for both the network and the isolated node j with which it is associated.

PROOF OF LEMMA 5.3. Consider a fluid sample path $(f, \hat{f}, \tau, \bar{f})$ with zero initial condition and all $f_i(\cdot)$ being absolutely continuous. Suppose there exists a constant $T > 0$ such that $r(T) \ge 1$. Obviously, $\bar{r}(T) \ge 1$.

We note that for this FSP, in addition to Lemma 4.1, all statements of Lemma 4.2 hold with $T_1 = T_2 = 0$.

We recall that the function $\bar{r}(\cdot)$ is nondecreasing Lipschitz continuous and therefore is absolutely continuous.

Consider the subset

$$B_T^{(1)} \doteq \{t \in [0, T] \mid t \text{ is regular}, \bar{r}'(t) > 0\}.$$

The following properties either follow directly from the definition of a regular point and Lemmas 4.1 and 4.2, or are easily verified:

(a) For any $t \in B_T^{(1)}$, for any $j \in \tilde{J}(t)$ and $(ik) \in \tilde{G}_j(t)$:

$$r'_{ik}(t) = r'(t) = \bar{r}'(t) > 0$$

and

$$\tau'_{ik}(t) = 1 - \alpha_i r'(t) < \infty.$$

(b) The image

$$B_r^{(1)} \doteq \bar{r}(B_T^{(1)}) \subseteq [0, \bar{r}(T)]$$

[where $\bar{r}(T) \ge r(T) \ge 1$] has Lebesgue measure $\bar{r}(T)$:

$$\Lambda(B_r^{(1)}) = \bar{r}(T).$$

(c1) We have

$$\frac{d(\Lambda\bar{r})}{d\Lambda}(t) = r'(t), \qquad t \in B_T^{(1)}, \tag{5.23}$$

where $\Lambda\bar{r}$ is the measure on $B_T^{(1)}$ defined as

$$(\Lambda\bar{r})(U) \doteq \Lambda(\bar{r}(U))$$

for any Lebesgue-measurable $U \subseteq B_T^{(1)}$.

(c2) We have

$$\frac{d(\Lambda\bar{r}^{-1})}{d\Lambda}(x) = \frac{1}{r'(\bar{r}^{-1}(x))}, \qquad x \in B_r^{(1)}, \tag{5.24}$$

where $(\Lambda\bar{r}^{-1})(U) = \Lambda(\bar{r}^{-1}(U))$ for Lebesgue-measurable $U \subseteq B_r^{(1)}$.

(d) Let us write $B_T \doteq B_T^{(1)} \setminus \bigcup_j H_j$ and $B_r \doteq \bar{r}(B_T) \subseteq B_r^{(1)}$. Then properties (c1) and (c2) imply that

$$\Lambda(B_T) = \Lambda(B_T^{(1)}) \quad \text{and} \quad \Lambda(B_r) = \Lambda(B_r^{(1)}) = \bar{r}(T),$$

and the relations (5.23) and (5.24) still hold for the measures restricted to the subsets B_T and B_r.

(e) The set $\tilde{N}(t)$ [defined by (4.31)] is nonempty for every $t \in B_T$.

We can write

$$\begin{aligned}
J_T(f) &= \sum_{i=1}^{N} \int_0^T L_i(f_i'(s))\,ds \\
&\geq \sum_{i=1}^{N} \int_{\{t\in[0,T] | \tau_i'(t)<\infty\}} L_i(f_i'(\tau_i(t)))\tau_i'(t)\,dt \\
&\geq \int_{B_T} dt \sum_{i\in\tilde{N}(t)} L_i(f_i'(\tau_i(t)))\tau_i'(t) \\
&= \int_{B_T} dt \sum_{i\in\tilde{N}(t)} (1-\alpha_i r'(t)) L_i(f_i'(\tau_i(t))) \\
&= \int_{B_r} dy \frac{1}{r'(\bar{r}^{-1}(y))} \sum_{i\in\tilde{N}(\bar{r}^{-1}(y))} (1-\alpha_i r'(\bar{r}^{-1}(y))) L_i(f_i'(\tau_i(\bar{r}^{-1}(y)))) \\
&= \int_{B_r} c(y)\,dy, \qquad (5.25)
\end{aligned}$$

where in (5.25) we have used the following notation. For every $y \in B_r$, we consider the node $j = \tilde{j}(\bar{r}^{-1}(y))$ and the subset of flows $\tilde{N}_j = \tilde{N}(\bar{r}^{-1}(y)) \subseteq N_j$, and denote

$$c(y) \doteq \frac{1}{\gamma} \sum_{i\in\tilde{N}_j} (1-\alpha_i\gamma) L_i(x_i),$$

$$\gamma = r'(\bar{r}^{-1}(y)) > 0 \quad \text{and} \quad x_i = f_i'(\tau_i(\bar{r}^{-1}(y))) > 0,\ i \in \tilde{N}_j.$$

It follows from Lemma 4.2(e) that, for any $y \in B_r$, condition (5.3) holds [which is equivalent to (5.2)]. The definition of the set $\tilde{N}_j$ implies condition (5.1). This means (by Lemma 5.2) that $c(y) \geq J_{*,j}$, and therefore $c(y) \geq J_*$ for any $y \in B_r$. Using this lower bound in (5.25) and the fact that $\bar{r}(T) \geq 1$, we finally obtain

$$J_T(f) \geq J_*.$$

On the other hand, for the optimal network simple element f^0, we have $J_{T^0}(f^0) = J_*$ and, according to Lemma 5.1, for any FSP with $f = f^0$, we have $r^0(T^0) \geq 1$, where $r^0(\cdot)$ is the r-component of this FSP. Therefore, the infimum in (5.21) is indeed attained on f^0, which completes the proof. [It is easy to verify using scaling property of fluid sample paths, proved later in the paper, that actually the equality $r^0(T^0) = 1$ must hold.] $\square$

6. Proof of the large deviations lower bound (LDLB) [Theorem 3.2(ii)]. Note that if $J_* = \infty$ the lower bound is trivial. Suppose, $J_* < \infty$. Consider the optimal network simple element f^0 defined earlier in the paper, associated with a fixed node j. A fixed constant $T^0 > 0$ is one of the parameters of f^0. Consider any fixed $t > T^0$, and let us fix s, $T^0 < s < t$, such that $c = s/T^0$ is close to 1. Consider $f^* = \Gamma^{1/c} f^0$, which is the version of f^0 scaled up by factor c, and note that its cost is

$$J_s(f^*) = cJ_*,$$

and

$$\sum_{i \in \tilde{N}_j} [f_i^*(s - c\alpha_i) - f_i^*(0)] = \mu_j s, \tag{6.1}$$

where j and $\tilde{N}_j$ are parameters of the f^0. Let us fix $\varepsilon = \varepsilon(c) > 0$ small enough so that $c(1 - \varepsilon) > 1$. For $\delta > 0$, we denote by

$$U_\delta \doteq \{h \in \mathcal{D}^N \mid \|h - f^*\|_s < \delta\}$$

the (open) δ-neighborhood of f^* (in $\|\cdot\|_s$ metric). Equation (6.1) and the trivial fact that the maximum amount of work the server j can perform in the interval $[t - s, t]$ is $\mu_j s$, imply the following observation. If we choose $\delta = \delta(\varepsilon) > 0$ small enough, then, for any DSP with the realization f of the input flows such that $(\theta_{t-s} f - (\theta_{t-s} f)(0)) \in U_\delta$, there exists at least one $i \in \tilde{N}_j$ and at least one class i customer that arrived in the network in the interval $[t - s, t - c\alpha_i \times (1 - \varepsilon)]$, which is *not* served by the node j by time t. Therefore, for any f such that $(\theta_{t-s} f - (\theta_{t-s} f)(0)) \in U_\delta$, we must have, for at least one $i \in \tilde{N}_j$,

$$w_i(t) \geq c\alpha_i(1 - \varepsilon) > \alpha_i,$$

implying

$$r(t) > 1.$$

Thus, we can write

$$\begin{aligned}\liminf_{n\to\infty}\frac{1}{n}\log P_*\left(\frac{1}{n}\mathbf{r}(nt)>1\right) &\geq \liminf_{n\to\infty}\frac{1}{n}\log P\big(\theta_{t-s}\Gamma^n\mathbf{f}-(\theta_{t-s}\Gamma^n\mathbf{f})(0)\in U_\delta\big)\\ &\geq -\inf\{J_s(f)\mid f\in U_\delta\}\\ &\geq -J_s(f^*)\\ &= -cJ_*.\end{aligned}$$

(The second inequality above follows from the LDP lower bound for the sequence $\{\Gamma^n\mathbf{f}\}$, and the third one holds because, obviously, $f^*\in U_\delta$.) Since we can choose s such that $c=s/T^0$ is arbitrarily close to 1, Theorem 3.2(ii) is proved.

7. LWDF network: family of fluid sample paths.

7.1. *More definitions and basic properties.* The properties of the family of fluid sample paths we describe in this subsection (and their proofs) are analogous to those in [30].

Let us denote by Φ the set of all possible fluid sample paths and by $\Phi_d\subset\Phi$ the set of all possible discrete sample paths.

Consider a fluid sample path $\phi=(f,\widehat{f},\tau,\bar{f})\in\Phi$. The following set of functions $s(t)$ we will call the *state* of the FSP ϕ at time $t\geq 0$:

$$s(t)\doteq(\psi,\hat{\psi},\kappa,\bar{\psi}), \tag{7.1}$$

where

$$\kappa=(\kappa_1,\dots,\kappa_N)\in\mathbb{R}^N_-$$

with

$$\kappa_i=\tau_i(t)-t=-w_i(t);$$

$$\psi=\big(\psi_i=\big(\psi_i(\xi),\ \xi\in(-\infty,\infty)\big),i\in N\big)\in\mathcal{I}^N_{+,0}$$

with

$$\psi_i(\xi)=\begin{cases}0, & \xi<\kappa_i,\\ f_i(t+\xi)-f_i\big((t+\kappa_i)-\big), & \kappa_i\leq\xi\leq 0,\\ f_i(t)-f_i\big((t+\kappa_i)-\big), & \xi>0;\end{cases} \tag{7.2}$$

$$\hat{\psi}=\big((\hat{\psi}_{ik},\ k=1,\dots,K_i),i\in N\big)\in\mathbb{R}^{|G|}_+$$

with

$$\hat{\psi}_{ik}=\widehat{f}_{ik}(t)-f_i\big((t+\kappa_i)-\big);$$

$$\bar{\psi}=\big((\psi_{ik},k=1,\dots,K_i),i\in N\big)\in\mathbb{R}^{|G|}_+$$

with

$$\psi_{ik} = f_{ik}(t) - f_i\big((t+\kappa_i)-\big).$$

We introduce the *norm* of the state $s(t)$ of a FSP as follows:

$$\begin{aligned}\|s(t)\| &\doteq \sum_{i\in N}\big[\big(\psi_i(0) - \hat{\psi}_{i,K_i}\big) + (-\kappa_i)\big]\\ &\equiv \sum_{i\in N}\big[\big(f_i(t) - \widehat{f}_{i,K_i}(t)\big) + w_i(t)\big].\end{aligned} \tag{7.3}$$

[The values $\psi_i(0)$, $\hat{\psi}_{i,K_i}$, and κ_i depend also on t of course. To avoid very cumbersome notation, we do not show this dependence explicitly.]

Let Θ be the set of possible states $s(t)$ of an FSP.

The set of all possible states of a discrete sample path we denote by $\Theta_d \subseteq \Theta$. (The definition of the state of an FSP induces the definition of that of a DSP, since each DSP is also a FSP.) We note for future reference that a state $x \in \Theta_d$ of a DSP satisfies the additional condition:

$$\psi \in \hat{\mathcal{S}}^N_{+,0}, \tag{7.4}$$

where

$$\hat{\mathcal{S}}_{+,0} \doteq \{h \in \mathcal{S}_{+,0} \mid h \text{ is constant in } [0,\infty)\}. \tag{7.5}$$

We will denote by

$$\Phi(x) \doteq \{\phi \in \Phi \mid s(0) = x\}$$

the subset of FSP with initial state x.

The set Φ possesses some simple scaling (similarity), truncation and compactness properties.

LEMMA 7.1. *For any $c > 0$ and x,*

$$\textit{if } \phi \in \Phi(x), \ \textit{then } \Gamma^c\phi \in \Phi(\Gamma^c x).$$

Consequently,

$$\Gamma^c[\Phi(x)] = \Phi(\Gamma^c x) \quad \textit{and} \quad \Gamma^c\Phi = \Phi.$$

The proof follows directly from the construction of an FSP and the fact that operator A_d is scalable.

LEMMA 7.2. *Let $f^{(1)}, f^{(2)} \in \mathcal{I}^N_{+,0}$ be such that, for some $T \ge 0$, $\zeta^T f^{(1)} = \zeta^T f^{(2)}$. Let an initial state $x = (\psi, \hat{\psi}, \kappa, \bar{\psi})$ of an FSP be fixed which is consistent with $f^{(1)}$ and $f^{(2)}$, namely $\zeta^0 f^{(1)} = \zeta^0 f^{(2)} = \zeta^0\psi$. Then*

$$\zeta^T A\big(f^{(1)}, \hat{\psi}, \kappa, \bar{\psi}\big) = \zeta^T A\big(f^{(2)}, \hat{\psi}, \kappa, \bar{\psi}\big).$$

The proof follows directly from the constructions of an FSP and operator A.

LEMMA 7.3. *Consider a sequence of FSP* $\{\phi^{(n)} = (f^{(n)}, \widehat{f}^{(n)}, \tau^{(n)}, \bar{f}^{(n)}) \in \Phi, n = 1, 2, \ldots\}$, *with the initial states* $s^{(n)}(0)$. *Then the following properties hold*:

(a) *The convergence*

$$\phi^{(n)} \Rightarrow \phi, \qquad n \to \infty,$$

implies $\phi \in \Phi$. (*We remind the reader that the weak convergence* "$\Rightarrow$" *is understood componentwise*.)

(b) *Suppose that*

$$\zeta_0 f^{(n)} - f^{(n)}(0) \to h \in \mathcal{C}^N_{(0),+,0} \qquad u.o.c.$$

and, *for some constant* $c > 0$,

$$\|s^{(n)}(0)\| \leq c.$$

Then there exist $\phi = (f, \widehat{f}, \tau, \bar{f}) \in \Phi$ *and a subsequence* $\{\phi^{(n(l))}, l = 1, 2, \ldots\} \subseteq \{\phi^{(n)}\}$ *such that*

$$\phi^{(n(l))} \Rightarrow \phi, \qquad l \to \infty, \tag{7.6}$$

and, *moreover*,

$$\zeta_0 f - f(0) = h \quad \text{and} \quad \|s(0)\| \leq c, \tag{7.7}$$

where $s(0)$ *is the initial state of* ϕ.

PROOF. (a) For each $\phi^{(n)}$, let us choose a sequence $(\phi^{(n,m)}, m = 1, 2, \ldots)$ of DSP which "defines" this $\phi^{(n)}$, that is,

$$\phi^{(n,m)} \Rightarrow \phi^{(n)}, \qquad m \to \infty.$$

We choose a countable dense subset of $\mathbb{R}$ such that all components of all the functions ϕ, $\phi^{(n)}$ and $\phi^{(n,m)}$ are continuous in those points. Using Cantor's diagonal procedure we can find a sequence of pairs $(n(l), m(l)), l = 1, 2, \ldots,$ such that

$$\phi^{(n(l),m(l))} \to \phi, \qquad l \to \infty,$$

where the last convergence is pointwise in all the points of the chosen subset. The latter convergence in turn implies

$$\phi^{(n(l),m(l))} \Rightarrow \phi, \qquad l \to \infty,$$

which proves statement (a).

(b) We can always choose a subsequence $(\phi^{(n(l))})$ such that the convergence (7.6) takes place for some ϕ. According to statement (a), $\phi \in \Phi$. Additional conditions (7.7) on ϕ are verified trivially. $\square$

Consider a fixed fluid sample path $\phi = (f, \hat{f}, \tau, \bar{f}) \in \Phi$, and let $s(t)$ denote its state at time $t \geq 0$. For a fixed $a \geq 0$, consider an FSP $\phi^a = (f^a, \hat{f}^a, \tau^a, \bar{f}^a)$ constructed from ϕ as follows:

$$\begin{aligned}
f_i^a &= \theta_a f_i - f_i(\tau_i(a)-), && i \in N,\\
\hat{f}_{ik}^a &= \zeta_0 \theta_a \hat{f}_{ik} - f_i(\tau_i(a)-), && (ik) \in G,\\
\tau_{ik}^a &= \zeta_0 \theta_a \tau_{ik} - a, && (ik) \in G,\\
f_{ik}^a &= \zeta_0 \theta_a f_{ik} - f_i(\tau_i(a)-), && (ik) \in G.
\end{aligned}$$

LEMMA 7.4. *For any fluid sample path $\phi \in \Phi$ and any $a \geq 0$ the following properties hold*:

$$\phi^a \in \Phi(s(a))$$

and

$$s^a(t) = s(a+t), \qquad t \geq 0,$$

where $s^a(t)$ is the state of ϕ^a at time t.

PROOF. Similar to the way it is done in the proof of Lemma 7.3, we choose a sequence of discrete sample path $\phi^{(n)}$ defining the FSP ϕ. Considering the evolution of each $\phi^{(n)}$ starting time a, we obtain a sequence of DSP which defines ϕ^a. We omit details. $\square$

7.2. *Properties related to path costs.* The central results of this subsection are Lemmas 7.8 and 7.9, which we will need to apply Wentzell–Freidlin theory in the proof of the large deviations upper bound in Section 10. As an interesting by-product, we prove a *strong stability* property in Lemma 7.7.

LEMMA 7.5. *Let constants $C_1 \geq 0$, $C_2 \geq 0$ and $\varepsilon \geq 0$ such that*

$$\sum_{i \in N_j} (\lambda_i + \varepsilon) < \mu_j \qquad \forall\, j \in J \tag{7.8}$$

be fixed. Then there exist constants $0 \leq T_2 \leq T_3 < \infty$ (depending on the system parameters, C_1, C_2 and ε) and a constant $C_3 > 0$ (depending only on the system parameters and ε) such that the following holds.

If a fluid sample path $(f, \widehat{f}, \tau, \bar{f})$ is such that

$$\|s(0)\| \leq C_1, \tag{7.9}$$

$$\zeta_0 f_i - f_i(0) \in \mathcal{A}_{(0),+,0} \qquad \forall\, i \in N \tag{7.10}$$

and

$$\sum_{i \in N} \int_0^\infty f_i'(t) I\{f_i'(t) > \lambda_i + \varepsilon\}\, dt \leq C_2, \tag{7.11}$$

then the following properties hold:

(a) *for all* $i \in N$,

$$\widehat{f}_{i,K_i}(T_2) \geq f_i(0) \quad \textit{and} \quad \tau_i(T_2) \geq 0; \tag{7.12}$$

(b) *there exists* $t_* \in [T_2, T_3]$ *such that*

$$\|s(t_*)\| = 0 \quad \textit{and} \quad \|s(t)\| \leq C_2 C_3 \qquad \forall\, t \geq t_*. \tag{7.13}$$

The key element of the proof of Lemma 7.5 is the following lemma.

LEMMA 7.6. *Suppose the constants* $C_1 \geq 0$, $C_2 \geq 0$ *and* $\varepsilon \geq 0$ *satisfying* (7.8) *are fixed. Then there exist constants* $\delta > 0$ *and* $\Delta > 0$ *such that the following property holds*:

Consider a fluid sample path $(f, \widehat{f}, \tau, \bar{f})$ *satisfying* (7.9)–(7.12) *for some* $T_2 \geq 0$. *Then*

$$\Lambda(V_2) \leq C_2/\Delta, \tag{7.14}$$

where

$$V_2 \doteq \{t \in V_1 \mid r'(t) > -\delta\}, \tag{7.15}$$

$$V_1 \doteq \left\{ t \in [T_2, \infty) \Big\backslash \bigcup_j H_j \;\Big|\; t \textit{ is regular}, r(t) > 0 \right\}. \tag{7.16}$$

PROOF. Consider a point $t \in V_1$. The set $\tilde{N}(t)$ is nonempty, and we have $\tau_i'(t) = 1 - \alpha_i r'(t)$ for each $i \in \tilde{N}(t)$. Then for $j = \tilde{j}(t)$, using Lemma 4.2(e) and the equivalence of (5.3) and (5.2), we can write

$$r'(t) = \frac{\sum_{i \in \tilde{N}(t)} f_i'(\tau_i(t)) - \mu_j}{\sum_{i \in \tilde{N}(t)} \alpha_i f_i'(\tau_i(t))}. \tag{7.17}$$

If $f_i'(\tau_i(t)) \leq \lambda_i + \varepsilon$ for all $i \in \tilde{N}(t)$, then

$$r'(t) \leq -\delta \doteq -\min_{j \in J} \frac{\mu_j - \sum_{i \in N_j} (\lambda_i + \varepsilon)}{\sum_{i \in N_j} \alpha_i (\lambda_i + \varepsilon)} < 0.$$

We see that

$$V_2 \subseteq V_3 \doteq \{t \in V_1 \mid \exists\, i \in \tilde{N}(t), f_i'(\tau_i(t)) > \lambda_i + \varepsilon\}.$$

To estimate the Lebesgue measure of V_3, let us consider the function

$$F(t) = \sum_{i \in N} \int_0^{\tau_i(t)} f_i'(s) I\{f_i'(s) > \lambda_i + \varepsilon\}\, ds, \qquad t \geq T_2.$$

This function is nondecreasing Lipschitz continuous, $F(0) = 0$, and $\lim_{t\to\infty} F(t) \le C_2$. It is easy to see that almost everywhere (a.e.) in $[T_2, \infty)$,

$$F'(t) = R_1(t) \doteq \sum_{i\in\tilde N(t),\, f_i'(\tau_i(t))>\lambda_i+\varepsilon} f_i'(\tau_i(t))\tau_i'(t). \tag{7.18}$$

We claim that there exists a constant $\Delta > 0$ such that, in any point $t \in V_3$,

$$R_1(t) \ge \Delta. \tag{7.19}$$

To show this, let us use the following notation:

$$j = \tilde j(t), \qquad x_i = f_i'(\tau_i(t)),$$

$$\tilde N^{(1)}(t) = \{i \in \tilde N(t) \mid x_i > \lambda_i + \varepsilon\}, \qquad \tilde N^{(2)}(t) = \tilde N(t) \setminus \tilde N^{(1)}(t).$$

Then we rewrite $R_1(t)$ as

$$R_1(t) = \sum_{i\in\tilde N^{(1)}(t)} x_i\bigl(1 - \alpha_i r'(t)\bigr),$$

and write

$$R_2(t) \doteq \sum_{i\in\tilde N^{(2)}(t)} x_i\bigl(1 - \alpha_i r'(t)\bigr).$$

Consider two cases: $r'(t) \le 0$ and $r'(t) > 0$.

If $r'(t) \le 0$, we have

$$R_1(t) \ge \sum_{i\in\tilde N^{(1)}(t)} x_i \ge \min_{i\in N}(\lambda_i + \varepsilon).$$

If $r'(t) > 0$, we observe that

$$r'(t) = \frac{\sum_{i\in\tilde N^{(2)}(t)} x_i - R_2(t)}{\sum_{i\in\tilde N^{(2)}(t)} \alpha_i x_i},$$

which implies that

$$R_2(t) < \sum_{i\in\tilde N^{(2)}(t)} x_i.$$

However,

$$R_1(t) + R_2(t) = \sum_{i\in\tilde N(t)} f_i'(\tau_i(t))\tau_i'(t) = \sum_{(ik)\in\tilde G_j(t)} \widehat{f}'_{ik}(t) = \mu_j$$

and so

$$R_1(t) > \mu_j - \sum_{i\in\tilde N^{(2)}(t)} x_i \ge \mu_j - \sum_{i\in N_j}(\lambda_i + \varepsilon) \ge \min_{j\in J}\Bigl[\mu_j - \sum_{i\in N_j}(\lambda_i + \varepsilon)\Bigr] > 0.$$

We have proved claim (7.19).

Thus, from (7.18) and (7.19) we see that

$$F(t) \geq \Lambda([T_2, t] \cap V_3)\Delta,$$

where Δ depends only on the system parameters and ε. This and the fact that $\lim F(t) \leq C_2$ imply

$$\Lambda(V_3) \leq C_2/\Delta,$$

and we are done since $V_2 \subseteq V_3$. $\square$

PROOF OF LEMMA 7.5.

(a) Consider any fixed flow i and the first node $j = \hat{j}(i, 1)$ on its route. Let us choose $T_2^{(i,1)} > 0$ such that

$$C_1 + T_2^{(i,1)}\left[\sum_{i \in N_j} (\lambda_i + \varepsilon)\right] + C_2 < T_2^{(i,1)}\mu_j.$$

Then it follows from Lemma 4.1(c) that, for some $t^{(i,1)} \in [0, T_2^{(i,1)}]$,

$$\widehat{f}_{i1}(t^{(i,1)}) = f_{i1}(t^{(i,1)}) = f_i(t^{(i,1)}) \geq f_i(0)$$

and

$$\tau_{i1}(t^{(i,1)}) = t^{(i,1)} \geq 0,$$

which means that

$$\widehat{f}_{i1}(T_2^{(i,1)}) \geq f_i(0) \quad \text{and} \quad \tau_{i1}(T_2^{(i,1)}) \geq 0.$$

Now consider the second node $j = \hat{j}(i, 2)$ (if any) of the flow i route. We know that $f_{i2}(T_2^{(i,1)}) = \widehat{f}_{i1}(T_2^{(i,1)})$ and observe that

$$\|s(T_2^{(i,1)})\| \leq \|s(0)\| + T_2^{(i,1)}N + T_2^{(i,1)}\left[\sum_{i \in N_j} (\lambda_i + \varepsilon)\right] + C_2.$$

Using the argument analogous to the one we used above, we obtain the existence of $T_2^{(i,2)} \geq T_2^{(i,1)}$ such that, for some $t^{(i,2)} \in [T_2^{(i,1)}, T_2^{(i,2)}]$, we have

$$\widehat{f}_{i2}(t^{(i,2)}) = f_{i2}(t^{(i,2)}) \quad \text{and} \quad \tau_{i2}(t^{(i,2)}) = \tau_{i1}(t^{(i,2)}),$$

which implies

$$\widehat{f}_{i2}(T_2^{(i,2)}) \geq \widehat{f}_{i1}(T_2^{(i,1)}) \geq f_i(0) \quad \text{and} \quad \tau_{i2}(T_2^{(i,2)}) \geq \tau_{i1}(T_2^{(i,1)}) \geq 0.$$

Continuing the same way, we obtain by induction that there exists $T_2^{(i)} \geq 0$ such that, uniformly on all fluid sample paths satisfying conditions of the lemma,

$$\widehat{f}_{i,K_i}(T_2^{(i)}) \geq f_i(0) \quad \text{and} \quad \tau_{i,K_i}(T_2^{(i)}) = \tau_i(T_2^{(i)}) \geq 0.$$

Choosing such $T_2^{(i)}$ for each flow i, we see that the statement (a) of the lemma holds with

$$T_2 = \max_{i \in N} T_2^{(i)}.$$

(b) Let the T_2 chosen above be fixed. Let us choose constants $\delta > 0$, $\Delta > 0$ and the corresponding sets V_1 and V_2 as in Lemma 7.6. For any t_1 and t_2, $T_2 \le t_1 \le t_2 < \infty$, we can write

$$r(t_2) \le r(t_1) + \int_{t_1}^{t_2} r'(t) I\{t \text{ is regular}\}\, dt. \tag{7.20}$$

[The inequality is because the RHS of (7.20) does not include the nonpositive increment of the singular component of r.] If we take into account the fact that in any regular point t, $r(t) = 0$ implies $r'(t) = 0$, we can rewrite the RHS of (7.20) as

$$r(t_1) + \int_{[t_1,t_2]\cap V_1} r'(t)\, dt = r(t_1) + \int_{[t_1,t_2]\cap V_2} r'(t)\, dt + \int_{[t_1,t_2]\cap(V_1\setminus V_2)} r'(t)\, dt.$$

So finally we obtain

$$r(t_2) \le r(t_1) + \frac{C_2}{\Delta} d_{\max} - \Lambda\big([t_1, t_2] \cap (V_1 \setminus V_2)\big)\delta, \tag{7.21}$$

where $d_{\max} \doteq 1/(\min_i \alpha_i)$.

Notice that

$$r(T_2) \le \big(\|s(0)\| + T_2 N\big) d_{\max} \le (C_1 + T_2 N) d_{\max}.$$

Let us choose $T_3 \ge T_2$ large enough so that

$$(C_1 + T_2 N) d_{\max} + \frac{C_2}{\Delta} d_{\max} - \delta\bigg(T_3 - T_2 - \frac{C_2}{\Delta}\bigg) < 0.$$

If $r(t)$ would be positive everywhere in $[T_2, T_3]$, then we would obtain $r(T_3) < 0$, which is of course impossible. So, there exists $t_* \in [T_2, T_3]$ such that $r(t_*) = 0$ and therefore $\|s(t_*)\| = 0$. Setting $t_2 = t \ge t_*$ and $t_1 = t_*$ in (7.21), we obtain

$$r(t) \le C_2 \frac{d_{\max}}{\Delta}, \qquad t \ge t_*,$$

which implies

$$\|s(t)\| \le C_2 \frac{d_{\max}}{\Delta}\Big(\max_i \alpha_i\Big) N + C_2 \frac{d_{\max}}{\Delta}\Big(\max_i \alpha_i\Big)\bigg[\sum_{i \in N} (\lambda_i + \varepsilon)\bigg] + C_2,$$

and this completes the proof of statement (b). $\square$

As a corollary of Lemma 7.5 we obtain the following result.

LEMMA 7.7. *Let a constant* $C_1 \geq 0$ *be fixed. Then there exists a constant* $T_3 \geq 0$ *such that the following holds. If a fluid sample path satisfies the conditions*

$$\|s(0)\| \leq C_1,$$

$$\zeta_0 f_i - f_i(0) \in \mathcal{A}_{(0),+,0} \qquad \forall i \in N$$

and

$$f_i'(t) \leq \lambda_i, \qquad t \geq 0, \tag{7.22}$$

then

$$\|s(t)\| = 0, \qquad t \geq T_3. \tag{7.23}$$

The property described in Lemma 7.7 can be called *strong stability* of the fluid process. It has been shown recently by Andrews [1] that, very surprisingly, for some queueing networks (FIFO network in [1]) the statement of the lemma holds if $f_i'(t) \equiv \lambda_i, t \geq 0$, but does *not* hold if this condition is relaxed to the inequality (7.22).

LEMMA 7.8. *Let constants* $C_1 > 0$ *and* $C_5 > 0$ *be fixed. Then there exists a constant* $T_3 \geq 0$ *such that, for any fixed* $T > T_3$, *the following holds: If a fluid sample path* ϕ *is such that*

$$\|s(0)\| \leq C_1$$

for each i, $f_i(\cdot)$ *is absolutely continuous in* $[0, T]$ *and*

$$\|s(t)\| > 0, \qquad t \in [0, T],$$

then

$$J_T(f) > C_5.$$

PROOF. Let us fix $\varepsilon > 0$ such that (7.8) holds. Also, let us fix $T_1 = 0$, and

$$C_2 = C_5 \max_{i \in N} \frac{\lambda_i + \varepsilon}{L_i(\lambda_i + \varepsilon)}. \tag{7.24}$$

For the constants C_1, C_2, ε and T_1, let us choose T_2 and T_3 as in Lemma 7.5.

We will prove that T_3 chosen above does satisfy the statement of the lemma. The proof is by contradiction. Let us fix arbitrary $T \geq T_3$. Consider a fluid sample path ϕ such that the conditions of the lemma are satisfied, and assume that

$$J_T(f) \leq C_5. \tag{7.25}$$

Without loss of generality (due to Lemma 7.2), we can assume that in the interval $[T,\infty)$ all functions f_i are linear with $f_i'(\cdot)\equiv\lambda_i$, and therefore $J_{\infty,T}(f)=0$. We obtain

$$\sum_{i\in N}\int_0^\infty f_i'(t)I\{f_i'(t)>\lambda_i+\varepsilon\}\,dt=\sum_{i\in N}\int_0^T f_i'(t)I\{f_i'(t)>\lambda_i+\varepsilon\}\,dt\le C_2,$$

where the inequality is easily obtained from (7.25) and (7.24), by using the facts that each $L_i(\cdot)$ is convex and it attains its minimum 0 in the single point λ_i.

Then, by Lemma 7.5, $\|s(t_*)\|=0$ for some $t_*\in[T_2,T_3]$, which contradicts our choice of the fluid sample path ϕ. □

LEMMA 7.9. *Let a constant $T_4>T^0$ be fixed.* (*Recall that T^0 is one of the parameters of the chosen optimal network simple element f^0.*) *For any fixed $T\ge T_4$ and $\varepsilon^*\ge 0$, consider the subset $\hat{\Phi}_{T,\varepsilon^*}\subseteq\Phi$ of fluid sample paths such that the following hold*:

(i) $\|s(0)\|\le\varepsilon^*$;
(ii) $\sup_{[0,T]}r(t)\ge 1$;
(iii) *for each i, $f_i(\cdot)$ is absolutely continuous in* $[0,T]$.

Then

$$\inf_{T\ge T_4}\inf_{\phi\in\hat{\Phi}_{T,\varepsilon^*}}J_T(f)\uparrow\inf_{\phi\in\hat{\Phi}_{T_4,0}}J_{T_4}(f)=J_*\qquad \text{as } \varepsilon^*\downarrow 0. \tag{7.26}$$

PROOF. First, we notice that the equality in (7.26) follows from the definition of J_* and the fact that $T_4>T^0$. The uniform convergence in (7.26) is proved by contradiction. If (7.26) did not hold, then we would find a sequence of fluid sample paths $\{\phi^{(m)}, m=1,2,\dots\}$ such that $\phi^{(m)}\in\hat{\Phi}_{T^{(m)},1/m}$ for some $T^{(m)}\ge 0$, and $J_{T^{(m)}}(f^{(m)})\le\varepsilon_4<J_*$ for some fixed $0<\varepsilon_4<\infty$. Due to Lemma 7.4 (and the definition of J_*), we can assume without loss of generality that this sequence is such that

$$\inf_{t\in[0,T^{(m)}]}\|s^{(m)}(t)\|>0\qquad\forall m.$$

Then, using Lemma 7.8, we observe that the sequence $\{T^{(m)}\}$ is bounded. It is also easy to see that $\{T^{(m)}\}$ must stay bounded away from 0. Then we can further assume (again, without loss of generality) that each $\phi^{(m)}$ is such that $r^{(m)}(T^{(m)})=1$, and, for each i,

$$\frac{d}{dt}f_i^{(m)}(t)\equiv\lambda_i,\qquad t\ge T^{(m)}.$$

(Note that all functions $f_i^{(m)}$ chosen this way, are absolutely continuous in $[0,\infty)$.) From this sequence $\{\phi^{(m)}\}$ we can choose a subsequence such that $T^{(m)}\to T>0$, and which weakly converges to a fluid sample path ϕ. [From the fact that $J_{T^{(m)}}(f^{(m)})$ stays bounded, it follows that, for each i, the functions $f_i^{(m)}(\cdot)$ remain uniformly bounded on finite intervals.] Moreover, since for an arbitrary fixed $T_5 > T$ the cost function $J_{T_5}(\cdot)$ has compact level sets, the subsequence can be chosen such that $\zeta_0 f^{(m)} - f^{(m)}(0) \to h \in \mathcal{C}^N_{(0),+,0}$ u.o.c., and $J_T(h) \le \varepsilon_4$. This immediately implies that the limiting ϕ is such that $f = h$, $\phi \in \hat{\Phi}_{T,0}$ and $J_T(f) \le \varepsilon_4$. The latter inequality is a contradiction with the definition of J_*. $\square$

8. LWDF network: properties of a sequence of scaled processes. In this section we consider the sequence of scaled *processes* $\{\phi^{(n)}, n = 1, 2, \ldots\}$, with $\phi^{(n)} = \Gamma^n \phi$, where ϕ is the original process describing the evolution of the LWDF network in the interval $[0, \infty)$. We assume that each process $\phi^{(n)}$ has a *nonrandom* initial state $x \in \Theta_d$.

As before, suppose that Assumption 3.1 for the input flows holds. (We note again that this assumption is weaker than Assumption 3.2.)

We denote by $P_x^{(n)}$ the distribution of the process $\phi^{(n)}$ with a fixed initial state $x \in \Theta_d$; and denote by $E_x^{(n)}$ the expectation with respect to $P_x^{(n)}$.

LEMMA 8.1. *Let positive constants $0 < \delta < \varepsilon^* < C < \infty$ and $T > 0$ be fixed. Consider the following subsets*:

$$X_0 \doteq \{x \in \Theta_d \mid \|x\| \le \varepsilon^*\}, \qquad Y_0 \doteq \{\phi \in \Phi \mid s(0) \in X_0\} \equiv \{\phi \in \Phi \mid \|s(0)\| \le \varepsilon^*\}.$$

Let Y be one of the subsets Y_1, Y_2, Y_3, defined as

$$Y_1 = \left\{\phi \in \Phi \;\Big|\; \inf_{t \in [0,T]} \|s(t)\| \ge \delta\right\},$$

$$Y_2 = \left\{\phi \in \Phi \;\Big|\; \sup_{t \in [0,T]} r(t) \ge 1\right\},$$

$$Y_3 = \left\{\phi \in \Phi \;\Big|\; \sup_{t \in [0,T]} \|s(t)\| \ge C\right\}.$$

Then

$$\limsup_{n \to \infty} \frac{1}{n} \log\left[\sup_{x \in X_0} P_x^{(n)}(Y)\right] \le - \inf_{\phi \in Y_0 \cap Y} J_T(f). \tag{8.1}$$

PROOF. Let us write $Z_d \doteq \Phi_d \cap Y_0 \cap Y$. Consider the set

$$F_d \doteq \{h = \zeta_0^T f - f(0) \mid f = f(\phi),\ \phi \in Z_d\}.$$

For any fixed $x \in X_0$ we can write

$$P_x^{(n)}(Y) \le P\{\zeta_0^T \mathbf{f}^{(n)} - \mathbf{f}^{(n)}(0) \in F_d\}.$$

Using the LDP upper bound for the sequence of input flows, we obtain

$$\limsup_{n\to\infty} \frac{1}{n} \log P\{\zeta_0^T \mathbf{f}^{(n)} - \mathbf{f}^{(n)}(0) \in F_d\}$$
$$\leq - \inf_{h\in\overline{F_d}} J_T(h) = - \inf_{h\in\overline{F_d}\cap\zeta_0^T \mathcal{A}_{(0),+,0}} J_T(h) \leq - \inf_{\phi\in Y_0\cap Y} J_T(f),$$

where $\overline{F_d}$ is the closure of F_d (in the uniform convergence topology). The last inequality follows from the inclusion

$$\overline{F_d} \cap \zeta_0^T \mathcal{A}_{(0),+,0} \subseteq \{\zeta_0^T f \mid f = f(\phi), \phi \in Y_0 \cap Y\},$$

which is easily proved using Lemma 7.3(b). $\square$

9. LWDF network with Markov input flows: process definition and stability. Now suppose that the Markov Assumption 3.2 holds for the input flow processes.

The system evolution is described by a random process with realizations being discrete sample paths. We consider the *extended state* of the network at time t, $\bar{s}(t) = (s(t), \sigma(t))$, where $s(t)$ is the state of the DSP, and

$$\sigma(t) = (\sigma_1(t), \dots, \sigma_N(t)),$$

where $\sigma_i(t)$ is the index of the state of the modulating Markov process for input flow i. We see that the *process* $\bar{\mathbf{s}} = (\bar{\mathbf{s}}(t), t \geq 0)$ takes values in the space

$$\Xi_d \subseteq \hat{\mathcal{S}}_{+,0}^N \times \mathbb{R}_+^{|G|} \times \mathbb{R}_+^N \times \mathbb{R}_+^{|G|} \times \Sigma, \tag{9.1}$$

where Σ is the finite set of possible values of $\sigma(\cdot)$ [and $\hat{\mathcal{S}}_{+,0}$ is defined in (7.5)]. To be more precise, Ξ_d is the set of possible states $\bar{s}$ which are reachable from any zero state [such that $s(t) = 0$] by a sample path f of the input flows, with component functions f_i having arbitrary positive jump sizes (i.e., customer sizes).

By definition, the norm of the extended state $\|\bar{s}(t)\| \doteq \|s(t)\|$.

On the space Ξ_d we define a metric which is the sum of the metrics on the component spaces, with the standard Euclidian distance metric on each $\mathbb{R}_+$ and Σ, and the following metric $\hat{m}_s$ on each $\hat{\mathcal{S}}_{+,0}$:

$$\hat{m}_s(h_1, h_2) \doteq m_s(\tilde{h}_1, \tilde{h}_2) + |\Delta(h_1) - \Delta(h_2)|, \tag{9.2}$$

where $\tilde{h}_i = (\tilde{h}_i(\xi), \xi \in [-1, 1])$, $i = 1, 2$, is defined by

$$\tilde{h}_i(\xi) = \begin{cases} 0, & \xi = -1, \\ h_i(\log(\xi + 1)), & -1 < \xi \leq 0, \\ h_i(0), & 0 < \xi \leq 1, \end{cases} \tag{9.3}$$

m_s is the standard Skorohod J_1-metric [5, 29] on the space $D([-1, 1])$ of RCLL functions in $[-1, 1]$; and $\Delta(h_i)$ is the number of jumps of function h_i.

It is not hard to verify that space Ξ_d equipped with such a metric is a locally compact separable metric space. The Borel σ-algebra on Ξ_d is equal to the σ-algebra induced by the product σ-algebra generated by cylinder subsets on the component spaces in the RHS of (9.1). (See [5], proof of Theorem 14.5.)

REMARK. The topology on the state space Ξ_d we introduced above is quite strong and may seem unusual. In fact, it is a form of *the* most natural (and conventional) topology on a space of states of a multiclass queueing network. Two states are "close" if, in both of them, there is the same number of customers of each type in each node, and their corresponding arrival and residual service times are close. In addition, this topology puts our process into the framework of [21, 22], which implies useful connections between stochastic and topological notions of stability (see [21], Section 3).

We note, however, that all the results in this paper would hold even if we were to introduce a weaker topology on Ξ_d. (In particular, we do not need Ξ_d to be separable and locally compact.) So, for example, all our results still hold if we exclude $|\Delta(h_1) - \Delta(h_2)|$ from the RHS of (9.2).

Consider the sequence of scaled processes

$$\big(\bar{\mathbf{s}}^{(n)} \doteq \Gamma^n \bar{\mathbf{s}}, n = 1, 2, \dots\big).$$

Each process $\bar{\mathbf{s}}^{(n)}$ is a strong Markov process with state space Ξ_d, and with right-continuous sample paths.

In what follows, with a slight abuse of notation, we will write $P_x(\cdot) \doteq P\{\cdot|\bar{\mathbf{s}}^{(n)}(0) = x\}$, and E_x will denote the corresponding expectation.

THEOREM 9.1. *Markov process $\bar{\mathbf{s}}$ is positive Harris recurrent* (*see the definition in* [22]), *and therefore has unique stationary distribution*.

PROOF. Assumption 3.2 implies that the functional strong law of large numbers holds for each input flow i. Namely, with probability 1,

$$\zeta_0 \mathbf{f}_i^{(n)} - \mathbf{f}_i^{(n)}(0) \to (\lambda_i t,\ t \ge 0).$$

Also, for any fixed $T_* \ge 0$, the family of random variables $\{\mathbf{f}_i^{(n)}(T_*) - \mathbf{f}_i^{(n)}(0), n = 1, 2, \dots\}$ is uniformly integrable. Finally, it is easy to see that, for any $c > 0$, $\{\|\bar{s}\| \le c\} \subset \Xi_d$ is a closed *petite* subset. (See the definition in [22]. This follows from the simple structure of our process: if $\|\mathbf{s}(0)\| \le c$, then there exists fixed time t_*, depending on c, such that $\mathbf{s}(t_*) = 0$ and $\sigma(t_*)$ takes some fixed value in Σ with a positive probability uniformly bounded away from 0.)

Using the facts listed in the paragraph above, along with Lemmas 7.3 and 7.7 [with condition (7.22) specialized to $f_i'(t) = \lambda_i,\ t \ge 0$], we can apply the *fluid limit* technique (see [7–9, 27, 30]), to prove that there exist constants $T_* > 0$ and $\varepsilon_* > 0$ such that

$$\limsup_{n\to\infty} \sup_{x \in \Xi_d, \|x\| \le 1} E_x \big\|\mathbf{s}^{(n)}(T_*)\big\| \le 1 - \varepsilon_*. \tag{9.4}$$

This, implies (see [8, 27]) that the Markov process $\bar{\mathbf{s}}$ is positive Harris recurrent. $\square$

In addition, from the property (9.4), using the Dynkin inequality (see its statement in [20] and its application in the proof of Theorem 2.1(ii) [23]), we obtain the following lemma.

LEMMA 9.2. *Consider the sequence of scaled Markov processes* ($\bar{\mathbf{s}}^{(n)}$, $n = 1, 2, \ldots$). *Let constants* $0 < \delta < C^*$ *be fixed. Consider the stopping time*

$$\beta_1^{(n)} = \inf\{t \geq 0 \mid \|\bar{\mathbf{s}}^{(n)}\| \leq \delta\}.$$

Then

$$\limsup_{n\to\infty} \sup_{x\in\Xi_d, \|x\|\leq C^*} E_x \beta_1^{(n)} \leq \Delta_* C^* < \infty,$$

where $\Delta_* = T_*/\varepsilon_*$, *and* T_* *and* ε_* *are those from* (9.4).

10. LWDF network with Markov input flows: proof of the large deviations upper bound. To prove the upper bound, in this section we use classical Wentzell–Freidlin constructions [13], which allow one to establish the large deviations principle for a sequence of stationary distributions via the LDP for the sequence of processes on a finite time interval. The key element is the representation of the stationary distribution of the process via the stationary distribution of a discrete time sampled chain associated with a sequence of stopping times (see Lemma 10.1).

Suppose, $J_* < \infty$. (In fact this is always the case under Assumption 3.2, but we keep the proof slightly more general due to reasons discussed in Section 3.3.)

Let us fix arbitrary $\delta_* > 0$.

Let us fix arbitrary $T_4 > T^0$. According to Lemma 7.9, we can choose $\varepsilon^* > 0$ small enough so that, for any $T \geq T_4$,

$$\inf_{\phi\in\hat{\Phi}_{T,\varepsilon^*}} J_T(f) > J_* - \delta_*. \tag{10.1}$$

(See the definition of $\hat{\Phi}_{T,\varepsilon^*}$ in Lemma 7.9.)

Let us choose ε and δ such that $0 < \delta < \varepsilon < \varepsilon^*$.

Due to Lemma 7.8, there exists $T_3 > 0$ sufficiently large so that, for any $T \geq T_3$,

$$\inf_{\phi\in\Phi, \|s(0)\|\leq\varepsilon^*, \inf_{[0,T]}\|s(t)\|\geq\delta} J_T(f) > J_* + \delta_*. \tag{10.2}$$

Let us fix arbitrary $T \geq T_4 \vee T_3$ and choose $C > \varepsilon^*$ large enough so that

$$\inf_{\phi\in\Phi, \|s(0)\|\leq\varepsilon^*, \sup_{[0,T]}\|s(t)\|\geq C} J_T(f) > J_* + \delta_*. \tag{10.3}$$

We can always do that because

$$\|s(T)\| \leq \|s(0)\| + NT + \sum_i [f_i(T) - f_i(0)].$$

Finally, let us fix arbitrary $C^* > C$.

Consider the sequence of scaled Markov processes ($\bar{\mathbf{s}}^{(n)}$, $n = 1, 2, \dots$). Each process $\bar{\mathbf{s}}^{(n)}$ is positive Harris recurrent since it is just a scaled version of $\bar{\mathbf{s}}$.

We put $\eta_0^{(n)} = 0$ and define the sequence of stopping times

$$0 = \eta_0^{(n)} \le \beta_1^{(n)} \le \eta_1^{(n)} \le \beta_2^{(n)} \le \cdots$$

as follows:

$$\beta_m^{(n)} = \inf\{t \ge \eta_{m-1}^{(n)} \mid \|\mathbf{s}^{(n)}(t)\| \le \delta\}, \qquad m \ge 1,$$
$$\eta_m^{(n)} = \inf\{t \ge \beta_m^{(n)} \mid \|\mathbf{s}^{(n)}(t)\| \ge \varepsilon\}, \qquad m \ge 1.$$

We observe that since customer sizes (in the original unscaled process $\bar{\mathbf{s}}$) are uniformly bounded and (with probability 1) no two customers arrive at the same time, for all $n \ge n_0$ (with some fixed n_0) and all $m \ge 1$,

$$\eta_m^{(n)} < \infty \quad \text{and} \quad \|\mathbf{s}^{(n)}(\eta_m^{(n)})\| \in [\varepsilon, \varepsilon^*] \qquad \text{w.p.1.}$$

Let us denote by $\pi^{(n)}$ the stationary distribution of the process $\bar{\mathbf{s}}^{(n)}$.

Note that, for all large n, the process $\bar{\mathbf{s}}^{(n)}$ sampled at the stopping times $\eta_m^{(n)}$, $m = 1, 2, \dots$, that is, the discrete time process

$$\hat{\mathbf{s}}^{(n)} = \{\hat{\mathbf{s}}_m^{(n)} \doteq \bar{\mathbf{s}}^{(n)}(\eta_m^{(n)}), m = 1, 2, \dots\},$$

is a Markov chain with state space

$$\hat{\Xi}_d = \{x \in \Xi_d \mid \|x\| \in [\varepsilon, \varepsilon^*]\}.$$

It is easy to verify that this Markov chain is positive Harris recurrent. Indeed, its entire state space is a *small* set. (See definitions in [20, 24]. The proof uses a simple observation similar to the one used in the proof of Theorem 9.1 to show that the set $\{\|\bar{s}\| \le c\}$ is petite.) Using Nummelin splitting (see [24], Section 4.4), this allows one to view $\hat{\mathbf{s}}^{(n)}$ as a regenerative process.

Let us denote by $\hat{\pi}^{(n)}$ the unique stationary distribution of $\hat{\mathbf{s}}^{(n)}$. The following lemma, describing the relation between the stationary distribution $\pi^{(n)}$ of the process and the stationary distribution $\hat{\pi}^{(n)}$ of the sampled chain, is quite standard. (See, e.g., the discussion and references in [21], page 510.) Since we could not find a reference which would apply to our specific setting, we present a proof in the Appendix .

LEMMA 10.1. *For any measurable $B \subseteq \Xi_d$,*

$$\pi^{(n)}(B) = \frac{\int_{\hat{\Xi}_d} \hat{\pi}^{(n)}(dx) E_x \int_0^{\eta_1^{(n)}} I\{\bar{\mathbf{s}}^{(n)}(t) \in B\}\, dt}{\int_{\hat{\Xi}_d} \hat{\pi}^{(n)}(dx) E_x \eta_1^{(n)}}. \tag{10.4}$$

We will consider (10.4) with the fixed subset $B = \{x \in \Xi_d \mid r = r(x) > 1\}$ and evaluate asymptotics of both the denominator and the numerator as $n \to \infty$.

First, it is very easy to show that there exists $\varepsilon_9 > 0$ such that

$$\liminf_{n\to\infty} \inf_{x\in\hat{\Xi}_d} E_x[\eta_1^{(n)} - \beta_1^{(n)}] > \varepsilon_9$$

and therefore the lim inf of the denominator of (10.4) is lower bounded by ε_9.

Consider the following additional stopping times:

$$\begin{aligned}
\eta^{(n),r} &= \inf\{t \geq 0 \mid \mathbf{r}^{(n)}(t) \geq 1\},\\
\eta^{(n),C} &= \inf\{t \geq 0 \mid \|\mathbf{s}^{(n)}(t)\| \geq C\},\\
\eta^{(n),r,C} &= \eta^{(n),r} \wedge \eta^{(n),C}.
\end{aligned}$$

For any fixed $x \in \hat{\Xi}_d$, we can write

$$\begin{aligned}
&E_x \int_0^{\eta_1^{(n)}} I\{\mathbf{r}^{(n)}(t) > 1\}\, dt\\
&\quad\leq E_x[I\{\eta^{(n),r} \leq \beta_1^{(n)}\}(\beta_1^{(n)} - \eta^{(n),r})]\\
&\quad\leq E_x[I\{\eta^{(n),r,C} \leq \beta_1^{(n)}\}(\beta_1^{(n)} - \eta^{(n),r,C})]\\
&\quad\leq P_x\{\eta^{(n),r,C} \leq \beta_1^{(n)}\} \sup_{\|y\|\leq C^*} E_y \beta_1^{(n)}.
\end{aligned}$$

We know from Lemma 9.2 that

$$\limsup_{n\to\infty} \sup_{\|y\|\leq C^*} E_y \beta_1^{(n)} \leq \Delta_* C^*$$

(with Δ_* defined in Lemma 9.2).

Finally, we have the estimate

$$\begin{aligned}
&P_x(\eta^{(n),r,C} \leq \beta_1^{(n)})\\
&\quad\leq P_x(\beta_1^{(n)} \geq T) + P_x(\eta^{(n),r,C} \leq T)\\
&\quad\leq P_x(\beta_1^{(n)} \geq T) + P_x(\eta^{(n),C} \leq T) + P_x(\eta^{(n),r} \leq T).
\end{aligned}$$

We have [due to Lemma 8.1 and our choice of ε^*, T and C, such that (10.1)–(10.3) hold]

$$\limsup_{n\to\infty} \frac{1}{n} \log\left[\sup_{x\in\hat{\Xi}_d} \Big(P_x(\beta_1^{(n)} \geq T) + P_x(\eta^{(n),C} \leq T)\Big)\right] \leq -(J_* + \delta_*)$$

and

$$\limsup_{n\to\infty} \frac{1}{n} \log\left[\sup_{x\in\hat{\Xi}_d} P_x(\eta^{(n),r} \leq T)\right] \leq -(J_* - \delta_*).$$

Since $\delta^* > 0$ can be chosen arbitrarily small, we have proved the desired upper bound

$$\limsup_{n\to\infty} \frac{1}{n} \log \pi^{(n)}(\{r > 1\}) \le -J_*.$$

If $J_* = \infty$, the same proof holds. We only need to replace the right-hand sides of the inequalities (10.1)–(10.3), by some fixed (arbitrarily large) constant $C_6 > 0$. □

11. The LWWF (GLQF) discipline optimality for the unfinished work process. In paper [31], for a single-server system, a result analogous to optimality of the LWDF is presented, which applies to the unfinished work process. Namely, Theorem 7.2 of [31] states that the largest weighted (unfinished) work first (LWWF) discipline, which is essentially the GLQF, is optimal. That result also extends to a network setting. We state the generalized result in this section. We do not present the proof because it is just a simplified version of the proof of Theorem 3.2.

Let $q_i(t)$ denote the total amount of unfinished class i work in the network at time t. [In the notation introduced earlier in the paper, $q_i(t) = f_i(t) - \widehat{f}_{i,K_i}(t)$.] Let us also denote by $q_{ik}(t)$ the amount of unfinished (ik)-work in node $\hat{j}(i,k)$ at time t [$q_{ik}(t) = f_{ik}(t) - \widehat{f}_{ik}(t)$]. As before, let the set of positive constants $\alpha_i > 0$, $i \in N$, be fixed, and let us denote by $\rho(t) \doteq \max_i q_i(t)/\alpha_i$ the maximal weighted unfinished work at time t.

DEFINITION 11.1 (The network LWWF discipline). The network LWWF discipline is a nonpreemptive, work-conserving discipline that chooses for service in any node j the head-of-the-line customer of type $(ik) \in G_j$ for which q_i/α_i is maximal (among those with $q_{ik} > 0$). If multiple $(ik) \in G_j$ belong to the same class i, then the type with the greatest k (furthest on the route) is chosen.

THEOREM 11.2. *For the constant $J_*^q > 0$, defined in statement* (iii) *below, the following hold*:

(i) *Consider the network with the LWWF scheduling discipline, and suppose that Assumption* 3.2 *holds. Then a stationary random process describing network evolution exists, is unique in distribution and is such that the* (*stationary*) *distribution of the maximal weighted unfinished work $\rho(0)$ satisfies the condition*

$$\limsup_{n\to\infty} \frac{1}{n} \log P\left(\frac{1}{n}\rho(0) > 1\right) = -J_*^q. \tag{11.1}$$

(ii) *Suppose that Assumption* 3.1 *holds. Then there exists $T^0 \in (0,\infty)$ such that, for any queueing discipline $G \in \overline{\mathcal{G}}$ and any $t > T^0$, we have the following lower bound*:

$$\liminf_{n\to\infty} \frac{1}{n} \log P_*\left(\frac{1}{n}\rho(nt) > 1\right) \ge -J_*^q. \tag{11.2}$$

(iii) *The constant* J_*^q *is defined as*

$$J_*^q = \min_{j \in J} J_{*,j}^q,$$

where $J_{*,j}^q$ *is determined by the following finite dimensional optimization problem for the node* j *in isolation*:

$$J_{*,j}^q = \min_{\tilde{N}_j \subseteq N_j;\ (x_i,\ i \in \tilde{N}_j)} \frac{1}{\gamma} \sum_{i \in \tilde{N}_j} L_i(x_i), \tag{11.3}$$

subject to

$$x_i > 0, \qquad i \in \tilde{N}_j,$$

$$0 < \gamma = \frac{\sum_{i \in \tilde{N}_j} x_i - \mu_j}{\sum_{i \in \tilde{N}_j} \alpha_i} < \min_{i \in \tilde{N}_j} \frac{\lambda_i}{\alpha_i}$$

and

$$\lambda_i/\alpha_i \le \gamma \qquad \forall\, i \in N_j \setminus \tilde{N}_j.$$

The form of an optimal network simple element f^0 for this problem, which is associated with a node j on which the minimum of $J_{*,j}^q$ is attained and the corresponding optimal parameters $\tilde{N}_j, x_i,\ i \in \tilde{N}_j$, is as follows:

$$f_i^0(t) \doteq \begin{cases} \lambda_i t, & t \in [0, \infty),\ \text{for } i \in N \setminus \tilde{N}_j, \\ x_i t, & t \in [0, T^0],\ \text{for } i \in \tilde{N}_j, \\ x_i T^0 + \lambda_i (t - T^0), & t \in [T^0, \infty),\ \text{for } i \in \tilde{N}_j, \end{cases} \tag{11.4}$$

where $T^0 \doteq 1/\gamma$.

REMARK. Theorem 11.2 also holds if the LWWF discipline is modified as follows: *A type* $(ik) \in G_j$ is picked for which $\tilde{q}_{ik}/\alpha_i$ is maximal, where

$$\tilde{q}_{ik}(t) \doteq \sum_{m \le k} q_{im}(t),$$

that is, $\tilde{q}_{ik}(t)$ *is the amount of work of class* i *which is* "*upstream*" *from the type* (ik) *on the class* i *route*. The LWWF discipline modified this way is more "decentralized" and may be more attractive in applications. Indeed, the information on the amount of class i unfinished work can be "shipped" downstream on the route with class i customers: there is no need for a feedback from the destination node.

12. Conclusions. Using large deviations techniques, we have proved that, in a queueing network of arbitrary topology, the LWDF is an optimal discipline to satisfy asymptotic ("tail") constraints on the end-to-end delay distributions. In the center of our analysis is the proof of a remarkable critical node property: there exists a most likely path to build large maximal weighted delay in the network, which is a most likely path to do so in one of the nodes in isolation. In other words, with the LWDF discipline, in spite of the complex interaction between nodes, the worst (weighted) delay distribution tail (among all the flows) is just the same as it would be if we would replace the network by a collection of its nodes operating in isolation.

We believe these results provide important insight into the following practical questions. How large end-to-end delays are built in a network? What are desirable features of a network scheduling algorithm trying to minimize end-to-end delays? We also think that in many practical situations the LWDF discipline can be directly used for network scheduling.

As far as technique is concerned, we think that our approach to the crucial problem of finding the lowest cost fluid paths in a network, based on an "infinitesimal interval" argument, may be useful for other models too.

It was already mentioned that some of our assumptions (e.g., the specifics of the Markov assumption on the input flows) can be significantly generalized. The proofs will still work with appropriate adjustments. Also, as we explained before, in the special case of a feedforward network, the additional Markov assumption can be dropped.

Relaxing the Markov assumption on the input flows to just a stationarity condition in the case of general network topology is an interesting subject for future research.

APPENDIX

PROOF OF LEMMA 10.1. The strong law of large numbers (SLLN) applies to the process $\bar{\mathbf{s}}$, for example, because it is a simple regenerative process, with regeneration times being the times of hitting one of the zero states [such that $s(t) = 0$]. So, for any measurable B, with probability 1,

$$\lim_{t\to\infty} \frac{1}{t} \int_0^t I\{\bar{\mathbf{s}}^{(n)}(\xi) \in B\}\, d\xi = \pi^{(n)}(B). \tag{A.1}$$

Consider the discrete time Markov chain $\tilde{\mathbf{s}}^{(n)} = \{\tilde{\mathbf{s}}_m^{(n)},\ m = 1, 2, \dots\}$, where $\tilde{\mathbf{s}}_m^{(n)} = (\bar{\mathbf{s}}^{(n)}(t),\ t \in [\eta_m^{(n)}, \eta_{m+1}^{(n)}])$. [To define state space of this chain rigorously we can, for example, include the duration $t_m = \eta_{m+1}^{(n)} - \eta_m^{(n)}$ into the state, use the time change $\xi = t - \eta_m^{(n)}$, let $\bar{\mathbf{s}}^{(n)}(\xi) \equiv \bar{\mathbf{s}}^{(n)}(t_m)$ for $\xi > t_m$ and define a σ-algebra as the one generated by cylinder sets. We hope that the above "loose," but intuitive, definition does not cause confusion.] This Markov chain is also positive Harris recurrent—again its entire state space is a *small* set [20, 24]. Therefore, it has a unique stationary distribution, which we denote by $\tilde{\pi}^{(n)}$. Obviously, distribution $\hat{\pi}^{(n)}$ is a projection of $\tilde{\pi}^{(n)}$.

The left-hand side of (A.1) can be rewritten as

$$\lim_{m\to\infty} \frac{\sum_1^m g(\tilde{\mathbf{s}}_l^{(n)})}{\sum_1^m \eta_{l+1}^{(n)} - \eta_l^{(n)}},$$

where $g(\cdot)$ is the function of the state defined as

$$g(\tilde{\mathbf{s}}_l^{(n)}) = \int_{\eta_l^{(n)}}^{\eta_{l+1}^{(n)}} I\{\bar{\mathbf{s}}^{(n)}(\xi) \in B\}\, d\xi.$$

The SLLN applies to the Markov chain $\tilde{\mathbf{s}}^{(n)}$, because using Nummelin splitting (see [24], Section 4.4) it can be represented as a simple regenerative process. Thus, with probability 1, as $m \to \infty$,

$$\frac{1}{m}\sum_1^m g(\tilde{\mathbf{s}}_l^{(n)})$$

and

$$\frac{1}{m}\sum_1^m (\eta_{l+1}^{(n)} - \eta_l^{(n)})$$

converge to the numerator and denominator of (10.4), respectively. The proof is complete. $\square$

Aknowledgments. I am grateful to many of my colleagues at Bell Labs for creating a wonderful research environment. Special thanks are due to Matthew Andrews for very useful discussions of his recent work [1–3].

REFERENCES

[1] ANDREWS, M. (2000). Instability of FIFO in Session-Oriented Networks. In *Symposium on Discrete Algorithms (SODA)*. ACM–SIAM, New York.

[2] ANDREWS, M. (2000). Probabilistic end-to-end delay bounds for earliest deadline first scheduling. In *IEEE INFOCOM 2000*. IEEE, New York.

[3] ANDREWS, M., AWERBUCH, B., FERNANDEZ, A., KLEINBERG, J., LEIGHTON, T. and LIU, Z. (1996). Universal stability results for greedy contention-resolution protocols. In *Proceedings of the 37th Annual Symposium on Foundations of Computer Science* 380–389. IEEE, New York.

[4] BERTSIMAS, D., PASCHALIDIS, I. C. and TSITSIKLIS, J. N. (1998). Asymptotic buffer overflow probabilities in multiclass multiplexers: An optimal control approach. *IEEE Trans. Automat. Contr.* **43** 315–335.

[5] BILLINGSLEY, P. (1968). *Convergence of Probability Measures*. Wiley, New York.

[6] BRAMSON, M. (2001). Stability of earliest-due-date, first-served queueing networks. *Queueing Syst. Theory Appl.* **39** 79–102.

[7] CHEN, H. (1995). Fluid approximations and stability of multiclass queueing networks: Work-conserving disciplines. *Ann. Appl. Probab.* **5** 637–665.

[8] DAI, J. G. (1995). On the positive Harris recurrence for open multiclass queueing networks: A unified approach via fluid limit models. *Ann. Appl. Probab.* **5** 49–77.

[9] DAI, J. G. and MEYN, S. P. (1995). Stability and Convergence of moments for multiclass queueing networks via fluid limit models. *IEEE Trans. Automat. Contr.* **40** 1889–1904.

[10] DEMBO, A. and ZEITOUNI, O. (1998). *Large Deviations Techniques and Applications*, 2nd ed. Springer, New York.

[11] DOYTCHINOV, B., LEHOCZKY, J. and SHREVE, S. (2001). Real-time queues in heavy traffic with earliest-deadline-first queue discipline. *Ann. Appl. Probab.* **11** 332–378.

[12] ELWALID, A. and MITRA, D. (1999). Design of generalized processor sharing schedulers which statistically multiplex heterogeneous QoS classes. In *IEEE INFOCOM'99* **3** 1220–1230. IEEE, New York.

[13] FREIDLIN, M. I. and WENTZELL, A. D. (1998). *Random Perturbations of Dynamical Systems*, 2nd ed. Springer, New York.

[14] GEORGIADIS, L., GUERIN, R. and PAREKH, A. (1997). Optimal multiplexing on a single link: Delay and buffer requirements. *IEEE Trans. Inform. Theory* **43** 1518–1535.

[15] KUMARAN K., MARGRAVE G., MITRA, D. and STANLEY, K. (2000). Novel techniques for the design and control of generalized processor sharing schedulers for multiple QoS classes. In *IEEE INFOCOM 2000*. IEEE, New York.

[16] LIEBEHERR, J., WREGE, D. and FERRARI, D. (1996). Exact admission control for networks with a bounded delay service. *IEEE/ACM Trans. Networking* **4** 885–901.

[17] LOYNES, R. M. (1962). The stability of a queue with non-independent inter-arrival and service times. *Proc. Cambridge Philos. Soc.* **58** 497–520.

[18] MAGLARAS, C. and VAN MIEGHEM, J. A. (2001). Queueing systems with leadtime constraints: A fluid-model approach for admission and sequencing control. Preprint, Northwestern Univ.

[19] MAJEWSKI, K. (1996). Large deviations of feedforward queueing networks. Ph.D. dissertation, Univ. Munich.

[20] MEYN, S. P. and TWEEDIE, R. L. (1993). *Markov Chains and Stochastic Stability*. Springer, New York.

[21] MEYN, S. P. and TWEEDIE, R. L. (1993). Stability of Markovian processes. II. Continuous time processes and sampled chains. *Adv. in Appl. Probab.* **25** 487–517.

[22] MEYN, S. P. and TWEEDIE, R. L. (1993). Stability of Markovian processes. III. Foster–Lyapunov criteria for continuous time processes. *Adv. in Appl. Probab.* **25** 518–548.

[23] MEYN, S. P. and TWEEDIE, R. L. (1994). State dependent criteria for convergence of Markov chains. *Ann. Appl. Probab.* **4** 149–168.

[24] NUMMELIN, E. (1984). *General Irreducible Markov Chains and Non-Negative Operators*. Cambridge Univ. Press.

[25] PLAMBECK, E., KUMAR, S. and HARRISON, J. M. (2001). A multiclass queue in heavy traffic with throughput time constraints: Asymptotically optimal dynamic controls. *Queueing Syst. Theory Appl.* **39** 23–54.

[26] RAMANAN, K. and DUPUIS, P. (1998). Large deviation properties of data streams that share a buffer. *Ann. Appl. Probab.* **8** 1070–1129.

[27] RYBKO, A. N. and STOLYAR, A. L. (1992). Ergodicity of stochastic processes describing the operation of open queueing networks. *Probl. Inf. Transm.* **28** 199–220.

[28] SHWARTZ, A. and WEISS, A. (1995). *Large Deviations for Performance Analysis*: *Queues*, *Communication and Computing*. Chapman and Hall, New York.

[29] SKOROHOD, A. V. (1956). Limit theorems for stochastic processes. *Theory Probab. Appl.* **1** 261–290.

[30] STOLYAR, A. L. (1995). On the stability of multiclass queueing networks: A relaxed sufficient condition via limiting fluid processes. *Markov Process. Related Fields* **1** 491–512.

[31] STOLYAR, A. L. and RAMANAN, K. (2001). Largest weighted delay first scheduling: Large deviations and optimality. *Ann. Appl. Probab.* **11** 1–48.

[32] VAN MIEGHEM, J. A. (2003). Due date scheduling: Asymptotic optimality of generalized longest queue and generalized largest delay rules. *Oper. Res.* **51** 113–122.

LUCENT TECHNOLOGIES
BELL LABORATORIES
600 MOUNTAIN AVENUE, 2C-322
MURRAY HILL, NEW JERSEY 07974
E-MAIL: stolyar@research.bell-labs.com