The Annals of Applied Probability
2001, Vol. 11, No. 1, 1–48

LARGEST WEIGHTED DELAY FIRST SCHEDULING: LARGE DEVIATIONS AND OPTIMALITY

BY ALEXANDER L. STOLYAR AND KAVITA RAMANAN

Bell Labs

We consider a single server system with N input flows. We assume that each flow has stationary increments and satisfies a sample path large deviation principle, and that the system is stable. We introduce the largest weighted delay first (LWDF) queueing discipline associated with any given weight vector $\alpha = (\alpha_1, \ldots, \alpha_N)$. We show that under the LWDF discipline the sequence of scaled stationary distributions of the delay $\hat{w}_i$ of each flow satisfies a large deviation principle with the rate function given by a finite-dimensional optimization problem. We also prove that the LWDF discipline is optimal in the sense that it maximizes the quantity

$$\min_{i=1,\ldots,N}\left[\alpha_i \lim_{n\to\infty} \frac{-1}{n} \log P(\hat{w}_i > n)\right],$$

within a large class of work conserving disciplines.

1. Introduction.

1.1. *Quality of service requirements and the LWDF discipline.* A very important and challenging problem in the design of high-speed communication networks is that of providing quality of service (QoS) guarantees, usually specified in terms of loss probabilities or delays of packets in the network. The control of delays is often of crucial importance, especially for real-time applications like video. For example, the QoS requirements for each traffic flow (customer class) can be specified in terms of a deadline T_i and an allowed violation probability δ_i. More precisely, if there are N classes of users and $\hat{w}_i$ is the stationary customer delay of the ith class, then one would like to determine if there exists a policy that would meet the given QoS constraints

$$P(\hat{w}_i > T_i) \le \delta_i \quad \text{for } i = 1, \ldots, N. \tag{1.1}$$

In this paper we restrict ourselves to a single node and address a related asymptotic question. Suppose a set of positive constants (weights) $\alpha_1, \alpha_2, \ldots, \alpha_N$ is fixed. For a nonnegative random variable X, let us use the notation

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

Received October 1999; revised February 2000.

AMS 2000 *subject classifications*. 60F10, 90B12, 60K25.

Key words and phrases. Queueing theory, queueing delay, large deviations, rate function, optimality, fluid limit, control, scheduling, quality of service (QoS), earliest deadline first (EDF), LWDF.

assuming the limit exists. We ask the following question. Is there a discipline G such that the *asymptotic* QoS requirements

$$\beta(\hat{w}_i^G) \geq \alpha_i^{-1} \quad \text{for } i = 1, \dots, N \tag{1.2}$$

are satisfied, where $\hat{w}_i^G$ is the stationary delay for class i with discipline G? [Observe that when T_i is large and δ_i is small, the problem (1.2) can be related to (1.1) by setting $\alpha_i = -T_i/\log \delta_i$ for all i.] We address the problem (1.2) by considering the following equivalent *optimization problem:*

$$\max_G \min_{i=1,\dots,N} \alpha_i \beta(\hat{w}_i^G). \tag{1.3}$$

The problems are equivalent in the sense that (1.2) has a solution if and only if the maximum in (1.3) is greater than or equal to 1.

We introduce the largest weighted delay first (LWDF) queueing (scheduling) discipline, which is parametrized by a weight vector $(\alpha_1, \dots, \alpha_N)$. Roughly speaking, this discipline always chooses for service the longest waiting customer from the queue i for which the current *weighted delay* $\hat{w}_i(t)/\alpha_i$ is maximal. Under the assumption that the input flows have stationary increments and satisfy a sample path large deviation principle, and the system is stable, we prove that *LWDF is an optimal solution to the problem* (1.3). This implies that if there is any discipline G which solves problem (1.2), then LWDF does so.

Note that if all weights α_i are equal, LWDF reduces to the simple FIFO discipline. Also note that *LWDF is invariant with respect to the stochastic structure of the input flows*, in the sense that the algorithm itself relies *only* on the weights α_i which characterize the QoS requirements of the flows, and not on the structure of the input flows.

To provide some intuition behind our result, note that if for a given discipline G all $\beta(\hat{w}_i^G)$ are well defined, then

$$\min_{i=1,\dots,N} \alpha_i \beta(\hat{w}_i^G) = \beta\left(\max_{i=1,\dots,N} \frac{\hat{w}_i^G}{\alpha_i}\right). \tag{1.4}$$

Thus the problem (1.3) is reduced to that of maximizing (over all disciplines G) the rate of decay of the tail of the stationary distribution of the *maximal weighted delay* $\hat{r}^G \doteq \max_{i=1,\dots,N} \hat{w}_i^G/\alpha_i$. We solve this problem using large deviations techniques. First, we establish a large deviations principle (LDP) for the sequence of scaled stationary maximal weighted delays under the LWDF discipline. In particular, we characterize J_*, defined to be the quantity in (1.4) with G equal to the LWDF discipline, in terms of a finite-dimensional optimization problem. Then we establish optimality of the LWDF discipline by showing that the value of (1.4) for any other discipline G is bounded above by J_*.

The same technique used to establish the LDP for the maximal weighted delay can also be used to derive the LDP for each sequence of scaled stationary delays $\hat{w}_i$ under the LWDF discipline. Also it is clear that a Priority discipline can be viewed as the "limit" of the LWDF discipline with parameters

$\alpha_i = \varepsilon^{N-i}$, $\varepsilon \downarrow 0$. (The lower flow index means higher priority.) This allows a derivation of the LDP for the priority discipline from that of LWDF. We present the corresponding results in Section 6.4.

In addition, in Section 7 we state an analogous result for the unfinished work processes. In this case an optimal discipline is the largest weighted (unfinished) work first (LWWF), which is defined in Definition 7.1. These results can be obtained analogously (and in fact more easily) than the results for the delays.

The QoS criterion we consider is expressed in terms of probabilistic bounds on delays of packets within the network. The performance of delays in a queueing system is often harder to analyze than that of queue lengths due to the fact that systems with delays often do not admit a finite Markovian state representation. However, as mentioned earlier, delays are often the crucial measure of performance, especially for real time traffic like video, and also in certain wireless settings [1]. In [11, 15] it was shown that the earliest deadline first (EDF) discipline is optimal in the context of deterministic QoS guarantees (or worst case delay bounds) for a single node. It is generally believed that deterministic QoS requirements lead to an overly conservative admission policy, and a consequent decrease in system throughput [8]. This motivates our consideration of probabilistic QoS guarantees, which provide a very advantageous trade-off of a little quality of service for a large capacity gain.

There is a vast body of literature on QoS guarantees in communication networks (see [3, 7, 13, 17, 19, 20, 22] for just a few examples). In much of the work just cited, a policy is fixed a priori and the focus is then on performance analysis. Our perspective is different. We fix the desired performance objective, and then try to determine the policy that is best suited to achieve it. Other work that takes this perspective includes [12, 14, 18]. However these papers focus on optimality in the heavy traffic regime. In contrast, this paper provides a proof of optimality with respect to a large deviations criterion in a queueing context. Previous analyses of optimization with respect to a large deviations criteria have concentrated more on *unconstrained* diffusion processes [6, 9]. Analysis of queueing processes, and in particular analysis of their large deviation properties, is usually further complicated by the discontinuities introduced by the state space constraints. Our optimality proof entails a large deviations analysis of the LWDF discipline. Although the existence of large deviation principles for sequences of scaled queue length processes have been established for a relatively large class of queueing networks [5], the problem of determining the rate function and characterizing optimal paths has remained elusive for the vast majority of cases where there are multiple classes, even for a single node. Most of the results in the single node case have been limited to two classes [20, 19], where the analysis relies heavily on the relative simplicity of planar geometry, with the multidimensional results being rather few and restricted to queue lengths [7].

The outline of the paper is as follows. The basic notation used throughout the paper is provided in Section 1.2. In Section 2.1 we state our basic assumptions on the input flows and the class of disciplines considered, and then formulate the main result in Theorem 2.2. In Section 2.2 we introduce some basic queueing operators that are used throughout the paper. The LWDF discipline is defined and analyzed in Section 3. It is first defined for discrete inputs in Section 3.1 and then extended in Sections 3.2 and 3.3 to include fluid inputs. In Section 4 we analyze the infinite-dimensional variational problem that characterizes the rate function for the sequence of scaled stationary maximal weighted delays under the LWDF discipline, although the proof of the LDP is not given till Section 6. Theorem 4.4 shows that the variational problem can be reduced to a finite-dimensional one using the fact that the optimal paths have a very simple structure. In Section 5 we prove a result which is not used directly in the derivation of the main results of the paper, but nevertheless provides key insight into the optimality property of the LWDF scheduling discipline. In Section 6 we present the proof of the main theorem, and derive some important corollaries. Section 7 contains a result similar to our main result, Theorem 2.2, that is valid for unfinished work processes. We conclude in Section 8 with discussions of some open problems, future work and some practical implications of our results.

1.2. *Basic notation and definitions.* Let $\mathscr{D}$ [respectively $\mathscr{D}_+$] be the space of RCLL functions (i.e., right continuous functions with left limits) on $(-\infty, \infty)$ [respectively, $[0, \infty)$]. Unless otherwise specified, we assume $\mathscr{D}$ and $\mathscr{D}_+$ are endowed with the topology of uniform convergence on compact sets (u.o.c.). For any $s \geq 0$ and $f \in \mathscr{D}^N$, we define the norm

$$\|f\|_s \doteq \max_{i=1,\dots,N} \sup_{-s\leq t\leq s} |f_i(s)|.$$

For $f \in \mathscr{D}_+^N$ the norm $\|f\|_s$ is defined similarly with $-s$ replaced by 0 in the above display. Thus convergence in $\mathscr{D}^N$ or $\mathscr{D}_+^N$ is equivalent to convergence in the corresponding norm $\|\cdot\|_s$ for all $s > 0$. For $f \in \mathscr{D}^N$ we use $f(t-)$ to denote the left limit $\lim_{u\uparrow t} f(u)$ of the function f at the point t. As measurable spaces, we always assume that $\mathscr{D}$ and $\mathscr{D}_+$ are endowed with the σ-algebra generated by the cylinder sets. We now define some subspaces of $\mathscr{D}$ and $\mathscr{D}_+$ that will be used often in the sequel. Let $\mathscr{D}_{+,0}$ be the subset of functions h in $\mathscr{D}_+$ such that $h(0) = 0$. Let $\mathscr{S}$ be the space of functions in $\mathscr{D}$ which are nondecreasing, piecewise constant, and have only a finite number of jumps on any finite time interval, and let $\mathscr{S}_+$ (respectively, $\mathscr{S}_{+,0}$) be the subset of *nonnegative* functions in $\mathscr{D}_+$ (respectively, $\mathscr{D}_{+,0}$) which are nondecreasing, piecewise constant, and have only a finite number of jumps on any finite time interval. We use $\mathscr{I}$ to denote the subset of nondecreasing functions in $\mathscr{D}$, and $\mathscr{I}_+$ (respectively, $\mathscr{I}_{+,0}$) the subset of *nonnegative* nondecreasing functions in $\mathscr{D}_+$ (respectively, $\mathscr{D}_{+,0}$). Let $\mathscr{C}$ be the subset of continuous functions in $\mathscr{I}_{+,0}$, and let $\mathscr{C}_a$ be the subset of absolutely continuous functions in $\mathscr{C}$. We assume that the subspaces inherit the topology and σ-algebra of the original space. Given any space $\mathscr{S}$, $\mathscr{S}^N$ represents the N times product space with the product topology and product σ-algebra defined in the natural way. We define $\mathbb{R}_+^N \doteq \{x \in \mathbb{R}^N\colon x_i \geq 0, \text{ for } i = 1, \dots, N\}$.

In this paper we use the notation

$$h_n \Rightarrow h \quad \text{as } n \to \infty,$$

for $h, h_n \in \mathscr{I}_+$ to denote convergence at every point of continuity of h. Although we do not use this fact, we note that (since all h_n and h are nondecreasing) this convergence is equivalent to convergence in the Skorohod M_1-topology [23].

Let $\mathbf{\Omega} \doteq (\Omega, \mathscr{F}, P)$ 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 $B \subset \Omega$. If $B \in \mathscr{F}$, then $P_*(B) = P(B)$. Given any subset B of a topological space, we use $\overline{B}$ and B° to denote its closure and interior, respectively. The infimum of a function over an empty set is interpreted as ∞. We use $\mathscr{Q}$ to denote the set of rational numbers. We let $a \wedge b$ denote the minimum of the scalars a and b.

We follow the convention of using bold font for stochastic processes and roman font for deterministic processes.

We now give the definition of a large deviation principle ([4], page 5). Let $\mathscr{X}$ be a topological space and $\mathscr{B}$ a σ-algebra on $\mathscr{X}$. Note that $\mathscr{B}$ is not necessarily the Borel σ-algebra.

DEFINITION 1.1 (LDP). A sequence of random variables $\{\mathbf{X}_n\}$ on Ω taking values in a topological space $(\mathscr{X}, \mathscr{B})$ is said to satisfy the LDP with rate function I if for all $\Theta \in \mathscr{B}$,

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

and

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

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

We define the scaling operator Γ^c, $c > 0$, for elements $Y \in \mathscr{D}^N$ (or $\mathscr{D}_+^N$) as follows:

$$(\Gamma^c Y)(t) \doteq \frac{1}{c} Y(ct). \tag{1.5}$$

For a scalar b $\Gamma^c b \doteq b/c$ and for the pair (b, Y), $\Gamma^c(b, Y) \doteq (b/c, \Gamma^c Y)$. 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 $f \in S$ and $c > 0$,

$$\Gamma^c A(f) = A(\Gamma^c f). \tag{1.6}$$

2. Basic model. In Section 2.1 we state our assumptions and the main result, and in Section 2.2 we introduce some basic operators that will be used throughout the paper.

2.1. *Assumptions and main results.* We consider a multiclass single server queueing system with N input flows that satisfy the following assumptions.

ASSUMPTION 2.1 (Input flows).

(i) *Each flow* $\mathbf{f}_i$ *is a random process on* $\mathbf{\Omega}$ *with stationary increments that takes values in* $\mathscr{S}$.

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

(iii) *For each* i, *and every* $T < \infty$, *the sequence of processes* $\{\mathbf{f}_i^n - \mathbf{f}_i^n(0), n = 1, 2, \ldots\}$ *restricted to* $[0, T]$ *satisfies a LDP with rate function* J_T^i *given by*

$$J_T^i(f) \doteq \begin{cases} \int_0^T L_i(\dot{f}(s))\, ds, & \text{if } f \in \mathscr{C}_a, \\ \infty, & \text{otherwise,} \end{cases} \tag{2.1}$$

where $\mathbf{f}^n = \Gamma^n \mathbf{f}$ *is the scaled version of* $\mathbf{f}$ *as defined in* (1.5), L_i *is a convex function taking values in* $[0, \infty]$ *such that* $L_i(\lambda_i) = 0$ *for some* $\lambda_i \in (0, \infty)$, $L_i(x) > 0$ *for* $x \neq \lambda_i$, *and* $\lim_{x \to \infty} L_i(x)/x = \infty$.

(iv) *The server is not overloaded. In other words,*

$$\sum_1^N \lambda_i < 1.$$

The process $\mathbf{f}_i(t)$ represents the cumulative amount of work of class i (measured in terms of the required service time) 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 time" of that customer equal to the size of the jump $\mathbf{f}_i(t) - \mathbf{f}_i(t-)$.

Consider the class $\mathscr{G}$ of queueing (or scheduling) disciplines G such that:

1. G is work conserving.
2. Scheduling decisions at any time t are *independent* of both the future of the process and the history of the process *before the beginning of the current busy period*.

Let G be an arbitrary discipline in $\mathscr{G}$. Given an input flow sample path $f \in \mathscr{S}^N$, let $\hat{\tau}_i^G(t)$ be the arrival time of the "longest waiting" class i customer present in the system at time t, including the customer(s) being served at time t. [We will show later that this function is well defined for almost all sample paths of the input flow process $\mathbf{f} = (\mathbf{f}_1, \ldots, \mathbf{f}_N)$.] By convention we assume that $\hat{\tau}_i^G$ is right continuous, and set $\hat{\tau}_i^G(t) = t$ if there is no class i customer in the system at time t. We refer to $\hat{\tau}_i^G \doteq (\hat{\tau}_i^G(t), -\infty < t < \infty)$ as the class i *backlog* process. Suppose we are given the set of positive *weights*,

$$0 < \alpha_1 \leq \alpha_2 \leq \cdots \leq \alpha_N.$$

Then we define the class i delay $\hat{w}_i^G$ and *weighted delay* $\hat{r}_i^G$ processes and the *maximal weighted delay* $\hat{r}^G$ process in terms of the class i backlog process as follows. For $t \geq 0$ and $i = 1, \dots, N$,

$$\hat{w}_i^G(t) \doteq t - \hat{\tau}_i^G(t), \tag{2.2}$$

$$\hat{r}_i^G(t) \doteq \hat{w}_i^G(t)/\alpha_i, \tag{2.3}$$

$$\hat{r}^G(t) \doteq \max_i \hat{r}_i^G(t). \tag{2.4}$$

We are interested in identifying a discipline that is optimal in the sense that it maximizes the exponential decay rate of the stationary distribution of the maximal weighted delay $\hat{r}^G(\cdot)$. Our main result is that the largest weighted delay first (LWDF) discipline introduced in Section 3 is optimal in a large deviations sense as formulated below in Theorem 2.2.

THEOREM 2.2. *There exists $J_* < \infty$ such that the following hold.*

(i) *For the LWDF scheduling discipline defined in Section* 3, *the upper bound*

$$\limsup_{n\to\infty} \frac{1}{n} \log P\left(\frac{1}{n}\hat{\mathbf{r}}(0) > 1\right) \leq -J_* \tag{2.5}$$

holds, where $\hat{r}(0)$ is the stationary maximal weighted delay associated with the LWDF discipline.

(ii) *For any $G \in \mathscr{G}$, we have the lower bound*

$$\liminf_{n\to\infty} \frac{1}{n} \log P_*\left(\frac{1}{n}\hat{\mathbf{r}}^G(0) > 1\right) \geq -J_*, \tag{2.6}$$

where $\hat{r}^G(0)$ is the stationary maximal weighted delay associated with the discipline G.

(iii) *Moreover, J_* solves the following finite-dimensional optimization problem:*

$$J_* = \min_{j;\, x_1, \dots, x_j} \frac{1}{\gamma} \sum_{i=1}^{j} (1 - \alpha_i \gamma) L_i(x_i), \tag{2.7}$$

subject to

$$j \in \{1, \dots, N\}, \qquad x_i > 0, \qquad \sum_{i=1}^{j} x_i > 1$$

and

$$\frac{1}{\alpha_{j+1}} < \gamma = \frac{\sum_{i=1}^{j} x_i - 1}{\sum_{i=1}^{j} \alpha_i x_i} \leq \frac{1}{\alpha_j},$$

with $\alpha_{N+1} \doteq \infty$ by convention.

REMARK. A rigorous definition of the stationary maximal weighted delay $\hat{\mathbf{r}}^G$ for an arbitrary discipline $G \in \mathscr{G}$, and the proof of the theorem is presented in Section 6. As we will see in Section 6, $\hat{\mathbf{r}}^G(\cdot)$ need not in general be a measurable function on the probability space. To allow for this generality, we use the notion of inner measure in the lower bound [Theorem 2.2(ii) above].

2.2. *Operators associated with a queueing discipline.* Consider a fixed discipline $G \in \mathscr{G}$. Suppose the server is empty at time 0, which means that the left limit of the unfinished work in the system is equal to 0. Let a path $f \in \mathscr{S}_+^N$ describe the input process starting at time 0. [If $f_i(0) > 0$, it means that a type i customer arrived at time 0 requiring a service time $f_i(0)$.] By assumption (2) in the definition of the class $\mathscr{G}$, for each i, $\hat{\tau}_i^G(0) = 0$ and the "evolution" of $\hat{\tau}_i^G(\cdot)$ from time 0 onwards depends only on the path f, and not on the system behavior before $t = 0$. Given discipline G, we define $\widehat{A}^G$ to be the operator

$$\widehat{A}^G\colon \mathscr{S}_+^N \mapsto \mathscr{I}_{+,0}^N$$

that maps f into the corresponding $\hat{\tau}^G \doteq ((\hat{\tau}_i^G(t), t \geq 0), i = 1, \dots, N)$. It is clear that for any $f \in \mathscr{S}_+^N$, $\hat{\tau}^G = \widehat{A}^G f$ is such that for $i = 1, \dots, N$,

$$\hat{\tau}_i^G(0) = 0 \quad \text{and} \quad \hat{\tau}_i^G(t) \leq t \ \text{ for } t \geq 0.$$

Once again by assumption (2) in the definition of the class $\mathscr{G}$, the discipline G is completely characterized by its "behavior" within a busy period. However the points in time when new busy periods of the system start do *not* depend on the queueing discipline, since all disciplines in $\mathscr{G}$ are work conserving. Thus the operator $\widehat{A}^G$ associated with a discipline G completely characterizes the evolution of all $\hat{\tau}_i^G(t)$ for all $t \in (-\infty, \infty)$. Moreover, if the discipline $G \in \mathscr{G}$ is *nonpreemptive*, then the operator $\widehat{A}^G$ completely characterizes the discipline itself.

We also introduce a related operator,

$$\widehat{R}^G\colon \mathscr{S}_+^N \mapsto \mathscr{D}_{+,0},$$

which maps an input flow path $f \in \mathscr{S}_+^N$ into the corresponding maximum weighted delay $\hat{r}^G \doteq (\hat{r}^G(t), t \geq 0)$. Clearly, each operator $\widehat{R}^G$ is uniquely determined by the corresponding operator $\widehat{A}^G$.

3. The largest weighted delay first discipline.

In this section we define the largest weighted delay first (LWDF) discipline. In Section 3.1 we define the LWDF discipline for discrete input paths taking values in $\mathscr{S}_{+,0}^N$. In Section 3.2 we introduce the notion of a *virtual backlog* process, which is then used in Section 3.3 to extend the definition of LWDF to fluid input flow paths $f \in \mathscr{C}^N$.

3.1. *The LWDF discipline for discrete inputs.* In Section 2.1 we introduced the backlog, delay, weighted delay and maximal weighted delay associated with any queueing discipline G and input flows in $\mathscr{S}_+^N$. Recall that a discipline is said to be nonpreemptive if the service of a customer cannot be interrupted before its service has been completed. We now define the largest weighted delay first (LWDF) discipline for input flows in $\mathscr{S}_+^N$.

DEFINITION 3.1 [The largest weighted delay first (LWDF) discipline]. The LWDF discipline is a nonpreemptive, work conserving discipline that always chooses for service the longest waiting (i.e., head-of-the-line) customer of the flow i which has the maximal weighted delay $\hat{r}_i(t) = \hat{r}(t)$. In case of a tie, by convention the LWDF discipline chooses the class with the highest index.

Note that LWDF is first-in-first-out within each class. We drop the subscript G for the LWDF discipline and denote the operators $\widehat{A}^G$ and $\widehat{R}^G$ associated with the LWDF discipline simply by $\widehat{A}$ and $\widehat{R}$, respectively.

3.2. *Virtual delays for LWDF.* In the last section we introduced the operator $\widehat{A}$ that characterizes the LWDF discipline. We now introduce the related operator

$$A\colon \mathscr{S}_+^N \mapsto \mathscr{I}_{+,0}^N,$$

which maps $f \in \mathscr{S}_+^N$ (i.e., input flows starting at time 0) into the *virtual backlog* process $\tau \in \mathscr{I}_{+,0}^N$ defined below. The main reason for introducing the virtual backlog process is that it enables one to extend the definition of the LWDF discipline to continuous input sample paths $f \in \mathscr{C}^N$, as shown in the next section. To define the operator A, as in the definition of the operator $\widehat{A}$ we first assume that the system is empty at time 0– and consider a fixed $f \in \mathscr{S}_+^N$ which describes the input flows starting at time 0. Consider $\hat{\tau} \doteq \widehat{A}f$, and the corresponding processes $\hat{w}_i(\cdot)$, $\hat{r}_i(\cdot)$ for $i = 1, \dots, N$, and $\widehat{R}f = \hat{r} = \max_i \hat{r}_i$.

We define the *virtual backlog* $\{\tau_i(t), t \geq 0\}$ for the flow i as follows. At any time t when the system is empty, $\tau_i(t) = \hat{\tau}_i(t) = t$. Within a busy period of the system, let

$$t_0 < t_1 < \cdots < t_k$$

be a (finite) sequence of time instants when a new customer arrival or a service completion occurs. Then within this busy period the functions $\tau_i(\cdot)$ are nondecreasing piecewise constant (namely, constant in each interval $[t_m, t_{m+1}]$) and are defined by the following recursion. For every $i = 1, \dots, N$, let

$$\tau_i(t_0) \doteq \hat{\tau}_i(t_0) = t_0, \tag{3.1}$$

$$\tau_i(t_{m+1}) \doteq \max\{\tau_i(t_m), t_{m+1} - \alpha_i \hat{r}(t_{m+1})\}. \tag{3.2}$$

We will refer to the recursion (3.2) as the "discrete system jump rule" (for the LWDF discipline). It simply describes how far ahead each $\tau_i(\cdot)$ is moved after a service of a customer is completed.

We defined the *virtual delay, virtual weighted delay* and the *virtual maximal weighted delay* in the natural way as follows. For $t \geq 0$ let

$$w_i(t) \doteq t - \tau_i(t), \tag{3.3}$$

$$r_i(t) \doteq w_i(t)/\alpha_i, \tag{3.4}$$

$$r(t) \doteq \max_i r_i(t). \tag{3.5}$$

We let R denote the operator which maps $f \in \mathscr{S}_+^N$ into $r \in \mathscr{D}_{+,0}$.

The backlog $\hat{\tau}_i(\cdot)$ and the virtual backlog $\tau_i(\cdot)$ are clearly closely related. The following lemma summarizes the relation between the two.

LEMMA 3.2. *The virtual delay process τ satisfies the following properties:*

(a) *For all $t \geq 0$,*

$$\tau_i(t) \leq \hat{\tau}_i(t) \quad \text{for } i = 1, \dots, N, \tag{3.6}$$

which implies that $w_i(t) \geq \hat{w}_i(t)$ and $r_i(t) \geq \hat{r}_i(t)$. In addition, we have

$$r_1(t) \geq r_2(t) \geq \cdots \geq r_N(t), \tag{3.7}$$

and

$$\tau_1(t) \geq \tau_2(t) \geq \cdots \geq \tau_N(t). \tag{3.8}$$

(b) *If the service of a customer of type i starts at time t, then*

$$\tau_i(t) = \hat{\tau}_i(t), \tag{3.9}$$

and consequently $w_i(t) = \hat{w}_i(t)$ and $r_i(t) = \hat{r}_i(t)$. Moreover,

$$r_i(t) = r(t) = \hat{r}(t) = \hat{r}_i(t), \tag{3.10}$$

$$r_1(t) \geq r_2(t) \geq \cdots \geq r_N(t), \tag{3.11}$$

$$\tau_1(t) \geq \tau_2(t) \geq \cdots \geq \tau_N(t), \tag{3.12}$$

$$\hat{k}(t) = \arg\max_j \hat{r}_j(t) \subseteq k(t) = \arg\max_j r_j(t), \tag{3.13}$$

with the set $k(t)$ having the form

$$k(t) = \{1, 2, \dots, i\}. \tag{3.14}$$

PROOF. Property (3.6) of statement (a) follows directly from (3.2). As far as the proof of statement (b) is concerned, we only need to prove it within a system's busy period, which is easily done by induction of the "event" time instants $t_m, m = 0, 1, \dots,$ using again the recurrence (3.2). Then we get (3.7) and (3.8) from (3.11) and (3.12). □

3.3. *Extension of operator A to fluid inputs.* We now extend the domain of the operator A (characterizing the LWDF discipline) that was defined on $\mathscr{S}_+^N$ in the last section to include the space of continuous (or "fluid") input flows $f \in \mathscr{C}^N$. As shown in Theorem 3.3, the operator is naturally extended to the space $\mathscr{C}^N$ by continuity in the product topology induced by the "at the continuity points" ($\Rightarrow$) convergence.

THEOREM 3.3. *Given $f \in \mathscr{C}^N$, there exists a unique*

$$\tau = \{(\tau_i(t), t \geq 0), i = 1, \dots, N\} \in \mathscr{I}_{+,0}^N$$

such that the following statements hold:

(a) *For any sequence $\{f^{(n)}: f^{(n)} \in \mathscr{S}_+^N, n = 1, 2, \dots\}$ such that*

$$f^{(n)} \Rightarrow f, \tag{3.15}$$

we have

$$\tau^{(n)} \doteq Af^{(n)} \Rightarrow \tau. \tag{3.16}$$

(b) *Each function $\tau_i(\cdot)$ is nondecreasing, right continuous, with*

$$\tau_i(0) = 0 \quad \textit{and} \quad \tau_i(t) \leq t \ \textit{ for } t \geq 0. \tag{3.17}$$

(c) *Define*

$$\hat{\tau}_i(t) \doteq \sup\big(\xi \mid f_i(\xi) = f_i(\tau_j(t-))\big) \wedge t$$

and let $w_i(\cdot), r_i(\cdot), r(\cdot)$ be defined in terms of τ, and $\hat{w}_i(\cdot), \hat{r}_i(\cdot)$ and $\hat{r}(\cdot)$ in terms of $\hat{\tau}$ as for the discrete system by the set of equations (2.2)–(2.4) *and* (3.3)–(3.5)*. Then τ satisfies the following additional conditions:*

(i) *For $i = 1, \dots, N$ and all $t \geq 0$,*

$$\tau_i(t) \leq \hat{\tau}_i(t). \tag{3.18}$$

(ii) *The following "conservation law" holds for any $t \geq 0$,*

$$\sum_{i=1}^{N} f_i(\tau_i(t)) = t + \left[\inf_{0 \leq \xi \leq t} \left(\sum_{i=1}^{N} f_i(\xi) - \xi\right)\right] \wedge 0. \tag{3.19}$$

(iii) *If $t \in \arg\min_{0 \leq \xi \leq t}(\sum_{i=1}^N f_i(\xi) - \xi)$ and $\sum_{i=1}^N f_i(t) - t \leq 0$ (i.e., the fluid system at time t is empty), then for every $i = 1, \dots, N$,*

$$\tau_i(t) = t. \tag{3.20}$$

(iv) *If for some t, $r_i(t) < r_j(t)$ and $\tau_j(t) = \hat{\tau}_j(t)$, then there exists $\varepsilon > 0$ such that for any $\theta \in [t, t + \varepsilon]$,*

$$\tau_i(\theta) = \tau_i(t). \tag{3.21}$$

(v) *The following "fluid system jump rule" holds for any $t \geq 0$, and any $i \in \{1, \dots, N\}$,*

$$r_i(t) = r_i(t-) \wedge \hat{r}(t). \tag{3.22}$$

(vi) *For any* $t \geq 0$,

$$r_1(t) \geq r_2(t) \geq \cdots \geq r_N(t), \tag{3.23}$$

$$\tau_1(t) \geq \tau_2(t) \geq \cdots \geq \tau_N(t), \tag{3.24}$$

and for any $t \geq 0$, *there exists* j *such that*

$$r(t) = r_j(t) = \hat{r}_j(t) = \hat{r}(t). \tag{3.25}$$

Theorem 3.3 is a direct consequence of the following two propositions.

PROPOSITION 1. *Consider* $f \in \mathscr{C}^N$ *and any sequence* $\{f^{(n)}: f^{(n)} \in \mathscr{S}_+^N\}$ *such that*

$$f^{(n)} \Rightarrow f.$$

Then there exists a subsequence $\{f^{(n_l)}\} \subseteq \{f^{(n)}\}$ *and* $\tau \in \mathscr{I}_{+,0}^N$ *satisfying properties* (3.17)–(3.25) *such that*

$$\tau^{(n_l)} \Rightarrow \tau, \tag{3.26}$$

where $\tau^{(n_l)} \doteq Af^{(n_l)}$.

PROPOSITION 2. *Given* $f \in \mathscr{C}^N$, *an element* $\tau \in \mathscr{I}_{+,0}^N$ *satisfying properties* (3.17)–(3.25) *is unique.*

PROOF OF PROPOSITION 1. Given $f \in \mathscr{C}^N$ and a sequence $\{f^{(n)}: f^{(n)} \in \mathscr{S}_+^N\}$ such that $f^{(n)} \Rightarrow f$, let $\tau^{(n)} \doteq Af^{(n)}$. Since each function $\tau_i^{(n)}(\cdot)$ is nondecreasing, and for any fixed t satisfies

$$\tau_i^{(n)}(t) \leq t, \tag{3.27}$$

there exists a subsequence $\{n_l\} \subseteq \{n\}$ such that

$$\tau_i^{(n_l)}(q) \to \bar{\tau}_i(q), \qquad q \in \mathscr{Q}, q \geq 0, \tag{3.28}$$

where recall that $\mathscr{Q}$ is the set of rational numbers. Clearly $\bar{\tau}_i(\cdot)$ is a nondecreasing function since each $\tau_i^{(n_l)}(\cdot)$ is. We extend the definition of $\tau_i(\cdot)$ to all reals by right-continuity, that is for $t \geq 0$ we set

$$\tau_i(t) \doteq \inf_{q > t, q \in \mathscr{Q}} \bar{\tau}_i(q). \tag{3.29}$$

It follows immediately that $\tau_i(\cdot)$ is nondecreasing, and

$$\tau_i^{(n_l)}(t) \to \tau_i(t)$$

at each point of continuity t of $\tau_i(\cdot)$, and thus we obtain

$$\tau^{(n_l)} \Rightarrow \tau. \tag{3.30}$$

It remains to show that τ so defined satisfies (3.17) to (3.25). Property (3.18) follows trivially from the definition of $\hat{\tau}$. Properties (3.23) and (3.24) of τ follow from the construction of τ and the corresponding properties (3.7) and (3.8) of the prelimiting functions $\tau^{(n_l)}$.

Before we proceed, let us notice that property (3.19) (along with all the properties described in Lemma 3.2) holds with f and τ replaced by every *prelimit* pair of functions $f^{(n_l)}$ and $\tau^{(n_l)}$, for any time t when the service of a customer starts or, trivially, when the (fluid) system is empty. It is easy to see that the set of such time points is asymptotically dense as $n_l \to \infty$, in the sense that the following property holds. *For any n_l and $t \geq 0$, we define $s^{(n_l)}(t)$ to be the first time instant in (t, ∞) at which either a new service of a customer starts or the system is empty. Then*

$$\lim_{n_l \to \infty} s^{(n_l)}(t) = t \quad \forall t \geq 0. \tag{3.31}$$

Indeed, if this were not true, there must exist $\varepsilon > 0$ such that for all sufficiently large n_l, nowhere in $(t, t+\varepsilon)$ does a new service start or the system become empty. This would mean that for all sufficiently large n_l a single customer is served in that interval, which implies that for all sufficiently large n_l at least one of the functions $f_i^{(n_l)}(\cdot)$ has a jump of size at least ε in the interval $[0, t]$. This contradicts the u.o.c. convergence

$$f_i^{(n_l)}(\cdot) \to f_i(\cdot) \quad \text{for every } i.$$

This proves property (3.31).

For notational convenience, let us denote by $\widehat{V}$ the "completed work" mapping which maps each $f \in \mathscr{I}^N_{+,0}$ into the function $\widehat{V} f \in \mathscr{D}_+$ defined by the right-hand side (RHS) of (3.19). It is well known that $\widehat{V} f \in \mathscr{C}$ for any $f \in \mathscr{I}^N_{+,0}$, and the mapping $\widehat{V}$ is continuous on the elements $f \in \mathscr{C}^N$.

We now prove property (3.19). Using the fact that the RHS of (3.19) is equal to $\widehat{V} f$ and is therefore a continuous function, and using property (3.31), we infer that (3.19) holds at every time t where all $\tau_i(\cdot)$ are continuous. Then using right-continuity of all $\tau_i(\cdot)$, and again the continuity of the RHS of (3.19), we get (3.19) for all $t \geq 0$.

We introduce some more notation to prove the remaining properties. For a nondecreasing function $h \in \mathscr{I}_+$ a time $t \geq 0$ will be called a *right point of growth* of h if

$$h(t) < h(\xi) \quad \text{for all } \xi > t.$$

The proof of property (3.20) is by contradiction. Suppose, at some fixed time t the fluid system is empty, but $\tau_{i_*}(t) < t$ for some i_*. If t is *not* a right point of growth of any of the functions $f_i(\cdot)$, then it is easy to see that for any $\varepsilon > 0$, for all sufficiently large n_l, there exists a time $x^{(n_l)} \in (t, t+\varepsilon)$ when the *prelimit* system with index n_l must be empty, which means $\tau_i^{(n_l)}(x^{(n_l)}) = x^{(n_l)}$, $\forall\, i$. This would imply $\tau_i^{(n_l)}(t+\varepsilon) \geq t$, $\forall\, i$, and therefore $\tau_i(t+\varepsilon) \geq t$, $\forall\, i$. By right-continuity of each $\tau_i(\cdot)$ we get $\tau_i(t) \geq t, \forall\, i$, a contradiction.

Now suppose t is a right point of growth of at least one function $f_i(\cdot)$. Notice that for every i, $f_i(\tau_i(t)) = f_i(t)$ due to the conservation law and the fact that the (fluid) system is empty at time t. From the latter observation and the conservation law we see that for at least one fixed k, and any $\varepsilon > 0$, we have $\tau_k(t+\varepsilon) > t$ [and $\tau_k(t+\varepsilon) \le t+\varepsilon$, of course]. On the other hand for any small $\delta_1 > 0$, there exists $\delta > 0$ such that for any $\xi \in (t, t+\delta)$, $\tau_{i_*}(\xi) \le \tau_{i_*}(t) + \delta_1$. This means that the sequence of prelimit systems is such that for all large n_l, there exists a time $y^{(n_l)} \in (t, t+\varepsilon)$ such that the service of a type k customer (that arrived after time t) starts, and yet $\tau_{i_*}^{(n_l)}(y^{(n_l)}) \le \tau_{i_*}(t) + 2\delta_1$. Since for a fixed small $\delta_1 > 0$ (and the corresponding $\delta > 0$), we can choose $\varepsilon \in (0, \delta)$ arbitrarily small, this leads to a contradiction with the LWDF scheduling rule, namely with the property (3.10) saying that if (in a prelimit system with index n_l) a type k customer is chosen for service at time ξ, then $r_k^{(n_l)}(\xi) = \max_i r_i^{(n_l)}(\xi)$. This completes the proof of property (3.20).

We now prove property (3.22). First, let us observe that if at time t the fluid system is empty; that is, we are in the conditions of statement (c)(iii) of the theorem, then (3.20) holds and therefore (3.22) holds trivially. So we only need to consider the case when the fluid system is nonempty at time t, that is,

$$\sum f_i(\tau_i(t)) = \widehat{V}[f](t) < \sum f_i(t).$$

Then it follows from the definition of the operator $\widehat{V}$ that t is a right point of growth of $\widehat{V}[f]$. This in turn implies that for at least one m, $\tau_m(t+\varepsilon) > \hat{\tau}_m(t)$ for any $\varepsilon > 0$; which means that $\tau_m(t) \ge \hat{\tau}_m(t)$ by the right-continuity of $\tau_m(\cdot)$. Thus we get the existence of m such that $\tau_m(t) = \hat{\tau}_m(t)$. Next, we observe that this m can be chosen such that

$$\hat{r}_m(t) = \max_i \hat{r}_i(t) = \hat{r}(t).$$

[If this were not true, we would get a contradiction to the LWDF scheduling rule similar to the one we got in the proof of property (3.20).]

We show below that for every $i = 1, \dots, N$,

$$r_i(t) \le \hat{r}(t) \quad \forall\, i. \tag{3.32}$$

Indeed, for any $\varepsilon > 0$, for all large n_l there exists a time point $y^{(n_l)} \in (t, t+\varepsilon)$ such that the service of a class m customer starts at $y^{(n_l)}$, and therefore

$$r_m^{(n_l)}(y^{(n_l)}) = r^{(n_l)}(y^{(n_l)}),$$

which implies that for any i,

$$\tau_i^{(n_l)}(y^{(n_l)}) \ge t - r^{(n_l)}(y^{(n_l)})\alpha_i.$$

Since $\varepsilon > 0$ can be arbitrarily small, we see that for any $\delta > 0$,

$$\tau_i(t+\delta) \ge \liminf_{n_l\to\infty} \tau_i^{(n_l)}(t+\delta) \ge t - \hat{r}(t)\alpha_i.$$

Since $\delta > 0$ is arbitrary, we get

$$\tau_i(t) = \lim_{\delta\downarrow 0} \tau_i(t+\delta) \ge t - \hat{r}(t)\alpha_i,$$

which proves (3.32).

It remains to show that

$$\text{if } r_i(t-) \geq \hat{r}(t) \text{ then } r_i(t) \geq \hat{r}(t) \tag{3.33}$$

and

$$\text{if } r_i(t-) < \hat{r}(t) \text{ then } r_i(t) = r_i(t-). \tag{3.34}$$

Property (3.33) for $i \leq m$ is trivial because (3.23) implies

$$r_i(t) \geq r_m(t) = \hat{r}(t).$$

To prove (3.33) for $i > m$, we will need the following property of a prelimit system (with index n_l) which follows from the discrete system "jump rule" (3.2).

Suppose for some fixed i and $m, i > m$, and for a fixed $t \geq 0$, there is a class i customer which arrived at some time $\xi < t$, has not started service by time t, and $r_i^{(n_l)}(t) > (t - \xi)/\alpha_m$. Then for any time $n \geq t$,

$$\tau_m^{(n_l)}(\eta) \leq \xi \quad \text{implies} \quad \tau_i^{(n_l)}(\eta) \leq t - \alpha_i \frac{t - \xi}{\alpha_m}.$$

Returning to the proof of (3.33) for $i > m$, we observe that if $r_i(t-) \geq \hat{r}(t)$ for some $i > m$, then for any small $\varepsilon > 0$, and for all sufficiently large n_l and $\eta \geq t$,

$$\tau_m^{(n_l)}(\eta) \leq t - \alpha_m \hat{r}(t)(1 - \varepsilon) \quad \text{implies} \quad \tau_i^{(n_l)}(\eta) \leq t - \alpha_i \hat{r}(t)(1 - 2\varepsilon).$$

Let us fix a small $\varepsilon > 0$. Then we can choose $\delta > 0$ small enough so that for all large n_l,

$$\tau_m^{(n_l)}(t + \delta) \leq t - \alpha_m \hat{r}(t)(1 - \varepsilon),$$

which means that

$$\tau_i^{(n_l)}(t + \delta) \leq t - \alpha_i \hat{r}(t)(1 - 2\varepsilon),$$

that is,

$$r_i^{(n_l)}(t + \delta) \geq \hat{r}(t)(1 - 2\varepsilon).$$

This in turn implies $r_i(t) \geq \hat{r}(t)(1 - 2\varepsilon)$. Since $\varepsilon > 0$ is arbitrary, we get $r_i(t) \geq \hat{r}(t)$, and the proof of (3.33) is complete.

To prove (3.34) we first observe that

$$r_i(t-) < \hat{r}(t) = r_m(t)$$

is only possible if $i > m$. Arguing very similarly to the proof of (3.33), we can see that for a sufficiently small fixed $\varepsilon > 0$, the prelimit system for all large n_l must be such that $\tau_i^{(n_l)}(\cdot)$ is constant in the interval $(t - \varepsilon, t + \varepsilon)$ [because, in this interval, there always will be type m customers which according to the LWDF scheduling rule must be served before $\tau_i^{(n_l)}(\cdot)$ can "move" forward]. This means that $\tau_i(\cdot)$ is constant in $(t - \varepsilon, t + \varepsilon)$, and therefore $r_i(\cdot)$ is continuous at t. This completes the proof of property (3.22).

We now show that (3.21) holds. First notice that $r_i(t) < r_j(t)$ is only possible if $i > j$ and there is nonzero unfinished work in the fluid system. The proof then uses almost exactly the same argument as that used above to prove (3.34).

Finally (3.25) follows directly from (3.22). □

PROOF OF PROPOSITION 2. Given $f \in \mathscr{C}^N$, suppose there exist two distinct τ and τ^* satisfying all the listed conditions, and define the process $\Delta \doteq \tau^* - \tau$. Then we claim that the following two observations are true.

Observation 1. For any $t \geq 0$, if $\Delta_N(t) > 0$ then there exists $\varepsilon > 0$ such that the difference $\Delta_N(\cdot)$ is nonincreasing in the interval $[t, t+\varepsilon]$.

Observation 2. For any $t \geq 0$, if $\Delta_N(t-) > 0$, then

$$0 \leq \Delta_N(t) \leq \Delta_N(t-). \tag{3.35}$$

Likewise, for any $t \geq 0$, if $\Delta_N(t-) \leq 0$, then

$$\Delta_N(t-) \leq \Delta_N(t) \leq 0. \tag{3.36}$$

We now show by contradiction that the two observations imply that $\Delta_N(\cdot)$ must be zero everywhere, and relegate the proofs of the observations to the end. Note that $\Delta_N(0) = 0$. Suppose the function $\Delta_N(\cdot)$ is positive somewhere on $[0, \infty)$, then there must exist $a > 0$ and $t \in [0, \infty)$ such that either $\Delta_N(t)$ or $\Delta_N(t-)$ is equal to

$$a \doteq \sup_{0 \leq \xi \leq t} \Delta_N(\xi) > 0,$$

and for all $\xi < t$

$$\Delta_N(\xi) < a. \tag{3.37}$$

Now Observation 2 implies that one cannot have $\Delta_N(t-) < a$, since this would imply that $\Delta_N(t) = a > 0$, which contradicts (3.36) if $\Delta_N(t-) < 0$, and contradicts (3.35) if $0 \leq \Delta_N(t-) < a$. Thus we infer that

$$\Delta_N(t-) = \lim_{s \uparrow t} \Delta_N(s) = a. \tag{3.38}$$

Along with (3.37), this implies that there exists $\delta > 0$ and $b < a$ such that $\Delta_N(t-\delta) = b$, and $\Delta_N(\xi) > 0$ for $\xi \in [t-\delta, t)$. From Observations 1 and 2 it follows that $\Delta_N(t)$ is nonincreasing on $[t-\delta, t]$. Indeed, let

$$s \doteq \sup\{u \in [t-\delta, t]\colon \Delta_N(\cdot) \text{ is nonincreasing on } [t-\delta, u)\}.$$

Then $s > t - \delta$ by Observation 1, and by Observation 2 $\Delta_N(\cdot)$ is nonincreasing on $[t-\delta, s]$ [since $\Delta_N(\cdot)$ cannot jump up at s]. From this we deduce that $s = t$, since if $s < t$ the fact that $\Delta_N(s) > 0$ along with Observation 1 contradicts the definition of s as the supremum. This implies that $\Delta_N(t-) \leq b$, which leads to a contradiction due to (3.38). Thus we have established that $\Delta_N(\cdot)$ must be nonpositive everywhere on $[0, \infty)$. By symmetry (interchanging the roles of τ and τ^*), it is clear that in fact $\Delta_N(\cdot) \equiv 0$ or, equivalently, for all $t \geq 0$,

$$\tau_N(t) = \tau_N^*(t). \tag{3.39}$$

It is easy to verify that the same argument can now be used inductively to establish that $\tau_i(t) = \tau_i^*(t)$ for $i = N-1, N-2, \dots, 1$, which establishes the theorem. It only remains to prove the two observations.

PROOF OF OBSERVATION 1. Suppose there exists $t \geq 0$ such that

$$\tau_N(t) < \tau_N^*(t), \tag{3.40}$$

or equivalently

$$r_N(t) > r_N^*(t). \tag{3.41}$$

By the conservation law (3.19) it follows that

$$\sum_{i=1}^{N} f_i(\tau_i(t)) = \sum_{i=1}^{N} f_i(\tau_i^*(t)).$$

Thus if $\tau_i(t) < \tau_i^*(t)$ for every $i = 1, \dots, N$, then there exists $\varepsilon > 0$ such that $f_i(\tau_i(t) + u) = f_i(\tau_i(t))$ for all $i = 1, \dots, N$ and $u \in [0, \varepsilon]$. However, since by (3.25) there exists some i for which $\tau_i(t) = \hat{\tau}_i(t)$, this would contradict the definition of $\hat{\tau}_i$. Thus there must exist $j < N$ such that

$$r_j(t) \leq r_j^*(t). \tag{3.42}$$

The ordering property (3.23) then dictates that

$$r_N(t) \leq r_j(t). \tag{3.43}$$

From the last four displays and the definition of r^* we conclude that

$$r_N^*(t) < r_i^*(t) \leq r^*(t). \tag{3.44}$$

Observation 1 then follows from (3.25), (3.21) and the fact that τ_N is nondecreasing. □

PROOF OF OBSERVATION 2. If both $\tau_N(\cdot)$ and $\tau_N^*(\cdot)$ are continuous at time t, then Observation 2 is trivial. The rest of the proof relies on the following two basic properties of the backlog process, which were established in Theorem 3.3. First, if τ_N^* jumps, then due to (3.22), (3.25) and the ordering property (3.23), it follows that

$$r_N^*(t) = r_{N-1}^*(t) = \cdots = r_1^*(t). \tag{3.45}$$

Second, since by (3.25) there always exists one component i such that $\tau_i(t)$ is a right point of growth for f_i, one can never have $\tau_i(t) < \tau_i^*(t)$ for every $i = 1, \dots, N$, as this would imply that

$$\sum_{i=1}^{N} f_i(\tau_i(t)) < \sum_{i=1}^{N} f_i(\tau_i^*(t)),$$

which violates the conservation law (3.19). These properties clearly also hold with τ_i's interchanged with τ_i^*'s.

We now proceed with the proof of (3.35). First consider the case when τ_N^* jumps at time t. Then τ_N must jump at t as well because if not, by the ordering property, (3.45) and the assumption that $\Delta_N(t-) \geq 0$, it follows that

$$r_1^*(t) = r_2^*(t) = \cdots = r_N^*(t) < r_N^*(t-) \leq r_N(t-) = r_N(t) \leq r_{N-1}(t) \cdots \leq r_1(t).$$

This implies that for every $i = 1, \ldots, N$,

$$\tau_i^*(t) > \tau_i(t),$$

which is impossible as argued above. So τ_N must also jump at t, which implies that (3.45) also holds with r^* replaced by r. If $r_N(t) < r_N^*(t)$, then once again the last display holds, and if $r_N^*(t) < r_N(t)$ then the last display holds with $>$ replaced by $<$. Since both these cases contradict the conservation law, we deduce that $\Delta_N(t) = 0$, which proves (3.35).

Now suppose τ_N^* does not jump at t, but τ_N does. Then since τ_N is nondecreasing, $\Delta_N(t) \leq \Delta_N(t-)$. Moreover, if $\tau_N(t) > \tau_N^*(t)$ then

$$r_1(t) = r_2(t) = \cdots r_N(t) < r_N^*(t) < \cdots r_1^*(t),$$

which leads to a contradiction as above. This implies that $\Delta_N(t) \geq 0$, which establishes (3.35). The property (3.36) follows from symmetry. This concludes the proof of Observation 2, and therefore of the proposition. □

Having established Theorem 3.3, we can now define the operators A and R for input flows in $\mathscr{C}^N$.

DEFINITION 3.4. Given $f \in \mathscr{C}^N$, let Af be equal to the unique τ described in the statement of Theorem 3.3. [Note that by Theorem 3.3, $\tau = Af$ if and only if τ satisfies (3.17)–(3.25).] In addition, we define the operator R: $\mathscr{C}^N \to \mathscr{D}_{+,0}$ in the natural way by setting for $t \geq 0$,

$$Rf(t) \doteq \max_i \frac{t - [Af]_i(t)}{\alpha_i}.$$

REMARK 3.5. (i) Note that in Section 3.2 we defined the operators A and R on $\mathscr{S}_+^N$, and in this section we defined the operators on the domain $\mathscr{C}^N$. Since $\mathscr{S}_+^N \cap \mathscr{C}^N$ is the zero function, for which both definitions can be seen to coincide, the operators A and R are well defined on $\mathscr{S}_+^N \cup \mathscr{C}^N$.

(ii) It is easy to verify from the definition that the operators A and R are scalable [see definition (1.6)].

We end this section with a result that is essentially a corollary of Theorem 3.3. Even though it is not used in the proofs of our main results, it is important since the result can be used directly to derive the LDP (via the contraction principle [4, 21]) for the maximal weighted delay process $\mathbf{r}$ for a system with random fluid input flows.

THEOREM 3.6. *The mapping*

$$A\colon \mathscr{C}^N \mapsto \mathscr{I}^N_{+,0}$$

is continuous in the product topology induced by the "⇒" convergence.

PROOF. Consider $f \in \mathscr{C}^N$ and let $\{f^k\}$, $f^k \in \mathscr{C}^N$, be a sequence such that

$$f^k \Rightarrow f \quad \text{as } k \to \infty.$$

Let $\tau \doteq Af$ and $\tau^k \doteq Af^k$. We need to prove that

$$\tau^k \Rightarrow \tau. \tag{3.46}$$

If (3.46) were false, then we could choose a subsequence $\{f^{k_l}\}$ such that

$$\tau^{K_l} \Rightarrow \tau_* \neq \tau. \tag{3.47}$$

For each τ^{k_l}, by the definition of A, there exists a sequence $\{f^{k_l,n}\colon f^{k_l,n} \in \mathscr{S}^N_+, n = 1, 2, \ldots\}$ such that

$$f^{k_l,n} \Rightarrow f^{k_l} \quad \text{and} \quad \lim_{n\to\infty} Af^{k_l,n} = \tau^{k_l},$$

where recall that here A is defined on discrete input paths via equations (3.1) and (3.2). Then using Cantor's diagonal procedure (similarly to the way it is done in [24]), we can construct a sequence $\{f^{k_l(m),n(m)},\ m = 1, 2, \ldots\}$ such that

$$f^{k_l(m),n(m)} \Rightarrow f \quad \text{but } Af^{k_l(m),n(m)} \Rightarrow \tau_* \neq \tau.$$

However, this contradicts the definition of τ as the *unique* limit of Af^m for any $f^m \Rightarrow f$. □

4. Most likely paths for large stationary delays under the LWDF discipline. It is well known that large deviations techniques can be used to convert the problem of characterizing asymptotic limits of probabilities into that of solving variational problems. The optimal solutions of the variational problems shed insight into the most probable way in which rare events occur. In Section 6 we show that the optimal decay rate J_* of the tail of the stationary maximal weighted delay process associated with the LWDF discipline (see Theorem 6.8) is given by the following variational problem:

$$J_* = \inf_{s>0,\, f\in\mathscr{C}^N_a\colon\, Rf(s)\geq 1} J_s(f), \tag{4.1}$$

where for every $s \geq 0$,

$$J_s(f) \doteq \sum_{i=0}^{N} J^i_s(f_i), \tag{4.2}$$

and J^i_s is defined in (2.1). In this section we characterize the optimal path f of the variational problem that achieves the cost J_*. We use this characterization in Section 6 to prove the LD lower bound (2.6) of our main result Theorem 2.2. In the next three lemmas we show that the infimization in (4.1) can without loss of generality be restricted to successively simpler sets. In Theorem 4.4 we show that the infimization is actually attained on a "simple element" (whose definition is given below), and characterize J_* in terms of a finite-dimensional optimization problem.

LEMMA 4.1. *Let J_* be as defined in* (4.1). *Then*

$$J_* = \inf_{s>0,\, f\in\mathscr{C}_a^N} J_s(f), \tag{4.3}$$

subject to

$$Rf(s) = 1$$

and

$$\sum_{i=1}^{N} f_i(t) > t \quad \textit{for } t \in (0, s]. \tag{4.4}$$

PROOF. As stated in Remark 3.5, the operator R is scalable. Thus given any path $f \in \mathscr{C}_a^N$ such that $Rf(s) = c > 1$ for some $s > 0$, there exists a scaled down path of no greater cost for which equality holds. Specifically, consider the path $\tilde{f} \doteq \Gamma^c f$. Then $\tilde{f} \in \mathscr{C}_a^N$, $R\tilde{f}(s/c) = 1$ by the scalability of R, and

$$J_{s/c}(\tilde{f}) = \sum_{i=1}^{N} \int_0^{s/c} L_i(\dot{\tilde{f}}_i(u))\, du = \frac{1}{c}\sum_{i=1}^{N} \int_0^{s} L_i(\dot{f}_i(u))\, du = \frac{1}{c} J_s(f).$$

Thus we can restrict the infimization in (4.1) to the set of paths for which $Rf(s) = 1$. Now given any path $f \in \mathscr{C}_a^N$ with $Rf(s) = 1$, we can without any increase in cost replace it by the increments of f after the queue is empty for the last time. More precisely, replace f by the path $\tilde{f}$ defined by

$$\tilde{f}_i(t) \doteq \begin{cases} f_i(t + s_1), & \text{for } t \in [0, s - s_1], \\ f_i(s - s_1) + \lambda_i(t - s + s_1), & \text{for } t \in [s - s_1, \infty), \end{cases}$$

for $i = 1, \dots, N$, where

$$s_1 \doteq \max\left\{\xi \le s\colon \sum_i f_i(\xi) = \xi + \left[\inf_{0\le\eta\le\xi}\left(\sum_{i=1}^{N} f_i(\eta) - \eta\right)\right] \wedge 0\right\},$$

and recall that λ_i is the mean input rate for flow i that was introduced in Assumption 2.1.

It can be easily verified that $R\tilde{f}(s - s_1) = 1$ and $J_{s-s_1}(\tilde{f}) \le J_s(f)$, which establishes (4.3). □

Now consider any $f \in \mathscr{C}_a^N$ such that $Rf(T) = 1$ for some $T > 0$ and condition (4.4) is satisfied with $s = T$. Let $r \doteq Rf$. For $i = 1, \dots, N$ define

$$T_i \doteq \sup\{t \le T\colon r_i(t) = r(t)\}. \tag{4.5}$$

Due to the ordering property,

$$r_1(t) \ge r_2(t) \ge \cdots \ge r_N(t) \quad \text{for } t \ge 0, \tag{4.6}$$

proved in Theorem 3.3, these time instants are well defined and satisfy

$$T_{N+1} \doteq 0 \leq T_N \leq T_{N-1} \leq \cdots \leq T_1 = T.$$

The property (3.22) that $r_i(t) = r(t)$ if $r(\cdot)$ jumps at t in fact shows that the supremum in (4.5) is attained, and therefore that an equivalent definition of T_i is

$$T_i = \min\{t \leq T: \tau_i(t) = \tau_i(T)\}. \tag{4.7}$$

We define the piecewise linearization f^* of f as follows. Let $\tau \doteq Af$. Then for $j = 1, \ldots N$, we set

$$f_j^*(t) \doteq \begin{cases} \dfrac{(t - \tau_j(T_{i+1}))f_j(\tau_j(T_i)) + (\tau_j(T_i) - t)f_j(\tau_j(T_{i+1}))}{\tau_j(T_i) - \tau_j(T_{i+1})}, \\ \qquad\qquad\qquad \text{if } t \in [\tau_j(T_{i+1}), \tau_j(T_i)), \quad i \geq j, \\ f_j(\tau_j(T_j)) + \lambda_j(t - \tau_j(T_j)), \qquad \text{if } t \in [\tau_j(T_j), \infty). \end{cases} \tag{4.8}$$

LEMMA 4.2. *Consider any $f \in \mathscr{C}_a^N$ such that $Rf(T) = 1$ for some $T > 0$ and condition* (4.4) *is satisfied with $s = T$. Let f^* be its piecewise linearization defined above. Moreover, define $\tau^* \doteq Af^*$ and $r^* \doteq Rf^*$. Then the following properties hold:*

(i) *For every $j, i \in \{1, \ldots, N+1\}$,*

$$\tau_j^*(T_i) = \tau_j(T_i), \tag{4.9}$$

and for each $i \in \{1, \ldots, N\}$ such that $T_{i+1} < T_i$, and $t \in [T_{i+1}, T_i]$, the functions

$$r_1^*(t) = \cdots = r_i^*(t) = r^*(t) \tag{4.10}$$

are equal and linear with derivative

$$\gamma_i \doteq \frac{r(T_i) - r(T_{i+1})}{T_i - T_{i+1}} > 0. \tag{4.11}$$

(ii) *Furthermore*

$$J_T(f^*) \leq J_T(f). \tag{4.12}$$

PROOF. By construction for each $j \in \{1, \ldots, N\}$

$$f_j^*(\tau_j(T_i)) = f_j(\tau_j(T_i)) \quad \text{for } i = N+1, \ldots, j, \tag{4.13}$$

and the function $f_j^*(\cdot)$ is linear in each interval $[\tau_j(T_{i+1}), \tau_j(T_i)]$ for $i = N, \ldots, j$. Given $j \in \{1, \ldots, N\}$, let the process $\tau_j^*(\cdot) \in \mathscr{I}_{+,0}^N$ be such that

$$\tau_j^*(T_i) = \tau_j(T_i) \quad \text{for } i = N+1, N, \ldots, j, \tag{4.14}$$

τ_j^* is linear in each interval $[T_{i+1}, T_i]$ for $i = N, \ldots, j$, and for $t \in [T_j, T]$,

$$\tau_j^*(t) = \tau_j(T_j) = \tau_j(T). \tag{4.15}$$

We now show that τ^* defined above satisfies conditions (3.17)–(3.25), which by Definition 3.4 automatically establishes that $\tau^* = Af^*$. Property (3.17) follows trivially from the corresponding property for τ. Let $r_i^*(t) = (t - \tau_i^*(t))/\alpha_i$ for $t \geq 0$ and $i = 1, \ldots, N$. We claim that for $t \in [T_{i+1}, T_i]$,

$$r_1^*(t) = \cdots = r_i^*(t) = r^*(t), \tag{4.16}$$

and if $j > i$ and $T_j < T_i$, then for $t \in (T_{i+1}, T_i]$,

$$r_j^*(t) < r^*(t).$$

The above two properties clearly hold for $t = T_{i+1}$ and $t = T_i$, and so by the linearity of $r_j^*(\cdot)$ it also holds for all $t \in [T_{i+1}, T_i]$. This establishes properties (3.20)–(3.25). For every $i = 1, \ldots, N$, the fact that f_j is nondecreasing and $\tau_j(T_i) \leq T_i$, shows that

$$\sum_{j=1}^N f_j^*(T_i) \geq \sum_{j=1}^N f_j^*(\tau_j(T_i)) = \sum_{j=1}^N f_j(\tau_j(T_i)) = T_i.$$

Thus the linearity of $\tau_j^*(\cdot)$ on $[T_{i+1}, T_i]$ dictates that

$$\sum_{j=1}^N f_j^*(t) \geq t \quad \text{for } t \in [T_{i+1}, T_i] \tag{4.17}$$

and

$$\sum_{j=1}^N f_j^*(\tau_j^*(t)) = t. \tag{4.18}$$

Properties (4.17) and (4.18) verify the conservation law (3.19).

Finally, we note that (4.12) follows from the definitions of f^* and J and the fact that the functions L_i, $i = 1, \ldots, N$ are convex. □

Consider $f \in \mathscr{C}_a^N$ as defined in Lemma 4.2, and let f^* be its piecewise linearization as defined above. For $k \in \{1, \ldots, N\}$ such that $T_{k+1} < T_k$, and $i \leq k$, let $d_i^*(k) = \dot{f}_i^*(\tau_i(T_k)-)$ and let γ_i be as in (4.11). Then for k such that $T_{k+1} < T_k$ the unit cost c_k of raising the maximal weighted delay within the time interval $[T_{k+1}, T_k]$ is

$$\begin{aligned}
&\frac{1}{r(T_k) - r(T_{k+1})} \sum_{i=1}^N \int_{\tau_i(T_{k+1})}^{\tau_i(T_k)} L_i(\dot{f}_i^*(s))\, ds \\
&\quad = \frac{1}{r(T_k) - r(T_{k+1})} \sum_{i=1}^k [\tau_i(T_k) - \tau_i(T_{k+1})] L_i(d_i^*(k)),
\end{aligned}$$

where we have used the fact that $\dot{f}_i^*(s) = \lambda_i$ for $s > \tau_i(T_i)$ and $L_i(\lambda_i) = 0$. Using (4.10) and the last display, we see that the unit cost c_k is given by

$$(4.19)\quad c_k \doteq \begin{cases} \dfrac{1}{\gamma_k} \sum_{i \le k} (1 - \alpha_i \gamma_k) L_i(d_i^*(k)), & \text{if } T_{k+1} < T_k, \\ \infty, & \text{otherwise (by convention).} \end{cases}$$

Note that c_k represents the unit cost of raising r in the time interval $[T_{k+1}, T_k]$, in which only the first k flows are served. If j is such that c_j is the minimum among all the unit costs, then it is intuitively believable that f^* can be replaced by another trajectory f^0, such that only the first j classes are served while the maximal weighted delay attains level 1, and the cost of f^0 does not exceed that of f^*. This is the content of Lemma 4.3, and is a consequence of the fact that the parameters γ_j, $d_j^*(k)$, and α_j that characterize the piecewise linearizations f^* satisfy certain useful relations derived below. Define $r^* \doteq Rf^*$. Then for j such that $T_{j+1} < T_j$, from definitions (4.11), (4.9) and (4.7) it follows that for $t \in (T_{j+1}, T_j)$,

$$\dot{r}_{j+1}^*(t) < \dot{r}_j^*(t) = \gamma_j$$

and

$$\dot{\tau}_{j+1}^*(t) = 0.$$

This implies that $\dot{w}_{j+1}^*(t) = 1$, or equivalently $\dot{r}_{j+1}^*(t) = 1/\alpha_{j+1}$, which when substituted to the last but one display yields

$$(4.20)\qquad 1/\alpha_{j+1} < \gamma_j,$$

which in turn means that for all $i \geq j$,

$$(4.21)\qquad 1/\alpha_{i+1} < \gamma_j.$$

Obviously for every $i = 1, \ldots, N$ and $t \geq 0$ such that the derivative $\dot{\tau}_i^*(t)$ exists,

$$(4.22)\qquad \dot{\tau}_i^*(t) \geq 0,$$

which implies that for all $i \leq j$,

$$(4.23)\qquad \frac{1}{\alpha_i} \geq \gamma_j.$$

Finally from (4.10), (4.11), and the conservation law (3.19), we see that in each interval $[T_{j+1}, T_j]$ such that $T_{j+1} < T_j$,

$$(4.24)\qquad \sum_{i \le j} (1 - \alpha_i \gamma_j) d_i^*(j) = 1.$$

We now state and prove Lemma 4.3.

LEMMA 4.3. *Given $f \in \mathscr{C}_a^N$ as defined in Lemma 4.2, let f^* be its piecewise linearization and let c_k be the associated unit costs defined in* (4.19). *Let $j = \arg\min_{k=1,\dots,N} c_k$ (where the maximum index is chosen in case of a tie), let $\gamma \doteq \gamma_j$ and $x_i \doteq d_i^*(j)$ for $i = 1,\dots,N$. Define the trajectory f^0 by*

$$f_i^0(t) \doteq \begin{cases} \lambda_i t, & \text{for } i > j, t \in [0,\infty), \\ x_i t, & \text{for } i \le j, t \in [0, T_i^0], \\ f_i^0(T_i^0) + \lambda_i(t - T_i^0), & \text{for } i \le j, t \in [T_i^0, \infty), \end{cases} \tag{4.25}$$

where $T^0 \doteq 1/\gamma$ and for $i = 1,\dots,j$, $T_i^0 \doteq (1 - \alpha_i\gamma)T^0$. Then $\tau^0 \doteq Af^0$ and $r^0 \doteq Rf^0$ satisfy

$$\tau_i^0(t) = 0 \quad \text{for } t \in [0, T^0] \text{ and } i > j, \tag{4.26}$$

$$r_1^0(t) = \cdots = r_j^0(t) = r^0(t) = \gamma t \quad \text{for } t \in [0, T^0], \tag{4.27}$$

$$r^0(T^0) = 1 \quad \text{and} \quad J_{T^0}(f^0) \le J_T(f^*) \le J_T(f). \tag{4.28}$$

PROOF. By Definition 3.4 it suffices to show that any τ^0 satisfying properties (4.26) and (4.27) satisfies conditions (3.17)–(3.25). From (4.21), (4.24) and (4.23) it follows that

$$\sum_{i=1}^{j} x_i > 1, \tag{4.29}$$

$$\frac{1}{\alpha_{j+1}} < \gamma \doteq \frac{\sum_{i\le j} x_i - 1}{\sum_{i \le j}\alpha_i x_i} \le \frac{1}{\alpha_j}, \tag{4.30}$$

with $1/\alpha_{N+1} = 0$ convention. Now if τ^0 satisfies (4.26) and (4.27), then due to the above relations we obtain for $t \in (0, T^0)$,

$$\dot r_i^0(t) = 1/\alpha_i < \gamma = \dot r^0(t) \quad \text{for } i > j \tag{4.31}$$

and

$$\dot r_i^0(t) = \gamma = \dot r^0(t) \quad \text{and} \quad \dot\tau_i^0(t) \ge 0 \quad \text{for } i > j, \tag{4.32}$$

which establishes (3.17) and (3.20)–(3.25). Moreover the conservation law (3.19) holds since

$$\sum_{i=1}^{N} \dot f_i^0(\dot\tau_i^0(t)) = \sum_{i=1}^{j} x_i \dot\tau_i^0(t) = \sum_{i=1}^{j} x_i(1 - \alpha_i\gamma) = 1, \tag{4.33}$$

where the last equality follows from (4.30). This establishes that $\tau^0 = Af^0$.

Finally, to prove (4.28) we write

$$\begin{aligned} J_{T^0}(f^0) = c_j &\le c_j \sum_{1\le i\le N} 0 \wedge (r^*(T_i) - r^*(T_{i+1})) \\ &\le \sum_{1 \le i \le N} 0 \wedge c_i(r^*(T_i) - r^*(T_{i+1})) \\ &\le J_T(f^*) \le J_T(f). \end{aligned}$$

□

In summary, together the last three lemmas in this section show that given any $f \in \mathscr{C}_a^N$ for which the maximal weighted delay exceeds one at some finite time, we can find a "simple element" of no greater cost for which the maximal weighted delay attains the level 1. Moreover, Lemma 4.3 also shows that the backlog $\tau^0 = Af^0$ associated with any such "simple element" f^0 has the simple structure dictated by (4.26) and (4.27). This motivates the following definition.

DEFINITION. *A simple element* f^0 is a function defined by (4.25) for some parameters $(j, x\colon j \in \{1, \dots, N\}, x \in \mathbb{R}_+^j)$, which satisfy the constraints (4.29) and (4.30).

We conclude this section with Theorem 4.4, which shows that the variational problem (4.1) is minimized by a simple element.

THEOREM 4.4. *The infimum* J_* *of the variational problem defined in* (4.1) *is the solution to the following finite-dimensional optimization problem:*

$$J_* = \min_{j; x_1, \dots, x_j} \frac{1}{\gamma} \sum_{i=1}^{j} (1 - \alpha_i \gamma) L_i(x_i), \tag{4.34}$$

subject to

$$j \in \{1, \dots, N\}, \qquad x_i > 0, \qquad \sum_{i=1}^{j} x_i > 1$$

and

$$\frac{1}{\alpha_{j+1}} < \gamma = \frac{\sum_{i=1}^{j} x_i - 1}{\sum_{i=1}^{j} \alpha_i x_i} \le \frac{1}{\alpha_j},$$

where $\alpha_{N+1} \doteq \infty$ *by convention.*

Moreover, the infimum in (4.1) *is attained on simple elements* f^0 *associated with parameters* $(j, x_1, \dots, x_j)$ *that solve* (4.34).

REMARK. The structure of a simple element f^0 is illustrated by Figure 1. (The picture is for the case $N = 3$ and $j = 2$.) For a subset of flows of the form $\{1, \dots, j\}$ for some $j, 1 \le j \le N$, the input rate has a value $x_i > \lambda_i$ in the time interval $[0, T - \alpha_i]$ where $T = 1/\gamma$; and after time $T - \alpha_i$ the rate switches to the mean λ_i. The remaining flows $\{j + 1, \dots, N\}$ (if any) always have the mean input rate λ_i. The flows from the latter subset are not served at all.

PROOF OF THEOREM 4.4. Without loss of generality assume that all α_j are distinct. We proved in Lemma 4.3 that any element $f \in \mathscr{C}_a^N$ such that $r(T) \doteq Rf(T) \ge 1$ for some $T < \infty$ can be replaced by a *simple* element f^0 such that the cost of $r^0 \doteq Rf^0$ reaching 1 does not exceed the cost of $r = Rf$ to reach 1. Moreover there is a one-to-one correspondence between simple elements and parameters $(j, x\colon j \in \{1, \dots, N\}, x \in \mathbb{R}_+^j)$ satisfying (4.29) and (4.30). Furthermore, the unit cost of increase of r for a simple element is given by

$$c = c(x_1, \dots, x_j; j) = \frac{1}{\gamma} \sum_{i=1}^{j} (1 - \alpha_i \gamma) L_i(x_i). \tag{4.35}$$

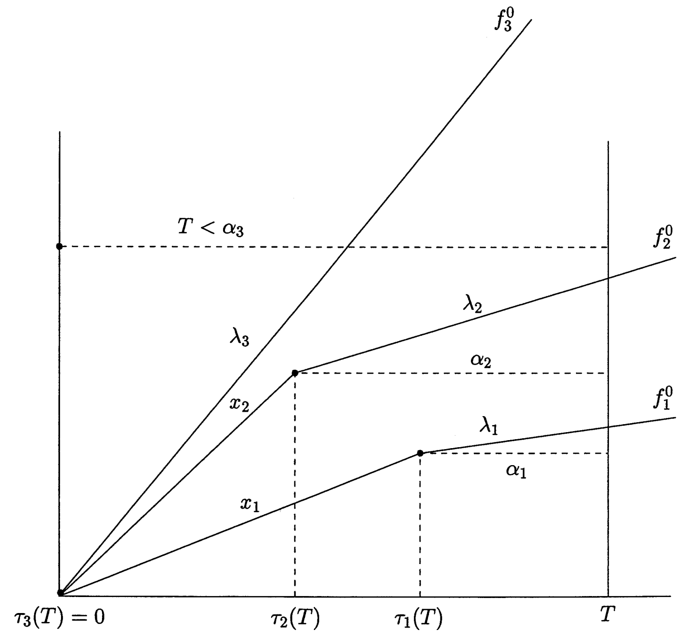

FIG. 1. *The structure of the simple element* $f^0(N = 3, j = 2)$.

Thus, to show that the infimum of the variational problem (4.1) is attained on a simple element it suffices to show that there exists $j \in \{1, \dots, N\}$ and $x \in \mathbb{R}_+^j$ satisfying (4.29) and (4.30) which *minimize* the cost c defined in (4.35).

We claim that for every vector $x = (x_1, \dots, x_N)$ such that

$$x_i \geq 0 \quad \text{for } i = 1, \dots N \tag{4.36}$$

and

$$\sum_{i=1}^{N} x_i > 1, \tag{4.37}$$

there exist a unique $j = j(x)$ and corresponding $\gamma = \gamma(x)$ such that (4.29) and (4.30) are satisfied. To prove the claim, we rewrite (4.30) as

$$\frac{\alpha_j}{\alpha_{j+1}} < z \doteq \frac{\sum_{i \leq j} x_i - 1}{\sum_{i \leq j} \frac{\alpha_i}{\alpha_j} x_i} \leq 1,$$

and notice that $z = z(j)$ satisfies

$$z(j) < 1 \quad \text{for } j = \min\left\{k \colon \sum_{1}^{k} x_i > 1\right\}, \tag{a}$$

$$z(j) \leq \frac{\alpha_j}{\alpha_{j+1}} \quad \text{implies} \quad z(j+1) \leq 1, \tag{b}$$

(c) $$z(j) > \frac{\alpha_j}{\alpha_{j+1}} \quad \text{implies} \quad z(j+1) > 1.$$

Now for each x described by (4.36) and (4.37), we define

$$c(x) \doteq c(x_1, \ldots, x_j; j) \tag{4.38}$$

with $j = j(x)$ and the RHS defined in (4.35). Let us denote

$$c^* \doteq \inf_{x \in \mathbb{R}_+^N:\ \sum_{i=1}^N x_i > 1} c(x), \tag{4.39}$$

and consider a sequence $\{x^{(l)}, l = 1, 2, \ldots,\}$ such that

$$\lim_{l \to \infty} c(x^{(l)}) = c^*. \tag{4.40}$$

It is easy to see from (4.35) and (4.37) that $\gamma(x^{(l)})$ remains uniformly bounded away from both 0 and infinity, that is, there exist $\varepsilon_1, \varepsilon_2$ such that for $l = 1, \ldots,$

$$0 < \varepsilon_1 < \gamma(x^{(l)}) < \varepsilon_2 < \infty. \tag{4.41}$$

Let us choose a subsequence (we will keep the same notation $\{x^{(l)}\}$ for it) such that

(1) $$\gamma(x^{(l)}) \to \gamma^* > 0,$$

(2) $$j(x^{(l)}) = \bar{j} \quad \text{is fixed.}$$

Then it is easy to see that for any $i \le \bar{j}$, $x_i^{(l)}$ must stay bounded away from infinity. Let us choose a further subsequence which we continue to denote by $x^{(l)}$ such that

$$x_i^{(l)} \to x_i^* \quad \text{for } i = 1, \ldots, \bar{j}. \tag{4.42}$$

We see that conditions (4.29) and (4.30) are satisfied for $x_i^*, i = 1, \ldots, j^*$, where

$$j^* = \begin{cases} \bar{j}, & \text{if } \gamma^* > 1/\alpha_{\bar{j}+1}, \\ \bar{j} + 1, & \text{if } \gamma^* = 1/\alpha_{\bar{j}+1}, \end{cases}$$

$$x_{j^*}^* = x_{\bar{j}+1}^* = \lambda_{j^*} \quad \text{if } j^* = \bar{j} + 1,$$

and also

$$c(x_1^*, \ldots, x_{j^*}^*; j^*) = c^*.$$

The proof is complete. □

5. Optimality of LWDF for a system with fluid inputs. In this section we show that in a system with *fluid* inputs, the LWDF discipline is optimal in the class of work-conserving disciplines. Namely, LWDF maximizes the (minimal) cost of $r(t)$ reaching level 1. The result is very simple, and although it is not used directly to prove the lower bound (2.6) in our main result Theorem 2.2, it provides the key intuition without invoking all the technical machinery required to prove it in full generality. *New definitions introduced in this section are confined to this section only, and not used in the rest of the paper.*

Consider the class of operators $\{A^G\}$ (with G being a queueing discipline defined for a "fluid" input system only) such that each operator A^G maps a fluid input $f \in \mathscr{C}^N$ into a virtual backlog path $\tau^G = A^G f \in \mathscr{I}^N_{+,0}$, that satisfies for $t \geq 0$,

$$\tau_i^G(t) \leq t \quad \text{for } i = 1, \dots, N,$$

and the "work conservation" condition

$$\sum_i f_i\big(\tau_i^G(t)\big) = t + \left[\min_{0\leq\xi\leq t}\left(\sum_i f_i(\xi) - \xi\right)\right] \wedge 0.$$

Let us denote

$$\overline{J}(A^G) \doteq \inf_{s>0,\, f\in\mathscr{C}^N:\ R^G f(s)\geq 1} \widehat{J}_s(f), \tag{5.1}$$

where $\widehat{J}_s(f)$ is an *arbitrary* nonnegative "cost function," and

$$R^G f(t) \doteq \max_i \frac{t - \tau_i^G(t)}{\alpha_i}.$$

LEMMA 5.1. *Suppose there exists an operator A^0 from the class defined above, and $T^0 > 0$ such that the minimum cost $\overline{J}(A^0)$ is attained on the path $f^0 \in \mathscr{C}^N$ such that $R^0 f^0(T^0) = 1$,*

$$\sum_i f_i^0(t) > t \quad \textit{for } t \in [0, T^0],$$

all functions $f_i^0(\cdot)$ are strictly increasing in $[0, T^0]$, and for some subset $K \subseteq \{1, 2, \dots, N\}$,

$$T^0 - \tau_i^0(T^0) = \alpha_i \quad \textit{for } i \in K,$$
$$\tau_i^0(T^0) = 0 \quad \textit{for } i \notin K.$$

Then for any operator A^G in the class,

$$\overline{J}(A^G) \leq \overline{J}(A^0).$$

PROOF. Consider the minimal cost element f^0, and let $\tau^0 \doteq A^0 f^0$ and $r^0 \doteq R^0 f^0$. Consider any other operator A^G in the class, and let $\tau^G \doteq A^G f^0$ and $r^G \doteq R^G f^0$. We now show that $r^G(T^0) \geq 1$ irrespective of the particular discipline G chosen. By assumption, each function f_i^0 is continuous, strictly increasing in $[0, T^0]$ and satisfies the inequality

$$\sum_i f_i^0(t) \geq t \quad \text{for } t \in [0, T^0].$$

Then by the work conservation condition,

$$\sum_{i \in K} f_i^0\big(\tau_i^G(T^0)\big) \leq \sum_{i=1}^N f_i^0\big(\tau_i^G(T^0)\big) = T^0 = \sum_{i=1}^N f_i^0\big(\tau_i^0(T^0)\big) = \sum_{i \in K} f_i^0\big(\tau_i^0(T^0)\big),$$

where the last equality follows from the definition of f^0. Since each $f_i^0(\cdot)$ is nonnegative and strictly increasing on $[0, T_0]$, we conclude that for at least one $i \in K$ it must be that

$$\tau_i^G(T^0) \leq \tau_i^0(T^0).$$

Therefore,

$$r^G(T^0) \geq r_i^G(T^0) \geq r_i^0(T^0) = r^0(T^0) = 1.$$

In other words, for any operator A^G the minimal cost of $r^G(\cdot)$ reaching level 1 is *at most* $\overline{J}(A^0)$. □

If we replace the cost function $\widehat{J}_s$ by the cost function J_s defined in (4.2), then it is clear from Theorem 4.4 and Lemma 4.3 that the conditions of Lemma 5.1 are satisfied with A° being the operator A associated with the LWDF discipline and $K = \{1, 2, \dots, j\}$ for some $j \in \{1, \dots, N\}$. If the class of queueing disciplines $\mathscr{G}$ were restricted to disciplines for which the same extension procedure carried out in Theorem 3.3 can be used to obtained a unique and well-defined operator A^G on the fluid inputs (as is the case for LWDF, FIFO, LIFO, processor sharing, generalized processor sharing, priority, and most other "non-pathological" disciplines), then we could use Lemma 5.1 in the proof of Theorem 2.2 (ii) directly. However, our class $\mathscr{G}$, and therefore the result of Theorem 2.2, is more general. As a consequence the proof of the lower bound in Theorem 2.2 is more involved, even though its key idea is the simple argument used in the proof of Lemma 5.1.

6. Rigorous statement and Proof of Theorem 2.2. In this section we make precise the statement of the main result given in Theorem 2.2 and then prove it. In Section 6.1 we introduce the stationary state process of the system, which is independent of the particular work-conserving discipline used. This, along with the operator $\widehat{R}^G$ defined earlier, is then used in Section 6.2 to rigorously define the stationary maximal weighted delay $\hat{\mathbf{r}}^G(0)$ associated with any queueing discipline $G \in \mathscr{G}$ and stochastic input flows $\mathbf{f}$ that satisfy Assumption 2.1. Since we allow for arbitrary disciplines $G \in \mathscr{G}$, which are not necessarily even scalable, $\hat{\mathbf{r}}^G(0)$ need not always be measurable. This necessitates the use of the inner measure in the statement of the lower bound in Theorem 2.2. For the LWDF discipline however, the stationary process $\hat{\mathbf{r}}(0)$

is measurable, and in Section 6.3 we show that the asymptotic rate of decay of its tails is indeed characterized by the variational problem considered in Section 4. The main result Theorem 2.2 follows directly from Theorems 6.8 and 4.4. We derive some important corollaries of Theorem 2.2 in Section 6.4.

6.1. *The stationary state process.* In this section we introduce an operator H that associates to any input path $f \in \mathscr{I}^N$ the state of the queueing system. The state is defined in such a way that it depends only on the input path f and is independent of the particular work conserving queueing discipline $G \in \mathscr{G}$ used. Roughly speaking, the state of the system at time t is the history of the input flows from the beginning of the current busy period. Using measurability properties of the operator H and the fact that the stochastic input flows satisfy Assumption 2.1, and therefore have stationary increments, we then use a Loynes-type construction to define the stationary state process $\boldsymbol{\theta} \doteq H\mathbf{f}$. Finally, in Lemma 6.6 we also formulate a "kind of large deviation principle" for a sequence of scaled state distributions.

We introduce the space of system states,

$$\Psi \doteq \{\psi = (b, g)\colon b \in \mathbb{R}_+, g \in \mathscr{S}_+^N\}.$$

We endow Ψ with the natural topology generated by u.o.c. convergence, and with the σ-algebra generated by cylinder subsets. Let Θ be the space of RCLL functions on $(-\infty, \infty)$ taking values in Ψ. We define the deterministic mapping H: $\mathscr{S}^N \to \Theta$ that takes input flow paths $f \in \mathscr{S}^N$ to the system state sample paths as follows. Given any $f \in \mathscr{S}^N$, for each $t \in (-\infty, \infty)$, $Hf(t) \doteq (b(t), g) \in \Psi$, where

$$b(t) \doteq t - \sup\left\{s \le t\colon \left(\sum_{i=1}^N f_i(s-) - s\right) = z(t)\right\}, \tag{6.1}$$

$$z(t) \doteq \begin{cases} \inf\limits_{s\le t}\left(\sum\limits_{i=1}^N f_i(s) - s\right), & \text{if } \inf\limits_{s\le t}\left(\sum\limits_{i=1}^N f_i(s) - s\right) > -\infty, \\ \sum\limits_{i=1}^N f_i(t-) - t, & \text{otherwise} \end{cases} \tag{6.2}$$

and

$$g_i(s) \doteq \begin{cases} f_i(s + t - b(t)) - f_i((t - b(t))-), & 0 \le s \le b(t), \\ f_i(t) - f_i((t - b(t))-), & s > b(t). \end{cases} \tag{6.3}$$

When f is the realization of an input flow to a single server queueing system (with service rate 1), $b(t)$ represents the time elapsed from the start of the busy period in progress at time t, and g captures the history of the input flows from the start of the current busy period. Note that b and g have this intuitive interpretation only when $\inf_{s\le t}(\sum_{i=1}^N f_i(s) - s)$ is finite for all t. However, since (as we show in Lemma 6.2) this is true almost surely for a sample path of the input flows $\mathbf{f}$ satisfying Assumption 2.1, this does not pose a problem in the definition of the process $H\mathbf{f}$ given below.

REMARK 6.1. Note that the expressions (6.1), (6.2) and (6.3) are in fact well defined for $f \in \mathscr{I}^N$, and hence can be used to define a more general mapping,

$$H(t)\colon \mathscr{I}^N \to \mathbb{R}_+ \times \mathscr{I}_+^N,$$

$H(t) = (b(t), g)$, where $b(t)$ and g have the same interpretation as above. Thus the results of this subsection are valid for a system with more general input processes that take values in $\mathscr{I}^N$ rather than just $\mathscr{S}^N$. This more general interpretation of the mapping H is used in Appendix B to prove its measurability.

In the following lemma we show that the pathwise mapping H defined above makes $\boldsymbol{\theta} = H\mathbf{f}$ a well-defined stationary process.

LEMMA 6.2. *Let H be the operator introduced above and let* $\mathbf{f}$ *satisfy Assumption* 2.1. *Then the following properties hold:*

(i) *The mapping H is measurable.*
(ii) $\boldsymbol{\theta} \doteq H\mathbf{f}$ *is a stationary RCLL process taking values in* Ψ.

The proof is relegated to Appendix B.

For any $t \geq 0$ we define the operator $H_+(t)\colon \mathscr{S}_{+,0}^N \to \Psi$ in a natural way so that is maps the input flows $f \in \mathscr{S}_{+,0}^N$, *after time* 0 into $H_+(t)f \in \Psi$, the state of the system at time t *given that the system is empty at time* 0. In other words, this mapping is defined by the same expressions (6.1), (6.2) and (6.3), with the additional constraints $t \geq 0$ and $s \geq 0$. Note that just as for the mapping $H(t)$, the definition of $H_+(t)$ can also be generalized to yield a mapping $H_+(t)\colon \mathscr{I}_{+,0}^N \to R_+ \times \mathscr{I}_+^N$.

LEMMA 6.3. *The mapping $H_+(t)$ is measurable for any $t \geq 0$.*

PROOF. Measurability of the operator $H_+(t)$ follows from measurability of the operator H. The details are provided in Appendix B.

We now introduce some definitions required to formulate a kind of large deviation principle for the sequence of scaled distributions $\{\Gamma^n \boldsymbol{\theta}(0), n = 1, 2, \ldots\}$, where $\boldsymbol{\theta}(0) = H\mathbf{f}(0)$ and Γ^n is the scaling operator defined in (1.5).

DEFINITION 6.4. Given any $f \in \mathscr{C}_a^N$ and $r \geq 0$, we define its r-modification f^r to be such that for $i = 1, \ldots, N$,

$$f_i^r(t) \doteq \begin{cases} \lambda_i t, & \text{for } 0 \leq t \leq r, \\ \lambda_i r + f_i(t-r), & \text{for } t > r. \end{cases} \tag{6.4}$$

DEFINITION 6.5. Given any subset $B \subset \Psi$ and any $s > 0$, we define the set

$$\Phi_s(B) \doteq \{f \in \mathscr{C}_a^N\colon \ \begin{array}{l}\text{for every } r > 0, \text{ there exists } \delta > 0 \text{ such that} \\ \text{if } h \in U_\delta^{r+s}, \text{ then } [H + (r+s)]h \in B\}\end{array} \tag{6.5}$$

where

$$U_\delta^{(r+s)} \doteq \{h \in \mathscr{S}^N_{+,0}\colon \|h - f^r\|_{r+s} < \delta\}, \tag{6.6}$$

and f^r is as defined in (6.4).

Note that in the above definition $B \subset \Psi$ need not be measurable.

LEMMA 6.6. *Let J_s be as defined in* (4.2).

(i) *For any measurable subset $B \subset \Psi$,*

$$\limsup_{n\to\infty} \frac{1}{n} \log P(\Gamma^n \boldsymbol{\theta}(0) \in B) \le - \inf_{s>0,\, f\in \overline{H_+^{-1}(s)B}} J_s(f). \tag{6.7}$$

(ii) *For any* (*not necessarily measurable*) *subset $B \subset \Psi$,*

$$\liminf_{n\to\infty} \frac{1}{n} \log P_*(\Gamma^n \boldsymbol{\theta}(0) \in B) \ge - \inf_{s>0,\, f\in \Phi_s(B)} J_s(f), \tag{6.8}$$

where P_* denote inner measure with respect to the probability P.

The proof is given in Appendix A.

REMARK 6.7. We call Lemma 6.6 a "kind of LDP," because formally it is not a LDP. The upper bound is (almost) the usual large deviation upper bound. However, the lower bound is more general than usual large deviations lower bounds, in particular due to the use of the inner measure. As mentioned earlier, the reason we need this generality is because we apply this lower bound in Section 6.3.2 to the case when the discipline $G \in \mathscr{G}$ is arbitrary, and therefore the associated subset B may be non-measurable.

6.2. *Stationary Maximal Weighted Delay* $\hat{\mathbf{r}}^G(0)$. We are now in a position to define precisely the stationary maximal weighted delay $\hat{\mathbf{r}}^G(0)$ that was used in the statement of our main result, Theorem 2.2.

Consider a discipline $G \in \mathscr{G}$ and the associated operators $\widehat{A}^G$ and $\widehat{R}^G$. The stationary maximal weighted delay $\hat{\mathbf{r}}^G = (\hat{\mathbf{r}}^G(t), -\infty < t < \infty)$ is defined by

$$\hat{\mathbf{r}}^G(t) \doteq [\widehat{R}^G \mathbf{g}](\mathbf{b}(t)),$$

where $(\mathbf{b}(t), \mathbf{g}) = \boldsymbol{\theta}(t) = [H\mathbf{f}](t) \in \Psi$.

Note that we do not refer to $\hat{\mathbf{r}}^G$ as a stationary "process" because although $\hat{\mathbf{r}}^G$ is a well-defined function on the probability space (and with probability 1 defines the process that corresponds intuitively with our notion of the stationary maximal weighted delay), it *need not be measurable.* If the function is measurable, the $\hat{\mathbf{r}}^G$ is indeed a stationary random process. As shown in Appendix B, the operator $\widehat{A}$ (and therefore $\widehat{R}$) associated with the LWDF discipline is measurable, $\hat{r}$ is a measurable function and thus the probability $P\{\hat{\mathbf{r}}(0)/n > 1\}$ is well defined.

The following connection between the operators $\widehat{R}^G$ and H_+ is quite straightforward and will be useful later in the paper. For any $f \in \mathscr{S}^N_{+,0}$ and any $t \geq 0$,

$$[\widehat{R}^G f](t) = [\widehat{R}^G g](b(t)), \tag{6.9}$$

where $(b(t), g) = H_+(t)f$.

We now introduce another pair of "unfinished work" operators V and V_+ which are, like H and H_+, discipline independent. For $f \in \mathscr{I}^N$ are define

$$Vf(t) \doteq \sum_{i=1}^N f_i(t) - t - \inf_{-\infty<u\leq t}\left[\sum_{i=1}^N f_i(u) - u\right], \tag{6.10}$$

to be the "unfinished work" in the system. Similarly, we define the operator $V_+ \colon \mathscr{I}^N_+ \to \mathscr{D}_+$ by

$$V_+ f(t) \doteq \sum_{i=1}^N f_i(t) - t - \inf_{0\leq u\leq t}\left[\sum_{i=1}^N f_i(u) - u\right] \wedge 0. \tag{6.11}$$

Let $\mathbf{v} \doteq V(\mathbf{f})$, where $\mathbf{f}$ is an input process satisfying Assumption 2.1. Note that $\mathbf{v}(t)$ depends only on the system state $\boldsymbol{\theta}(t) = [H\mathbf{f}](t)$. Therefore,

$$[V\mathbf{f}](t) = [V_+\mathbf{g}](\mathbf{b}(t)),$$

where for each $t \in (-\infty, \infty)$, $(\mathbf{b}(t), \mathbf{g}) = \boldsymbol{\theta}(t)$. Thus $\mathbf{v} = (\mathbf{v}(t),\ t \in \mathbb{R})$ is a stationary process.

6.3. *Proof of main result.* In this section we state and prove Theorem 6.8, which establishes the LDP for and proves optimality of the LWDF discipline. As shown in the theorem, the variational problem studied in Section 4 characterizes the exponential rate of decay of the tails of the stationary maximal weighted delay $\hat{\mathbf{r}}(0)$. In Section 6.3.1 and 6.3.2 we prove the upper and lower bounds, respectively.

THEOREM 6.8. *Let*

$$J_* \doteq \inf_{s>0,\, f\in\mathscr{C}^N_a:\ Rf(s)\geq 1} J_s(f). \tag{6.12}$$

Suppose $\hat{\mathbf{r}}(0)$ *and* $\hat{\mathbf{r}}^G(0)$ *is the stationary maximal weighted delay associated with the LWDF scheduling discipline, and with a discipline* $G \in \mathscr{G}$ *respectively. Then the following three properties hold:*

(i)

$$\limsup_{n\to\infty} \frac{1}{n} \log P\left(\frac{1}{n}\hat{\mathbf{r}}(0) > 1\right) \leq -J_*. \tag{6.13}$$

(ii) *For any* $G \in \mathscr{G}$,

$$\liminf_{n\to\infty} \frac{1}{n} \log P_*\left(\frac{1}{n}\hat{\mathbf{r}}^G(0) > 1\right) \geq -J_*. \tag{6.14}$$

(iii) *Moreover,* J_* *solves the finite-dimensional optimization problem* (4.34).

6.3.1. *Proof of upper bound in Theorem 6.8.* The upper bound follows quite easily by substituting the set

$$B \doteq \{\psi = (b, g) \in \Psi : [\widehat{R}(b)](g) > 1\}$$

into the upper bound in Lemma 6.6, where $\widehat{R}$ is the maximal weighted delay operator associated with the LWDF queueing discipline. Note that the set B is measurable due to Lemma 10.4 Furthermore,

$$\overline{H_+^{-1}(s)B} = \overline{\{f \in \mathscr{S}_{+,0}^N : [\widehat{R}f](s) > 1\}} \subseteq \overline{\{f \in \mathscr{S}_{+,0}^N : [Rf](s) > 1\}},$$

where the second inclusion is a consequence of the fact that for all $s \geq 0$, $Rf(s) \geq \widehat{R}f(s)$, which follows from the definition of R given in (3.1), (3.2) and (3.5). Using properties of the operator A defined on $\mathscr{S}_+^N$ and extended to $\mathscr{C}^N$ (see Lemma 3.2 and Theorem 3.3), which imply corresponding properties for the operator R, it can be easily verified that for any $s \geq 0$,

$$\overline{\{f \in \mathscr{S}_{+,0}^N : [Rf](s) > 1\}} \cap \mathscr{C}_a^N \subseteq \{f \in \mathscr{C}_a^N : \lim_{t \uparrow s} [Rf](t) \geq 1\}.$$

We denote the set on the right-hand side by $\mathscr{M}(s)$. We now show that for any $s > 0$ the infimum of $J_s(\cdot)$ over trajectories f in $\mathscr{M}(s)$ is equal to its infimum over a smaller set $\mathscr{M}_1(s)$ defined below. Consider an element $f \in \mathscr{M}(s)$ such that $Rf(s-) \geq 1$, but $Rf(s) < 1$. Then we claim that f can be replaced by another element $\tilde{f} \in \mathscr{M}(s)$ such that the function $\tilde{r}(\cdot) \doteq R[\tilde{f}](\cdot)$ is continuous at s; that is,

$$\tilde{r}(s) = \tilde{r}(s-) \geq 1 \tag{6.15}$$

and

$$J_s \tilde{f} \leq J_s(f). \tag{6.16}$$

As shown below, this follows from the properties of the operator R, which are in turn implied by the properties of A. From (3.22) it follows that $r(\cdot)$ can only jump down. In addition, the definition of $\hat{r}$ implies that if $r(\cdot)$ jumps at that time s, then there exist i and $\varepsilon_i > 0$ such that $f_i(\cdot)$ does not increase (and is hence constant) in the time interval $[\tau_i(s-), \tau_i(s-) + \varepsilon_i]$. Let $\mathscr{I}$ be the subset of $\{1, \dots, N\}$ for which the above property holds, and let ε be the minimum of ε_i, $i \in \mathscr{I}$. Now we define a new function $\tilde{f}$ as follows. For $i \in \mathscr{I}$,

$$\tilde{f}_i(t) \doteq \begin{cases} f_i(t), & \text{for } t \in [0, \tau_i(s-)], \\ f_i(\tau_i(s-)) + \lambda_i(t - \tau_i(s-)), & \text{for } t \in (\tau_i(s-), \tau_i(s-) + \varepsilon], \\ f_i(t) + \lambda_i \varepsilon, & \text{for } t \in (\tau_i(s-) + \varepsilon, \infty) \end{cases}$$

and $\tilde{f}_i(\cdot) \doteq f_i(\cdot)$ for $i \notin \mathscr{I}$. Then it is easy to check that $\tilde{f}$ satisfies properties (6.15) and (6.16).

Now define

$$\mathscr{M}_1(s) \doteq \{f \in \mathscr{C}_a^N : [Rf](s) \geq 1\} \subseteq \mathscr{M}(s),$$

and note that the construction above implies that

$$\inf_{s>0} \inf_{f \in \mathscr{M}(s)} J_s(f) = \inf_{s>0} \inf_{f \in \mathscr{M}_1(s)} J_s(f) = J_*.$$

The last display, along with the large deviation upper bound (6.7) for the state process, yields

$$\begin{aligned}
\limsup_{n\to\infty} \frac{1}{n} \log P\left(\frac{1}{n}\hat{\mathbf{r}}(0) > 1\right) &= \limsup_{n\to\infty} \frac{1}{n} \log P(\Gamma^n \boldsymbol{\theta}(0) \in B) \\
&\le - \inf_{s>0,\, f \in \overline{H_+^{-1}(s)B}} J_s(f) \\
&\le - \inf_{s>0,\, f \in \mathscr{M}(s)} J_s(f) \qquad (6.17) \\
&= - \inf_{s>0,\, f \in \mathscr{M}_1(s)} J_s(f) \\
&= -J_*.
\end{aligned}$$

This establishes the upper bound. Note that in Section 4 we already characterized J_* and found the optimal path f^0 on which the minimum cost J_* is attained. Thus part (iii) of Theorem 2.2 follows from Theorem 4.4.

6.3.2. *Proof of lower bound in Theorem 6.8.* Consider an arbitrary work-conserving discipline $G \in \mathscr{G}$. Consider the set

$$B \doteq \{\psi = (b, g) \in \Psi\text{: for every } c > 0, [\Gamma^c \widehat{R}^G \Gamma^{1/c} g](b) > 1\}. \qquad (6.18)$$

Note that this set may or may not be measurable depending on the particular discipline G chosen. Given $G \in \mathscr{G}$, if the operator $\widehat{A}^G$ (and consequently $\widehat{R}^G$) is *scalable* [see definition (1.6)], then the above definition reduces to the simpler form

$$B = \{\psi = (b, g) \in \Psi\text{: } [\widehat{R}^G g](b) > 1\}.$$

The scaling property holds for disciplines like LWDF, FIFO, LIFO, priority, GPS and most other conventional disciplines. However, since we want to consider any arbitrary discipline $G \in \mathscr{G}$ which may not be scalable, we use the more complicated expression (6.18).

Let f^0 be an optimal path for the variational problem (6.12). As described in Section 4, this path is a simple element characterized by parameters j, x: $j \in \{1, \dots, N\}$, $x \in \mathbb{R}_+^j$), subject to the constraints (4.29) and (4.30). As in Section 4 here too we denote $T^0 = 1/\gamma$, where γ is defined in terms of x by (4.30). Fix $c > 1$ arbitrarily close to 1, and define $f^* \doteq \Gamma^{1/c} f^0 \in \mathscr{C}_a^N$. Then

$$f_i^*(t) = \begin{cases} \lambda_i(t), & \text{for } i > j, t \in [0, \infty), \\ x_i t, & \text{for } i \le j, t \in [0, cT_i^0], \\ x_i c T_i^0 + \lambda_i(t - cT_i^0), & \text{for } i \le j, t \in [cT_i^0, \infty), \end{cases}$$

where for $i \le j$,

$$T_i^0 \doteq (1 - \alpha_i \gamma) T^0.$$

Moreover, note that by Lemma 4.3 for the input flow f^*, $\tau^* \doteq Af^*$ satisfies for $t \in [0, cT^0]$,

$$\tau_i^*(t) = \begin{cases} t(1-\gamma\alpha_i), & \text{for } i \le j, \\ 0, & \text{for } i > j, \end{cases}$$

or alternatively that the delay satisfies

$$w_i^*(t) = \begin{cases} \gamma\alpha_i t, & \text{for } i \le j, \\ t, & \text{for } i > j. \end{cases}$$

We now claim that

$$f^* \in \Phi_{cT^0}(B), \tag{6.19}$$

where Φ_s is defined in (6.5). Recall the queueing discipline independent mapping V_+ defined in (6.11), which maps input flows after time 0 to the unfinished work, given the system is empty at time $0-$. It is well known that for $f \in \mathscr{C}^N$, $V_+ f$ is a continuous function of time t, and if $f^{(n)} \to f$, $f^{(n)} \in \mathscr{I}_+^N$, then $V_+ f^{(n)} \to V_+ f$ (where convergence is u.o.c.). For f^* defined above, let $v^* \doteq V_+ f^*$. Then we see that

$$v^*(cT_0) = \left[\sum_{i\le j} \lambda_i \gamma\alpha_i + \sum_{i>j} \lambda_i\right] cT^0 = \sum_{i \le j} \lambda_i \alpha_i c + \sum_{i>j} \lambda_i c T^0. \tag{6.20}$$

Now consider a sequence $\{f^{(n)}\colon f^{(n)} \in \mathscr{S}_{+,0}^N,\ n = 1, \ldots, \}$ such that $f^{(n)} \to f^*$, and for $G \in \mathscr{G}$ let $\{\hat{w}^{G,(n)}, n = 1, \ldots, \}$ be the sequence of delays determined by the corresponding sequence of backlog paths $\widehat{A}^G f^{(n)}$. Then it follows from (6.20), properties of the mapping V_+ and the fact that f^* is continuous and strictly increasing, that for every $G \in \mathscr{G}$ there exists at least one flow $i \in \{1, \ldots, j\}$ such that

$$\limsup_{n\to\infty} \hat{w}_i^{G,(n)}(cT_0) \ge c\alpha_i,$$

which implies that

$$\limsup_{n\to\infty} \hat{r}^{G,(n)}(cT_0) \ge c, \tag{6.21}$$

where $\hat{r}^{G,(n)} = \widehat{R}^G f^{(n)}$. Let $s = cT^0$. Then the above display implies that for $r = 0$, there exists $\delta > 0$ such that if

$$h \in U_\delta^{r+s} \doteq \{h \in \mathscr{S}_{+,0}^N\colon \|h - [f^*]^r\|_{r+s} < \delta\},$$

then

$$H_+(r+s)h \in B.$$

The generalization of the above property for $r \ge 0$ is quite straightforward. This proves the claim (6.19).

We now apply the lower bound in Lemma 6.6 to infer that for any $G \in \mathscr{G}$,

$$\begin{aligned}\liminf_{n\to\infty} \frac{1}{n} \log P_*\left(\frac{1}{n}\hat{\mathbf{r}}^G(0) > 1\right) &= \liminf_{n\to\infty} \frac{1}{n} \log P_*(\Gamma^n[\boldsymbol{\theta}(0)] \in B) \\ &\geq - \inf_{s>0,\, f\in\Phi_s(B)} J_s(f) \\ &\geq -J_{cT^0}(f^*) \\ &= -cJ_*.\end{aligned}$$

Since $c > 1$ can be chosen arbitrarily close to 1 we obtain part (ii) of Theorem 6.8. □

6.4. *Corollaries of the main theorem.* The following three theorems are essentially just corollaries (or by-products of the proof) of Theorem 6.8. However, we formulate them as separate theorems because of the importance of their results.

THEOREM 6.9. *Let $\hat{\mathbf{w}}_m$ be the stationary class m delay associated with the LWDF discipline. Then there exist $J_*^{(m)} \in (0, \infty)$ such that for $m = 1, \dots, N$,*

$$\lim_{n\to\infty} \frac{1}{n} \log P(\hat{\mathbf{w}}_m(0) > n) = -\frac{J_*^{(m)}}{\alpha_m}.$$

Moreover, there exists $k \in \{1, \dots, N\}$ such that

$$0 < J_* = J_*^{(1)} = \cdots = J_*^{(k)} < \infty \tag{6.22}$$

and if $k < N$,

$$J_*^{(k)} < J_*^{(k+1)} \leq \cdots \leq J_*^{(N)} < \infty, \tag{6.23}$$

where $J_*^{(m)}$ is the optimal value of the following optimization problem [*which is the problem* (4.34) *with the additional constraint $j \geq m$*]:

$$J_*^{(m)} = \min_{j;\, x_1, \dots, x_j} \frac{1}{\gamma} \sum_{i=1}^{j} (1 - \alpha_i \gamma) L_i(x_i), \tag{6.24}$$

subject to

$$j \in \{m, \dots, N\}, \qquad x_i > 0, \qquad \sum_{i=1}^{j} x_i > 1$$

and

$$\frac{1}{\alpha_{j+1}} < \gamma = \frac{\sum_{i=1}^{j} x_i - 1}{\sum_{i=1}^{j} \alpha_i x_i} \leq \frac{1}{\alpha_j}.$$

(*Recall that $\alpha_{N+1} \doteq \infty$ by convention.*)

OUTLINE OF PROOF. First, $\hat{\mathbf{w}}_m(t)$, $m = 1, \dots N$, and $\hat{\mathbf{r}}(t)$ are measurable random variables for any t. Therefore the inner measure P_* in the lower bound (6.14) can be replaced by the probability P.

The optimal path f^0 used in the proof of upper and lower bounds in Theorem 6.8 can be directly used to get upper and lower bounds for the tail of the distribution of $\hat{\mathbf{w}}_\mathbf{1}(0)$, namely, to show the existence of

$$\lim_{n\to\infty} \frac{1}{n} \log P(\hat{\mathbf{w}}_\mathbf{m}(0) > n) = -J_*^{(m)}/\alpha_m \tag{6.25}$$

for $m = 1$, and to establish that $J_*^{(1)} = J_*$. However, the asymptotics of

$$\frac{1}{n} \log P(\hat{\mathbf{w}}_m(0) > n)$$

for *each* m could be analyzed analogously to the way

$$\frac{1}{n} \log P(\hat{\mathbf{r}}(0) > n)$$

was analyzed in Section 4, by finding the form of the optimal (fluid) path for each m. It is not hard to show that the optimal path for flow m is a simple element satisfying the constraint $j \geq m$. This yields the existence of the limit (6.25) for each m, and the characterization of $J_*^{(m)}$ as described above. This characterization implies the "ordering" inequalities (6.22) and (6.23). □

THEOREM 6.10. *Suppose for some work-conserving discipline $G \in \mathscr{G}$, for $i = 1, \dots, N$, there exists $J_*^{(i)}(G) \in (0, \infty)$ such that*

$$\lim_{n\to\infty} \frac{1}{n} \log P(\hat{\mathbf{w}}_\mathbf{i}^G(0) > n) = -\frac{J_*^{(i)}(G)}{\alpha_i}.$$

Then

$$\min_{i=1,\dots,N} J_*^{(i)}(G) \leq J_*,$$

where J_* is defined in Theorem 6.8.

The proof follows directly from Theorem 6.8.

As mentioned in the Introduction, the priority discipline can be viewed as the "limit" of the LWDF discipline with parameters $\alpha_i = \varepsilon^{N-i}$, $\varepsilon \downarrow 0$. (The lower flow index means higher priority.) This allows a derivation of the LDP for the priority discipline from that of LWDF, as shown in the following theorem. Note that for the priority discipline one can without loss of generality consider just the case of two flows. (For the case of Markov input flows this theorem can also be derived as a special case of the results of [20] for the two-flow system with the generalized processor sharing discipline. Notice that our assumptions on the input flows are more general.)

THEOREM 6.11. *Consider a priority system with two input flows* 1 *and* 2 *satisfying Assumption* 2.1, *with flow* 1 *having nonpreemptive priority over flow* 2. *Then the limit*

$$\lim_{n\to\infty} -\frac{1}{n} \log P(\hat{\mathbf{w}}_{\mathbf{2}}(0) > n) \tag{6.26}$$

exists and is equal to the optimal value of the following optimization problem:

$$\min_{x_1,\, x_2} \frac{x_2 L_1(x_1) + (1-x_1) L_2(x_2)}{x_1 + x_2 - 1},$$

subject to

$$1 \geq x_1 > 0, \qquad x_2 > 0, \qquad x_1 + x_2 > 1.$$

The statement of the above result can be derived formally from Theorem 6.9 by setting $N = 2$, $\alpha_2 = 1$, and letting $\alpha_1 \downarrow 0$. However, rather than rigorously justifying the limit transition, it would probably be easier to derive the result directly using the same approach we used to prove Theorem 6.9 for the LWDF discipline.

7. Analogous result for the unfinished work. A result analogous to Theorem 2.2 holds for the stationary unfinished work processes q_i. We only formulate the result below in Theorem 7.2 without the proof, because the analysis leading to this result is very similar and much simpler since the fluid process for the unfinished work is continuous. The only significant difference between the results for the unfinished work and the corresponding results for the delays, is that the property analogous to the "ordering property" (3.23) for the weighted delays [and the consequent ordering property of the rate functions (6.23)] does not hold for the unfinished work. This difference is reflected in the form of the corresponding finite dimensional optimization problem (7.3).

Let the positive weights $\alpha_1, \dots, \alpha_N$ be fixed. Denote by $\mathbf{q}_i^G$ the stationary unfinished work process for class i, under the discipline $G \in \mathscr{G}$. Also, denote by $\boldsymbol{\rho} \doteq \max_i \mathbf{q}_i^G/\alpha_i$ the stationary maximal weighted unfinished work.

DEFINITION 7.1 [The largest weighted (unfinished) work first (LWWF) discipline]. The LWWF discipline is a nonpreemptive, work-conserving discipline that always chooses for service the longest waiting (head-of-the-line) customer of the flow i for which the weighted unfinished work is maximal, that is, $\mathbf{q}_i(t)/\alpha_i = \boldsymbol{\rho}(t)$. In case of a tie, by convention the LWWF discipline chooses the class with the highest index.

THEOREM 7.2. *There exists $J_*^q < \infty$ such that the following holds:*

(i) *For the LWWF scheduling discipline*

$$\limsup_{n\to\infty} \frac{1}{n} \log P\left(\frac{1}{n}\boldsymbol{\rho}(0) > 1\right) \leq -J_*^q, \tag{7.1}$$

where $\boldsymbol{\rho}(0)$ *is the stationary maximal weighted unfinished work associated with the LWWF discipline.*

(ii) *For any* $G \in \mathcal{G}$,

$$\liminf_{n\to\infty} \frac{1}{n} \log P_*\left(\frac{1}{n}\boldsymbol{\rho}^G(0) > 1\right) \geq -J_*^q, \tag{7.2}$$

where $\boldsymbol{\rho}^G(0)$ *is the stationary maximal weighted unfinished work associated with the discipline* G.

(iii) *Moreover,* J_*^q *solves the following finite-dimensional optimization problem:*

$$J_*^q = \min_{K\subseteq\{1,\dots,N\};(x_i,\, i\in K)} \frac{1}{\gamma} \sum_{i\in K} L_i(x_i), \tag{7.3}$$

subject to

$$x_i > 0, \qquad i \in K, \qquad \sum_{i\in K} x_i > 1,$$

$$\frac{1}{\gamma}\left[\sum_{i\in K} x_i - 1\right] = \sum_{i\in K} \alpha_i$$

and

$$\lambda_i/\alpha_i < \gamma, \qquad i \notin K.$$

8. Conclusions. In this paper we introduce a new scheduling discipline called largest weighted delay first (LWDF) and prove that it is optimal in the sense that it maximizes the asymptotic rate of decay of the tails of the stationary maximal weighted delay within a rather general class $\mathcal{G}$ of scheduling disciplines. Even the two restrictions imposed on the class $\mathcal{G}$ appear to be mainly technical. For instance, it is only natural to expect any optimal discipline to be work conserving. We also state an analogous optimality result for the stationary maximal weighted unfinished work and the corresponding discipline LWWF.

Our results suggest that for large delays and small allowed violation probabilities, the LWDF discipline with weights $\alpha_i = -T_i/\log \delta_i$ would be a nearly optimal discipline to use in order to satisfy the QoS constraints (1.1). Thus, whenever it is feasible to satisfy these constraints, one would expect that LWDF would do so. However, in cases when it is not feasible, the LWDF policy will most likely violate the QoS constraints of most users. This has significant implications for flow admission control. Most importantly, LWDF allows one to detect "in real time" the infeasibility of satisfying QoS requirements. In addition even when the QoS constraints are infeasible, LWDF has the property of fairness, in that it tries to equalize for all users the ratio of the logarithm of actual violation probability to the logarithm of the desired probability. However, this notion of fairness may or may not be desirable. In certain cases, it may be preferable to satisfy the QoS for as many users as possible, while penalizing the rest. In such situations, one may want to use LWDF for only a subset of flows, while giving lower priority to other flows.

Note that our result concerns the stationary delays, rather than the stationary waiting times of individual customers. Nevertheless, one would expect that the asymptotics of the tails of stationary distributions of both these processes typically coincide. Indeed, we believe that similar results for the waiting times can be derived from our results.

An interesting and very challenging open problem is that of determining optimal disciplines in the network context.

APPENDIX A

In this section we outline the key steps used to prove Lemma 6.6. These follow the approach used in [16, 17].

A straightforward adaptation of the argument in [17], Proposition 7.2, yields the following lemma. Consider the unfinished work process $\mathbf{v} \doteq V\mathbf{f}$. (Recall the definition of the unfinished work mapping V given in Section 6.2.)

LEMMA A.1. *The unfinished work* $\mathbf{v}(0)$ *has exponential tail; that is,*

$$\limsup_{n\to\infty} \frac{1}{n} \log P\left(\frac{\mathbf{v}(0)}{n} > 1\right) = -c < 0.$$

The following lemma can be proved in the same way as [17], Proposition 8.1.

LEMMA A.2. *Consider a system with zero initial state* (*at time* 0) $\boldsymbol{\theta}(0) = (0; 0)$, *and let the input flow* $\mathbf{h}$ *be the increment of the original input flow* $\mathbf{f}$ *after time* 0, *so that for* $t \ge 0$ *and* $i = 1, \dots, N$,

$$\mathbf{h}_i(t) = \mathbf{f}_i(t) - \mathbf{f}_i(0). \tag{A.1}$$

Then for every $s \ge 0$ *and any measurable subset* $B \subset \Psi$,

$$\limsup_{n\to\infty} \frac{1}{n} \log P\{\Gamma^n[\boldsymbol{\theta}(ns)] \in B\} \le - \inf_{h \in \overline{H_+^{-1}(s)B}} J_s(h),$$

and for every $s \ge 0$ *and any* (*not necessarily measurable*) *subset* $B \subset \Psi$,

$$\liminf_{n\to\infty} \frac{1}{n} \log P_*\{\Gamma^n[\boldsymbol{\theta}(ns)] \in B\} \ge - \inf_{h \in \Phi_s(B)} J_s(h).$$

Finally, the following lemma can be established using an adaptation of the proof of Theorem 8.2 in [17].

LEMMA A.3. *Let the random initial state* $\boldsymbol{\theta}(0)$ *of the process* $\boldsymbol{\theta}$ *be such that the distribution of the unfinished work* $\mathbf{v}(0)$ *has exponential tail, and let* $\mathbf{h}$ *be as defined in* (A.1). *Then for any measurable subset* $B \subseteq \Psi$,

$$\limsup_{s\to\infty} \limsup_{n\to\infty} \frac{1}{n} \log P\{\Gamma^n[\boldsymbol{\theta}(ns)] \in B\} \le - \inf_{s>0,\, h \in \overline{H_+^{-1}(s)B}} J_s(h),$$

and any (not necessarily measurable) subset $B \subseteq \Psi$,

$$\liminf_{s\to\infty} \liminf_{n\to\infty} \frac{1}{n} \log P_*\{\Gamma^n[\boldsymbol{\theta}(ns)] \in B\} \geq - \inf_{s>0,\, h\in\Phi_s B} J_s(h).$$

Lemma 6.6 follows from Lemma A.3 and the fact that $\boldsymbol{\theta}$ is stationary.

APPENDIX B

Recall the notation that $\mathscr{F}(S)$ refers to the σ-algebra generated by the cylinder subsets of the function space S, and that $\mathscr{Q}$ is the set of rational numbers. It is well known that $\mathscr{I} \in \mathscr{F}(\mathscr{D})$ and $\mathscr{I}_+$, $\mathscr{I}_{+,0} \in \mathscr{F}(\mathscr{D}_+)$. The following lemma is quite straightforward, but since we could not find a direct reference, we include it here for the sake of completeness.

LEMMA B.1. *Recall the definition of the spaces* $\mathscr{D}, \mathscr{D}_+$, $\mathscr{S}, \mathscr{S}_+$ *and* $\mathscr{S}_{+,0}$ *given in Section* 1.2. *Then*

$$\mathscr{S} \in \mathscr{F}(\mathscr{D}), \quad \mathscr{S}_+ \in \mathscr{F}(\mathscr{D}_+), \quad \mathscr{S}_{+,0} \in \mathscr{F}(\mathscr{D}_+). \tag{B.1}$$

PROOF. Given any positive $q, \varepsilon, \delta \in \mathscr{Q}$ and $l \in \{1, 2, \ldots\}$, we define the set $M(q, \varepsilon, \delta, l) \in \mathscr{Q}^{2l}$ by

$$M(q,\varepsilon,\delta,l) \doteq \left\{ \begin{aligned} &m = ((m_{11}, m_{12}, (m_{21}, m_{22}, \ldots, (m_{l1}, m_{l2})):\\ &\qquad -q < m_{11} < m_{12} < \cdots < m_{l1} < m_{l2} < q,\\ &\qquad m_{i2} - m_{i1} < \delta \quad \text{for } i = 1, \ldots, l,\\ &\qquad \text{and } m_{i+1,1} - m_{i,2} > \varepsilon \quad \text{for } i = 1, \ldots, l-1 \end{aligned} \right\}$$

Let

$$\begin{aligned} K(q,\varepsilon,\delta,l) \doteq \mathscr{I} &\cap \left[\bigcup_{m\in M(q,\varepsilon,\delta,l)} \{h \in \mathscr{D}: h(-q) = h(m_{11}-)\} \right] \\ &\cap \{h \in \mathscr{D}: h(m_{l2}) = h(q-)\} \\ &\cap \left[\bigcap_{i=1,\ldots,l-1} \{h \in \mathscr{D}: h(m_{i2}) = h(m_{i+1,1})\} \right]. \end{aligned}$$

Clearly $K(q, \varepsilon, \delta, l) \in \mathscr{F}(\mathscr{D})$. Thus since

$$\mathscr{S} = \bigcap_{q>0,\, q\in\mathscr{Q}} \bigcup_{\varepsilon>0,\, \varepsilon\in\mathscr{Q}} \bigcap_{\delta>0,\, \delta\in\mathscr{Q}} \bigcup_{l=1,2,\ldots} K(q,\varepsilon,\delta,l),$$

it follows that $\mathscr{S} \in \mathscr{F}\mathscr{D}$. The measurability of $\mathscr{S}_+$ and $\mathscr{S}_{+,0}$ can be proved analogously. □

We remarked in Section 6 that the mappings

$$H(t)\colon \mathscr{S}^N \to \Psi \quad \text{and} \quad H_+(t)\colon \mathscr{S}^N_{+,0} \to \Psi$$

can be considered more generally as the mappings

$$H(t)\colon \mathscr{I}^N \to \mathbb{R}_+ \times \mathscr{I}_+^N \quad \text{and} \quad H_+(t)\colon \mathscr{I}_{+,0}^N \to \mathbb{R}_+ \times \mathscr{I}_+^N,$$

respectively, defined by the same expressions (6.1), (6.2) and (6.3). Due to Lemma B.1, to prove the measurability of the former mappings it obviously suffices to prove measurability of the latter ones. In the following lemma we do just that.

LEMMA B.2. *For any fixed $t \in \mathbb{R}$, the mappings*

$$H(t)\colon \mathscr{I}^N \to \mathbb{R}_+ \times \mathscr{I}_+^N \quad \textit{and} \quad H_+(t)\colon \mathscr{I}_{+,0}^N \to \mathbb{R}_+ \times \mathscr{I}_+^N$$

are measurable.

PROOF. For a fixed t, the mapping $H(t)$ maps an element $f = (f_1, \dots, f_N) \in \mathscr{I}^N$ into an element $\psi = (b(t); g_1, \dots, g_N) \in \mathbb{R}_+ \times \mathscr{I}_+^N$. Define the mappings $h\colon \mathscr{I}^N \to \mathscr{D}$ and $z\colon \mathscr{I}^N \to \mathscr{D}$ by

$$[hf](s) \doteq \sum_{i=1}^N f_i(s) - s$$

and

$$[zf](s) = \begin{cases} \inf_{u \le s}[hf](u), & \text{if } \inf_{u \le s}[hf](u) > -\infty, \\ \sum_{i=1}^N f_i(s-) - s, & \text{otherwise.} \end{cases}$$

Then both h and z are measurable mappings. For a fixed $u \ge 0$,

$$\{f \in \mathscr{I}^N\colon b(t) > u\} = \bigcup_{\delta_2 > \delta_1 > 0,\, \delta_1,\, \delta_2 \in \mathscr{Q}} \left\{ \inf_{s \le t-u-\delta_2} [hf](s) < \inf_{t-u-\delta_1 \le s \le t} [hf](s) \right\},$$

where each subset in the union is measurable because for a fixed f, $[hf](\cdot)$ is a RCLL function. Thus for each t the mapping of f into $b(t)$ is measurable.

Now fix $t > 0$ and i and consider $g_i = (g_i(u), u \ge 0)$. (Recall that although not explicitly notated, the function g itself depends on the time parameter, which is fixed to be t here.) For fixed $u \ge 0$ and $c \in \mathbb{R}$,

$$\{g_i(u) < c\} = \{u \ge b(t), f_i(t) - f_i((t - b(t))-) < c\}$$
$$\cup \{u < b(t), f_i(u + t - b(t)) - f_i((t - b(t))-) < c\}.$$

However, for any fixed $s \in \mathbb{R}$ both $f_i(s - b(t))$ and $f_i((s - b(t))-)$ are measurable. Indeed, for any $c_1 \in \mathbb{R}$,

$$\{f_i((s - b(t))-) > c_1\} = \bigcup_{q \ge 0,\, q \in \mathscr{Q}} \{f_i(s - q) > c_1, b(t) < q\},$$

and similarly,

$$\{f_i(s - b(t)) < c_1\} = \bigcup_{q \ge 0,\, q \in \mathscr{Q}} \{f_i(s - q) < c_1, b(t) \ge q\}.$$

This proves the measurability of $g_i(u)$ for any fixed $u \ge 0$ and i, which automatically implies the measurability of $(g_1, \dots, g_N)$. Thus we have shown that

$H(t)$ is measurable for each $t \geq 0$. Since $H_+(t)$ can be identified with the restriction of $H(t)$ to the measurable subset $\{f \in \mathscr{I}^N\colon f_i(u) = 0, u \leq 0, i = 1, \ldots, N\}$, For each $t \geq 0$ the measurability of $H_+(t)$ follows from that of $H(t)$. This also establishes the measurability of H and H_+ since we consider the σ-algebra generated by the cylinder sets. □

We now consider the operator $\widehat{A}$ defined on $\mathscr{S}_+^N$ (for the LWDF discipline) in Section 3 and denote

$$\widehat{A}(t)f \doteq [\widehat{A}f](t).$$

LEMMA B.3. *For any $t \geq 0$, the mapping*

$$\widehat{A}(t)\colon \mathscr{S}_+^N \to \mathbb{R}_+^N$$

is measurable.

PROOF. Let $f = (f_1, \ldots, f_N) \in \mathscr{S}_+^N$.

STEP 1. Consider $f_i \in \mathscr{S}_+$. By the definition of $\mathscr{S}_+$, the function f_i is uniquely defined by the two sequences,

$$x^i \doteq (x_1^i, x_2^i, \ldots) \in \mathbb{R}_+^\infty \quad \text{and} \quad y^i \doteq (y_1^i, y_2^i, \ldots) \in \mathbb{R}_+^\infty,$$

where $x_1^i, x_2^i, \ldots$ are the strictly positive jump sizes ordered by increasing time of the jumps, $y_1^i \geq 0$ is the time to the first jump and $y_2^i, \ldots$ are strictly positive time intervals between consecutive jumps. [We adopt the convention that if f_i has a jump at time zero, then $x_1^i = f_i(0)$ and $y_1^i = 0$.] It is easy to see that the mapping $X_i\colon \mathscr{S}_+ \to \mathbb{R}_+^\infty \times \mathbb{R}_+^\infty$ that takes f_i to (x^i, y^i) is measurable. We drop the fixed superscript i in the sequences x^i and y^i in the rest of this paragraph. Then for any $c > 0$, $\{y_1 > c\} = \{f_i(c) = 0\}$ and so y_1 is measurable. Consequently,

$$x_1 = f_i(0)1_{\{y_1=0\}} + [f_i(y_1) - f_i(y_1-)]1_{\{y_1>0\}}$$

is measurable. Similarly, since

$$\{y_1 + y_2 > c\} = \{f_i(c) \leq x_1\},$$

y_2 is measurable, and the fact that

$$x_2 = f_i(y_1 + y_2) - f_i((y_1 + y_2)-)$$

establishes the measurability of x_2. Proceeding iteratively we see that the mapping X_i is measurable. Consequently the mapping

$$X\colon \mathscr{S}_+^N \to (\mathbb{R}_+^\infty)^{(2N)}$$

that maps $f \in \mathscr{S}_+^N$ into

$$(x, y) = \left((x^1, x^2, \ldots, x^N), (y^1, y^2, \ldots, y^N)\right)$$

is measurable.

For future use, we define

$$\bar{y}^i_j \doteq \sum_{k=1}^{j} y^i_k$$

to be the time of arrival of the jth class i customer. From the above argument, it easily follows that for $j = 1, \ldots$ and $i \in \{1, \ldots, N\}$, $\bar{y}^i_j$ is measurable.

STEP 2. Now we consider the mapping Y which maps $f \in \mathscr{S}^N_+$ into $Yf \doteq (m, \xi) \in \{1, \ldots, N\}^\infty \times \mathbb{R}^\infty_+$ as follows. Denote $m = (m_1, m_2, \ldots)$ and $\xi = (\xi_1, \xi_2, \ldots)$. Then m_j represents the class of the jth customer to be served, and $\xi_j \geq 0$ the time at which the service of that customer starts. Note that Y is well defined for the LWDF discipline. We now use induction to prove that the mapping Y is measurable by showing that (m_j, ξ_j) is measurable for $j = 1, 2, \ldots$. Let $(x, y) = Xf$. First note that (m_1, ξ_1) is measurable because for any $i \in \{1, \ldots N\}$ and time $s \geq 0$,

$$\{m_1 = i, \xi_1 < s\} = \bigcap_{l>i} \left\{ y^i_1 < s, y^i_1 < y^l_1 \right\} \bigcap_{l<i} \left\{ y^i_1 \leq y^l_1 \right\}.$$

Now suppose that (m_j, ξ_j) are measurable for $j = 1, \ldots, k$. We will show that then $(m_{k+1,} \xi_{k+1})$ are also measurable. We first introduce some notation. Let $\bar{i}^k = (i_1, \ldots, i_k)$ denote a generic element of $\{1, \ldots, N\}^k$ (used to represent the vector of the classes of the first k customers that departed the system). Then define the functions η: $\{1, \ldots, N\} \times \{1, \ldots, N\}^k \to Z_+$ and z: $\{1, \ldots, N\}^k \to \mathbb{R}_+$ by

$$\eta(i, \bar{i}^k) \doteq \sum_{j=1}^{k} 1_{\{i_j = i\}}$$

and

$$z(\bar{i}^k) \doteq x^{i_k}_{\eta(i_k, \bar{i}^k)}.$$

Observe that $\eta(i, \bar{i}^k)$ represents the number of class i customers among the first k customers that departed the system, and $z(\bar{i}^k)$ represents the service time of the kth customer departed the system. Then for any $i \in \{1, \ldots, N\}$ and $s \geq 0$ we can write

$$\{m_{k+1} = i, \xi_{k+1} < s\} = \bigcup_{\bar{i}^k \in \{1, \ldots, N\}^k} K_{\bar{i}^k},$$

where

$$K_{\bar{i}^k} \doteq \left[\bigcap_{j \leq k} \{m_j = i_j\} \right] \cap \{\xi_k < s - z(\bar{i}^k)\} \cap (M_1 \cup M_2),$$

with

$$\begin{aligned} M_1 &\doteq \{\bar{y}^i_{\eta(i,\bar{i}^k)+1} \le \xi_k + z(\bar{i}^k)\} \\ &\quad \cap \left[\bigcap_{l>i} \left\{ \frac{\xi_k + z(\bar{i}^k) - \bar{y}^i_{\eta(i,\bar{i}^k)+1}}{\alpha_i} > \frac{\xi_k + z(\bar{i}^k) - \bar{y}^l_{\eta(l,\bar{i}^k)+1}}{\alpha_l} \right\} \right] \\ &\quad \bigcap_{l<i} \left\{ \frac{\xi_k + z(\bar{i}^k) - \bar{y}^i_{\eta(i,\bar{i}^k)+1}}{\alpha_i} \ge \frac{\xi_k + z(\bar{i}^k) - \bar{y}^l_{\eta(l,\bar{i}^k)+1}}{\alpha_l} \right\} \end{aligned}$$

and

$$\begin{aligned} M_2 &\doteq \left\{ \bar{y}^i_{\eta(i,\bar{i}^k)+1} > \xi_k + z(\bar{i}^k) \right\} \\ &\quad \bigcap_{l>i} \left\{ \bar{y}^i_{\eta(i,\bar{i}^k)+1} < \bar{y}^l_{\eta(l,\bar{i}^k)+1} \right\} \\ &\quad \bigcap_{l<i} \left\{ \bar{y}^i_{\eta(i,\bar{i}^k)+1} \le \bar{y}^l_{\eta(l,\bar{i}^k)+1} \right\}. \end{aligned}$$

Since $K_{\bar{i}^k}$ is measurable for any $\bar{i}^k \in \{1, \ldots, N\}^k$, this proves the induction step, and therefore the measurability of $(m, \xi) = Yf$.

STEP 3. We now prove the measurability of $\hat{\tau}_i(t)$ for fixed i and $t \ge 0$. Let $(x, y) = Xf$ and $(m, \xi) = Yf$. Let $\chi_i(t)$ denote the number of class i customers that have departed the system by time t. The function $\chi_i(t)$ is measurable, because

$$\begin{aligned} \{\chi_i(t) = j\} = \bigcup_{k \ge 1,\, \bar{i}^k \in \{1,\ldots,N\}^k,\, \eta(i,\bar{i}^{k-1}) = j} & \big[\{\xi_k + z(\bar{i}^k) > t\} \\ & \cap \{\xi_{k-1} + z(\bar{i}_{k-1}) \le t\} \big]. \end{aligned}$$

The observation that for any $s \ge 0$,

$$\{\hat{\tau}_i(t) < s\} = \begin{cases} \mathscr{S}^N_+, & \text{if } s > t, \\ \bigcup_{j=0,1,\ldots} \{\chi_i(t) = j,\, \bar{y}^i_{j+1} < s\}, & \text{if } s \le t, \end{cases}$$

then completes the proof. □

LEMMA B.4. *For any $a > 0$, the subset*

$$B \doteq \{\psi = (b, g) \in \Psi : [\widehat{R}(b)](g) > a\}$$

is measurable. (*Recall that $\widehat{R}$ is the mapping associated with the LWDF discipline.*)

PROOF. Note that for any fixed $g \in \mathscr{S}_+^N$, the subset $\{b \geq 0 | [\widehat{R}(b)](g) > a\}$ is open, because $\widehat{R}g$ is a nonnegative piecewise linear RCLL function with finite number of linear "pieces" in any finite interval, and this function can only jump down; that is,

$$[\widehat{R}g](t) \leq [\widehat{R}g](t-).$$

Then we can write

$$B = \bigcup_{0<q_1<q_2;\, q_1,\, q_2 \in \mathscr{Q}} \times (q_1, q_2) \Bigg[\bigcup_{\varepsilon>0,\, \varepsilon\in\mathscr{Q}} \; \bigcap_{q_1<q<q_2,\, q\in\mathscr{Q}} \big\{ g \in \mathscr{S}_+^N \colon \big[\widehat{R}g\big](q) > a + \varepsilon \big\} \Bigg].$$

Since each subset under the first union is measurable, we get the measurability of B. $\square$

Acknowledgments. We thank our colleagues Matthew Andrews, Debasis Mitra, Anwar Elwalid, Krishnan Kumaran, Marty Reiman, Alan Weiss and Phil Whiting at Bell Labs for very useful discussions during the course of this work.

REFERENCES

[1] ANDREWS, M., KUMARAN, K., RAMANAN, K., STOLYAR, A. and WHITING, P. (1999). Data rate scheduling algorithms and capacity estimates for the CDMA forward link. Bell Labs Technical memorandum.

[2] ANDREWS, M. and ZHANG, L. (1999). Minimizing end-to-end delay in high-speed networks with a simple coordinated schedule. *IEEE INFOCOM'99* 380–388.

[3] CHANG, C.-S. (1994). Stability, queue length and delay of deterministic and stochastic queueing networks. *IEEE Trans. Automat. Control* **39** 913–931.

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

[5] DUPUIS, P. and ELLIS R. S. (1997). *A Weak Convergence Approach to the Theory of Large Deviations*. Wiley, New York.

[6] DUPUIS, P. and KUSHNER, H. (1989). Minimizing escape probabilities: a large deviation approach. *SIAM J. Control Optim.* **2** 432–445.

[7] DUPUIS, P. and RAMANAN, K. (1998). A Skorokhod problem formulation and large deviation analysis of a processor sharing model. *Queueing Systems Theory Appl.* **28** 109–124.

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

[9] FLEMING, W. H. and SOUGANIDIS, P. E. (1986). PDE-viscosity solution approach to some problems of large deviations. *Ann. Scuola Norm. Sup. Pisa Cl. Sci.* **4** 171–192.

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

[11] 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.

[12] HARRISON, J. M. (1998). Heavy traffic analysis of a system with parallel servers: asymptotic optimality of discrete review policies. *Ann. Appl. Probab.* **8** 822–849.

[13] KELLY, F. P. (1991). Effective bandwidths at multi-class queues. *Queueing Systems Theory Appl.* **9** 5–16.

[14] KUSHNER, H. J. Control of trunk line systems in heavy traffic. *SIAM J. Control Option* **33** 765–803.

[15] 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.

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

[17] MAJEWSKI, K. (1996). Large deviations of feedforward queueing networks. Ph.D. thesis, Univ. Münich.

[18] MARTINS, L. F. and KUSHNER, H. J. (1990). Routing and singular control for queueing networks in heavy traffic. *SIAM J. Control Optim.* **28** 1209–1233.

[19] O'CONNELL, N. (1996). Large deviations for queue lengths at a multi-buffered resource. Technical report HPL-BRIMS-96-010, BRIMS Hewlett-Packard Labs, Bristol, England.

[20] PASCHALIDIS, I. C. (1996). Large deviations in high speed communication networks. Ph.D. dissertation, MIT.

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

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

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

[24] 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.

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

BELL LABS
LUCENT TECHNOLOGIES
600 MOUNTAIN AVENUE, 2C-319
MURRAY HILL, NEW JERSEY 07974
E-MAIL: kavita@research.bell-labs.com