Mathematics of Operations Research

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

Fluid Limits for Shortest Remaining Processing Time Queues

Douglas G. Downhttp://www.cas.mcmaster.ca/~downd/, H. Christian Gromollhttp://www.faculty.virginia.edu/gromoll/, Amber L. Puhahttp://www.csusm.edu/puha,

MATHEMATICS OF OPERATIONS RESEARCH
Vol. 34, No. 4, November 2009, pp. 880–911
ISSN 0364-765X | EISSN 1526-5471 | 09 | 3404 | 0880
DOI 10.1287/moor.1090.0409

Fluid Limits for Shortest Remaining Processing Time Queues

Douglas G. Down
Department of Computing and Software, McMaster University, Hamilton, Ontario L8S 4K1, Canada,
downd@mcmaster.ca, http://www.cas.mcmaster.ca/~downd/

H. Christian Gromoll
Department of Mathematics, University of Virginia, Charlottesville, Virginia 22903,
gromoll@virginia.edu, http://www.faculty.virginia.edu/gromoll/

Amber L. Puha
Department of Mathematics, California State University San Marcos,
San Marcos, California 92096, apuha@csusm.edu, http://www.csusm.edu/puha

We consider a single-server queue with renewal arrivals and i.i.d. service times in which the server uses the shortest remaining processing time policy. To describe the evolution of this queue, we use a measure-valued process that keeps track of the residual service times of all buffered jobs. We propose a fluid model (or formal law of large numbers approximation) for this system and, under mild assumptions, prove the existence and uniqueness of fluid model solutions. Furthermore, we prove a scaling limit theorem that justifies the fluid model as a first-order approximation of the stochastic model. The state descriptor of the fluid model is a measure-valued function whose dynamics are governed by certain inequalities in conjunction with the standard workload equation. In particular, these dynamics determine the evolution of the left edge (infimum) of the state descriptor's support, which yields conclusions about response times. We characterize the evolution of this left edge as an inverse functional of the initial condition, arrival rate, and service time distribution. This characterization reveals the manner in which the growth rate of the left edge depends on the service time distribution. By considering varying examples, the authors show that the rate can vary from logarithmic to polynomial.

Key words: shortest remaining processing time; fluid model; fluid limit; measure-valued process; response time
MSC2000 subject classification: Primary: 60K25; secondary: 60F17, 68M20, 90B22
OR/MS subject classification: Primary: Queues—limit theorems, queues—approximations; secondary: probability—stochastic modeling applications
History: Received June 22, 2008; revised February 20, 2009. Published online in *Articles in Advance* October 20, 2009.

1. Introduction. Consider a single-server queue operating under the shortest remaining processing time (SRPT) scheduling policy. The SRPT scheduling policy gives preemptive priority to the job in the system with the shortest remaining processing time. Note that to implement this policy, it is assumed that the service times of jobs are known on arrival.

Interest in the SRPT policy stretches back to the first optimality result from Schrage [17], who showed that SRPT minimizes the number of jobs in the system at any point in time (see also Smith [20]). This was done with no distributional assumptions on the underlying arrival and service processes. Expressions for the mean response time for a single-server M/G/1/SRPT queue were earlier developed by Schrage and Miller [18], with extended results available in Schassberger [16] and Perera [13] (a nice survey from the same time period is Schreiber [19]). Within these references, one can find expressions for various performance measures, all of which depend on the entire service time distribution through nested integrals and are thus somewhat difficult to work with, particularly if one wishes to make comparisons with other policies.

Recently, there has been renewed interest in the SRPT policy, mainly in computer science. For example, Bansal and Harchol-Balter [1] are interested in the issue of fairness for SRPT (Bansal and Harchol-Balter [1] is also a good source for a more extended list of prior work on SRPT). More recent work has attempted to provide a framework for comparing policies in the M/G/1 setting; see, for example, Wierman and Harchol-Balter [21].

There has also been a recent body of work on the tail behavior of single-server queues under SRPT; see, for example, Núñez Queija [11] and Nuyens and Zwart [12]. They discuss the advisability of implementing SRPT using large deviations techniques.

Down and Wu [3] use diffusion limits to show certain optimality properties of a multilayered round-robin routing policy for a system of parallel servers, each operating under SRPT. This is done under the assumption of a finitely supported service time distribution, mainly because of the absence of such limits for more general service time distributions.

In this paper, the goal is to take first steps toward developing a general diffusion limit by characterizing the fluid limits (functional law of large numbers approximations) for a single-server SRPT queue. Because SRPT is not a head-of-the-line policy, we use a state descriptor that tracks the remaining service times of all jobs in the system. Under mild conditions, we develop fluid limits for the measure-valued state descriptor that puts a unit of mass at each element in the set of remaining service times. Such an approach is in the spirit of Gromoll et al. [8], Puha et al. [15], Doytchinov et al. [5], and Kruk et al. [10], who consider single-server queues operating under the processor sharing (PS) and earliest deadline first (EDF) policies, respectively. The analysis here is more akin to that in Doytchinov et al. [5] and Kruk et al. [10] for EDF. In part, this is because of the observation that for both SRPT and EDF, there is only one job in service at any point in time, which contrasts with PS where all jobs receive simultaneous service. There is some additional similarity between SRPT and EDF priority schemes because they give preemptive priority to the job that has the smallest residual service time and the smallest current lead time, respectively. Indeed the analysis here also makes use of a frontier process similar to the one considered in Doytchinov et al. [5] and Kruk et al. [10]. However, because the lead times in Doytchinov et al. [5] and Kruk et al. [10] are required to be independent of the service times and decrease constantly at rate one, there are significant differences between the analysis in this paper and the work in Doytchinov et al. [5] and Kruk et al. [10].

Under mild conditions that include a finite limiting arrival rate and a finite first moment for the limiting service time distribution, we prove that there is a unique fluid limit. This fluid limit is a measure-valued function that has a nondecreasing left edge, the infimum of the measure's support. This is a direct reflection of the SRPT scheduling policy, which gives preemptive priority to the job with the shortest remaining service time. In particular, work in the fluid limit does not accumulate below the left edge. In this paper, we analyze the behavior of the fluid limit by defining a fluid model (§2.2), proving that under mild conditions the fluid limit is a fluid model solution (Theorem 3.3), and analyzing the behavior of fluid model solutions (Theorems 3.1 and 3.2 and Corollaries 3.1, 3.2, and 3.3).

Of particular interest is the behavior of the left edge of a fluid model solution as a function of time. The results presented here include an explicit description of the unique fluid model solution (Theorems 3.1 and 3.2) that characterizes the left edge as the right-continuous inverse of a simple functional (18) of the fluid model data. This characterization allows us to prove that under mild conditions, critical fluid model solutions corresponding to service time distributions with unbounded support converge to the zero measure as time tends to infinity (Corollary 3.1). This is somewhat surprising because the fluid analog of the workload is constant for critical fluid model solutions, i.e., the workload does not decrease with time. The characterization of the left edge is applied in some specific examples to determine the rate at which the left edge increases as time increases and, hence, the rate at which the critical fluid model empties (see §3.2). Interestingly, the rate depends on the fluid model data and, in particular, on the tail behavior of the limiting service time distribution. Corollary 3.2 characterizes the limiting behavior of critical fluid model solutions corresponding to service time distributions with bounded support.

The paper is organized as follows. In §2, we define our stochastic and fluid models for an SRPT queue. Section 3 contains the statements of our main results. Section 4 contains the proofs of the results concerning fluid model solutions, and the remainder of the paper is devoted to the proof of Theorem 3.3, the fluid limit theorem.

1.1. Notation. The following notation will be used throughout the paper. Let $\mathbb{N}$ denote the set of positive integers and let $\mathbb{R}$ denote the set of real numbers. For $a, b \in \mathbb{R}$, we write $a \vee b$ for the maximum of a and b, $a \wedge b$ for the minimum of a and b, a^+ and a^- for the positive and negative parts of a, respectively, $\lfloor a \rfloor$ for the largest integer less than or equal to a, and $\lceil a \rceil$ for the smallest integer greater than or equal to a. The nonnegative real numbers $[0, \infty)$ will be denoted by $\mathbb{R}_+$. By convention, a sum of the form $\sum_{i=n}^{m}$ with $n > m$ or a sum over an empty set of indices equals zero. The sets (a, b), $[a, b)$, and $(a, b]$ are empty for $a, b \in [0, \infty]$ with $a \geq b$ and, unless otherwise specified, the infimum of the empty set equals ∞. For a function $g\colon \mathbb{R}_+ \to \mathbb{R}$, let $\|g\|_\infty = \sup_{x \in \mathbb{R}_+} |g(x)|$ and $\|g\|_K = \sup_{x \in [0, K]} |g(x)|$ for each $K \geq 0$. We define the positive and negative parts of such a function g by $g^+(x) = g(x) \vee 0$ and $g^-(x) = (-g(x)) \vee 0$ for all $x \in \mathbb{R}_+$.

For a Borel set $B \subset \mathbb{R}_+$, we denote the indicator of the set B by 1_B. In addition, for $\epsilon > 0$,

$$B^\epsilon = \left\{ x \in \mathbb{R}_+ \colon \inf_{y \in B} |x - y| < \epsilon \right\}. \tag{1}$$

We also define the real-valued function $\chi(x) = x$ for $x \in \mathbb{R}_+$. For a topological space A, denote by $C^+(A)$ the set of nonnegative, continuous, real-valued functions defined on A, and denote by $C_b^+(A)$ the functions in $C^+(A)$ that are bounded.

Let $\mathbf{M}$ denote the set of finite, nonnegative Borel measures on $\mathbb{R}_+$ and let $\mathbf{M}^a$ denote those elements of $\mathbf{M}$ that do not charge the origin. Consider $\xi \in \mathbf{M}$ and a Borel measurable function $g\colon \mathbb{R}_+ \to \mathbb{R}$, which is integrable with respect to ξ. We define $\langle g, \xi \rangle = \int_{\mathbb{R}_+} g(x)\xi(dx)$. The set $\mathbf{M}$ is endowed with the weak topology, that is, for

$\xi_n, \xi \in \mathbf{M}$, $n \in \mathbb{N}$, we have $\xi_n \xrightarrow{w} \xi$ if and only if $\langle g, \xi_n \rangle \to \langle g, \xi \rangle$ as $n \to \infty$ for all $g\colon \mathbb{R}_+ \to \mathbb{R}$ that are bounded and continuous. With this topology, $\mathbf{M}$ is a Polish space (Prohorov [14]). We denote the zero measure in $\mathbf{M}$ by $\mathbf{0}$ and the measure in $\mathbf{M}$ that puts one unit of mass at the point $x \in \mathbb{R}_+$ by δ_x. For $x \in \mathbb{R}_+$, the measure δ_x^+ is δ_x if $x > 0$ and $\mathbf{0}$ otherwise.

We say that a measure $\xi \in \mathbf{M}$ has a finite first moment if $\langle \chi, \xi \rangle < \infty$. Let $\mathbf{M}_\chi$ denote the set of all such measures and let $\mathbf{M}_0 = \mathbf{M}_\chi \cap \mathbf{M}^a$. It will be convenient to extend the notion of uniform integrability for random variables (and their associated distributions) to elements of $\mathbf{M}$. Call a sequence $\{\xi_n\} \subset \mathbf{M}$ uniformly integrable if $\langle \chi, \xi_n \rangle < \infty$ for all n, and

$$\lim_{x\to\infty} \sup_n \langle \chi 1_{[x,\infty)}, \xi_n \rangle = 0.$$

It is easy to show that if $\{\xi_n\} \subset \mathbf{M}$ is uniformly integrable and $\xi_n \xrightarrow{w} \xi$, then $\langle \chi, \xi \rangle < \infty$ and $\langle \chi, \xi_n \rangle \to \langle \chi, \xi \rangle$.

We use "$\Rightarrow$" to denote convergence in distribution of random elements of a metric space. Following Billingsley [2], we use $\mathbf{P}$ and $\mathbf{E}$, respectively, to denote the probability measure and expectation operator associated with whatever space the relevant random element is defined on. Unless otherwise specified, all stochastic processes used in this paper are assumed to have paths that are right continuous with finite left limits (r.c.l.l.). For a Polish space $\mathcal{S}$, we denote by $D([0,\infty), \mathcal{S})$ the space of r.c.l.l. functions from $[0,\infty)$ into $\mathcal{S}$, endowed with the Skorohod J_1-topology (Ethier and Kurtz [6]).

2. Stochastic and fluid models for an SRPT queue.

2.1. Stochastic model.

Our stochastic model of an SRPT queue consists of the following: a random initial condition $\mathcal{Z}(0) \in \mathbf{M}$ specifying the state of the system at time zero, stochastic primitives $E(\cdot)$ and $\{v_k\}_{k\in\mathbb{N}}$ describing the arrival of jobs and their service times to the queue, and a measure-valued state descriptor $\mathcal{Z}(\cdot)$ describing the time evolution of the system. These are defined next.

2.1.1. Initial condition.

The initial condition specifies the number $Z(0)$ of jobs in the queue at time zero as well as the initial service time of each job. Assume that $Z(0)$ is a nonnegative integer-valued random variable that is finite almost surely. The initial service times are the first $Z(0)$ elements of a sequence $\{\tilde{v}_j\}_{j\in\mathbb{N}}$ of strictly positive, finite random variables. We sometimes refer to jobs in the system at time zero as initial jobs. The initial job with service time $\tilde{v}_j$, $j \le Z(0)$ is called job j.

A convenient way to express the initial condition is to define an initial random measure $\mathcal{Z}(0) \in \mathbf{M}$ by

$$\mathcal{Z}(0) = \sum_{j=1}^{Z(0)} \delta_{\tilde{v}_j},$$

which equals $\mathbf{0}$ if $Z(0) = 0$. Our assumptions imply that $\mathcal{Z}(0)$ satisfies

$$\mathbf{P}(\langle 1, \mathcal{Z}(0) \rangle \vee \langle \chi, \mathcal{Z}(0) \rangle < \infty) = 1. \tag{2}$$

In particular, the number of initial jobs and the initial workload are finite almost surely, and so $\mathcal{Z}(0) \in \mathbf{M}_0$ almost surely.

2.1.2. Stochastic primitives.

The stochastic primitives consist of an exogenous arrival process $E(\cdot)$ and a sequence of initial service times $\{v_k\}_{k\in\mathbb{N}}$. The arrival process $E(\cdot)$ is a rate $\alpha \in (0,\infty)$ delayed renewal process. For $t \in [0,\infty)$, $E(t)$ represents the number of jobs that arrive to the queue during the time interval $(0, t]$. Jobs arriving after time zero are indexed by integers $j > Z(0)$. For $t \in [0,\infty)$, let

$$A(t) = Z(0) + E(t). \tag{3}$$

Then, job $j \in \mathbb{N}$ arrives at time $T_j = \inf\{t \in [0,\infty)\colon A(t) \ge j\}$. Hence, for $i < j$, $T_i \le T_j$ and we say that job i arrives before job j.

For each $k \in \mathbb{N}$, the random variable v_k represents the initial service time of the $(Z(0)+k)$th job. That is, job $j > Z(0)$ has initial service time $v_{j-Z(0)}$. Assume that the random variables $\{v_k\}_{k\in\mathbb{N}}$ are strictly positive and form an independent and identically distributed sequence with common Borel distribution ν on $\mathbb{R}_+$. Assume that the mean $\langle \chi, \nu \rangle \in (0,\infty)$ and let $\beta = \langle \chi, \nu \rangle^{-1}$. Define the traffic intensity $\rho = \alpha/\beta$.

It will be convenient to combine the stochastic primitives into a single, measure-valued load process.

Definition 2.1. The load process is given by

$$\mathcal{V}(t) = \sum_{k=1}^{E(t)} \delta_{v_k}, \quad \text{for } t \in [0, \infty).$$

Then, $\mathcal{V}(\cdot) \in D([0, \infty), \mathbf{M})$.

2.1.3. Evolution of the residual service times. In an SRPT queue, the smallest nonzero residual service time decreases at rate one until it becomes zero or a job arrives that has a smaller initial service time, at which time the rate changes to zero and the new smallest nonzero residual service time begins decreasing at rate one. We adopt the convention that in case of a tie, the residual service time of the job that arrived first (that is, the job with smaller index) begins decreasing at rate one. These dynamics are captured by the unique solution to the following set of equations.

For $x, y \in \mathbb{R}_+$, let $\varphi(x, y) = 1$ if $x = 0$ and $y = 1$, and zero otherwise. For $j \in \mathbb{N}$, let

$$w_j = \begin{cases} \tilde{v}_j, & 1 \le j \le Z(0), \\ v_{j-Z(0)}, & j > Z(0). \end{cases} \tag{4}$$

For $t \in [0, \infty)$, let $\mathcal{Z}_0(t) = \mathbf{0}$ and, for all $t \in [0, \infty)$ and $j \in \mathbb{N}$, define

$$w_j(t) = w_j - \int_{T_j}^{T_j \vee t} \varphi(\langle 1_{[0, w_j(s))}, \mathcal{Z}_{A(s)}(s) \rangle, \langle 1_{[0, w_j(s)]}, \mathcal{Z}_j(s) \rangle)\, ds, \tag{5}$$

$$\mathcal{Z}_j(t) = \sum_{\ell=1}^{j} \delta^+_{w_\ell(t)}. \tag{6}$$

Because $Z(0) < \infty$ and $E(t) < \infty$ for all $t \in [0, \infty)$ almost surely, Equations (4)–(6) have a unique right-continuous solution $\{w_j(\cdot)\}_{j \in \mathbb{N}}$ almost surely. For $j \in \mathbb{N}$ and $t \in [0, \infty)$, $w_j(t)$ is the residual service time at time t of job j.

The unique solution of (4)–(6) satisfies the following properties. First, because $\varphi(\cdot, \cdot) \ge 0$, $w_j(\cdot)$ is continuous and nonincreasing for each $j \in \mathbb{N}$. Furthermore, for all $j \in \mathbb{N}$, $0 \le w_j(t) \le w_j$ for all $t \in [0, \infty)$. Also, for each $j \in \mathbb{N}$, $\mathcal{Z}_j(\cdot) \in D([0, \infty), \mathbf{M})$, and $\mathcal{Z}_j(\cdot)$ does not charge the origin at any time.

We adopt the following terminology and definitions. For $j \in \mathbb{N}$, we say that job j is in the system at time $t \in [0, \infty)$ if $t \ge T_j$ and $w_j(t) > 0$. Hence, if there are no jobs in the system at time t, $\mathcal{Z}_{A(t)}(t) = \mathbf{0}$ and the system is empty at time t. For $t \in [0, \infty)$ and $j \in \mathbb{N}$, let

$$\phi_j(t) = \begin{cases} 0, & \text{if } t \in [0, T_j), \\ \varphi(\langle 1_{[0, w_j(t))}, \mathcal{Z}_{A(t)}(t) \rangle, \langle 1_{[0, w_j(t)]}, \mathcal{Z}_j(t) \rangle), & \text{if } t \in [T_j, \infty). \end{cases} \tag{7}$$

For $t \in [0, \infty)$ and $j \in \mathbb{N}$, we refer to $\phi_j(t)$ as the *instantaneous rate of service* allocated to job j at time t. If the system is not empty at time $t \in [0, \infty)$, $\phi_j(t) = 1$ for exactly one j such that $1 \le j \le A(t)$, and is zero for all other indices. Given $t \in [0, \infty)$ such that the system is not empty at time t, we refer to $1 \le j \le A(t)$ such that $\phi_j(t) = 1$ as the *job in service* at time t. Thus, whenever the system is not empty, there is exactly one job in service and the server is *busy*. Otherwise, the system is empty, no jobs are in service, and the server is *idle*. Once a job is in service, it remains in service until either its residual service time reaches zero or a job enters the system with initial service time strictly smaller than the residual service time of the job in service. Specifically, if $\phi_j(s) = 1$ for some $s \in [0, \infty)$, then by right continuity, there exists a time $t > s$ such that $\phi_j(u) = 1$ for all $u \in [s, t)$. We say that a nonempty time interval $[s, t) \subset [0, \infty)$ is a *busy period* if for each $u \in [s, t)$, there exists $1 \le j \le A(u)$ such that $\phi_j(u) = 1$. Finally, if, at time $t \in [0, \infty)$, $0 < w_i(t) < w_j(t)$ for some $1 \le i, j \le A(t)$, then because $w_j(\cdot)$ is continuous and $w_i(\cdot)$ is nonincreasing, $\phi_j(s) = 0$ for all $s \in [t, D_i)$, where $D_i = \inf\{u \in [T_i, \infty) : w_i(u) = 0\}$. That is, job j cannot enter or resume service until job i departs the system.

2.1.4. Measure-valued state descriptor. For $t \in [0, \infty)$, define the state descriptor by

$$\mathcal{Z}(t) = \sum_{j=1}^{A(t)} \delta^+_{w_j(t)}. \tag{8}$$

Note that $\mathcal{Z}(t) = \mathcal{Z}_{A(t)}(t)$ for all $t \in [0, \infty)$.

2.2. Fluid model. The fluid model has two parameters: $\alpha \in (0, \infty)$ and a Borel probability measure ν on $\mathbb{R}_+$ that does not charge the origin and satisfies $\langle \chi, \nu \rangle < \infty$. These parameters are limits of parameters in the stochastic model, where α corresponds to the rate at which jobs arrive to the system and the probability measure ν corresponds to the distribution of the i.i.d. service times for those jobs. The *traffic intensity parameter* ρ is given by $\rho = \alpha/\beta$, where $\beta = 1/\langle \chi, \nu \rangle$. The pair (α, ν) is referred to as the data for the fluid model. The adjectives *strictly subcritical*, *subcritical*, *critical*, *supercritical*, and *strictly supercritical* are used to refer to data that satisfy $\rho < 1$, $\rho \le 1$, $\rho = 1$, $\rho \ge 1$, and $\rho > 1$, respectively.

Given a measure-valued function ζ: $[0, \infty) \to \mathbf{M}$, for each $t \in [0, \infty)$, let

$$l(\zeta, t) = \sup\{x \in \mathbb{R}_+ \colon \langle 1_{[0,x)}, \zeta(t) \rangle = 0\}, \tag{9}$$

which is the infimum of the support of $\zeta(t)$. Note that for $t \in [0, \infty)$, $l(\zeta, t)$ equals infinity if $\zeta(t) = \mathbf{0}$ and equals zero if $\langle 1_{[0,x)}, \zeta(t) \rangle > 0$ for all $x \in (0, \infty)$. When it is understood which measure-valued function $\zeta(\cdot)$ is under consideration, the dependence on $\zeta(\cdot)$ is suppressed by using the abbreviated notation $l(t)$ in place of $l(\zeta, t)$. We refer to $l(\cdot)$ as the *left edge* of the measure-valued function $\zeta(\cdot)$.

Definition 2.2. Let (α, ν) be fluid model data and let $\xi \in \mathbf{M}_0$. A measure-valued function ζ: $[0, \infty) \to \mathbf{M}$ is a fluid model solution for the data (α, ν) and the initial measure ξ if each of the following hold:

(C1) $\zeta(\cdot)$ is right continuous;

(C2) for all $t \in [0, \infty)$,

$$\langle \chi, \zeta(t) \rangle = [\langle \chi, \xi \rangle + (\rho - 1)t]^+; \tag{10}$$

(C3) for all $t \in [0, \infty)$ and for all $g \in C_b^+(\mathbb{R}_+)$,

$$\langle g 1_{(l(t), \infty)}, \xi + \alpha t \nu \rangle \le \langle g, \zeta(t) \rangle \le \langle g 1_{[l(t), \infty)}, \xi + \alpha t \nu \rangle, \tag{11}$$

where $[\infty, \infty) = (\infty, \infty) = \varnothing$.

Note that the upper bound in (11) implies that $\zeta(t)$ cannot have an atom at zero. That is,

$$\sup_{t \in [0, \infty)} \langle 1_{\{0\}}, \zeta(t) \rangle = 0. \tag{12}$$

This is immediate if $l(t) > 0$ and follows by bounded convergence if $l(t) = 0$ because neither ξ nor ν charges the origin. Together with (10) and (11), this implies that $\zeta(0) = \xi$.

Condition (C1) is natural because a fluid model solution can be viewed as a formal functional law of large numbers approximation of the measure-valued state descriptor for the stochastic model, which is right continuous. Condition (C2) is the standard workload equation and is also natural because the SRPT policy is work conserving.

Condition (C3) is specific to SRPT. It implies that for each $t \in [0, \infty)$, $\zeta(t)$ has no support below $l(t)$ and agrees with the measure $\xi + \alpha t \nu$ above $l(t)$. If we intuitively regard the fluid model as a deterministic system that receives αt units of mass during each time interval $(0, t]$, where arriving mass is instantaneously distributed over $\mathbb{R}_+$ according to the distribution ν and processed according to the SRPT discipline, then (C3) can be interpreted as follows. Mass arriving below level $l(t)$ at time t is instantaneously flushed out of the system, while mass arriving above $l(t)$ by time t receives no processing by time t. Hence, mass that is at $l(t)$ at time t is being processed at time t. This reflects the fact that in an SRPT queue, jobs with the shortest remaining processing time are served first. The inequalities allow for the possibility of atoms in ξ and ν. In particular, discrete distributions are also included in this analysis.

3. Results.

3.1. Characterization of the left edge. Analysis of the left edge of fluid model solution depends on some details of the relationship between the data (α, ν) and the initial measure ξ. For data (α, ν), let

$$x_1 = \sup\{x \in \mathbb{R}_+ \colon \alpha \langle \chi 1_{[0,x]}, \nu \rangle < 1\}, \tag{13}$$

$$x_2 = \inf\{x \in \mathbb{R}_+ \colon \alpha \langle \chi 1_{[0,x]}, \nu \rangle > 1\}. \tag{14}$$

Because ν does not charge the origin, then $x_1 > 0$. In addition, $x_1 \le x_2$. If $\rho < 1$ or $\rho = 1$ and ν has unbounded support, then $x_1 = x_2 = \infty$. If $\rho > 1$ or $\rho = 1$ and ν has bounded support, then $x_1 < \infty$. In this case, it is possible that $\alpha \langle \chi 1_{[0,x_1)}, \nu \rangle < 1$ because ν may have an atom at x_1. In fact, if $\rho > 1$, it may be the case that

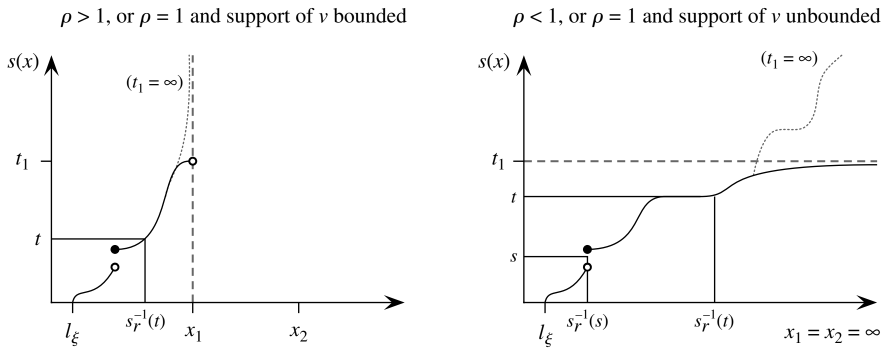

FIGURE 1. Possible relationships among (α, ν), ξ, x_1, x_2, t_1, and $s(\cdot)$ when $\xi \in \mathbf{M}_1$.

$\alpha\langle \chi 1_{[0, x_1]}, \nu\rangle > 1$. In such a case, $x_1 = x_2$. It is also possible that $\alpha\langle \chi 1_{[0, x_1]}, \nu\rangle = 1$. Then, if $\langle 1_{(x_1, x_1+\epsilon)}, \nu\rangle = 0$ for some $\epsilon > 0$, $x_1 < x_2$. In particular, if $\rho = 1$, then $x_2 = \infty$.

For $\xi \in \mathbf{M}$, define the left edge of ξ by

$$l_\xi = \sup\{x \in \mathbb{R}_+ \colon \langle 1_{[0, x)}, \xi\rangle = 0\}. \tag{15}$$

For data (α, ν), let

$$\mathbf{M}_1 = \{\xi \in \mathbf{M}_0 \colon \langle 1_{[0, x_1)}, \xi\rangle > 0\} = \{\xi \in \mathbf{M}_0 \colon l_\xi < x_1\}. \tag{16}$$

When $x_1 = \infty$, $\mathbf{M}_1$ is simply $\mathbf{M}_0$ without $\mathbf{0}$. For data (α, ν), let

$$\mathbf{M}_2 = \{\xi \in \mathbf{M}_0 \colon \xi \neq \mathbf{0} \text{ and } \langle 1_{[0, x_1)}, \xi\rangle = 0\} = \{\xi \in \mathbf{M}_0 \colon x_1 \leq l_\xi < \infty\}. \tag{17}$$

For data (α, ν) and $\xi \in \mathbf{M}_0$, let

$$s(x) = \frac{\langle \chi 1_{[0, x]}, \xi\rangle}{1 - \alpha\langle \chi 1_{[0, x]}, \nu\rangle} \quad \text{for all } x \in [0, x_1), \tag{18}$$

and define

$$t_1 = \sup\{s(x) \colon x \in [0, x_1)\}. \tag{19}$$

Note that $s(\cdot)$ is nondecreasing and right continuous, and may achieve the value t_1 for some $x \in [0, x_1)$. For example, $s(\cdot)$ achieves the value t_1 for some $x \in [0, x_1)$ if $\rho < 1$ and the union of the supports of ν and ξ is bounded. Also, note that $t_1 < \infty$ if and only if $\rho < 1$, or $\rho \geq 1$ and either $\xi \notin \mathbf{M}_1$ or $x_1 < \infty$ and $\alpha\langle \chi 1_{[0, x_1)}, \nu\rangle < 1$, that is, ν has an atom at x_1. Figure 1 illustrates some of the possible relationships among (α, ν), ξ, x_1, x_2, t_1, and $s(\cdot)$.

Let $s_r^{-1} \colon [0, t_1) \to \mathbb{R}_+$ be the right-continuous inverse of $s(\cdot)$ on $[0, t_1)$, which is given by

$$s_r^{-1}(t) = \inf\{x \in [0, x_1) \colon s(x) > t\} \quad \text{for all } t \in [0, t_1). \tag{20}$$

If $t_1 < \infty$, then for convenience, we extend $s_r^{-1}(\cdot)$ to be defined on all of $[0, \infty)$ by letting

$$s_r^{-1}(t) = \begin{cases} x_1, & \text{if } \xi \in \mathbf{M}_1 \text{ and } t \in [t_1, \infty), \\ l_\xi, & \text{if } \xi \notin \mathbf{M}_1 \text{ and } t = 0, \\ x_2 \wedge l_\xi, & \text{if } \xi \notin \mathbf{M}_1 \text{ and } t \in (0, \infty). \end{cases} \tag{21}$$

Note that if $\xi \notin \mathbf{M}_1$, then $t_1 = 0$. Also, note that when $t_1 < \infty$, $s_r^{-1}(t)$ is not necessarily finite for $t \in [t_1, \infty)$. For example, if either $\rho < 1$ or $\rho = 1$ and $\xi = \mathbf{0}$, then $t_1 < \infty$ and $s_r^{-1}(t) = \infty$ for all $t \in [t_1, \infty)$.

The following result characterizes the left-edge dynamics of fluid model solutions. It is proved in §4.1.

THEOREM 3.1. *Let (α, ν) be fluid model data and $\xi \in \mathbf{M}_0$. If $\zeta(\cdot)$ is a fluid model solution for the data (α, ν) and initial measure ξ, then $l(\zeta, t) = s_r^{-1}(t)$ for all $t \in [0, \infty)$.*

REMARK 3.1. In light of Theorem 3.1, for $x \in [0, x_1)$, $s(x)$ can be viewed as the fluid analog of the waiting time for a job of size x that is in the system at time 0. Furthermore, because service times become negligible on fluid scale, the fluid analog of the waiting time is synonymous with the fluid analog of the response time. Therefore, $s(\cdot)$ can also be viewed as the fluid analog of the response time.

3.2. Left-edge asymptotics for unbounded support. One corollary of Theorem 3.1 is that if (α, ν) are critical data and the support of ν is unbounded, then all fluid model solutions for all initial measures converge to the zero measure as time tends to infinity (see Corollary 3.1). This happens despite the fact that the first moment, or workload, is constant for the critical fluid model (cf. (C2)). In particular, for such data, the unique invariant state is the zero measure and so the invariant manifold consists of a single state. This presents challenges for developing a diffusion limit via state-space collapse arguments.

Corollary 3.1. *Let (α, ν) be critical fluid model data such that $x_1 = \infty$ and let $\xi \in \mathbf{M}_0$. If $\zeta(\cdot)$ is a fluid model solution for the data (α, ν) and initial measure ξ, then $\zeta(t) \xrightarrow{w} \mathbf{0}$ as $t \to \infty$.*

Proof. If $\xi = \mathbf{0}$, then by (C2), $\langle \chi, \zeta(t) \rangle = 0$ for all $t \in [0, \infty)$. Hence, by (12), $\zeta(\cdot) \equiv \mathbf{0}$ and the result follows. Otherwise, $\xi \neq \mathbf{0}$. Because $x_1 = \infty$, (13) and Lemma 4.2(iii) below imply that $t_1 = \infty$ and $\lim_{t\to\infty} s_r^{-1}(t) = \infty$. By Theorem 3.1, for $t \in [0, \infty)$ such that $s_r^{-1}(t) > 0$,

$$\langle 1, \zeta(t) \rangle = \langle 1_{[s_r^{-1}(t), \infty)}, \zeta(t) \rangle \leq \frac{\langle \chi 1_{[s_r^{-1}(t), \infty)}, \zeta(t) \rangle}{s_r^{-1}(t)} = \frac{\langle \chi, \xi \rangle}{s_r^{-1}(t)},$$

where the final equality is by (C2). Letting t tend to infinity in the above inequality completes the proof. □

The asymptotic behavior of critical fluid model solutions for data with $x_1 < \infty$ is addressed in Corollary 3.2. At this point, it is instructive to look at a few examples to see how the critical fluid model empties under different distributional assumptions. In particular, we demonstrate that the asymptotic behavior of the left-edge $l(\cdot)$ depends on the tail behavior of the distribution ν.

Example 3.1. Let $\xi \in \mathbf{M}_0$ be such that $\xi \neq \mathbf{0}$. Suppose that ν is an exponential distribution with rate $\lambda > 0$ and let $\alpha = \lambda$. Then, $\rho = 1$, $x_1 = \infty$, and, for $x \in \mathbb{R}_+$,

$$1 - \alpha \langle \chi 1_{[0,x]}, \nu \rangle = \lambda \int_x^\infty \lambda y e^{-\lambda y}\, dy = (\lambda x + 1) e^{-\lambda x}.$$

Therefore,

$$s(x) = \langle \chi 1_{[0,x]}, \xi \rangle \frac{e^{\lambda x}}{\lambda x + 1}, \qquad x \in \mathbb{R}_+.$$

Given $\epsilon \in (0, \lambda)$ and $\delta \in (0, 1)$, let $y \in \mathbb{R}_+$ be such that $e^{\epsilon x} > \lambda x + 1$ and $\langle \chi 1_{[0,x]}, \xi \rangle \geq \delta \langle \chi, \xi \rangle$ for all $x \geq y$. Then,

$$\delta \langle \chi, \xi \rangle e^{(\lambda - \epsilon)x} \leq s(x) \leq \langle \chi, \xi \rangle e^{\lambda x} \quad \text{for all } x \geq y.$$

The left-edge $l(\zeta, \cdot)$ of any fluid model solution $\zeta(\cdot)$ for data (α, ν) and initial measure ξ thus satisfies

$$\frac{1}{\lambda} \ln\left(\frac{t}{\langle \chi, \xi \rangle} \right) \leq l(\zeta, t) \leq \frac{1}{\lambda - \epsilon} \ln\left(\frac{t}{\delta \langle \chi, \xi \rangle} \right) \quad \text{for all } t \in [\langle \chi, \xi \rangle e^{\lambda y}, \infty).$$

Because $\epsilon \in (0, \lambda)$ is arbitrary,

$$l(\zeta, t) \sim \frac{1}{\lambda} \ln(t) \quad \text{as } t \to \infty.$$

Example 3.2. Let $b, k > 0$. Suppose that ν has Pareto density $f(x) = (k+1)b^{k+1}/x^{k+2}$ for $x \geq b$, which has mean $(k+1)b/k$. Let $\alpha = k/(k+1)b$. Then, $\rho = 1$, $x_1 = \infty$, and, for $x \geq b$,

$$1 - \alpha \langle \chi 1_{[0,x]}, \nu \rangle = k b^k \int_x^\infty y^{-k-1}\, dy = \left(\frac{b}{x} \right)^k.$$

Therefore,

$$s(x) = \langle \chi 1_{[0,x]}, \xi \rangle \left(\frac{x}{b} \right)^k \quad \text{for } x \geq b.$$

Let $\delta \in (0, 1)$ and let $y > b$ be such that $\langle \chi 1_{[0,y]}, \xi \rangle \geq \delta \langle \chi, \xi \rangle$. This implies that the left-edge $l(\zeta, \cdot)$ of any fluid model solution $\zeta(\cdot)$ for data (α, ν) and initial measure ξ satisfies

$$b \left(\frac{t}{\langle \chi, \xi \rangle} \right)^{1/k} \leq l(\zeta, t) \leq b \left(\frac{t}{\delta \langle \chi, \xi \rangle} \right)^{1/k}, \qquad t \in [\langle \chi, \xi \rangle (y/b)^k, \infty).$$

Because $\delta \in (0, 1)$ is arbitrary,

$$l(\zeta, t) \sim b \left(\frac{t}{\langle \chi, \xi \rangle} \right)^{1/k} \quad \text{as } t \to \infty.$$

These and an additional example are presented in Down et al. [4], with further discussion on the relationship between the expressions for the left edge and response times.

3.3. Explicit characterization of fluid model solutions. The definition of fluid model solutions given in §2.2 arises naturally by considering the limiting dynamics of the SRPT policy under fluid scaling of the stochastic model. However, this definition only determines fluid model solutions implicitly, as there is mutual dependence between $\zeta(t)$ and $l(t)$ in (11). Although our main interest is in the left-edge $l(t)$, it is important to have an explicit characterization of the entire fluid model solution $\zeta(t)$. This characterization is the key to proving existence and uniqueness of fluid model solutions, and also has some useful consequences described next.

THEOREM 3.2. *Let (α, ν) be fluid model data and $\xi \in \mathbf{M}_0$. A fluid model solution for the data (α, ν) and initial measure ξ exists and is unique. The unique solution $\zeta(\cdot)$ satisfies the following. For each $t \in [0, \infty)$ and $g \in C_b^+(\mathbb{R}_+)$,*

$$\langle g, \zeta(t)\rangle = \begin{cases} g(s_r^{-1}(t))a(t) + \langle g1_{(s_r^{-1}(t),\infty)}, \xi + \alpha t\nu\rangle, & \text{if } s_r^{-1}(t) < \infty, \\ 0, & \text{if } s_r^{-1}(t) = \infty, \end{cases} \tag{22}$$

where for each $t \in [0, \infty)$ such that $s_r^{-1}(t) < \infty$,

$$a(t) = \begin{cases} 0, & \text{if } s_r^{-1}(t) = 0, \\ \dfrac{1}{s_r^{-1}(t)}[\langle \chi 1_{[0, s_r^{-1}(t)]}, \xi + \alpha t\nu\rangle - t], & \text{if } s_r^{-1}(t) > 0. \end{cases} \tag{23}$$

Theorem 3.2 is proved in §4.2.

In the case of critical fluid model data (α, ν) such that ν has bounded support (that is, $x_1 < \infty$), $\mathbf{M}_2$ is nonempty. Furthermore, as a result of Theorem 3.2, each measure in $\mathbf{M}_2$ is an invariant state. In particular, the set of invariant states is $\mathbf{M}_2 \cup \{\mathbf{0}\}$. Corollary 3.2 of Theorem 3.2 states that fluid model solutions with any initial measure $\xi \in \mathbf{M}_0$ converge to the set of invariant states, and explicitly identifies the limiting invariant state.

COROLLARY 3.2. *Let (α, ν) be critical fluid model data such that $x_1 < \infty$ and let $\xi \in \mathbf{M}_0$. If $\zeta(\cdot)$ is a fluid model solution for the data (α, ν) and initial measure ξ, then, for all $g \in C_b^+(\mathbb{R}_+)$,*

$$\lim_{t\to\infty}\langle g, \zeta(t)\rangle = \frac{g(x_1)\langle \chi 1_{[0,x_1)}, \xi\rangle}{x_1} + \langle g1_{[x_1,\infty)}, \xi\rangle$$

and so the limiting invariant state is $\langle \chi 1_{[0,x_1)}, \xi\rangle x_1^{-1}\delta_{x_1} + 1_{[x_1,\infty)}\xi$.

Here, $1_{[x_1,\infty)}\xi$ is the measure in $\mathbf{M}$ such that $\langle g, 1_{[x_1,\infty)}\xi\rangle = \langle 1_{[x_1,\infty)}g, \xi\rangle$ for all bounded continuous functions g. Corollary 3.2 is proved in §4.3 as is Corollary 3.3 of Theorem 3.2.

COROLLARY 3.3. *Let (α, ν) be fluid model data and $\xi \in \mathbf{M}_0$. The unique fluid model solution $\zeta(\cdot)$ for data (α, ν) and initial measure ξ is continuous.*

3.4. Fluid limit theorem. This section presents the limit theorem that rigorously justifies the fluid model discussed above as an approximation of the original stochastic model. We first define a sequence of systems over which the limit is taken. Let $\mathcal{R}$ be a sequence of positive real numbers increasing to infinity. Consider an $\mathcal{R}$-indexed sequence of stochastic models, each defined as in §2.1. For each $r \in \mathcal{R}$, there is an initial condition $\mathcal{Z}^r(0)$; stochastic primitives $E^r(\cdot)$ and $\{v_k^r\}_{k\in\mathbb{N}}$ with parameters α^r, ν^r, β^r, and ρ^r and an arrival process $A^r(\cdot)$ with arrival times $\{T_j^r\}_{j\in\mathbb{N}}$; a corresponding measure-valued load process $\mathcal{V}^r(\cdot)$; and a state descriptor $\mathcal{Z}^r(\cdot)$. The stochastic elements of each model are defined on a probability space $(\Omega^r, \mathcal{F}^r, \mathbf{P}^r)$ with expectation operator $\mathbf{E}^r$. A fluid scaling (or law of large numbers scaling) is applied to each model in the $\mathcal{R}$-indexed sequence as follows. For each $r \in \mathcal{R}$ and $t \in [0, \infty)$, let

$$\bar{E}^r(t) = \frac{1}{r}E^r(rt), \qquad \bar{\mathcal{V}}^r(t) = \frac{1}{r}\mathcal{V}^r(rt), \qquad \text{and} \qquad \bar{\mathcal{Z}}^r(t) = \frac{1}{r}\mathcal{Z}^r(rt). \tag{24}$$

Let $\alpha \in (0, \infty)$ and define $\alpha(t) = \alpha t$ for all $t \in [0, \infty)$. Let $\nu \in \mathbf{M}_0$ be a probability measure. Then,

$$\langle 1_{\{0\}}, \nu\rangle = 0 \qquad \text{and} \qquad 0 < \langle \chi, \nu\rangle < \infty. \tag{25}$$

Let $\beta = \langle \chi, \nu \rangle^{-1}$ and $\rho = \alpha/\beta$. Define $\rho(t) = \rho t$ for all $t \in [0, \infty)$. For the sequence of exogenous arrival processes, assume that as $r \to \infty$,

$$\bar{E}^r(\cdot) \;\Rightarrow\; \alpha(\cdot). \tag{26}$$

For the sequence of service time distributions, assume that

$$\nu^r \xrightarrow{w} \nu \text{ and } \{\nu^r \colon r \in \mathcal{R}\} \text{ is uniformly integrable.} \tag{27}$$

This implies that $\beta^r \to \beta$ and, thus, $\rho^r \to \rho$ as $r \to \infty$.

For the sequence of fluid-scaled initial conditions $\{\bar{\mathcal{Z}}^r(0)\colon r > 0\}$, assume that as $r \to \infty$,

$$(\bar{\mathcal{Z}}^r(0), \langle \chi, \bar{\mathcal{Z}}^r(0)\rangle) \;\Rightarrow\; (\mathcal{Z}_0, \langle \chi, \mathcal{Z}_0\rangle), \tag{28}$$

where $\mathcal{Z}_0$ is a random measure satisfying

$$\mathbf{P}(\mathcal{Z}_0 \in \mathbf{M}_0) = 1. \tag{29}$$

In particular, $\mathcal{Z}_0$ has finite total mass, finite first moment, and does not charge the origin almost surely. The following result establishes the fluid approximation.

THEOREM 3.3. *Under the asymptotic assumptions* (26)–(29), *the sequence* $\{\bar{\mathcal{Z}}^r(\cdot)\colon r \in \mathcal{R}\}$ *converges in distribution on* $D([0,\infty), \mathbf{M})$ *to a measure-valued process* $\mathcal{Z}^*(\cdot)$ *such that* $\mathcal{Z}^*(0)$ *is equal in distribution to* $\mathcal{Z}_0$, *and almost surely* $\mathcal{Z}^*(\cdot)$ *is a fluid model solution for the data* (α, ν) *and initial condition* $\mathcal{Z}^*(0)$.

PROOF. This follows from Theorems 5.1 and 5.2 in §5 below. □

4. Proofs of properties of fluid model solutions.

4.1. Proof of Theorem 3.1. This section begins with Lemma 4.1, which summarizes some basic properties of the left edge of any fluid model solution. Then, Lemma 4.2 and Propositions 4.1 and 4.2 summarize some basic relations satisfied by $s_r^{-1}(\cdot)$. To conclude, these results are put together to prove Theorem 3.1.

LEMMA 4.1. *Suppose that* $\zeta(\cdot)$ *is a fluid model solution for data* (α, ν) *and initial measure* $\xi \in \mathbf{M}_0$.
(i) *If* $\xi \in \mathbf{M}_1$, *then* $l(t) < x_1$ *for all* $t \in [0, t_1)$ *and* $l(t) = x_1$ *for all* $t \in [t_1, \infty)$.
(ii) *If* $\xi \notin \mathbf{M}_1$, *then* $t_1 = 0$, $l(0) = l_\xi$, *and* $l(t) = x_2 \wedge l_\xi$ *for all* $t \in (0, \infty)$.
(iii) $l(\cdot)$ *is nondecreasing and right continuous for* $t \in [0, t_1)$.

PROOF. First, we derive some basic relationships satisfied by $l(\cdot)$. In (11), take a sequence $\{g_n\}_{n=1}^\infty \subset C_b^+(\mathbb{R}_+)$ such that $g_n \nearrow \chi$ as $n \to \infty$. Then, an application of the monotone convergence theorem yields that for all $t \in [0, \infty)$,

$$\langle \chi 1_{(l(t),\infty)}, \xi + \alpha t \nu\rangle \le \langle \chi, \zeta(t)\rangle \le \langle \chi 1_{[l(t),\infty)}, \xi + \alpha t \nu\rangle. \tag{30}$$

For $t \in [0, \infty)$ such that $\langle \chi, \zeta(t)\rangle > 0$, (30) together with (10) implies that

$$\langle \chi 1_{[0,l(t))}, \xi + \alpha t \nu\rangle \le t \le \langle \chi 1_{[0,l(t)]}, \xi + \alpha t \nu\rangle. \tag{31}$$

Inequalities (31) have an intuitive interpretation as follows. In the fluid limit, t units of work can be completed by time t. Thus, (31) reflects the fact that the jobs with the shortest remaining processing time are served first. However, (31) by itself does not necessarily uniquely determine $l(t)$ for each t because there may be intervals in $\mathbb{R}_+$ that do not intersect the union of the supports of ν and ξ, e.g., when ν and ξ are discrete distributions. In any case, from (31), we obtain that for $t \in [0, \infty)$ such that $\langle \chi, \zeta(t)\rangle > 0$,

$$\langle \chi 1_{[0,l(t))}, \xi\rangle \le t(1 - \alpha\langle \chi 1_{[0,l(t))}, \nu\rangle), \tag{32}$$

$$t(1 - \alpha\langle \chi 1_{[0,l(t)]}, \nu\rangle) \le \langle \chi 1_{[0,l(t)]}, \xi\rangle. \tag{33}$$

Parts (i) *and* (ii), $\rho < 1$: Suppose that $\rho < 1$. Note that in this case, $x_1 = x_2 = \infty$ and $t_1 = \langle \chi, \xi\rangle/(1-\rho) < \infty$. Hence, for $t \in [0, t_1)$, $\langle \chi, \xi\rangle + (\rho - 1)t > 0$ and for $t \in [t_1, \infty)$, $\langle \chi, \xi\rangle + (\rho - 1)t \le 0$. This together with (C2) implies that $\langle \chi, \zeta(t)\rangle > 0$ for all $t \in [0, t_1)$ and $\langle \chi, \zeta(t)\rangle = 0$ for $t \in [t_1, \infty)$. Therefore, $\zeta(t) \ne \mathbf{0}$ for all $t \in [0, t_1)$ and, by (12), $\zeta(t) = \mathbf{0}$ for all $t \in [t_1, \infty)$. Thus, if $\rho < 1$, then $l(t) < \infty$ for $t \in [0, t_1)$ and $l(t) = \infty$ for $t \in [t_1, \infty)$. Hence, (i) and (ii) hold for $\rho < 1$.

Part (i), $\rho \geq 1$: Suppose that $\rho \geq 1$ and fix $\xi \in \mathbf{M}_1$. Then, by (C2), $\langle \chi, \zeta(t) \rangle > 0$ for all $t \in [0, \infty)$. In particular, $l(t) < \infty$ for all $t \in [0, \infty)$ and (32) and (33) hold for all $t \in [0, \infty)$. Fix $t \in [0, \infty)$. If $x_1 = \infty$, then $\rho = 1$ and $t_1 = \infty$. Because $l(t) < \infty$, then $l(t) < x_1$. Thus, (i) holds if $x_1 = \infty$. Otherwise, $x_1 < \infty$. If $l(t) > x_1$, then, by (16), (32), and (13),

$$0 < \langle \chi 1_{[0, x_1)}, \xi \rangle \leq \langle \chi 1_{[0, l(t))}, \xi \rangle \leq t(1 - \alpha \langle \chi 1_{[0, l(t))}, \nu \rangle) \leq 0,$$

which is a contradiction. Then, $l(t) \leq x_1$. It suffices to show that $l(t) = x_1$ if and only if $t_1 < \infty$ and $t \in [t_1, \infty)$. First, suppose that $l(t) = x_1$. By (16) and (32),

$$0 < \langle \chi 1_{[0, x_1)}, \xi \rangle \leq t(1 - \alpha \langle \chi 1_{[0, x_1)}, \nu \rangle),$$

that is, $1 - \alpha \langle \chi 1_{[0, x_1)}, \nu \rangle > 0$. This together with (19) implies that $t_1 < \infty$; thus, using (32),

$$t_1 = \frac{\langle \chi 1_{[0, x_1)}, \xi \rangle}{1 - \alpha \langle \chi 1_{[0, x_1)}, \nu \rangle} = \frac{\langle \chi 1_{[0, l(t))}, \xi \rangle}{1 - \alpha \langle \chi 1_{[0, l(t))}, \nu \rangle} \leq t.$$

This completes the proof of the "only if" direction. For the "if" direction, suppose now that $t_1 < \infty$ and $t \in [t_1, \infty)$. Then, by (19), $1 - \alpha \langle \chi 1_{[0, x_1)}, \nu \rangle > 0$ and

$$t_1 = \frac{\langle \chi 1_{[0, x_1)}, \xi \rangle}{1 - \alpha \langle \chi 1_{[0, x_1)}, \nu \rangle}. \tag{34}$$

Furthermore, because $\rho \geq 1$ and $1 - \alpha \langle \chi 1_{[0, x_1)}, \nu \rangle > 0$, then $x_1 < \infty$. To obtain a contradiction, suppose that $l(t) < x_1$. Then, $s(l(t)) \leq t_1$ by (34) and $t \leq s(l(t))$ by (33) and, hence, $t \leq t_1$. However, $t \in [t_1, \infty)$ so $t = t_1$ and we have $l(t_1) < x_1$. Then, (33) and (34) imply that $s(l(t_1)) = t_1$ and, hence, the union of the supports of ν and ξ does not intersect $(l(t_1), x_1)$. Thus, by (C3), for some $a(t_1) > 0$,

$$\langle g, \zeta(t_1) \rangle = \langle g 1_{[x_1, \infty)}, \xi + \alpha t_1 \nu \rangle + g(l(t_1)) a(t_1) \quad \text{for all } g \in C_b^+(\mathbb{R}_+). \tag{35}$$

Then, by (C2), (35), and monotone convergence,

$$\langle \chi, \xi \rangle + (\rho - 1) t_1 = \langle \chi 1_{[x_1, \infty)}, \xi + \alpha t_1 \nu \rangle + l(t_1) a(t_1)$$

so that

$$\langle \chi 1_{[0, x_1)}, \xi + \alpha t_1 \nu \rangle - t_1 = l(t_1) a(t_1).$$

Then, by (34), $l(t_1) a(t_1) = 0$. However, $a(t_1) > 0$ so $l(t_1) = 0$ and, hence, by (C3) and the fact that neither ν nor ξ charges the origin, $a(t_1) = 0$, which is a contradiction. Hence, $l(t_1) = x_1$ and (i) holds for $\rho \geq 1$.

Part (ii), $\rho \geq 1$: Suppose that $\rho \geq 1$ and fix $\xi \notin \mathbf{M}_1$. Then, by (18) and (19), $t_1 = 0$. If $\rho = 1$ and $\xi = \mathbf{0}$, then, by (C2) and (12), $\zeta(t) = \mathbf{0}$ for all $t \in [0, \infty)$. Thus, $l(t) = \infty$ for all $t \in [0, \infty)$, as desired. Otherwise, either $\rho > 1$ or $\xi \neq \mathbf{0}$. Then, by (C2), $\langle \chi, \zeta(t) \rangle > 0$ for all $t \in (0, \infty)$. In particular, $l(t) < \infty$ and (32) and (33) hold for all $t \in (0, \infty)$. Because $\zeta(0) = \xi$, it is immediate that $l(0) = l_\xi$. Fix $t \in (0, \infty)$. If $l(t) > x_2 \wedge l_\xi$, then the left side of (32) is positive and the right side of (32) is negative, which is a contradiction. Thus, $l(t) \leq x_2 \wedge l_\xi$. If $l(t) < x_1$, then the left side of (33) is positive. However, $\xi \notin \mathbf{M}_1$ so $x_1 \leq l_\xi$. Hence, if $l(t) < x_1$, the right side of (33) is zero, which is a contradiction, and thus so $x_1 \leq l(t)$. If $x_1 = x_2 \wedge l_\xi$, it follows that $l(t) = x_2 \wedge l_\xi$. Otherwise, $x_1 < x_2 \wedge l_\xi$. Then, because $x_2 \wedge l_\xi \leq x_2$, $\alpha \langle \chi 1_{[0, x_2 \wedge l_\xi)}, \nu \rangle = 1$ so that $\alpha \langle \chi 1_{[x_2 \wedge l_\xi, \infty)}, \nu \rangle = \rho - 1$. Suppose that $x_1 \leq l(t) < x_2 \wedge l_\xi$. Then, because $\langle \chi, \zeta(t) \rangle > 0$, by (C2), (9), (C3), and (15),

$$\begin{aligned}
\langle \chi, \xi \rangle + \alpha t \langle \chi 1_{[x_2 \wedge l_\xi, \infty)}, \nu \rangle &= \langle \chi, \xi \rangle + (\rho - 1) t = \langle \chi, \zeta(t) \rangle \\
&= \langle \chi 1_{[l(t), \infty)}, \zeta(t) \rangle = \langle \chi 1_{[l(t), x_2 \wedge l_\xi)}, \zeta(t) \rangle + \langle \chi 1_{[x_2 \wedge l_\xi, \infty)}, \zeta(t) \rangle \\
&= \langle \chi 1_{[l(t), x_2 \wedge l_\xi)}, \zeta(t) \rangle + \langle \chi 1_{[x_2 \wedge l_\xi, \infty)}, \xi \rangle + \alpha t \langle \chi 1_{[x_2 \wedge l_\xi, \infty)}, \nu \rangle \\
&= \langle \chi 1_{[l(t), x_2 \wedge l_\xi)}, \zeta(t) \rangle + \langle \chi, \xi \rangle + \alpha t \langle \chi 1_{[x_2 \wedge l_\xi, \infty)}, \nu \rangle.
\end{aligned}$$

However, then $\langle \chi 1_{[l(t), x_2 \wedge l_\xi)}, \zeta(t) \rangle = 0$, which contradicts (9). Thus, $l(t) \geq x_2 \wedge l_\xi$ and so (ii) holds for $\rho \geq 1$.

Part (iii): Note that $[0, t_1)$ is empty when $t_1 = 0$. Henceforth, we assume $t_1 > 0$. By parts (i) and (ii), we may assume that $\xi \in \mathbf{M}_1$ and thus $l(t) < \infty$ for all $t \in [0, t_1)$. Hence, by (9), it follows that $\langle \chi, \zeta(t) \rangle > 0$ for all $t \in [0, t_1)$. Suppose that there exists $0 \leq s < t < t_1$ such that $l(t) < l(s)$. Then, by (31),

$$t \leq \langle \chi 1_{[0, l(t)]}, \xi + \alpha t \nu \rangle \leq \langle \chi 1_{[0, l(s))}, \xi + \alpha t \nu \rangle \leq s + (t - s) \alpha \langle \chi 1_{[0, l(s))}, \nu \rangle.$$

By part (i), $l(s) < x_1$ and so by (13), $\alpha\langle \chi 1_{[0,l(s))}, \nu\rangle < 1$, which is a contradiction. Thus, $l(\cdot)$ is nondecreasing on $[0, t_1)$. For the verification that $l(\cdot)$ is right continuous on $[0, t_1)$, let

$$z(t, x) = \langle 1_{[0,x]}, \zeta(t)\rangle \quad \text{for all } t \in [0, \infty) \text{ and } x \in \mathbb{R}_+.$$

For all $t \in (0, \infty)$, $z(t, x) = 0$ for $x < l(t)$ and $z(t, x) > 0$ for $x > l(t)$. Because $\zeta(\cdot)$ is right continuous, it follows that for each $t \in [0, t_1)$, $\lim_{s\searrow t} z(s, x) = z(t, x)$ for all $x \in \mathbb{R}_+$ that are continuity points for $z(t, \cdot)$. Because $l(\cdot)$ is nondecreasing on $[0, t_1)$, we have $l(t) \le l(t+)$ for all $t \in [0, t_1)$. Suppose that there exists a $t \in [0, t_1)$ such that $l(t) < l(t+)$. Then, because $l(\cdot)$ is nondecreasing on $[0, t_1)$, $l(t+) \le l(s)$ for all $s \in (t, t_1)$. In particular, $z(s, x) = 0$ for all $s \in (t, t_1)$ and $x \in (l(t), l(t+))$. However, $z(t, x) > 0$ for all $x \in (l(t), l(t+))$. Hence, given $x \in (l(t), l(t+))$ such that x is a continuity point for $z(t, \cdot)$, it follows that

$$0 < z(t, x) = \lim_{s\searrow t} z(s, x) = 0,$$

which is a contradiction (because there are at most countably many points of discontinuity for $z(t, \cdot)$). This completes the proof of part (iii). □

Inequalities (32) and (33) and right continuity of $l(\cdot)$ suggest that $l(\cdot)$ is related to the right-continuous inverse $s_r^{-1}(\cdot)$ of $s(\cdot)$. Lemma 4.2 states some of the basic properties of $s_r^{-1}(\cdot)$ on $[0, t_1)$. For this, we need the following definition:

$$x_0 = \inf\{x \in [0, x_1): s(x) \ge t_1\} \wedge x_1. \tag{36}$$

Lemma 4.2. *Let (α, ν) be fluid model data and $\xi \in \mathbf{M}_0$. Then,*
(i) *$s_r^{-1}(t) < x_1$ for all $t \in [0, t_1)$;*
(ii) *$s_r^{-1}(\cdot)$ is nondecreasing and right continuous on $[0, t_1)$;*
(iii) *$\lim_{t\nearrow t_1} s_r^{-1}(t) = x_0$;*
(iv) *$s(s_r^{-1}(t)) \ge t$ for all $t \in [0, t_1)$;*
(v) *$s(s_r^{-1}(t)-) \le t$ for all $t \in [0, t_1)$;*
(vi) *$\langle 1_{[s_r^{-1}(t), x)}, \xi + \alpha t\nu\rangle > 0$ for all $x \in (s_r^{-1}(t), x_1)$ and all $t \in [0, t_1)$.*

Proof. Property (i) is an immediate consequence of (19) and (20). Property (ii) holds because of (20) and the fact that $s(\cdot)$ is nondecreasing and right continuous on $[0, x_1)$. Property (iii) follows from (i), (ii), and (20). Properties (iv) and (v) follow from (20) and the fact that $s(\cdot)$ is nondecreasing and right continuous on $[0, x_1)$. Property (vi) follows from (20) and right continuity of $s(\cdot)$. To see this, fix $t \in [0, t_1)$. If $s(s_r^{-1}(t)) = t$, then, by (20), $s(x) > t$ for all $x \in (s_r^{-1}(t), x_1)$, which implies (vi). Otherwise, $s(s_r^{-1}(t)) > t$ and by (v), (20), and the fact that $s(\cdot)$ is nondecreasing, $s(\cdot)$ has a discontinuity at $s_r^{-1}(t)$, which corresponds to $\langle 1_{\{s_r^{-1}(t)\}}, \xi + \alpha t\nu\rangle > 0$. □

Proposition 4.1. *Let (α, ν) be fluid model data and $\xi \in \mathbf{M}_0 \backslash \{\mathbf{0}\}$.*
(i) *If either $t \in [0, t_1)$ or $t \in [t_1, \infty)$ and $\rho \ge 1$, then $s_r^{-1}(t) < \infty$.*
(ii) *If $t \in [t_1, \infty)$ and $\rho < 1$, then $s_r^{-1}(t) = \infty$.*

Proof. If $t_1 > 0$ and $t \in [0, t_1)$, then $\xi \in \mathbf{M}_1$ and the result follows from Lemma 4.2(i). Next, suppose that $t_1 < \infty$, $t \in [t_1, \infty)$ and $\rho \ge 1$. Then, by (19), $s(x_1-) < \infty$. Because $\rho \ge 1$, it follows that $x_1 < \infty$. If $\xi \in \mathbf{M}_1$, then, by (21), $s_r^{-1}(t) = x_1 < \infty$. If $\xi \in \mathbf{M}_2$, then $t_1 = 0$ and, by (21), $s_r^{-1}(t) \le l_\xi < \infty$. Hence, (i) holds. If $t_1 < \infty$, $t \in [t_1, \infty)$ and $\rho < 1$. Then, because $\xi \ne \mathbf{0}$, $\xi \in \mathbf{M}_1$ and by (13) and (21), $s_r^{-1}(t) = x_1 = \infty$. Hence, (ii) holds. □

Proposition 4.2. *Let (α, ν) be fluid model data and $\xi \in \mathbf{M}_0$. Then, for all $t \in [0, \infty)$ such that $s_r^{-1}(t) < \infty$,*

$$\langle \chi 1_{[0, s_r^{-1}(t))}, \xi + \alpha t\nu\rangle \le t \le \langle \chi 1_{[0, s_r^{-1}(t)]}, \xi + \alpha t\nu\rangle. \tag{37}$$

Proof. To verify this, fix $t \in [0, \infty)$ such that $s_r^{-1}(t) < \infty$. If $\xi = \mathbf{0}$, then, by (19), $t_1 = 0$ and by (21), $t \in (0, \infty)$ and $x_2 = s_r^{-1}(t) < \infty$. In particular, $\rho > 1$. By (14), $\alpha\langle \chi 1_{[0,x_2)}, \nu\rangle \le 1 \le \alpha\langle \chi 1_{[0,x_2]}, \nu\rangle$. In these inequalities, replace x_2 with $s_r^{-1}(t)$ and multiply by t to obtain (37) for $\xi = \mathbf{0}$.

Otherwise, $\xi \ne \mathbf{0}$ and by Proposition 4.1, either $t \in [0, t_1)$ or $t \in [t_1, \infty)$ and $\rho \ge 1$. If $t_1 > 0$ and $t \in [0, t_1)$, then by Lemma 4.2(i) and (13), the denominator of $s(s_r^{-1}(t))$ is positive. After rearranging terms in the inequalities in Lemma 4.2(iv) and (v), we obtain (37).

If $t_1 < \infty$, $t \in [t_1, \infty)$, and $\rho \ge 1$, then there are two subcases to consider: $\xi \in \mathbf{M}_1$ and $\xi \in \mathbf{M}_2$. If $\xi \in \mathbf{M}_1$, then $t_1 > 0$ by (19) and $s_r^{-1}(t) = x_1$ by (21), i.e., $x_1 < \infty$. Because (37) holds on $[0, t_1)$, Lemma 4.2(iii) implies that

$$\langle \chi 1_{[0,x_0)}, \xi + \alpha t_1\nu\rangle \le t_1 \le \langle \chi 1_{[0,x_1]}, \xi + \alpha t_1\nu\rangle.$$

If $x_0 = x_1$, then we can replace x_0 with x_1 in the above inequalities. If $x_0 < x_1$, then, by (36), $s(x) = t_1$ for all $x \in (x_0, x_1)$. By right continuity of $s(\cdot)$, $s(x_0) = t_1$ so that $\langle \chi 1_{[0,x_0]}, \xi + \alpha t_1 \nu\rangle = t_1$ and $\langle \chi 1_{(x_0,x_1)}, \xi + \alpha t_1 \nu\rangle = 0$. Hence, in both cases ($x_0 = x_1$ or $x_0 < x_1$),

$$\langle \chi 1_{[0,x_1)}, \xi + \alpha t_1 \nu\rangle \le t_1 \le \langle \chi 1_{[0,x_1]}, \xi + \alpha t_1 \nu\rangle.$$

By (13),

$$\langle \chi 1_{[0,x_1)}, \alpha(t-t_1)\nu\rangle \le t - t_1 \le \langle \chi 1_{[0,x_1]}, \alpha(t-t_1)\nu\rangle.$$

Adding the two previous displays and replacing x_1 with $s_r^{-1}(t)$ implies (37) for the first subcase $\xi \in \mathbf{M}_1$.

For the second subcase $\xi \in \mathbf{M}_2$, note that $t_1 = 0$ by (19). Then, (37) is immediate for $t = 0$ because (21) implies that $s_r^{-1}(0) = l_\xi$. For $t \in (0, \infty)$, (21) implies that $s_r^{-1}(t) = x_2 \wedge l_\xi$. Then, $\langle \chi 1_{[0, s_r^{-1}(t))}, \xi\rangle = 0$ and $s_r^{-1}(t) \in [x_1, x_2]$. Hence, by (13) and (14),

$$\alpha\langle \chi 1_{[0,s_r^{-1}(t))}, \nu\rangle \le \alpha\langle \chi 1_{[0,x_2)}, \nu\rangle \le 1 \le \alpha\langle \chi 1_{[0,x_1]}, \nu\rangle \le \alpha\langle \chi 1_{[0,s_r^{-1}(t)]}, \nu\rangle.$$

In these inequalities, multiply by t and add $\langle \chi 1_{[0,s_r^{-1}(t)]}, \xi\rangle$ or $\langle \chi 1_{[0,s_r^{-1}(t))}, \xi\rangle$ as appropriate to obtain (37). □

Proof of Theorem 3.1. Suppose that $\zeta(\cdot)$ is a fluid model solution for the data (α, ν) and initial measure $\xi \in \mathbf{M}_0$. By Lemma 4.1 and (21), $l(t) = s_r^{-1}(t)$ for $t \in [t_1, \infty)$. If $t_1 = 0$, then the proof is complete. Otherwise, $t_1 > 0$ and thus $\xi \in \mathbf{M}_1$. Then, it suffices to show that $l(t) = s_r^{-1}(t)$ for all $t \in [0, t_1)$. By Lemma 4.1(i), $l(t) < \infty$ for all $t \in [0, t_1)$ so that, by (12), $\langle \chi, \zeta(t)\rangle > 0$ for all $t \in [0, t_1)$. Furthermore, by Lemma 4.1(i) and (13), $1 - \alpha\langle \chi 1_{[0,l(t)]}, \nu\rangle > 0$ for all $t \in [0, t_1)$. This together with (32) and (33) implies that

$$\frac{\langle \chi 1_{[0,l(t))}, \xi\rangle}{1 - \alpha\langle \chi 1_{[0,l(t))}, \nu\rangle} \le t \le \frac{\langle \chi 1_{[0,l(t)]}, \xi\rangle}{1 - \alpha\langle \chi 1_{[0,l(t)]}, \nu\rangle} \quad \text{for all } t \in [0, t_1). \tag{38}$$

If $l(t) = 0$, then (38) implies that $t = 0$. Because $s_r^{-1}(0) = 0$, $l(0) = s_r^{-1}(0)$. Otherwise, $l(t) > 0$. By (18), (38) is equivalent to

$$s(l(t)-) \le t \le s(l(t)) \quad \text{for all } t \in [0, t_1). \tag{39}$$

Fix $t \in [0, t_1)$. From (39) and Lemma 4.1(i), for all $\epsilon \in (0, l(t) \wedge (t_1 - t))$, $t + \epsilon \le s(l(t+\epsilon))$ and $s(l(t) - \epsilon) \le t$. Hence, by (20), $l(t) - \epsilon \le s_r^{-1}(t) \le l(t+\epsilon)$ for all $\epsilon \in (0, l(t) \wedge (t_1 - t))$. Then, by Lemma 4.1(iii), it follows that $l(t) \le s_r^{-1}(t) \le l(t+) = l(t)$. Because $t \in [0, t_1)$ was arbitrary, the proof is complete. □

4.2. Proof of Theorem 3.2. In this section, Theorem 3.1 is used to prove Theorem 3.2.

Proof of Theorem 3.2. Given fluid model data (α, ν) and $\xi \in \mathbf{M}_0$ for each $t \in [0, \infty)$, define $\zeta(t)$ to be the unique finite Borel measure that satisfies (22). Note that by (23) and Proposition 4.2, $a(t) \ge 0$ for all $t \in [0, \infty)$ such that $s_r^{-1}(t) < \infty$. Thus, $\zeta(t) \in \mathbf{M}$ for each $t \in [0, \infty)$. Furthermore, $\zeta(0) = \xi$ because $s_r^{-1}(0) = l_\xi$. We begin by proving that $\zeta(\cdot)$ is a fluid model solution, which implies the existence of a fluid model solution that also satisfies (22). For this, we need to verify that $\zeta(\cdot)$ satisfies (C1)–(C3).

We begin by verifying that $\zeta(\cdot)$ satisfies (C1). First, suppose that $\xi \notin \mathbf{M}_1$. Then, it suffices to show that $\zeta(t) \stackrel{w}{\to} \xi$ as $t \to 0$. This is immediate by (21) if $l_\xi \le x_2$. Otherwise, $x_2 < l_\xi$. Then, the support of ξ does not intersect $[0, x_2]$ or (x_2, l_ξ) and the result follows from (22) and (23). For $\xi \in \mathbf{M}_1$, right continuity of $\zeta(\cdot)$ on $[t_1, \infty)$ follows from (21), (22), and (23). For $\xi \in \mathbf{M}_1$, right continuity of $\zeta(\cdot)$ on $[0, t_1)$ follows from Lemma 4.2(ii), (22), and (23) once we show that $a(\cdot)$ is right continuous on $[0, t_1)$. Because $\xi \in \mathbf{M}_1$, $t_1 > 0$. Due to Lemma 4.2(ii) and (23), the only issue is to verify right continuity of $a(t)$ at $t \in [0, t_1)$ such that $s_r^{-1}(t) = 0$. Because $s(0) = 0$, $t \in [0, t_1)$ and Lemma 4.2(iv) imply that $s_r^{-1}(t) = 0$ only if $t = 0$. Therefore, in order to complete the verification of right continuity of $\zeta(\cdot)$, it suffices to show that if $t_1 > 0$ and $s_r^{-1}(0) = 0$, then $\lim_{t \searrow 0} a(t) = 0$. Note that if $t_1 > 0$ and $s_r^{-1}(0) = 0$, then, by right continuity of $s_r^{-1}(\cdot)$, $\lim_{t\searrow 0} s_r^{-1}(t) = 0$. Because neither ξ nor ν charges the origin, it follows that

$$0 \le \liminf_{t\searrow 0} a(t) \le \limsup_{t\searrow 0} a(t) \le \limsup_{t\searrow 0} \frac{s_r^{-1}(t)\langle 1_{[0,s_r^{-1}(t)]}, \xi + \alpha t\nu\rangle}{s_r^{-1}(t)} = 0.$$

Therefore, $\zeta(\cdot)$ is right continuous.

Next, we verify that $\zeta(\cdot)$ satisfies (C2). For this, take a sequence $\{g_n\}_{n=1}^\infty \subset C_b^+(\mathbb{R}_+)$ in (22) such that $g_n \nearrow \chi$ as $n \to \infty$. Then, an application of the monotone convergence theorem yields that for each $t \in [0, \infty)$,

$$\langle \chi, \zeta(t)\rangle = \begin{cases} \langle \chi 1_{[0,s_r^{-1}(t)]}, \xi + \alpha t\nu\rangle - t + \langle \chi 1_{(s_r^{-1}(t),\infty)}, \xi + \alpha t\nu\rangle & \text{if } s_r^{-1}(t) < \infty, \\ 0 & \text{otherwise.} \end{cases} \tag{40}$$

For $t \in [0, \infty)$ such that $s_r^{-1}(t) < \infty$, (40) together with Proposition 4.2 implies that (10) holds at time t. For $t \in [0, \infty)$ such that $s_r^{-1}(t) = \infty$, (22) implies that $\zeta(t) = \mathbf{0}$. Hence, to verify that (10) holds for $t \in [0, \infty)$ such that $s_r^{-1}(t) = \infty$, we must show that $\langle \chi, \xi \rangle + (\rho - 1)t \leq 0$. By Proposition 4.1, there are two cases to consider: $t \in [t_1, \infty)$ and $\rho < 1$, and $\xi = \mathbf{0}$. If $t \in [t_1, \infty)$ and $\rho < 1$, then $x_1 = \infty$ and $t_1 < \infty$. By (18) and (19), $\langle \chi, \xi \rangle + (\rho - 1)t_1 = 0$. Thus, for $t \in [t_1, \infty)$, $\langle \chi, \xi \rangle + (\rho - 1)t \leq 0$ and (10) holds at time t. If $\xi = \mathbf{0}$, then $t_1 = 0$ by (19). For $t = 0$, (10) is immediate because $\xi = \mathbf{0}$. For $t \in (0, \infty)$, $x_2 \wedge l_\xi = s_r^{-1}(t) = \infty$ and thus $\rho \leq 1$. This together with $\xi = \mathbf{0}$ implies that (10) holds at time t. Therefore, (C2) holds.

Next, we verify (C3). For $t \in [0, \infty)$ such that $s_r^{-1}(t) = \infty$, (C3) is immediate. Fix $t \in [0, \infty)$ such that $s_r^{-1}(t) < \infty$. We first show that $l(t) = s_r^{-1}(t)$. For this, by (22), it suffices to show that either $a(t) > 0$ or $\langle 1_{(s_r^{-1}(t), x]}, \xi + \alpha t \nu \rangle > 0$ for all $x > s_r^{-1}(t)$. Because $s_r^{-1}(t) < \infty$, Proposition 4.1 implies that either $t \in [0, t_1)$, or $t \in [t_1, \infty)$ and $\rho \geq 1$. If $t \in [0, t_1)$, then by Lemma 4.2(i) $s_r^{-1}(t) < x_1$. Hence, by (20) and right continuity of $s(\cdot)$, either $s(s_r^{-1}(t)) > t$ or $s(s_r^{-1}(t)) = t$. If $s(s_r^{-1}(t)) > t$, then, by (23), $a(t) > 0$. Otherwise, $s(s_r^{-1}(t)) = t$ (and so $a(t) = 0$). Then, by (20), $s(x) > t$ for all $x \in (s_r^{-1}(t), x_1]$. In particular, $s(x) > s(s_r^{-1}(t))$ for all $x \in (s_r^{-1}(t), x_1]$. Hence, $\langle 1_{(s_r^{-1}(t), x]}, \xi + \alpha t \nu \rangle > 0$ for all $x \in (s_r^{-1}(t), x_1]$. Therefore, if $t \in [0, t_1)$, $l(t) = s_r^{-1}(t)$. Otherwise, $t \in [t_1, \infty)$ and $\rho \geq 1$. Then, by (21), $s_r^{-1}(t) \geq x_1$. Hence, $\alpha \langle \chi 1_{[0, s_r^{-1}(t)]}, \nu \rangle \geq 1$. If $\langle \chi 1_{[0, s_r^{-1}(t)]}, \xi \rangle > 0$ or $\alpha \langle \chi 1_{[0, s_r^{-1}(t)]}, \nu \rangle > 1$, then, by (23), $a(t) > 0$. Otherwise, $\langle \chi 1_{[0, s_r^{-1}(t)]}, \xi \rangle = 0$ and $\alpha \langle \chi 1_{[0, s_r^{-1}(t)]}, \nu \rangle = 1$ (and so $a(t) = 0$). Because $s_r^{-1}(t) \geq x_1$ and $\langle \chi 1_{[0, s_r^{-1}(t)]}, \xi \rangle = 0$, $\xi \notin \mathbf{M}_1$. Hence, by (21), (14), and (15), $\langle 1_{(s_r^{-1}(t), x]}, \xi + \alpha t \nu \rangle > 0$ for all $x > s_r^{-1}(t)$. Thus, if $t \in [t_1, \infty)$ and $\rho \geq 1$, $l(t) = s_r^{-1}(t)$, which completes the proof that $l(t) = s_r^{-1}(t)$.

Having established that $l(t) = s_r^{-1}(t)$ for $t \in [0, \infty)$ such that $s_r^{-1}(t) < \infty$, we return to the proof of (C3) for $t \in [0, \infty)$ such that $s_r^{-1}(t) < \infty$. Again, fix $t \in [0, \infty)$ such that $s_r^{-1}(t) < \infty$. Then, the first inequality in (11) follows from $l(t) = s_r^{-1}(t)$, (22), and the fact that $a(t) \geq 0$ (cf. Proposition 4.2). To verify that the second inequality in (11) holds, it suffices to show that

$$a(t) \leq \langle 1_{\{s_r^{-1}(t)\}}, \xi + \alpha t \nu \rangle. \tag{41}$$

By (23), (41) holds if $s_r^{-1}(t) = 0$. Thus, assume that $s_r^{-1}(t) > 0$. The first inequality in (37) together with (23) implies that

$$\begin{aligned} a(t) &= \frac{1}{s_r^{-1}(t)} [\langle \chi 1_{[0, s_r^{-1}(t)]}, \xi + \alpha t \nu \rangle - t] \\ &\leq \frac{1}{s_r^{-1}(t)} [\langle \chi 1_{[0, s_r^{-1}(t)]}, \xi + \alpha t \nu \rangle - \langle \chi 1_{[0, s_r^{-1}(t))}, \xi + \alpha t \nu \rangle] \\ &= \langle 1_{\{s_r^{-1}(t)\}}, \xi + \alpha t \nu \rangle. \end{aligned}$$

This completes the proof that $\zeta(\cdot)$ is a fluid model solution for (α, ν) and $\xi \in \mathbf{M}_0$.

It remains to prove uniqueness of fluid model solutions. For this, let $\zeta_1(\cdot)$ be a fluid model solution for the data (α, ν) and initial measure $\xi \in \mathbf{M}_0$. Then, by Theorem 3.1, $l_1(\cdot) = l(\zeta_1, \cdot) = s_r^{-1}(\cdot)$. We show that $\zeta_1(\cdot) = \zeta(\cdot)$, where $\zeta(\cdot)$ is given by (22). For this, fix $t \in [0, \infty)$. Because $l_1(t) = s_r^{-1}(t)$, if $s_r^{-1}(t) = \infty$, $\zeta_1(t) = \mathbf{0} = \zeta(t)$. Otherwise, $s_r^{-1}(t) < \infty$. Then, (C3) implies that $\zeta_1(t)$ has no support below $s_r^{-1}(t)$ and agrees with $\xi + \alpha t \nu$ above $s_r^{-1}(t)$. Thus, for some $a_1(t) \in \mathbb{R}_+$,

$$\langle g, \zeta_1(t) \rangle = g(s_r^{-1}(t)) a_1(t) + \langle g 1_{(s_r^{-1}(t), \infty)}, \xi + \alpha t \nu \rangle \quad \text{for all } g \in C_b^+(\mathbb{R}_+). \tag{42}$$

To completely characterize $\zeta_1(t)$, one simply needs to determine $a_1(t)$. To verify uniqueness, we must show that $a_1(t) = a(t)$. If $l_1(t) = s_r^{-1}(t) = 0$, then by (C3), $a_1(t) = 0$ because neither ν nor ξ charges the origin. Otherwise, $0 < l_1(t) = s_r^{-1}(t) < \infty$. Then, take a sequence $\{g_n\}_{n=1}^\infty \subset C_b^+(\mathbb{R}_+)$ in (42) such that $g_n \nearrow \chi$ as $n \to \infty$ to obtain

$$\langle \chi, \zeta_1(t) \rangle = s_r^{-1}(t) a_1(t) + \langle \chi 1_{(s_r^{-1}(t), \infty)}, \xi + \alpha t \nu \rangle. \tag{43}$$

Because $0 < l_1(t) = s_r^{-1}(t) < \infty$, $\langle \chi, \zeta_1(t) \rangle > 0$. Thus, by combining (43) and (C2), one can solve for $a_1(t)$ to verify that $a_1(t) = a(t)$. □

4.3. Proof of Corollaries 3.2 and 3.3. In this section, we first prove convergence to an invariant state for fluid model solutions corresponding to critical data (α, ν) such that ν has bounded support ($x_1 < \infty$) and initial measure $\xi \in \mathbf{M}_0$. Then, for general data (α, ν) and initial measure $\xi \in \mathbf{M}_0$, we prove continuity of fluid model solutions. These are obtained as consequences of Theorem 3.2.

Proof of Corollary 3.2. Fix $g \in C_b^+(\mathbb{R}_+)$. If $\xi \notin \mathbf{M}_1$, then, by (16), $l_\xi \geq x_1$ and thus, by (19), $t_1 = 0$. Because $\rho = 1$, $x_2 = \infty$. Hence, by (21), $s_r^{-1}(t) = l_\xi$ for all $t \in [0, \infty)$. If $l_\xi = \infty$, then $\xi = \mathbf{0}$ and, by Theorem 3.2, $\zeta(t) = \mathbf{0}$ for all $t \in [0, \infty)$ and the result follows. Otherwise, $l_\xi < \infty$. Then, by Theorem 3.2 and the equalities $\langle 1_{(x_1, \infty)}, \nu \rangle = 0$ and $\alpha \langle \chi 1_{[0, l_\xi]}, \nu \rangle = 1$, for all $t \in [0, \infty)$,

$$\langle g, \zeta(t) \rangle = \frac{g(l_\xi) \langle \chi 1_{[0, l_\xi]}, \xi \rangle}{l_\xi} + \langle g 1_{(l_\xi, \infty)}, \xi \rangle = \langle g 1_{[l_\xi, \infty)}, \xi \rangle = \langle g 1_{[x_1, \infty)}, \xi \rangle$$

and the result follows.

Otherwise, $\xi \in \mathbf{M}_1$. Because $x_1 < \infty$, Lemma 4.2(i) and (21) imply that $s_r^{-1}(t) < \infty$ for all $t \in [0, \infty)$. Consider the case $t_1 < \infty$. By Theorem 3.2, (21), and the equalities $\langle 1_{(x_1, \infty)}, \nu \rangle = 0$ and $\alpha \langle \chi 1_{[0, x_1]}, \nu \rangle = 1$, we have for all $t \geq t_1$,

$$\langle g, \zeta(t) \rangle = \frac{g(x_1) \langle \chi 1_{[0, x_1]}, \xi \rangle}{x_1} + \langle g 1_{(x_1, \infty)}, \xi \rangle = \frac{g(x_1) \langle \chi 1_{[0, x_1)}, \xi \rangle}{x_1} + \langle g 1_{[x_1, \infty)}, \xi \rangle.$$

Hence, the result holds if $t_1 < \infty$. Next, consider the case $t_1 = \infty$. By Lemma 4.2(i) and (iii), we have the two equalities

$$\lim_{t \to \infty} \langle g 1_{(s_r^{-1}(t), \infty)}, \xi \rangle = \langle g 1_{[x_1, \infty)}, \xi \rangle,$$

$$\lim_{t \to \infty} \frac{g(s_r^{-1}(t)) \langle \chi 1_{[0, s_r^{-1}(t)]}, \xi \rangle}{s_r^{-1}(t)} = \frac{g(x_1) \langle \chi 1_{[0, x_1)}, \xi \rangle}{x_1}.$$

Hence, by Theorem 3.2 and the fact $\langle 1_{[x_1, \infty)}, \nu \rangle = 0$, it suffices to show that

$$\lim_{t \to \infty} \left[\alpha \langle g 1_{(s_r^{-1}(t), x_1)}, \nu \rangle - \frac{g(s_r^{-1}(t))}{s_r^{-1}(t)} (1 - \alpha \langle \chi 1_{[0, s_r^{-1}(t)]}, \nu \rangle) \right] t = 0.$$

First, note that for all $t \in [0, \infty)$, we have the two inequalities

$$\langle g 1_{(s_r^{-1}(t), x_1)}, \nu \rangle \geq \frac{\inf\{g(x) \colon x \in (s_r^{-1}(t), x_1)\}}{x_1} \langle \chi 1_{(s_r^{-1}(t), x_1)}, \nu \rangle,$$

$$\langle g 1_{(s_r^{-1}(t), x_1)}, \nu \rangle \leq \frac{\sup\{g(x) \colon x \in (s_r^{-1}(t), x_1)\}}{s_r^{-1}(t)} \langle \chi 1_{(s_r^{-1}(t), x_1)}, \nu \rangle.$$

In addition, x_1 is not an atom of ν because $t_1 = \infty$. Thus, for all $t \in [0, \infty)$,

$$\alpha \langle \chi 1_{(s_r^{-1}(t), x_1)}, \nu \rangle = 1 - \alpha \langle \chi 1_{[0, s_r^{-1}(t)]}, \nu \rangle.$$

Thus, for all $t \in [0, \infty)$,

$$\begin{aligned} &\left[\frac{\inf\{g(x) \colon x \in (s_r^{-1}(t), x_1)\}}{x_1} - \frac{g(s_r^{-1}(t))}{s_r^{-1}(t)} \right] [1 - \alpha \langle \chi 1_{[0, s_r^{-1}(t)]}, \nu \rangle] t \\ &\quad \leq \left[\alpha \langle g 1_{(s_r^{-1}(t), x_1)}, \nu \rangle - \frac{g(s_r^{-1}(t))}{s_r^{-1}(t)} (1 - \alpha \langle \chi 1_{[0, s_r^{-1}(t)]}, \nu \rangle) \right] t \\ &\quad \leq \left[\frac{\sup\{g(x) \colon x \in (s_r^{-1}(t), x_1)\}}{s_r^{-1}(t)} - \frac{g(s_r^{-1}(t))}{s_r^{-1}(t)} \right] [1 - \alpha \langle \chi 1_{[0, s_r^{-1}(t)]}, \nu \rangle] t. \end{aligned}$$

Because

$$\lim_{t \to \infty} \left[\frac{\inf\{g(x) \colon x \in (s_r^{-1}(t), x_1)\}}{x_1} - \frac{g(s_r^{-1}(t))}{s_r^{-1}(t)} \right] = 0, \quad \text{and}$$

$$\lim_{t \to \infty} \left[\frac{\sup\{g(x) \colon x \in (s_r^{-1}(t), x_1)\}}{s_r^{-1}(t)} - \frac{g(s_r^{-1}(t))}{s_r^{-1}(t)} \right] = 0,$$

it suffices to show that $0 \leq [1 - \alpha \langle \chi 1_{[0, s_r^{-1}(t)]}, \nu \rangle] t \leq \langle \chi, \xi \rangle$ for all $t \in [0, \infty)$. For this, note that by Lemma 4.2(iv), for all $t \in [0, \infty)$,

$$0 \leq [1 - \alpha \langle \chi 1_{[0, s_r^{-1}(t)]}, \nu \rangle] t = \frac{\langle \chi 1_{[0, s_r^{-1}(t)]}, \xi \rangle t}{s(s_r^{-1}(t))} \leq \langle \chi, \xi \rangle. \quad \square$$

To prove Corollary 3.3, it suffices by (C1) to verify left continuity. For this, it will be necessary to consider the left-continuous inverse $s_\ell^{-1}(\cdot)$ of $s(\cdot)$ defined as follows. Given fluid model data (α, ν) and $\xi \in \mathbf{M}_0$,

$$s_\ell^{-1}(t) = \inf\{x \in [0, x_1): s(x) \geq t\} \quad \text{for all } t \in [0, t_1). \tag{44}$$

For ease of notation, we extend the definition of $s_\ell^{-1}(\cdot)$ to all of $[0, \infty)$. For this, recall the definition of $x_0 = \inf\{x \in [0, x_1): s(x) \geq t_1\} \wedge x_1$ given in (36). If $t_1 < \infty$, then let

$$s_\ell^{-1}(t) = \begin{cases} x_0, & \text{if } t = t_1, \\ x_0, & \text{if } \rho < 1,\ \xi \neq \mathbf{0}, \text{ and } t \in (t_1, \infty), \\ x_1, & \text{if } \rho \geq 1 \text{ or } \xi = \mathbf{0}, \text{ and } t \in (t_1, \infty). \end{cases} \tag{45}$$

Note that $s_\ell^{-1}(\cdot)$ is left continuous on $\mathbb{R}_+$. Also, note that if $\xi \notin \mathbf{M}_1$, then $t_1 = 0$ and $x_0 = 0$.

Lemma 4.3. *Let (α, ν) be fluid model data and $\xi \in \mathbf{M}_0$. Then,*
(i) *$s_\ell^{-1}(t) \leq s_r^{-1}(t)$ for all $t \in [0, \infty)$;*
(ii) *$s_r^{-1}(s) \leq s_\ell^{-1}(t)$ for $s \in [0, t)$, $t \in [0, t_1)$ and, if $t_1 < \infty$, $t = t_1$;*
(iii) *$s_r^{-1}(t-) = s_\ell^{-1}(t)$ for all $t \in (0, t_1)$ and, if $t_1 < \infty$, for $t = t_1$;*
(iv) *$\langle 1_{(s_\ell^{-1}(t), s_r^{-1}(t))}, \xi + \alpha t \nu\rangle = 0$ for all $t \in [0, \infty)$;*
(v) *$s(s_\ell^{-1}(t)) \geq t$ for all $t \in [0, t_1)$;*
(vi) *$\langle \chi 1_{[0, s_r^{-1}(t)]}, \xi + \alpha t \nu\rangle - t = 0$ for all $t \in [0, t_1)$ such that $s_\ell^{-1}(t) < s_r^{-1}(t)$ and, if $t_1 < \infty$ and $s_\ell^{-1}(t_1) < s_r^{-1}(t_1)$, for $t = t_1$ as well.*

Proof. Properties (i)–(iii) and (v) follow directly from the definitions, using Lemma 4.2(iii) for (iii) when $t = t_1$. We now verify (iv). First, consider $t \in [0, t_1)$. Because $s(x) = t$ for all $x \in [s_\ell^{-1}(t), s_r^{-1}(t))$, it is immediate that $\langle 1_{(s_\ell^{-1}(t), s_r^{-1}(t))}, \xi + \alpha t \nu\rangle = 0$. The proof is complete if $t_1 = \infty$. Otherwise, $t_1 < \infty$ and it suffices to consider $t \in [t_1, \infty)$. If $t_1 > 0$, then $\xi \in \mathbf{M}_1$ and $s(x) = t_1$ for all $x \in [x_0, x_1)$ and so $\langle 1_{(x_0, x_1)}, \xi + \alpha t \nu\rangle = 0$, which implies (iv). If $t_1 = 0$, then $\xi \notin \mathbf{M}_1$. Then, we have $s_\ell^{-1}(0) = 0$ and $s_r^{-1}(0) = l_\xi$, which implies (iv) for $t = 0$. For $t \in (0, \infty)$, $s_\ell^{-1}(t) = x_1$ and $s_r^{-1}(t) = x_2 \wedge l_\xi$, which also implies (iv) for $t \in (0, \infty)$. Hence, (iv) holds. To verify (vi), note that for $t \in [0, t_1)$ such that $s_\ell^{-1}(t) < s_r^{-1}(t)$, (v) and Lemma 4.2(v) combine to yield $t \leq s(s_\ell^{-1}(t)) \leq s(s_r^{-1}(t)-) \leq t$. Hence, $s(s_\ell^{-1}(t)) = t$, which in turn implies the desired result. If $t_1 < \infty$ and $s_\ell^{-1}(t_1) < s_r^{-1}(t_1)$, we have $x_0 < x_1$ and $s(x) = t_1$ for all $x \in [x_0, x_1)$ and, in particular, for $x = x_0$. □

Proof of Corollary 3.3. By (C1), it suffices to show that $\zeta(\cdot)$ is left continuous, i.e., that, for all $t \in (0, \infty)$ and $g \in C_b^+(\mathbb{R}_+)$,

$$\lim_{u \nearrow t} \langle g, \zeta(u)\rangle = \langle g, \zeta(t)\rangle. \tag{46}$$

By (22), (23), and (21), (46) holds for $t \in (t_1, \infty)$. Fix $t \in (0, t_1)$ and $g \in C_b^+(\mathbb{R}_+)$. By Lemma 4.3(ii), there are two cases to consider: $s_r^{-1}(u) = s_\ell^{-1}(t)$ for some $u < t$ and $s_r^{-1}(u) < s_\ell^{-1}(t)$ for all $u < t$. We first show that, in both cases,

$$\lim_{u \nearrow t} \langle g, \zeta(u)\rangle = \langle g 1_{(s_\ell^{-1}(t), \infty)}, \xi + \alpha t \nu\rangle + g(s_\ell^{-1}(t)) \frac{1}{s_\ell^{-1}(t)} [\langle \chi 1_{[0, s_\ell^{-1}(t)]}, \xi + \alpha t \nu\rangle - t]. \tag{47}$$

If $s_r^{-1}(u) = s_\ell^{-1}(t)$ for some $u < t$, then because it is nondecreasing, $s_r^{-1}(\cdot)$ is constant on $[u, t)$ by Lemma 4.3(ii). In this case, (47) follows immediately from (22) and (23). Otherwise, (22), (23), and Lemma 4.3(iii) imply that

$$\lim_{u \nearrow t} \langle g, \zeta(u)\rangle = \langle g 1_{[s_\ell^{-1}(t), \infty)}, \xi + \alpha t \nu\rangle + g(s_\ell^{-1}(t)) \frac{1}{s_\ell^{-1}(t)} [\langle \chi 1_{[0, s_\ell^{-1}(t))}, \xi + \alpha t \nu\rangle - t]. \tag{48}$$

Then, (47) follows by subtracting $g(s_\ell^{-1}(t))\langle 1_{\{s_\ell^{-1}(t)\}}, \xi + \alpha t \nu\rangle$ from the first term on the right side of (48) and adding the equivalent expression

$$g(s_\ell^{-1}(t)) \frac{1}{s_\ell^{-1}(t)} \langle \chi 1_{\{s_\ell^{-1}(t)\}}, \xi + \alpha t \nu\rangle$$

to the second term on the right side of (48).

With (47) established, there are two cases to consider: $s_\ell^{-1}(t) = s_r^{-1}(t)$ and $s_\ell^{-1}(t) < s_r^{-1}(t)$. If $s_\ell^{-1}(t) = s_r^{-1}(t)$, then (47) implies (46) immediately by (22) and (23). Otherwise, (47), Lemma 4.3(iv), and Lemma 4.3(vi) imply that

$$\lim_{u \nearrow t} \langle g, \zeta(u)\rangle = \langle g 1_{[s_r^{-1}(t), \infty)}, \xi + \alpha t \nu\rangle$$
$$= \langle g 1_{(s_r^{-1}(t), \infty)}, \xi + \alpha t \nu\rangle + g(s_r^{-1}(t)) \frac{1}{s_r^{-1}(t)} \langle \chi 1_{\{s_r^{-1}(t)\}}, \xi + \alpha t \nu\rangle.$$

Because $s_\ell^{-1}(t) < s_r^{-1}(t)$ and $s(\cdot)$ is nondecreasing, Lemma 4.2(v) and Lemma 4.3(v) imply that $s(s_r^{-1}(t)-) = t$. It follows that $\langle \chi 1_{[0, s_r^{-1}(t))}, \xi + \alpha t \nu\rangle - t = 0$ and thus (46) holds by (22) and (23).

If $t_1 = 0$ or $t_1 = \infty$, the proof is complete. Otherwise, $0 < t_1 < \infty$, and we must show that $\lim_{t \nearrow t_1} \langle g, \zeta(t)\rangle = \langle g, \zeta(t_1)\rangle$. If $s_\ell^{-1}(t_1) = \infty$, then, by Lemma 4.3(i), $s_r^{-1}(t_1) = \infty$ and thus $\zeta(t_1) = \mathbf{0}$ by (22). By Lemma 4.3(iii), $\lim_{t \nearrow t_1} s_r^{-1}(t) = \infty$. This together with (22) and (23) implies (46) for $t = t_1$. Otherwise, $s_\ell^{-1}(t_1) < \infty$ and (47) also holds for $t = t_1$ by the same argument given for $t \in (0, t_1)$. Then (46) follows by an argument similar to that used for $t \in (0, t_1)$. □

5. Proof of fluid limit theorem. Theorem 3.3 is proved in this section. We lay the foundation for the proof in §5.1. In §5.2, we verify tightness. Fluid limit points are characterized in §5.3.

5.1. Foundation for the proof. In this section, we set up the framework necessary for proving tightness and characterizing fluid limit points. We begin by defining some performance processes for the stochastic model. We then establish some relationships among these processes, arising from the dynamics of SRPT. We also state a weak law of large numbers for the measure-valued load process.

5.1.1. Performance processes. Here, we introduce the performance processes that play a key role in the analysis of the stochastic model of an SRPT queue. Denote by $L(\cdot)$ the *left-edge process* of the measure-valued state descriptor $\mathcal{Z}(\cdot)$, which is given by

$$L(t) = \sup\{x \in \mathbb{R}_+ \colon \langle 1_{[0,x)}, \mathcal{Z}(t)\rangle = 0\} \quad \text{for all } t \in [0, \infty). \tag{49}$$

Denote by $C(\cdot)$ the *current residual service time process*, which is given by

$$C(t) = \begin{cases} L(t) & \text{if } \mathcal{Z}(t) \neq \mathbf{0} \\ 0 & \text{otherwise} \end{cases} \quad \text{for all } t \in [0, \infty). \tag{50}$$

Define the *frontier process* $F(\cdot)$ by

$$F(t) = \sup_{0 \le s \le t} C(s) \quad \text{for all } t \in [0, \infty). \tag{51}$$

The *queue length process* is given by

$$Z(t) = \langle 1, \mathcal{Z}(t)\rangle \quad \text{for all } t \in [0, \infty).$$

For each $x \in \mathbb{R}_+$, the *truncated arrival process* $E(\cdot, x)$ and *truncated queue length process* $Z(\cdot, x)$ are, respectively, given by

$$E(t, x) = \langle 1_{[0,x]}, \mathcal{V}(t)\rangle \qquad \text{and} \qquad Z(t, x) = \langle 1_{[0,x]}, \mathcal{Z}(t)\rangle \quad \text{for all } t \in [0, \infty).$$

The *immediate workload process* $W(\cdot)$ is given by

$$W(t) = \langle \chi, \mathcal{Z}(t)\rangle \quad \text{for all } t \in [0, \infty).$$

Three related processes are defined as follows. For $t \in [0, \infty)$, let

$$V(t) = \langle \chi, \mathcal{V}(t)\rangle, \qquad X(t) = W(0) + V(t) - t, \qquad \text{and} \qquad Y(t) = \inf_{0 \le s \le t} X(s).$$

Our analysis of the SRPT queue also exploits properties of the *truncated immediate workload processes* $W(\cdot, x)$, $x \in \mathbb{R}_+$, where for each $x \in \mathbb{R}_+$,

$$W(t, x) = \langle \chi 1_{[0,x]}, \mathcal{Z}(t)\rangle \quad \text{for all } t \in [0, \infty).$$

For each $x \in \mathbb{R}_+$, three more truncated processes are defined as follows. For all $t \in [0, \infty)$,

$$\begin{aligned} V(t,x) &= \langle \chi 1_{[0,x]}, \mathcal{V}(t)\rangle, \qquad X(t,x) = W(0,x) + V(t,x) - t, \\ Y(t,x) &= \inf_{0\le s\le t} X(s,x). \end{aligned} \tag{52}$$

5.1.2. Dynamic inequalities and equations. Next, the SRPT dynamics are used to obtain equations and bounds that the performance processes satisfy. For future reference, these equations and bounds are collected and written out at the end for the fluid-scaled, $\mathcal{R}$-indexed sequence of systems.

5.1.2.1. Queue length.

LEMMA 5.1. *Almost surely, for all $s, t \in [0, \infty)$ such that $s \le t$, all $x, y \in \mathbb{R}_+$ such that $x \le y$, all $a > 0$, and all closed $B \subset \mathbb{R}_+$,*

$$Z(t) \le Z(0) + E(t), \tag{53}$$

$$\langle 1_{[x,\infty)}, \mathcal{Z}(t)\rangle \le \langle 1_{[x,\infty)}, \mathcal{Z}(0) + \mathcal{V}(t)\rangle, \tag{54}$$

$$\langle 1_{[x,y]}, \mathcal{Z}(t)\rangle \le \langle 1_{[x,y]}, \mathcal{Z}(0) + \mathcal{V}(t)\rangle + 1, \tag{55}$$

$$\langle 1_B, \mathcal{Z}(t)\rangle \le \langle 1_B, \mathcal{Z}(s)\rangle + E(t) - E(s) + 1, \tag{56}$$

$$\langle 1_B, \mathcal{Z}(s)\rangle \le \langle 1_B, \mathcal{Z}(t)\rangle + \langle 1_{[0,a]}, \mathcal{Z}(s)\rangle + \frac{t-s}{a} + 1. \tag{57}$$

PROOF. Inequalities (53) and (54) follow from Definition 2.1 and (8) because the residual service times are nonincreasing.

Inequality (55) holds because the residual service times are continuous and nonincreasing, and at most one job with initial service time not in $[x, y]$ has positive residual service in $[x, y]$ at time t. To prove this, fix $t \in [0, \infty)$ and $x, y \in \mathbb{R}_+$. Let $I^t_{x,y} = \{1 \le j \le A(t)\colon w_j \notin [x, y] \text{ and } w_j(t) \in [x, y] \cap (0, \infty)\}$, where $A(\cdot)$, $w_\cdot$ and $w_\cdot(\cdot)$ are defined by (3), (4), and (5), respectively. Then,

$$\langle 1_{[x,y]}, \mathcal{Z}(t)\rangle \le \langle 1_{[x,y]}, \mathcal{Z}(0) + \mathcal{V}(t)\rangle + |I^t_{x,y}|,$$

where $|\cdot|$ denotes cardinality of the set. Suppose that $|I^t_{x,y}| \ge 2$ and let $i, j \in I^t_{x,y}$ with $i < j$. If $w_i(T_j) \le w_j$, then $\phi_j(u) = 0$ for all $u \in [T_j, D_i]$ (see (7) and the discussion at the end of §2.1). Because $D_i > t$, this implies that $w_j(t) = w_j \notin [x, y]$, which is a contradiction. Alternatively, if $w_i(T_j) > w_j$, then $\phi_i(u) = 0$ for all $u \in [T_j, D_j]$. Because $D_j > t$, this implies that $w_i(t) = w_i(T_j) > w_j$. The residual service times are continuous and nonincreasing and thus $w_j > y$, which yields another contradiction. Thus, $|I^t_{x,y}| \le 1$.

Inequality (56) is proved similarly but with a slight modification. It holds because the residual service times are continuous and nonincreasing, and at most one job with residual service time not in B at time s has positive residual service in B at time t. To prove this, fix $s, t \in [0, \infty)$ with $s \le t$ and $B \subset \mathbb{R}_+$. Let $I^{s,t}_B = \{1 \le j \le A(s)\colon w_j(s) \notin B \text{ and } w_j(t) \in B \cap (0, \infty)\}$. Then,

$$\langle 1_B, \mathcal{Z}(t)\rangle \le \langle 1_B, \mathcal{Z}(s)\rangle + E(t) - E(s) + \left|I^{s,t}_B\right|.$$

Suppose that $|I^{s,t}_B| \ge 2$ and let $i, j \in I^{s,t}_B$ with $i < j$. If $w_i(s) \le w_j(s)$, then $\phi_j(u) = 0$ for all $u \in [s, D_i]$. Because $D_i > t$, this implies that $w_j(t) = w_j(s) \notin B$, which is a contradiction. One arrives at the analogous contradiction if $w_i(s) > w_j(s)$. Thus, $|I^{s,t}_B| \le 1$.

For (57), fix $s, t \in [0, \infty)$ with $s \le t$, $a > 0$, $B \subset \mathbb{R}_+$ and let

$$J^{s,t}_B = \{1 \le j \le A(s)\colon w_j(s) \in B \cap (a, \infty) \text{ and } w_j(t) \notin B \cap (0, \infty)\}.$$

Then,

$$\langle 1_B, \mathcal{Z}(s)\rangle \le \langle 1_B, \mathcal{Z}(t)\rangle + \langle 1_{[0,a]}, \mathcal{Z}(s)\rangle + |J^{s,t}_B|.$$

Assume $J_B^{s,t}$ has at least two elements, as the statement is trivial otherwise. Let $w_J = \min\{w_j(s): j \in J_B^{s,t}\}$ and let $i = \min\{j \in J_B^{s,t}: w_j(s) = w_J\}$. Then, for all $j \in J_B^{s,t}\backslash i$, $\phi_j(u) = 0$ for all $u \in [s, s+a]$ because $D_i > s+a$. That is, job $j \in J_B^{s,t}\backslash i$ cannot enter or resume service until job i has exited the system, which requires at least a units of time. Thus, job $j \in J_B^{s,t}\backslash i$ enters service in $(s+a, t)$. Iterating this argument and using the fact that all jobs in $J_B^{s,t}\backslash i$ either enter or resume service in (s,t) implies that the job in $J_B^{s,t}$ to enter or resume service last does so during the nonempty time interval $(s + (|J_B^{s,t}| - 1)a, t)$, which implies that $|J_B^{s,t}| \le (t-s)a^{-1} + 1$. □

5.1.2.2. Immediate workload.

LEMMA 5.2. *Almost surely, for all* $t \in [0, \infty)$,

$$W(t) = X(t) + [Y(t)]^-. \tag{58}$$

Almost surely, for all $t \in [0,\infty)$ *and* $x \in \mathbb{R}_+$,

$$\langle \chi 1_{(x,\infty)}, \mathscr{Z}(t)\rangle \le \langle \chi 1_{(x,\infty)}, \mathscr{Z}(0) + \mathscr{V}(t)\rangle. \tag{59}$$

In addition, almost surely, for all busy periods $[s,t) \subset [0,\infty)$,

$$W(t) = W(s) + V(t) - V(s) - (t-s). \tag{60}$$

Finally, almost surely, for each $x \in \mathbb{R}_+$ *and* $t \in [0,\infty)$,

$$W(t,x) \le W(0,x) + V(t,x) - V(\tau(t,x)-, x) + x - (t - \tau(t,x)), \tag{61}$$

where $\tau(t,x) = \sup\{0 \le s \le t: W(s,x) = 0\}$ *with the supremum of the empty set taken to be zero.*

PROOF. Because the server idles if and only if the system is empty, (58) is merely the classical characterization of the workload process for a work-conserving service discipline. Also, (59) follows from the fact that $\chi(\cdot)$ is nondecreasing and each residual service time is nonincreasing. In addition, (60) follows by subtracting the sum of (5) at time s over $1 \le k \le A(s)$ from the sum of (5) at time t over $1 \le k \le A(t)$, and using the fact that $\sum_{k=1}^{A(u)} \phi_k(u) = 1$ for all $u \in [s,t)$.

To verify (61), we consider three cases. Case 1 is when $W(t,x) = 0$. Then, $\tau(t,x) = t$ and (61) holds. Case 2 is when $W(t,x) > 0$ and $\tau(t,x) < t$. Then, $[\tau(t,x), t)$ is a busy period. For all $s \in [\tau(t,x), t)$, the job in service at time s has residual service time less than or equal to x. Hence, by the same reasoning used to verify (60),

$$W(t,x) = W(\tau(t,x), x) + V(t,x) - V(\tau(t,x), x) - (t - \tau(t,x)).$$

At this point, Case 2 splits into two subcases. Case 2(a) is when $\tau(t,x) = 0$. Then, $W(\tau(t,x),x) = W(0,x)$ and thus, because $V(\tau(t,x)-, x) \le V(\tau(t,x), x)$, (61) holds. Case 2(b) is when $\tau(t,x) > 0$. Then, either there exists $\epsilon > 0$ such that the residual service time of the job in service on the time interval $(\tau(t,x) - \epsilon, \tau(t,x)]$ decreases to x as time approaches $\tau(t,x)$, or jobs arrive at time $\tau(t,x)$ with initial service time in $(0,x]$. Thus, $W(\tau(t,x),x) \le V(\tau(t,x),x) - V(\tau(t,x)-,x) + x$ and (61) holds. Case 3 is when $W(t,x) > 0$ and $\tau(t,x) = t$. Then, by the last argument, $W(t,x) \le V(t,x) - V(t-,x) + x$ and (61) holds. □

5.1.2.3. Behavior above the frontier.

LEMMA 5.3. *Almost surely, for all* $t \in [0,\infty)$,

(i) $\langle 1_{[0,C(t))}, \mathscr{Z}(t)\rangle = 0$;

(ii) *for all* $j \in \mathbb{N}$ *such that* $F(t) < w_j$, $w_j(s) = w_j$ *for all* $s \in [0,t]$.

PROOF. Property (i) follows from (49) and (50). For a proof of property (ii), suppose that there exist $t \in [0,\infty)$ and $j \in \mathbb{N}$ such that $F(t) < w_j$ and $w_j(t) < w_j$ (recall that $w_j(\cdot)$ is nonincreasing). Then, by (5), $t > 0$ and there exists $0 \le a < b \le t$ such that $\phi_j(s) = 1$ for all $s \in [a,b)$ and $w_j(a) = w_j$. Thus, for all $s \in [a,b)$,

$$\langle 1_{[0,w_j(s))}, \mathscr{Z}(s)\rangle = 0.$$

Hence, $L(s) = w_j(s)$ for $s \in [a,b)$ and so $C(s) = w_j(s)$ for $s \in [a,b)$. However, then $F(t) \ge C(a) = w_j$, which is a contradiction. Hence, (ii) holds. □

The following result is an immediate consequence of Lemma 5.3 and (4)–(8).

COROLLARY 5.1. *Almost surely, for all measurable functions* $g: \mathbb{R}_+ \to \mathbb{R}_+$ *and* $t \in [0,\infty)$,

$$\langle g1_{(F(t),\infty)}, \mathscr{Z}(t)\rangle = \langle g1_{(F(t),\infty)}, \mathscr{Z}(0) + \mathscr{V}(t)\rangle, \tag{62}$$

$$\langle g1_{[F(t),\infty)}, \mathscr{Z}(t)\rangle \le \langle g1_{[F(t),\infty)}, \mathscr{Z}(0) + \mathscr{V}(t)\rangle. \tag{63}$$

5.1.2.4. Behavior below the frontier.

LEMMA 5.4. *Almost surely, for each* $t \in [0, \infty)$,

$$\langle \chi 1_{[0, F(t))}, \mathcal{Z}(t) \rangle \le V(t, F(t)) - V(\tau(t)-, F(t)) + F(t) - (t - \tau(t)), \tag{64}$$

where $\tau(t) = \sup\{0 \le s \le t: C(s) = F(s)\} \vee \sup\{0 \le s \le t: W(s) = 0\}$, *with the supremum of the empty set taken to be zero.*

PROOF. Let $t \in [0, \infty)$. If $\tau(t) = t$, (64) is trivial because $\langle \chi 1_{[0, F(t))}, \mathcal{Z}(t) \rangle = 0$ almost surely in that case. Henceforth, we assume that $\tau(t) < t$. Then, $C(s) < F(s)$ for all $s \in (\tau(t), t]$ and $W(s) > 0$ for all $s \in [\tau(t), t]$. Hence, $F(t) = F(\tau(t))$ and $[\tau(t), t)$ is a busy period. Therefore, by (60) and (62) and by combining like terms,

$$\begin{aligned} W(t) &= W(\tau(t)) + V(t) - V(\tau(t)) - (t - \tau(t)) \\ &= W(\tau(t), F(t)) + \langle \chi 1_{(F(t), \infty)}, \mathcal{Z}(0) \rangle + \langle \chi 1_{(F(t), \infty)}, \mathcal{V}(\tau(t)) \rangle \\ &\quad + V(t) - V(\tau(t)) - (t - \tau(t)) \\ &= W(\tau(t), F(t)) + \langle \chi 1_{(F(t), \infty)}, \mathcal{Z}(0) \rangle - V(\tau(t), F(t)) \\ &\quad + V(t, F(t)) + \langle \chi 1_{(F(t), \infty)}, \mathcal{V}(t) \rangle - (t - \tau(t)). \end{aligned}$$

Hence, by subtracting $\langle \chi 1_{(F(t), \infty)}, \mathcal{Z}(0) \rangle + \langle \chi 1_{(F(t), \infty)}, \mathcal{V}(t) \rangle$ from both sides and then using (62), we obtain

$$W(t, F(t)) = W(\tau(t), F(t)) + V(t, F(t)) - V(\tau(t), F(t)) - (t - \tau(t)).$$

We have $W(\tau(t), F(t)) = \langle \chi 1_{[0, F(t))}, \mathcal{Z}(\tau(t)) \rangle + F(t) \langle 1_{\{F(t)\}}, \mathcal{Z}(\tau(t)) \rangle$. Clearly, $\langle \chi 1_{[0, F(t))}, \mathcal{Z}(\tau(t)-) \rangle = 0$. Hence, $\langle \chi 1_{[0, F(t))}, \mathcal{Z}(\tau(t)) \rangle \le V(\tau(t), F(t)) - V(\tau(t)-, F(t))$. Then,

$$\begin{aligned} W(t, F(t)) &\le V(t, F(t)) - V(\tau(t)-, F(t)) \\ &\qquad + F(t) \langle 1_{\{F(t)\}}, \mathcal{Z}(\tau(t)) \rangle - (t - \tau(t)). \end{aligned}$$

If we subtract $F(t) \langle 1_{\{F(t)\}}, \mathcal{Z}(t) \rangle$ from both sides of this inequality, the result follows provided that $\langle 1_{\{F(t)\}}, \mathcal{Z}(\tau(t)) \rangle - \langle 1_{\{F(t)\}}, \mathcal{Z}(t) \rangle \le 1$. Because $C(s) < F(s)$ and $W(s) > 0$ for all $s \in (\tau(t), t]$, $\langle 1_{[0, F(t))}, \mathcal{Z}(s) \rangle > 0$ for all $s \in (\tau(t), t]$. Hence, any job with residual service time equal to $F(t)$ at time $s \in (\tau(t), t]$ is not in service at time s. At time $\tau(t)$, at most one job with residual service time equal to $F(t)$ is in service. Therefore, $\langle 1_{\{F(t)\}}, \mathcal{Z}(\tau(t)) \rangle \le \langle 1_{\{F(t)\}}, \mathcal{Z}(t) \rangle + 1$ and the desired bound follows. □

5.1.2.5. Bounds for the frontier process.

LEMMA 5.5. *Almost surely, for all* $x \in \mathbb{R}_+$,

$$W(t, x) = X(t, x) \quad \textit{for all } t \in [0, \tau_x),$$

where $\tau_x = \sup\{t \in [0, \infty): Y(t, x) \ge 0\}$.

PROOF. For $x \in \mathbb{R}_+$, let $\tilde{\tau}_x = \inf\{s \in [0, \infty): W(s, x) = 0\}$. Note that if $\tilde{\tau}_x = 0$, then, by (52), $\tau_x = 0$ and there is nothing to prove. Suppose $\tilde{\tau}_x > 0$. For all $x \in \mathbb{R}_+$ and $t \in [0, \tilde{\tau}_x)$, $W(t, x) > 0$ and thus $0 < L(t) \le x$. For all $x \in \mathbb{R}_+$ and $t \in [0, \tilde{\tau}_x)$, $C(t) \le x$ and, consequently, $F(t) \le x$. Therefore, by (62), almost surely for all $x \in \mathbb{R}_+$ and $t \in [0, \tilde{\tau}_x)$,

$$W(t) = W(t, x) + \langle \chi 1_{(x, \infty)}, \mathcal{Z}(0) \rangle + \langle \chi 1_{(x, \infty)}, \mathcal{V}(t) \rangle.$$

Because, for each $x \in \mathbb{R}_+$, $[0, \tilde{\tau}_x)$ is a busy period, (60) implies that almost surely for all $x \in \mathbb{R}_+$ and $t \in [0, \tilde{\tau}_x)$,

$$W(t) = W(0) + V(t) - t = \langle \chi, \mathcal{Z}(0) \rangle + \langle \chi, \mathcal{V}(t) \rangle - t.$$

Subtracting the first of the preceding two displays from the second yields that almost surely for all $x \in \mathbb{R}_+$ and $t \in [0, \tilde{\tau}_x)$,

$$W(t, x) = X(t, x).$$

Thus, almost surely, $X(t, x) > 0$ for all $x \in \mathbb{R}_+$ and $t \in [0, \tilde{\tau}_x)$. Hence, almost surely, $Y(t, x) \ge 0$ for all $x \in \mathbb{R}_+$ and $t \in [0, \tilde{\tau}_x)$. Therefore, almost surely, $\tilde{\tau}_x \le \tau_x$ for all $x \in \mathbb{R}_+$. If $\tilde{\tau}_x = \infty$, the proof is complete. If $\tilde{\tau}_x < \infty$, then $W(\tilde{\tau}_x, x) = 0$ by right continuity. Thus, no job arrives at time $\tilde{\tau}_x$ with service time less than or equal to x and, consequently, $X(\tilde{\tau}_x, x) = 0$. So, by (52) and because $V(\cdot, x)$ is piecewise constant almost surely, $Y(s, x) < 0$ for all $s > \tilde{\tau}_x$. Thus, almost surely, $\tilde{\tau}_x \ge \tau_x$. □

Corollary 5.2. *Almost surely, for all* $t \in [0, \infty)$,

$$F(t) \le \inf\{x \in \mathbb{R}_+ \colon Y(t, x) > 0\}. \tag{65}$$

Proof. Fix $t \in [0, \infty)$. If $\{x \in \mathbb{R}_+ \colon Y(t, x) > 0\} = \varnothing$, (65) is trivial because $\inf \varnothing = \infty$. Otherwise, there exists $x \in \mathbb{R}_+$ such that $Y(t, x) > 0$. Then, by (52), $\tau_x > t$ and thus, by Lemma 5.5, $W(s, x) = X(s, x) \ge Y(t, x) > 0$ for all $s \in [0, t]$. Hence, for all $s \in [0, t]$, $L(s) \le x$, which implies that $C(s) \le x$ and thus $F(t) \le x$. □

Corollary 5.3. *Let* $\tau = \sup\{s \in [0, \infty) \colon Y(s) > 0\}$. *Almost surely, for all* $t \in [0, \tau)$,

$$F(t) \ge \sup\{x \in \mathbb{R}_+ \colon Y(t, x) < 0\}. \tag{66}$$

Proof. Note that if $\{s \in [0, \infty) \colon Y(s) > 0\} = \varnothing$, then $\tau = 0$ and there is nothing to prove. Suppose $\tau > 0$ and fix $t \in [0, \tau)$. If $\{x \in \mathbb{R}_+ \colon Y(t, x) < 0\} = \varnothing$, (66) is trivial. Otherwise, there exists an $x \in \mathbb{R}_+$ such that $Y(t, x) < 0$. Then, $\tau_x < t < \tau$. Almost surely, $W(\tau_x, x) = 0$ because $\tau_x = \tilde{\tau}_x$ (see the proof of Lemma 5.5). Because $\tau_x < \tau$, $W(\tau_x) > 0$. Therefore, $x \le L(\tau_x) < \infty$ so that $x \le C(\tau_x) \le F(\tau_x)$. Because $\tau_x < t$, $F(t) \ge x$. □

5.1.2.6. Application to the fluid-scaled sequence of systems. Recall that we append a superscript r to each object associated with the rth model in the $\mathcal{R}$-indexed sequence, including the performance processes defined in §5.1.1 and the random times $\tau(\cdot, x)$, $\tau(\cdot)$ as well as τ defined in Lemmas 5.2, 5.4, and Corollary 5.3, respectively. A fluid scaling is applied to each model in the sequence. In addition to the scaling already defined in (24), define for each $r \in \mathcal{R}$, $t \in [0, \infty)$, and $x \in \mathbb{R}_+$,

$$\begin{gathered}
\bar{L}^r(t) = L^r(rt), \qquad \bar{C}^r(t) = C^r(rt), \qquad \bar{F}^r(t) = F^r(rt), \\
\bar{E}^r(t, x) = \frac{1}{r} E^r(rt, x), \\
\bar{Z}^r(t) = \frac{1}{r} Z^r(rt), \qquad \bar{Z}^r(t, x) = \frac{1}{r} Z^r(rt, x), \\
\bar{W}^r(t) = \frac{1}{r} W^r(rt), \qquad \bar{W}^r(t, x) = \frac{1}{r} W^r(rt, x), \\
\bar{V}^r(t) = \frac{1}{r} V^r(rt), \qquad \bar{V}^r(t, x) = \frac{1}{r} V^r(rt, x), \\
\bar{X}^r(t) = \frac{1}{r} X^r(rt), \qquad \bar{X}^r(t, x) = \frac{1}{r} X^r(rt, x), \\
\bar{Y}^r(t) = \frac{1}{r} Y^r(rt), \qquad \bar{Y}^r(t, x) = \frac{1}{r} Y^r(rt, x), \\
\bar{\tau}^r = \frac{1}{r} \tau^r, \qquad \bar{\tau}^r(t) = \frac{1}{r} \tau^r(rt), \qquad \bar{\tau}^r(t, x) = \frac{1}{r} \tau^r(rt, x).
\end{gathered} \tag{67}$$

Then, (2), Lemmas 5.1 and 5.2, Corollary 5.1, Lemma 5.4, and Corollaries 5.2 and 5.3 imply that, for each $r \in \mathcal{R}$, almost surely, for all $s, t \in [0, \infty)$ such that $s \le t$, all $x, y \in \mathbb{R}_+$ such that $x \le y$, all $a > 0$, all closed $B \subset \mathbb{R}_+$, and all measurable $g \colon \mathbb{R}_+ \to \mathbb{R}_+$,

$$\bar{Z}^r(0) < \infty, \tag{68}$$

$$\bar{W}^r(0) < \infty, \tag{69}$$

$$\bar{Z}^r(t) \le \bar{Z}^r(0) + \bar{E}^r(t), \tag{70}$$

$$\langle 1_{[x, \infty)}, \bar{\mathcal{Z}}^r(t) \rangle \le \langle 1_{[x, \infty)}, \bar{\mathcal{Z}}^r(0) + \bar{\mathcal{V}}^r(t) \rangle, \tag{71}$$

$$\langle 1_{[x, y]}, \bar{\mathcal{Z}}^r(t) \rangle \le \langle 1_{[x, y]}, \bar{\mathcal{Z}}^r(0) + \bar{\mathcal{V}}^r(t) \rangle + \frac{1}{r}, \tag{72}$$

$$\langle 1_B, \bar{\mathcal{Z}}^r(t) \rangle \le \langle 1_B, \bar{\mathcal{Z}}^r(s) \rangle + \bar{E}^r(t) - \bar{E}^r(s) + \frac{1}{r}, \tag{73}$$

$$\langle 1_B, \bar{\mathcal{Z}}^r(s) \rangle \le \langle 1_B, \bar{\mathcal{Z}}^r(t) \rangle + \langle 1_{[0, a]}, \bar{\mathcal{Z}}^r(s) \rangle + \frac{t - s}{a} + \frac{1}{r}, \tag{74}$$

$$\bar{W}^r(t) = \bar{X}^r(t) + [\bar{Y}^r(t)]^-, \tag{75}$$

$$\langle \chi 1_{(x,\infty)}, \bar{\mathcal{Z}}^r(t)\rangle \le \langle \chi 1_{(x,\infty)}, \bar{\mathcal{Z}}^r(0)+\bar{\mathcal{V}}^r(t)\rangle, \tag{76}$$

$$\bar{W}^r(t,x) \le \bar{W}^r(0,x)+\bar{V}^r(t,x)-\bar{V}^r(\bar{\tau}^r(t,x)-,x)+\frac{x}{r}-(t-\bar{\tau}^r(t,x)), \tag{77}$$

$$\langle g1_{(\bar{F}^r(t),\infty)}, \bar{\mathcal{Z}}^r(t)\rangle = \langle g1_{(\bar{F}^r(t),\infty)}, \bar{\mathcal{Z}}^r(0)+\bar{\mathcal{V}}^r(t)\rangle, \tag{78}$$

$$\langle g1_{[\bar{F}^r(t),\infty)}, \bar{\mathcal{Z}}^r(t)\rangle \le \langle g1_{[\bar{F}^r(t),\infty)}, \bar{\mathcal{Z}}^r(0)+\bar{\mathcal{V}}^r(t)\rangle, \tag{79}$$

$$\langle \chi 1_{[0,\bar{F}^r(t))}, \bar{\mathcal{Z}}^r(t)\rangle \le \bar{V}^r(t,\bar{F}^r(t))-\bar{V}^r(\bar{\tau}^r(t)-,\bar{F}^r(t))+\frac{\bar{F}^r(t)}{r}-(t-\bar{\tau}^r(t)), \tag{80}$$

$$\bar{F}^r(t) \le \inf\{x\in\mathbb{R}_+\colon \bar{Y}^r(t,x)>0\}, \tag{81}$$

and, for all $t\in[0,\bar{\tau}^r)$,

$$\bar{F}^r(t) \ge \sup\{x\in\mathbb{R}_+\colon \bar{Y}^r(t,x)<0\}. \tag{82}$$

5.1.3. Functional weak law of large numbers. The following result gives the limiting behavior under fluid scaling of the stochastic primitives. It is a special case, for example, of Lemma 5.1 in Gromoll and Williams [7] for the case of one route ($\mathbf{I}=1$) and follows by (25) and (26).

Proposition 5.1. *As* $r\to\infty$,

$$(\bar{\mathcal{V}}^r(\cdot),\bar{V}^r(\cdot)) \;\Rightarrow\; (\mathcal{V}^*(\cdot),V^*(\cdot)),$$

where $\mathcal{V}^*(t)=\alpha t\nu$ *and* $V^*(t)=\rho t$ *for all* $t\in[0,\infty)$.

5.2. Tightness. To prove the fluid limit result, we first need to establish that $\{\bar{\mathcal{Z}}^r(\cdot)\colon r\in\mathcal{R}\}$ is tight. There are two main steps: to verify a compact containment condition and verify that sample path oscillations are uniformly small.

5.2.1. Compact containment.

Lemma 5.6. *Let* $T>0$ *and* $\eta>0$. *There exists a compact* $\mathbf{K}\subset\mathbf{M}$ *such that*

$$\liminf_{r\to\infty}\mathbf{P}^r(\bar{\mathcal{Z}}^r(t)\in\mathbf{K} \text{ for all } t\in[0,T])\ge 1-\eta.$$

Proof. By (29), there exists an $M>0$ such that

$$\mathbf{P}(\langle 1,\mathcal{Z}_0\rangle\vee\langle\chi,\mathcal{Z}_0\rangle\ge M)\le\eta.$$

Fix such an M and let $K=(\alpha+\rho)T+1$. For each $r\in\mathcal{R}$, let

$$\Omega_1^r=\{\bar{Z}^r(0)\vee\bar{W}^r(0)<M\},$$

$$\Omega_2^r=\{\bar{E}^r(T)\vee\bar{V}^r(T)<K\},$$

$$\Omega_3^r=\{\bar{Z}^r(t)\le\bar{Z}^r(0)+\bar{E}^r(T) \text{ and } \bar{W}^r(t)\le\bar{W}^r(0)+\bar{V}^r(T) \text{ for all } t\in[0,T]\}.$$

Because $\xi\mapsto\langle 1,\xi\rangle$ is continuous on $\mathbf{M}$, (28) implies that

$$(\bar{Z}^r(0),\bar{W}^r(0)) \;\Rightarrow\; (\langle 1,\mathcal{Z}_0\rangle,\langle\chi,\mathcal{Z}_0\rangle),\quad \text{as } r\to\infty.$$

The set $\{(z,w)\in\mathbb{R}_+\times\mathbb{R}_+\colon z\vee w<M\}$ is open so, by the Portmanteau theorem,

$$\liminf_{r\to\infty}\mathbf{P}^r(\Omega_1^r)\ge\mathbf{P}(\langle 1,\mathcal{Z}_0\rangle\vee\langle\chi,\mathcal{Z}_0\rangle<M)\ge 1-\eta.$$

Similarly, $(\bar{E}^r(T),\bar{V}^r(T))\Rightarrow(\alpha T,\rho T)$ as $r\to\infty$, by Proposition 5.1. Thus, by choice of K, $\liminf_{r\to\infty}\mathbf{P}^r(\Omega_2^r)=1$. By (70) and (76), $\mathbf{P}^r(\Omega_3^r)=1$ for all $r\in\mathcal{R}$. Hence,

$$\liminf_{r\to\infty}\mathbf{P}^r(\Omega_1^r\cap\Omega_2^r\cap\Omega_3^r)\ge 1-\eta.$$

Let $\mathbf{K}$ be the closure in $\mathbf{M}$ of the set $\{\xi\in\mathbf{M}\colon\langle 1,\xi\rangle\vee\langle\chi,\xi\rangle\le M+K\}$. The set $\mathbf{K}$ is compact by Theorem 15.7.5 in Kallenberg [9]. Furthermore, on $\Omega_1^r\cap\Omega_2^r\cap\Omega_3^r$, $\bar{Z}^r(t)\le M+K$ and $\bar{W}^r(t)\le M+K$ for all $t\in[0,T]$. In particular, $\Omega_1^r\cap\Omega_2^r\cap\Omega_3^r\subset\{\bar{\mathcal{Z}}^r(t)\in\mathbf{K} \text{ for all } t\in[0,T]\}$. □

5.2.2. Asymptotic regularity near the origin. To control the oscillations of the measure-valued state descriptors, it is necessary to control the number of departures in a short period of time. A large number of departures can only occur if a large number of jobs build up arbitrarily close to the origin. Lemma 5.7 implies that, with high probability, the number of jobs in a sufficiently small region around the origin is uniformly small over a compact time interval.

LEMMA 5.7. *Let $T > 0$. For each $\epsilon, \eta \in (0,1)$, there exists an $a > 0$ such that*

$$\liminf_{r\to\infty} \mathbf{P}^r\left(\sup_{t\in[0,T]} \bar{Z}^r(t,a) \le \epsilon\right) \ge 1-\eta.$$

PROOF. Almost surely, $\mathscr{Z}_0 \in \mathbf{M}_0$ and thus $\langle 1_{\{0\}}, \mathscr{Z}_0\rangle = 0$. Hence, there exists $a_1 > 0$ such that

$$\mathbf{P}\left(\langle 1_{[0,a_1]}, \mathscr{Z}_0\rangle < \frac{\epsilon}{4}\right) \ge 1-\eta. \tag{83}$$

Furthermore, because ν does not have an atom at the origin, there is an $a_2 > 0$ such that

$$\alpha T\langle 1_{[0,a_2]}, \nu\rangle < \frac{\epsilon}{4}. \tag{84}$$

Let $a = a_1 \wedge a_2$ and define

$$\Omega_1^r = \left\{\bar{Z}^r(0,a) < \frac{\epsilon}{4}\right\},$$
$$\Omega_2^r = \left\{\bar{E}^r(T,a) < \frac{\epsilon}{4}\right\},$$
$$\Omega_3^r = \left\{\bar{Z}^r(t,a) \le \bar{Z}^r(0,a) + \bar{E}^r(T,a) + \frac{1}{r} \text{ for all } t\in[0,T]\right\}.$$

The set $\{\xi \in \mathbf{M}: \langle 1_{[0,a_1]}, \xi\rangle < \epsilon/4\}$ is open in the weak topology. Thus, by (28), (83), the Portmanteau theorem, and because $a \le a_1$,

$$\liminf_{r\to\infty} \mathbf{P}^r(\Omega_1^r) \ge 1-\eta.$$

Similarly, by Proposition 5.1, (84), the Portmanteau theorem, and since $a \le a_2$,

$$\liminf_{r\to\infty} \mathbf{P}^r(\Omega_2^r) = 1.$$

In addition, by (72) with $x = 0$ and $y = a$, and because $\langle 1_{[0,a]}, \bar{\mathscr{V}}^r(t)\rangle \le \bar{E}^r(T,a)$ for all $t \in [0,T]$, $\mathbf{P}^r(\Omega_3^r) = 1$ for all $r \in \mathscr{R}$. Hence,

$$\liminf_{r\to\infty} \mathbf{P}^r(\Omega_1^r \cap \Omega_2^r \cap \Omega_3^r) \ge 1-\eta.$$

On $\Omega_1^r \cap \Omega_2^r \cap \Omega_3^r$,

$$\sup_{t\in[0,T]} \bar{Z}^r(t,a) \le \frac{\epsilon}{4} + \frac{\epsilon}{4} + \frac{1}{r},$$

which is bounded above by ϵ for sufficiently large r. □

5.2.3. Oscillations.

DEFINITION 5.1. Let $\mathbf{d}$ be the metric on $\mathbf{M}$, given by

$$\mathbf{d}[\zeta,\xi] = \inf\{\epsilon > 0: \langle 1_B, \zeta\rangle \le \langle 1_{B^\epsilon}, \xi\rangle + \epsilon \text{ and } \langle 1_B, \xi\rangle \le \langle 1_{B^\epsilon}, \zeta\rangle + \epsilon \text{ for all closed } B \subset \mathbb{R}_+\}.$$

DEFINITION 5.2. Let $T > 0$ and $\delta \in [0,T]$. For each $\zeta(\cdot) \in \mathbf{D}([0,\infty),\mathbf{M})$, define a modulus of continuity on $[0,T]$ by

$$\mathbf{w}_T(\zeta(\cdot),\delta) = \sup\{\mathbf{d}[\zeta(t),\zeta(s)]: 0 \le s,t \le T, |t-s| < \delta\}.$$

LEMMA 5.8. *Let $T > 0$. For all $\epsilon, \eta \in (0,1)$, there exists $\delta > 0$ such that*

$$\liminf_{r\to\infty} \mathbf{P}^r(\mathbf{w}_T(\bar{\mathscr{Z}}^r(\cdot),\delta) \le \epsilon) \ge 1-\eta.$$

Proof. Fix $\epsilon, \eta \in (0,1)$. Let $a > 0$ be as in Lemma 5.7 with ϵ replaced by $\epsilon/3$ and define

$$\Omega_1^r = \left\{ \sup_{t\in[0,T]} |\bar{E}^r(t) - \alpha t| \le \frac{\epsilon}{6} \right\},$$

$$\Omega_2^r = \left\{ \sup_{t\in[0,T]} \bar{Z}^r(t,a) \le \frac{\epsilon}{3} \right\},$$

$$\Omega_3^r = \Bigg\{ \langle 1_B, \bar{\mathcal{Z}}^r(t)\rangle \le \langle 1_B, \bar{\mathcal{Z}}^r(s)\rangle + \bar{E}^r(t) - \bar{E}^r(s) + \frac{1}{r} \text{ and } \langle 1_B, \bar{\mathcal{Z}}^r(s)\rangle \le \langle 1_B, \bar{\mathcal{Z}}^r(t)\rangle + \langle 1_{[0,a]}, \bar{\mathcal{Z}}^r(s)\rangle + \frac{t-s}{a} + \frac{1}{r} \text{ for all closed } B\subset\mathbb{R}_+ \text{ and all } 0\le s\le t\le T \Bigg\}.$$

By (26), $\liminf_{r\to\infty} \mathbf{P}^r(\Omega_1^r) = 1$. Together with Lemma 5.7, (73), and (74), this implies that

$$\liminf_{r\to\infty} \mathbf{P}^r(\Omega_1^r \cap \Omega_2^r \cap \Omega_3^r) \ge 1 - \eta.$$

Thus, to complete the proof, it suffices to specify $\delta > 0$ such that

$$\Omega_1^r \cap \Omega_2^r \cap \Omega_3^r \subset \{\mathbf{w}_T(\bar{\mathcal{Z}}^r(\cdot), \delta) \le \epsilon\}$$

for all sufficiently large r.

Set $\delta = \epsilon/(3\alpha) \wedge \epsilon a/3$. Fix $r > 3/\epsilon$, $\omega \in \Omega_1^r \cap \Omega_2^r \cap \Omega_3^r$, and $0 \le s \le t \le T$ such that $t - s < \delta$. Once we verify that $\mathbf{d}[\bar{\mathcal{Z}}^r(t)(\omega), \bar{\mathcal{Z}}^r(s)(\omega)] \le \epsilon$, the proof is complete. By the definitions of δ, Ω_1^r, Ω_2^r, and Ω_3^r, for each closed $B \subset \mathbb{R}_+$,

$$\langle 1_B, \bar{\mathcal{Z}}^r(t)\rangle \le \langle 1_B, \bar{\mathcal{Z}}^r(s)\rangle + \alpha(t-s) + \frac{\epsilon}{6} + \frac{\epsilon}{6} + \frac{1}{r} \le \langle 1_{B^\epsilon}, \bar{\mathcal{Z}}^r(s)\rangle + \epsilon.$$

Similarly, for all closed $B \subset \mathbb{R}_+$,

$$\langle 1_B, \bar{\mathcal{Z}}^r(s)\rangle \le \langle 1_B, \bar{\mathcal{Z}}^r(t)\rangle + \frac{\epsilon}{3} + \frac{\delta}{a} + \frac{1}{r} \le \langle 1_{B^\epsilon}, \bar{\mathcal{Z}}^r(t)\rangle + \epsilon. \quad \square$$

5.2.4. Proof of tightness.

Theorem 5.1. *The sequence* $\{\bar{\mathcal{Z}}^r(\cdot)\colon r\in\mathcal{R}\}$ *is C-tight.*

Proof. For each $T > 0$ and $\delta \in [0,T]$, let $\mathbf{w}'_T(\cdot,\delta)$ be the modulus of continuity on $D([0,\infty),\mathbf{M})$ used in Corollary 3.7.4 of Ethier and Kurtz [6]. By Definition 5.2, $\mathbf{w}'_T(\zeta(\cdot),\delta) \le \mathbf{w}_{T+\delta}(\zeta(\cdot), 2\delta)$ for all $\zeta(\cdot) \in D([0,\infty),\mathbf{M})$. Thus, by Lemmas 5.6 and 5.8, $\{\bar{\mathcal{Z}}^r(\cdot)\colon r\in\mathcal{R}\}$ satisfies the compact containment and oscillation conditions of Corollary 3.7.4 in Ethier and Kurtz [6]. Thus, $\{\bar{\mathcal{Z}}^r(\cdot)\colon r\in\mathcal{R}\}$ is tight. Moreover, Definition 5.2 and Lemma 5.8 imply that any weak limit point $\mathcal{Z}^*(\cdot)$ is continuous almost surely. $\square$

5.3. Characterization of fluid limit points.

Let $\mathcal{Z}^*(\cdot)$ be a weak limit of $\{\bar{\mathcal{Z}}^r(\cdot)\colon r\in\mathcal{R}\}$ and let $\mathcal{Q}\subset\mathcal{R}$ be a subsequence such that

$$\bar{\mathcal{Z}}^q(\cdot) \;\Rightarrow\; \mathcal{Z}^*(\cdot) \quad \text{as } q\to\infty.$$

By Theorem 5.1, $\mathcal{Z}^*(\cdot)$ is continuous almost surely. Let $W^*(0) = \langle \chi, \mathcal{Z}^*(0)\rangle$. By (28) and Proposition 5.1, and because $\{(\bar{\mathcal{V}}^r(\cdot), \bar{V}^r(\cdot))\colon r\in\mathcal{R}\}$ converges in distribution to a deterministic process, we have the joint convergence

$$(\bar{\mathcal{V}}^q(\cdot), \bar{V}^q(\cdot), \bar{W}^q(0), \bar{\mathcal{Z}}^q(\cdot)) \;\Rightarrow\; (\mathcal{V}^*(\cdot), V^*(\cdot), W^*(0), \mathcal{Z}^*(\cdot)) \quad \text{as } q\to\infty.$$

By the Skorohod representation theorem, we assume that all random elements are defined on a common probability space $(\Omega^*, \mathbf{P}^*, \mathcal{F}^*)$ such that, almost surely as $q\to\infty$,

$$(\bar{\mathcal{V}}^q(\cdot), \bar{V}^q(\cdot), \bar{W}^q(0), \bar{\mathcal{Z}}^q(\cdot)) \to (\mathcal{V}^*(\cdot), V^*(\cdot), W^*(0), \mathcal{Z}^*(\cdot)) \tag{85}$$

uniformly on compact time intervals. Note that because the Skorohod representations in (85) are equal in distribution to the original processes, they still satisfy (68)–(82) almost surely, with the various functionals there defined analogously as for the original processes. In this section, we prove the following theorem.

Theorem 5.2. $\mathcal{Z}^*(\cdot)$ *is almost surely a fluid model solution for the data* (α,ν) *and initial measure* $\mathcal{Z}^*(0)$.

To prove Theorem 5.2, we will need to verify that $\mathcal{Z}^*(\cdot)$ almost surely satisfies (C1), (C2), and (C3). For this, let $\widetilde{\Omega}^* \in \mathcal{F}^*$ be such that $\mathbf{P}^*(\widetilde{\Omega}^*) = 1$ and, on $\widetilde{\Omega}^*$, that (85) holds, $\mathcal{Z}^*(\cdot)$ is continuous, $\mathcal{Z}^*(0) \in \mathbf{M}_0$ (which is possible because of (29)), and, finally, that (68)–(82) hold for all $q \in \mathcal{Q}$. Then, (C1) holds on $\widetilde{\Omega}^*$. It remains to show that $\mathcal{Z}^*(\cdot)$ satisfies (C2) and (C3) on $\widetilde{\Omega}^*$. This is accomplished in §§5.3.1 and 5.3.2, respectively. To simplify the presentation, $\omega \in \widetilde{\Omega}^*$ is fixed throughout §§5.3.1 and 5.3.2 and all random elements are evaluated at this fixed ω. Let $\xi = \mathcal{Z}^*(0)(\omega)$. Then, $\xi \in \mathbf{M}_0$. Using this ξ and the data (α, ν) given in (26) and (27), define x_1, x_2, l_ξ, $\mathbf{M}_1$, $\mathbf{M}_2$, $s(\cdot)$, t_1, $s_r^{-1}(\cdot)$, and $s_\ell^{-1}(\cdot)$, respectively, by (13), (14), (15), (16), (17), (18), (19), (20), (21), (44), and (45).

5.3.1. Verification of (C2). We begin with an observation. The convergence of the second and third components in (85) implies uniform integrability (for finite measures) of the sequences of measures $\{\overline{\mathcal{Z}}^q(0)\}_{q\in\mathcal{Q}}$ and $\{\overline{\mathcal{V}}^q(t)\}_{q\in\mathcal{Q}}$ for each $t \in [0,\infty)$. Together with (76), this implies that $\{\overline{\mathcal{Z}}^q(t)\}_{q\in\mathcal{Q}}$ is uniformly integrable for each $t \in [0,\infty)$. Thus, for each $t \in [0,\infty)$,

$$\langle \chi, \mathcal{Z}^*(t)\rangle = \lim_{M\nearrow\infty} \langle \chi \wedge M, \mathcal{Z}^*(t)\rangle = \lim_{M\nearrow\infty}\lim_{q\to\infty} \langle \chi\wedge M, \overline{\mathcal{Z}}^q(t)\rangle \le \limsup_{q\to\infty} \langle \chi, \overline{\mathcal{Z}}^q(t)\rangle < \infty.$$

Furthermore, for all $q \in \mathcal{Q}$, $t \in [0,\infty)$ and $M > 0$,

$$|\langle \chi, \overline{\mathcal{Z}}^q(t)\rangle - \langle \chi, \mathcal{Z}^*(t)\rangle| \le |\langle \chi\wedge M, \overline{\mathcal{Z}}^q(t)\rangle - \langle \chi\wedge M, \mathcal{Z}^*(t)\rangle| + \langle \chi 1_{(M,\infty)}, \overline{\mathcal{Z}}^q(t)\rangle + \langle \chi 1_{(M,\infty)}, \mathcal{Z}^*(t)\rangle.$$

Hence, for each $t \in [0,\infty)$,

$$\lim_{q\to\infty} \overline{W}^q(t) = \lim_{q\to\infty} \langle \chi, \overline{\mathcal{Z}}^q(t)\rangle = \langle \chi, \mathcal{Z}^*(t)\rangle. \tag{86}$$

We wish to show that $\langle \chi, \mathcal{Z}^*(t)\rangle = [W^*(0) + (\rho-1)t]^+$ for all $t \in [0,\infty)$. For this, it suffices to show that, for all $t \in [0,\infty)$,

$$\lim_{q\to\infty} \overline{W}^q(t) = [W^*(0) + (\rho-1)t]^+. \tag{87}$$

For this, define $X^*(\cdot)$ and $Y^*(\cdot)$ as follows: for $t \in [0,\infty)$,

$$X^*(t) = W^*(0) + V^*(t) - t = W^*(0) + (\rho-1)t, \tag{88}$$

and

$$Y^*(t) = \inf_{0\le s\le t} X^*(s) = \begin{cases} W^*(0), & \text{if } \rho \ge 1,\\ W^*(0) + (\rho-1)t, & \text{if } \rho < 1. \end{cases} \tag{89}$$

Lemma 5.9 is an immediate consequence of (85).

Lemma 5.9. *For each* $t \in [0,\infty)$,

$$\lim_{q\to\infty} \overline{X}^q(t) = X^*(t) \qquad \textit{and} \qquad \lim_{q\to\infty} \overline{Y}^q(t) = Y^*(t).$$

Combining Lemma 5.9 with (75) proves (87). Hence, (C2) holds.

5.3.2. Verification of (C3). Verifying that $\mathcal{Z}^*(\cdot)$ satisfies (C3) presents the greatest challenge. First, we develop some elementary results that will facilitate this. Then, we derive asymptotic bounds for the sequence of fluid-scaled frontier processes. This leads to the proof that $\mathcal{Z}^*(\cdot)$ satisfies (C3).

5.3.2.1. Fluid limits for truncated processes. For $x \in \mathbb{R}_+$, let $\rho_x = \alpha\langle \chi 1_{[0,x]}, \nu\rangle$ and $W^*(0,x) = \langle \chi 1_{[0,x]}, \mathcal{Z}^*(0)\rangle$. For $x \in \mathbb{R}_+$ and $t \in [0,\infty)$, let

$$V^*(t,x) = \rho_x t, \tag{90}$$

$$X^*(t,x) = W^*(0,x) + V^*(t,x) - t = W^*(0,x) + (\rho_x - 1)t, \tag{91}$$

$$Y^*(t,x) = \inf_{0\le s\le t} X^*(s,x). \tag{92}$$

Lemma 5.10.
(i) *For each* $x \in \mathbb{R}_+$ *such that* ξ *does not have an atom at* x, $\overline{W}^q(0,x)$ *converges to* $W^*(0,x)$ *as* $q \to \infty$.
(ii) *For each* $x \in \mathbb{R}_+$ *such that* ν *does not have an atom at* x, $\overline{V}^q(\cdot,x)$ *converges to* $V^*(\cdot,x)$ *uniformly on compact time intervals as* $q \to \infty$.

Proof. Note that (i) follows from the fact that $\bar{\mathscr{Z}}^q(0) \xrightarrow{w} \xi \in \mathbf{M}_0$ as $q \to \infty$. For a proof of (ii), fix $x \in \mathbb{R}_+$ such that ν does not have an atom at x. Because $\bar{\mathscr{V}}^q(t) \xrightarrow{w} \mathscr{V}^*(t)$, $\lim_{q\to\infty} \bar{V}^q(t,x) = V^*(t,x)$ for each $t \in [0,\infty)$. Because $V^*(\cdot, x)$ is continuous and nondecreasing, the convergence is uniform on compact time intervals. ☐

Corollary 5.4. *For each $t \in [0,\infty)$ and $x \in \mathbb{R}_+$ such that neither ξ nor ν has an atom at x,*

$$\lim_{q\to\infty} \bar{X}^q(t,x) = X^*(t,x)$$

and

$$\lim_{q\to\infty} \bar{Y}^q(t,x) = Y^*(t,x) = \begin{cases} W^*(0,x) + (\rho_x - 1)t, & x < x_1, \\ W^*(0,x), & x \geq x_1. \end{cases}$$

5.3.2.2. Asymptotics for the frontier. We wish to analyze the behavior of the frontier process $\bar{F}^q(\cdot)$ as $q \to \infty$. This is slightly delicate because it may not be the case that the frontier of $\bar{\mathscr{Z}}^q(\cdot)$ converges to the frontier $F^*(\cdot)$ of $\mathscr{Z}^*(\cdot)$, where for all $t \in [0,\infty)$,

$$L^*(t) = \sup\{x \in \mathbb{R}_+ : \langle 1_{[0,x)}, \mathscr{Z}^*(t)\rangle = 0\},$$

$$C^*(t) = \begin{cases} L^*(t), & \text{if } \mathscr{Z}^*(t) \neq \mathbf{0}, \\ 0, & \text{otherwise}, \end{cases}$$

$$F^*(t) = \sup_{0\leq s\leq t} C^*(s).$$

To see this, consider the following example. Suppose $\mathscr{Q} = \mathbb{N}$ and, for each $q \in \mathscr{Q}$, suppose that $[0,4]$ does not intersect the support of ν^q and

$$\mathscr{Z}^q(0) = q\delta_1 + \delta_2 + q\delta_3.$$

Then, $\mathscr{Z}^*(0) = \delta_1 + \delta_3$,

$$\bar{F}^q(t) = \begin{cases} 1, & 0 \leq t < 1, \\ 2, & 1 \leq t < 1 + \dfrac{2}{q}, \\ 3, & 1 + \dfrac{2}{q} \leq t < 4 + \dfrac{2}{q}, \end{cases} \qquad \text{and} \qquad F^*(t) = \begin{cases} 1, & 0 \leq t < 1, \\ 3, & 1 \leq t < 4. \end{cases}$$

Thus, $\bar{F}^q(\cdot)$ does not converge to $F^*(\cdot)$ in the Skorohod J_1-topology. The fact that $\mathscr{Z}^*(0)$ does not have any support in the interval $(1,3)$ allows the prelimits to have a very small but positive mass in $(1,3)$, which delays the time at which the frontier process achieves the value 3. In general, the prelimits for the frontier process can exhibit similar behavior whenever there are open intervals that do not intersect the union of the supports of ξ and ν. It is sufficient for our purposes to bound $\liminf_{q\to\infty} \bar{F}^q(\cdot)$ and $\limsup_{q\to\infty} \bar{F}^q(\cdot)$.

Theorem 5.3.
(i) *If $\xi \in \mathbf{M}_1$, then for all $t \in [0,\infty)$,*

$$s_\ell^{-1}(t) \leq \liminf_{q\to\infty} \bar{F}^q(t) \leq \limsup_{q\to\infty} \bar{F}^q(t) \leq s_r^{-1}(t). \tag{93}$$

(ii) *If $\xi \in \mathbf{M}_2$, then for all $t \in [0,\infty)$,*

$$s_\ell^{-1}(t) \leq \liminf_{q\to\infty} \bar{F}^q(t) \leq \limsup_{q\to\infty} \bar{F}^q(t) \leq l_\xi. \tag{94}$$

Remark 5.1. If $\xi \in \mathbf{M}_2$, then, by (21), $s_r^{-1}(t) = x_2 \wedge l_\xi$ for all $t \in (0,\infty)$. Hence, if $l_\xi \leq x_2$, then $s_r^{-1}(t) = l_\xi$ for all $t \in [0,\infty)$. Therefore, the upper bound in (94) is consistent with that in (93). However, if $x_2 < l_\xi$, then for all $t \in (0,\infty)$, $s_r^{-1}(t) < l_\xi$ and the upper bound in (94) is not equal to $s_r^{-1}(t)$. This results from the fact that, at time zero, the prelimit frontier processes can jump above x_2 and thereby remain larger than x_2 for all time. Hence, the frontier process fails to characterize the left edge of the fluid limit when $\xi \neq \mathbf{0}$ and $x_2 < l_\xi$.

PROOF. We begin with a proof of (i). If $\xi \in \mathbf{M}_1$, then $t_1 > 0$. First, fix $t \in [0, t_1)$. By Lemma 4.2(i), there is an $\epsilon > 0$ such that $s_r^{-1}(t) + \epsilon < x_1$ and neither ξ nor ν has an atom at $s_r^{-1}(t) + \epsilon$. By (20), $s(s_r^{-1}(t) + \epsilon) > t$. Hence, by (18) and (91), $X^*(t, s_r^{-1}(t) + \epsilon) > 0$, which implies that $Y^*(t, s_r^{-1}(t) + \epsilon) > 0$. By Corollary 5.4, there exists Q such that $q > Q$ implies $\bar{Y}^q(t, s_r^{-1}(t) + \epsilon) > 0$. By (81), $q > Q$ implies $\bar{F}^q(t) \le s_r^{-1}(t) + \epsilon$. Because ϵ may be chosen arbitrarily small, the third inequality in (93) holds at time t. If $s_\ell^{-1}(t) = 0$, the proof that (93) holds at time t is complete. Otherwise, $s_\ell^{-1}(t) > 0$. In this case, choose $\epsilon > 0$ such that $0 < s_\ell^{-1}(t) - \epsilon$ and neither ξ nor ν has an atom at $s_\ell^{-1}(t) - \epsilon$. By (44), $s(s_\ell^{-1}(t) - \epsilon) < t$. Hence, $X^*(t, s_\ell^{-1}(t) - \epsilon) < 0$, which implies that $Y^*(t, s_\ell^{-1}(t) - \epsilon) < 0$. Because $\xi \in \mathbf{M}_1$, then $W^*(0) > 0$. If $\rho \ge 1$, then $Y^*(t) = W^*(0) > 0$. Otherwise, $\rho < 1$ and $t < t_1 = W^*(0)/(1 - \rho)$, which also implies that $Y^*(t) > 0$. Because $Y^*(t, s_\ell^{-1}(t) - \epsilon) < 0$ and $Y^*(t) > 0$, then by Lemma 5.9 and Corollary 5.4, there exists Q such that $q > Q$ implies $\bar{Y}^q(t, s_\ell^{-1}(t) - \epsilon) < 0$ and $\bar{Y}^q(t) > 0$. Note that $\bar{Y}^q(t) > 0$ implies that $t < \bar{\tau}^q$. Thus, by (82), $q > Q$ implies $s_\ell^{-1}(t) - \epsilon \le \bar{F}^q(t)$. Hence, (93) holds at time t for $t \in [0, t_1)$. If $t_1 = \infty$, the proof of (i) is complete.

Otherwise, $t_1 < \infty$ and we must verify that (93) holds for $t \in [t_1, \infty)$. Fix $t \in [t_1, \infty)$. We begin by verifying that the first inequality in (93) holds at time t. If $t = t_1$, this follows because (93) holds for $t \in [0, t_1)$, $\bar{F}^q(\cdot)$ is nondecreasing, and $s_\ell^{-1}(\cdot)$ is left continuous. Otherwise, $t \in (t_1, \infty)$. If $\rho < 1$ or $x_0 = x_1$, then $s_\ell^{-1}(t) = s_\ell^{-1}(t_1)$ by (45). Because $\bar{F}^q(\cdot)$ is nondecreasing, the first inequality in (93) holds at time t. Otherwise, $\rho \ge 1$ and $x_0 < x_1$. Then, $s(x) = t_1$ for all $x \in [x_0, x_1)$. Hence, $X^*(t_1, x) = 0$ for all $x \in [x_0, x_1)$. Because $t \in (t_1, \infty)$ and $\rho_x - 1 < 0$ for all $x \in [x_0, x_1)$, then $X^*(t, x) < 0$ for all $x \in [x_0, x_1)$. Hence, $Y^*(t, x) < 0$ for all $x \in [x_0, x_1)$. Because $\rho \ge 1$, then $Y^*(t) = W^*(0) > 0$. Fix $x \in [x_0, x_1)$. By Lemma 5.9 and Corollary 5.4, there exists a Q such that $q > Q$ implies $\bar{Y}^q(t, x) < 0$ and $\bar{Y}^q(t) > 0$. Then, by (82), $q > Q$ implies $\bar{F}^q(t) \ge x$. Because $s_\ell^{-1}(t) = x_1$ by (45), this implies that the first inequality in (93) holds at time t. Next, we verify that the third inequality in (93) holds at time $t \in [t_1, \infty)$. Note that $s_r^{-1}(t) = x_1$. Hence, if $x_1 = \infty$, there is nothing to prove. Suppose that $x_1 < \infty$. Because $\xi \in \mathbf{M}_1$, then $W^*(0, x_1) > 0$. By definition of x_1, $\rho_{x_1} - 1 \ge 0$. Thus, $Y^*(t, x_1) = W^*(0, x_1) > 0$. Then, by Corollary 5.4, there exists Q such that $q > Q$ implies $\bar{Y}^q(t, x_1) > 0$. This together with (81) implies that $\bar{F}^q(t) \le x_1$ for all $q > Q$. Thus, the third inequality in (93) holds at time t. This completes the proof of (i).

To prove (ii), suppose that $\xi \in \mathbf{M}_2$. Then $t_1 = 0$, $x_1 < \infty$, and $\rho \ge 1$. We begin by verifying the first inequality in (94). For $t = 0$, the first inequality in (94) is trivial because, by (45), $s_\ell^{-1}(0) = 0$. Fix $t \in (0, \infty)$. Then, by (45), $s_\ell^{-1}(t) = x_1$. Note that $X^*(t) = W^*(0) + (\rho - 1)t$ and thus $Y^*(t) = W^*(0) > 0$ because $\rho \ge 1$ and $\xi \in \mathbf{M}_2$. Given $x \in [0, x_1)$, because $\xi \notin \mathbf{M}_1$, $W^*(0, x) = 0$. Thus, $X^*(t, x) = (\rho_x - 1)t$ for all $x \in [0, x_1)$. Because $(\rho_x - 1) < 0$ for all $x \in [0, x_1)$, then $Y^*(t, x) < 0$ for all $x \in [0, x_1)$. Fix $x \in [0, x_1)$. By Lemma 5.9 and Corollary 5.4, there exists Q such that $q > Q$ implies $\bar{Y}^q(t, x) < 0$ and $\bar{Y}^q(t) > 0$. Hence, by (82), for all $q > Q$, $\bar{F}^q(t) \ge x$. Because $x \in [0, x_1)$ was arbitrary, the first inequality in (94) holds at time t. To verify the third inequality in (94), note that because $\xi \in \mathbf{M}_2$, then $x_1 \le l_\xi < \infty$. Fix $t \in [0, \infty)$. Thus, $Y^*(t, x) = W^*(0, x) > 0$ for all $x > l_\xi$. Fix $x > l_\xi$. By Corollary 5.4, there exists Q such that $q > Q$ implies that $\bar{Y}^q(t, x) > 0$. By (81), $\bar{F}^q(t) \le x$ for all $q > Q$. Thus, (94) holds. $\square$

5.3.2.3. Analysis of $\mathcal{Z}^*(\cdot)$. We are now prepared to prove that $\mathcal{Z}^*(\cdot)$ satisfies (C3). The idea is that for $q \in \mathcal{Q}$, $g \in C_b^+(\mathbb{R}_+)$, and $t \in [0, \infty)$,

$$\langle g, \bar{\mathcal{Z}}^q(t)\rangle = \langle g 1_{[0, \bar{F}^q(t))}, \bar{\mathcal{Z}}^q(t)\rangle + \langle g 1_{[\bar{F}^q(t), \infty)}, \bar{\mathcal{Z}}^q(t)\rangle.$$

To obtain a lower bound, we can remove the first term on the right side, replace the closed interval in the second term with an open interval, and use (78). Then, by combining this with Theorem 5.3 and (85), a preliminary lower bound is obtained for the fluid limit (see Lemma 5.11). To obtain a suitable upper bound, we show that the first term on the right side tends to zero (see Lemma 5.12). With that accomplished, (79), Theorem 5.3, and (85) are used to obtain a preliminary upper bound (see Lemma 5.13). The preliminary bounds stated in Lemmas 5.11 and 5.13 are used in the proof of Lemma 5.14, where it is verified that $\mathcal{Z}^*(\cdot)$ satisfies (C3).

LEMMA 5.11. *Let $g \in C_b^+(\mathbb{R}_+)$.*
(i) *If $l_\xi < x_2$, then for all $t \in [0, \infty)$*

$$\langle g, \mathcal{Z}^*(t)\rangle \ge \langle g 1_{(s_r^{-1}(t), \infty)}, \xi + \alpha t \nu\rangle, \tag{95}$$

(ii) *If $l_\xi \ge x_2$, then for all $t \in [0, \infty)$,*

$$\langle g, \mathcal{Z}^*(t)\rangle \ge \langle g 1_{(l_\xi, \infty)}, \xi + \alpha t \nu\rangle. \tag{96}$$

Proof. If $\xi = \mathbf{0}$, then $l_\xi = \infty$, $l_\xi \geq x_2$, and (96) holds. Otherwise, $\xi \neq \mathbf{0}$ and $l_\xi < \infty$. Fix $g \in C_b^+(\mathbb{R}_+)$ and $t \in [0, \infty)$. If $t \in [t_1, \infty)$ and $\rho < 1$, then $x_2 = \infty$ and thus $l_\xi < x_2$. Also, by Propositon 4.1, $s_r^{-1}(t) = \infty$. Hence, (95) holds at time t. Therefore, it suffices to consider the case where either $t \in [0, t_1)$, or $t \in [t_1, \infty)$ and $\rho \geq 1$. Then, by Proposition 4.1, $s_r^{-1}(t) < \infty$. Let $\epsilon > 0$ be such that neither ν nor ξ has an atom at either $s_r^{-1}(t) + \epsilon$ or $l_\xi + \epsilon$, and define

$$b(t, \epsilon) = \begin{cases} s_r^{-1}(t) + \epsilon, & \text{if } l_\xi < x_2, \\ l_\xi + \epsilon, & \text{if } l_\xi \geq x_2. \end{cases}$$

Recall that if $\xi \in \mathbf{M}_2$ and $l_\xi < x_2$, then $s_r^{-1}(t) = l_\xi$ (see (21)). Hence, by Theorem 5.3, there exists Q such that $q > Q$ implies that $\bar{F}^q(t) < b(t, \epsilon)$. Because g is nonnegative, (78) implies that for all $q \in \mathcal{Q}$,

$$\langle g, \bar{\mathcal{Z}}^q(t) \rangle \geq \langle g 1_{(\bar{F}^q(t), \infty)}, \bar{\mathcal{Z}}^q(t) \rangle = \langle g 1_{(\bar{F}^q(t), \infty)}, \bar{\mathcal{Z}}^q(0) \rangle + \langle g 1_{(\bar{F}^q(t), \infty)}, \bar{\mathcal{V}}^q(t) \rangle.$$

Hence, for all $q > Q$,

$$\langle g, \bar{\mathcal{Z}}^q(t) \rangle \geq \langle g 1_{[b(t,\epsilon), \infty)}, \bar{\mathcal{Z}}^q(0) \rangle + \langle g 1_{[b(t,\epsilon), \infty)}, \bar{\mathcal{V}}^q(t) \rangle.$$

Because neither ν nor ξ has an atom at $b(t, \epsilon)$, (85) implies that

$$\lim_{q \to \infty} (\langle g 1_{[b(t,\epsilon), \infty)}, \bar{\mathcal{Z}}^q(0) \rangle + \langle g 1_{[b(t,\epsilon), \infty)}, \bar{\mathcal{V}}^q(t) \rangle) = \langle g 1_{[b(t,\epsilon), \infty)}, \xi + \alpha t \nu \rangle.$$

It follows that

$$\langle g, \mathcal{Z}^*(t) \rangle \geq \langle g 1_{[b(t,\epsilon), \infty)}, \xi + \alpha t \nu \rangle.$$

Because $\epsilon > 0$ can be made arbitrarily small, the result follows. □

Lemma 5.12. *Let* $g \in C_b^+(\mathbb{R}_+)$.

(i) *If* $l_\xi < x_2$ *and either* $t \in [0, t_1)$, *or* $t \in [t_1, \infty)$ *and* $\rho_{x_1} \leq 1$, *then*

$$\lim_{q \to \infty} \langle g 1_{[0, \bar{F}^q(t))}, \bar{\mathcal{Z}}^q(t) \rangle = 0.$$

(ii) *If* $\xi \notin \mathbf{M}_1$, *then for all* $t \in [0, \infty)$ *and* $x \in [0, x_2 \wedge l_\xi)$,

$$\lim_{q \to \infty} \langle g 1_{[0, x]}, \bar{\mathcal{Z}}^q(t) \rangle = 0.$$

Proof. We begin by proving (i). For this, fix $t \in [0, \infty)$. We first show that

$$\lim_{q \to \infty} \langle \chi 1_{[0, \bar{F}^q(t))}, \bar{\mathcal{Z}}^q(t) \rangle = 0. \tag{97}$$

If $t \in [t_1, \infty)$ and $\rho < 1$, then (97) is an immediate consequence of (86), (C2), (18), and (19). Henceforth, we assume that either $t \in [0, t_1)$, or $t \in [t_1, \infty)$ and $\rho \geq 1$. Then, because $\xi \neq \mathbf{0}$, Proposition 4.1 implies that $s_r^{-1}(t) < \infty$. Let $\epsilon > 0$ be such that neither ν nor ξ has an atom at $s_r^{-1}(t) + \epsilon$; if $t \in [0, t_1)$, then by Lemma 4.2(i), we may further require that $s_r^{-1}(t) + \epsilon < x_1$. Let $x = s_r^{-1}(t) + \epsilon$. By (80), Theorem 5.3, and monotonicity of $\bar{V}^q(t, \cdot) - \bar{V}^q(s-, \cdot)$ for each fixed $0 \leq s < t < \infty$, there exists Q_1 such that $q > Q_1$ implies

$$\langle \chi 1_{[0, \bar{F}^q(t))}, \bar{\mathcal{Z}}^q(t) \rangle \leq \bar{V}^q(t, x) - \bar{V}^q(\bar{\tau}^q(t)-, x) - (t - \bar{\tau}^q(t)) + \frac{x}{q}.$$

By Lemma 5.10(ii), there exists $Q_2 \geq Q_1$ such that $q > Q_2$ implies

$$\langle \chi 1_{[0, \bar{F}^q(t))}, \bar{\mathcal{Z}}^q(t) \rangle \leq \rho_x (t - \bar{\tau}^q(t)) + \epsilon - (t - \bar{\tau}^q(t)) + \frac{x}{q}.$$

If $t \in [0, t_1)$, then $x < x_1$ and thus $\rho_x - 1 < 0$. Otherwise, $t \in [t_1, \infty)$ and $\rho \geq 1$. Then, because $l_\xi < x_2$, (21) implies that $x > l_\xi \vee x_1$. Hence, $\rho_x - 1 \geq 0$. Then, for $q > Q_2$,

$$\langle \chi 1_{[0, \bar{F}^q(t))}, \bar{\mathcal{Z}}^q(t) \rangle \leq \begin{cases} \epsilon + \dfrac{x}{q}, & t \in [0, t_1), \\ (\rho_x - 1)t + \epsilon + \dfrac{x}{q}, & t \in [t_1, \infty). \end{cases}$$

If $t \in [t_1, \infty)$, then because $l_\xi < x_2$ and $\rho_{x_1} \le 1$, it follows by (21) that $\lim_{\epsilon \searrow 0} \rho_x - 1 = \rho_{l_\xi \vee x_1} - 1 = 0$. Thus, letting q tend to infinity and then letting ϵ decrease to zero completes the proof of (97).

To complete the proof of (i), fix $g \in C_b^+(\mathbb{R}_+)$ and $\eta > 0$. Because neither ν nor ξ has an atom at the origin, there exists $0 < \delta < 1$ such that

$$\langle 1_{[0,\delta]}, \xi + \alpha t \nu \rangle < \frac{\eta}{3\|g\|_\infty}.$$

Fix such a δ that, in addition, has the property that neither ν nor ξ has an atom at δ. Then, there exists Q such that $q > Q$ implies that

$$\langle 1_{[0,\delta]}, \bar{\mathscr{Z}}^q(0) + \bar{\mathscr{V}}^q(t) \rangle < \frac{\eta}{3\|g\|_\infty}, \qquad \frac{1}{q} < \frac{\eta}{3\|g\|_\infty}, \qquad \text{and} \qquad \langle \chi 1_{[0, \bar{F}^q(t))}, \bar{\mathscr{Z}}^q(t) \rangle < \frac{\eta\delta}{3\|g\|_\infty}.$$

By (72), with $x = 0$ and $y = \delta$, for all $q > Q$,

$$\begin{aligned} \langle g 1_{[0, \bar{F}^q(t))}, \bar{\mathscr{Z}}^q(t) \rangle &\le \|g\|_\infty \left(\langle 1_{[0,\delta]}, \bar{\mathscr{Z}}^q(t) \rangle + \frac{1}{\delta} \langle \chi 1_{[\delta, \bar{F}^q(t))}, \bar{\mathscr{Z}}^q(t) \rangle \right) \\ &\le \|g\|_\infty \left(\langle 1_{[0,\delta]}, \bar{\mathscr{Z}}^q(0) + \bar{\mathscr{V}}^q(t) \rangle + \frac{1}{q} + \frac{1}{\delta} \langle \chi 1_{[0, \bar{F}^q(t))}, \bar{\mathscr{Z}}^q(t) \rangle \right) \\ &< \eta. \end{aligned}$$

Hence, (i) holds.

Next, we prove (ii). For this, suppose that $\xi \notin \mathbf{M}_1$. Fix $t \in [0, \infty)$, $x \in [0, x_2 \wedge l_\xi)$, and $\epsilon > 0$. Assume first that neither ν nor ξ has an atom at x. Then, by (77), Lemma 5.10(i), the inequality $x < l_\xi$, and Lemma 5.10(ii), there exists Q such $q > Q$ implies

$$\bar{W}^q(t, x) \le \epsilon + \frac{x}{q} + \rho_x (t - \bar{\tau}^q(t, x)) + \epsilon - (t - \bar{\tau}^q(t, x)).$$

Because $x < x_2 \wedge l_\xi$, $\rho_x \le 1$. Hence, for all $q > Q$,

$$\bar{W}^q(t, x) \le \epsilon + \frac{x}{q} + \epsilon.$$

Letting $q \to \infty$ and then $\epsilon \to 0$ implies that $\lim_{q\to\infty} \bar{W}^q(t, x) = 0$. Because there are at most countably many atoms for ν and ξ in $[0, x_2 \wedge l_\xi)$ and $\bar{W}^q(t, \cdot)$ is nonnegative and nondecreasing,

$$\lim_{q\to\infty} \bar{W}^q(t, x) = 0 \quad \text{for all } x \in [0, x_2 \wedge l_\xi).$$

The result in (ii) follows from this by an argument analogous to that used in the previous paragraph to obtain (i) from (97). $\square$

Lemma 5.13. *Let* $g \in C_b^+(\mathbb{R}_+)$.

(i) *If* $l_\xi < x_2$ *and either* $t \in [0, t_1)$, *or* $t \in [t_1, \infty)$ *and* $\rho_{x_1} \le 1$, *then*

$$\langle g, \mathscr{Z}^*(t) \rangle \le \langle g 1_{[s_\ell^{-1}(t), \infty)}, \xi + \alpha t \nu \rangle. \tag{98}$$

(ii) *If* $l_\xi \ge x_2$, *then for all* $t \in [0, \infty)$,

$$\langle g, \mathscr{Z}^*(t) \rangle \le \langle g 1_{[x_2, \infty)}, \xi + \alpha t \nu \rangle. \tag{99}$$

Proof. We begin by proving (i). Fix $t \in [0, \infty)$. Because $l_\xi < x_2$, it follows that $\xi \ne \mathbf{0}$. If $t \in [t_1, \infty)$ and $\rho < 1$, then (18), (19), (C2), and (12) imply that $\mathscr{Z}^*(t) = \mathbf{0}$, so (98) holds. Otherwise, by Proposition 4.1(i) and Lemma 4.3(i), $s_\ell^{-1}(t) < \infty$. For all $q \in \mathcal{Q}$,

$$\langle g, \bar{\mathscr{Z}}^q(t) \rangle = \langle g 1_{[0, \bar{F}^q(t))}, \bar{\mathscr{Z}}^q(t) \rangle + \langle g 1_{[\bar{F}^q(t), \infty)}, \bar{\mathscr{Z}}^q(t) \rangle.$$

This together with (85) and Lemma 5.12(i) implies that

$$\langle g, \mathscr{Z}^*(t) \rangle \le \limsup_{q\to\infty} \langle g 1_{[\bar{F}^q(t), \infty)}, \bar{\mathscr{Z}}^q(t) \rangle. \tag{100}$$

By (79), for all $q \in \mathcal{Q}$,

$$\langle g1_{[\bar{F}^q(t),\infty)}, \bar{\mathcal{Z}}^q(t)\rangle \le \langle g1_{[\bar{F}^q(t),\infty)}, \bar{\mathcal{Z}}^q(0)\rangle + \langle g1_{[\bar{F}^q(t),\infty)}, \bar{\mathcal{V}}^q(t)\rangle. \tag{101}$$

If $t=0$, (98) follows because $s_\ell^{-1}(0)=0$. Otherwise, $t\in(0,\infty)$. Then, $s_\ell^{-1}(t)>0$. Let $0<\epsilon<s_\ell^{-1}(t)$ be such that neither ν nor ξ has an atom at $s_\ell^{-1}(t)-\epsilon$ and define $a(t,\epsilon)=s_\ell^{-1}(t)-\epsilon$. By (101) and Theorem 5.3, there exists Q such that $q>Q$ implies

$$\langle g1_{[\bar{F}^q(t),\infty)}, \bar{\mathcal{Z}}^q(t)\rangle \le \langle g1_{(a(t,\epsilon),\infty)}, \bar{\mathcal{Z}}^q(0)\rangle + \langle g1_{(a(t,\epsilon),\infty)}, \bar{\mathcal{V}}^q(t)\rangle. \tag{102}$$

Hence, because neither ν nor ξ has an atom at $a(t,\epsilon)$, (85) and (102) imply that

$$\limsup_{q\to\infty}\langle g1_{[\bar{F}^q(t),\infty)}, \bar{\mathcal{Z}}^q(t)\rangle \le \langle g1_{(a(t,\epsilon),\infty)}, \xi+\alpha t\nu\rangle.$$

Letting $\epsilon\searrow 0$, we obtain

$$\limsup_{q\to\infty}\langle g1_{[\bar{F}^q(t),\infty)}, \bar{\mathcal{Z}}^q(t)\rangle \le \langle g1_{[s_\ell^{-1}(t),\infty)}, \xi+\alpha t\nu\rangle. \tag{103}$$

Combining (100) and (103) yields (98).

Next, we prove (ii). Fix $t\in[0,\infty)$. Because $l_\xi\ge x_2$, $\xi\notin\mathbf{M}_1$. Hence, by (85) and Lemma 5.12(ii), for each $x<x_2$,

$$\langle g, \mathcal{Z}^*(t)\rangle \le \limsup_{q\to\infty}\langle g1_{(x,\infty)}, \bar{\mathcal{Z}}^q(t)\rangle.$$

Fix $x<x_2$ such that neither ξ nor ν has an atom at x. Given $\epsilon>0$, let $M>x$ be such that neither ξ nor ν has an atom at M and $\|g\|_\infty\langle 1_{[M,\infty)}, \xi+\alpha t\nu\rangle<\epsilon$. Let $y_1^1=x$ and $y_2^1=M$. For $n=2,3,4,\ldots$, let $\{y_k^n\}_{k=1}^{m(n)}\subset[x,M]$ be such that $\{y_k^{n-1}\}_{k=1}^{m(n-1)}\subset\{y_k^n\}_{k=1}^{m(n)}$, neither ξ nor ν has an atom at y_k^n for all $k=1,\ldots,m(n)$, and $0<y_{k+1}^n-y_k^n<2(M-x)/n$ for all $k=1,\ldots,m(n)-1$. Given $n\in\mathbb{N}$ and $y\in[x,M)$, let k be such that $y\in[y_k^n,y_{k+1}^n)$ and set

$$g^n(y)=\sup\{g(z)\colon z\in[y_k^n,y_{k+1}^n]\}.$$

Then, for each q and $n\in\mathbb{N}$,

$$\langle g1_{(x,\infty)}, \bar{\mathcal{Z}}^q(t)\rangle \le \sum_{k=1}^{m(n)} g^n(y_k^n)\langle 1_{[y_k^n,y_{k+1}^n)}, \bar{\mathcal{Z}}^q(t)\rangle + \|g\|_\infty\langle 1_{[M,\infty)}, \bar{\mathcal{Z}}^q(t)\rangle.$$

By (85) and (72), for each $k=1,\ldots,m(n)-1$,

$$\begin{aligned}\limsup_{q\to\infty}\langle 1_{[y_k^n,y_{k+1}^n)}, \bar{\mathcal{Z}}^q(t)\rangle &\le \limsup_{q\to\infty}\langle 1_{[y_k^n,y_{k+1}^n]}, \bar{\mathcal{Z}}^q(0)+\bar{\mathcal{V}}^q(t)\rangle + \frac{1}{q}\\ &= \langle 1_{[y_k^n,y_{k+1}^n]}, \xi+\alpha t\nu\rangle = \langle 1_{[y_k^n,y_{k+1}^n)}, \xi+\alpha t\nu\rangle.\end{aligned}$$

Also, by (71) and (85),

$$\limsup_{q\to\infty}\langle 1_{[M,\infty)}, \bar{\mathcal{Z}}^q(t)\rangle \le \limsup_{q\to\infty}\langle 1_{[M,\infty)}, \bar{\mathcal{Z}}^q(0)+\bar{\mathcal{V}}^q(t)\rangle = \langle 1_{[M,\infty)}, \xi+\alpha t\nu\rangle.$$

It follows that for all $n\in\mathbb{N}$,

$$\langle g, \mathcal{Z}^*(t)\rangle \le \langle g^n 1_{[x,M)}, \xi+\alpha t\nu\rangle + \epsilon.$$

Letting n tend to infinity and ϵ tend to zero,

$$\langle g, \mathcal{Z}^*(t)\rangle \le \langle g1_{[x,\infty)}, \xi+\alpha t\nu\rangle.$$

Finally, letting x increase to x_2 completes the proof. □

The following definitions are needed for the proof of the final lemma, Lemma 5.14.

5.3.2.4. Time-shifted fluid limits. For $s \in [0, \infty)$ and $q \in \mathcal{Q}$, define the time s shifted stochastic processes $\bar{E}_s^q(t) = \bar{E}^q(s+t) - \bar{E}^q(s)$, $\bar{\mathcal{V}}_s^q(t) = \bar{\mathcal{V}}^q(s+t) - \bar{\mathcal{V}}^q(s)$, $\bar{V}_s^q(t) = \bar{V}^q(s+t) - \bar{V}^q(s)$, $\bar{\mathcal{Z}}_s^q(t) = \bar{\mathcal{Z}}^q(s+t)$, and $\bar{W}_s^q(t) = \bar{W}^q(s+t)$ for all $t \in [0, \infty)$. In addition, for $s \in [0, \infty)$, define the time s shifted limit process $\mathcal{V}_s^*(t) = \mathcal{V}^*(s+t) - \mathcal{V}^*(s)$, $V_s^*(t) = V^*(s+t) - V^*(s)$, $\mathcal{Z}_s^*(t) = \mathcal{Z}^*(s+t)$, and $W_s^*(t) = W^*(s+t)$ for all $t \in [0, \infty)$. It is immediate that for $s \in [0, \infty)$, the time s shifted stochastic processes satisfy (25)–(27). In addition, by (85) and (86) for each $s \in [0, \infty)$ as $q \to \infty$,

$$(\bar{\mathcal{V}}_s^q(\cdot), \bar{V}_s^q(\cdot), \bar{W}_s^q(0), \bar{\mathcal{Z}}_s^q(\cdot)) \to (\mathcal{V}_s^*(\cdot), V_s^*(\cdot), W_s^*(0), \mathcal{Z}_s^*(\cdot)). \tag{104}$$

Then, for each $s \in [0, \infty)$ such that $\mathcal{Z}_s^*(0)$ does not charge the origin, it follows that the time s shifted stochastic processes satisfy (25)–(29). In particular, if $s \in [0, \infty)$ is such that $\mathcal{Z}_s^*(0)$ does not charge the origin, then any result proved for $\mathcal{Z}^*(\cdot)$ also holds for $\mathcal{Z}_s^*(\cdot)$.

Lemma 5.14. *$\mathcal{Z}^*(\cdot)$ satisfies* (C3).

Proof. For $t = 0$, (C3) is immediate because $\mathcal{Z}^*(0) = \xi$ and $L^*(0) = l_\xi$. Therefore, it suffices to consider $t \in (0, \infty)$. This is proved in three cases.

Case 1. Assume that $l_\xi \geq x_2$. By Lemmas 5.11(ii) and 5.13(ii), for all $t \in (0, \infty)$ and $g \in C_b^+(\mathbb{R}_+)$,

$$\langle g 1_{(l_\xi, \infty)}, \xi + \alpha t \nu \rangle \leq \langle g, \mathcal{Z}^*(t) \rangle \leq \langle g 1_{[x_2, \infty)}, \xi + \alpha t \nu \rangle. \tag{105}$$

If $x_2 = \infty$, then, by (105), $\mathcal{Z}^*(t) = \mathbf{0}$ and $L^*(t) = \infty$ for all $t \in (0, \infty)$ so that (C3) holds. Otherwise, $x_2 < \infty$. Then $\rho > 1$ and by (105), $x_2 \leq L^*(t)$ for all $t \in (0, \infty)$. Hence, (105) implies that for all $t \in (0, \infty)$ and $g \in C_b^+(\mathbb{R}_+)$,

$$\langle g, \mathcal{Z}^*(t) \rangle \leq \langle g 1_{[L^*(t), \infty)}, \xi + \alpha t \nu \rangle. \tag{106}$$

Letting $g \nearrow \chi$ in (106) and invoking the monotone convergence theorem yields, for all $t \in (0, \infty)$,

$$\langle \chi, \mathcal{Z}^*(t) \rangle \leq \langle \chi 1_{[L^*(t), \infty)}, \xi + \alpha t \nu \rangle.$$

This together with (C2) and the inequality $\rho > 1$ implies that $\alpha \langle \chi 1_{[0, L^*(t))}, \nu \rangle \leq 1$ for all $t \in (0, \infty)$. Hence, by (14), $L^*(t) \leq x_2$ for all $t \in (0, \infty)$. Thus, for all $t \in (0, \infty)$,

$$L^*(t) = x_2. \tag{107}$$

Case 1(a). Assume that $l_\xi = x_2$. Then, by (105) and (107), (C3) holds for all $t \in (0, \infty)$.

Case 1(b). Assume that $l_\xi > x_2$. For each $\epsilon > 0$, consider the time-shifted fluid limit point $\mathcal{Z}_\epsilon^*(\cdot) = \mathcal{Z}^*(\cdot + \epsilon)$. Then, by (107), for each $\epsilon > 0$, $L_\epsilon^*(t) = x_2$ for all $t \in [0, \infty)$; in particular, $L_\epsilon^*(0) = x_2$. Because $x_2 > 0$, $\mathcal{Z}_\epsilon^*(0)$ does not charge the origin for each $\epsilon > 0$. Hence, by the commentary on time-shifted fluid limit points (preceding the statement of Lemma 5.14 and Case 1(a), (C3) holds for $\mathcal{Z}_\epsilon^*(\cdot)$ for each $\epsilon > 0$. Then, for all $t \in (0, \infty)$, $\epsilon \in (0, t)$, and $g \in C_b^+(\mathbb{R}_+)$, we obtain

$$\langle g 1_{(x_2, \infty)}, \mathcal{Z}^*(\epsilon) + \alpha(t-\epsilon)\nu \rangle \leq \langle g, \mathcal{Z}_\epsilon^*(t-\epsilon) \rangle \leq \langle g 1_{[x_2, \infty)}, \mathcal{Z}^*(\epsilon) + \alpha(t-\epsilon)\nu \rangle.$$

Using the fact that for each $0 < \epsilon < t < \infty$ and $g \in C_b^+(\mathbb{R}_+)$, $\langle g, \mathcal{Z}_\epsilon^*(t-\epsilon) \rangle = \langle g, \mathcal{Z}^*(t) \rangle$, letting ϵ decrease to zero, and using continuity of $\mathcal{Z}^*(\cdot)$ completes the proof of (C3) for $t \in (0, \infty)$ when $l_\xi > x_2$.

Case 2. Assume that $l_\xi < x_2$ and $t \in (0, t_1)$, or $t \in [t_1, \infty)$ and $\rho_{x_1} \leq 1$. Fix $t \in (0, \infty)$. By Lemmas 5.11(i) and 5.13(i), for all $g \in C_b^+(\mathbb{R}_+)$,

$$\langle g 1_{(s_r^{-1}(t), \infty)}, \xi + \alpha t \nu \rangle \leq \langle g, \mathcal{Z}^*(t) \rangle \leq \langle g 1_{[s_\ell^{-1}(t), \infty)}, \xi + \alpha t \nu \rangle. \tag{108}$$

In the upper bound, we wish to replace $s_\ell^{-1}(t)$ with $s_r^{-1}(t)$. Because $s_\ell^{-1}(t) \leq s_r^{-1}(t)$, it suffices to show that

$$\langle 1_{[0, s_r^{-1}(t))}, \mathcal{Z}^*(t) \rangle = 0. \tag{109}$$

Then, to complete the proof of (C3) in Case 2, we must show that

$$L^*(t) = s_r^{-1}(t). \tag{110}$$

Indeed, the combination of (108), (109), and (110) imply (C3) at time t. We now proceed to verify (109) and (110).

Case 2(a). Assume that $t \in (0, t_1)$. Then, by Lemma 4.3(ii), $s_r^{-1}(t) \le s_\ell^{-1}(s)$ for $t < s < t_1$. This together with (108) at time s, for $t < s < t_1$ implies that $[0, s_r^{-1}(t))$ does not intersect the support of $\mathcal{Z}^*(s)$ for all $t < s < t_1$. Hence, by continuity of $\mathcal{Z}^*(\cdot)$, (109) holds. Then, (108) and (109) imply that for all $g \in C_b^+(\mathbb{R}_+)$,

$$\langle g 1_{(s_r^{-1}(t),\infty)}, \xi + \alpha t \nu\rangle \le \langle g, \mathcal{Z}^*(t)\rangle \le \langle g 1_{[s_r^{-1}(t),\infty)}, \xi + \alpha t \nu\rangle. \tag{111}$$

Hence, $s_r^{-1}(t) \le L^*(t)$. Suppose that $s_r^{-1}(t) < L^*(t)$. Then by (111), for all $g \in C_b^+(\mathbb{R}_+)$,

$$\langle g, \mathcal{Z}^*(t)\rangle = \langle g 1_{[L^*(t),\infty)}, \xi + \alpha t \nu\rangle. \tag{112}$$

Letting $g \nearrow \chi$ in (112) and using the monotone convergence theorem and then (C2) together with the fact that $\langle \chi, \xi\rangle + (\rho - 1)t > 0$ because $t \in (0, t_1)$, it follows that

$$\langle \chi, \xi\rangle + (\rho - 1)t = \langle \chi 1_{[L^*(t),\infty)}, \xi + \alpha t \nu\rangle. \tag{113}$$

Hence,

$$\langle \chi 1_{[0, L^*(t))}, \xi\rangle = t(1 - \langle \chi 1_{[0, L^*(t))}, \alpha \nu\rangle). \tag{114}$$

If $L^*(t) > x_1$, then the left side of (114) is positive because $t_1 > 0$ implies $\xi \in \mathbf{M}_1$. However, the right side is nonpositive, which is a contradiction. Hence, $s_r^{-1}(t) < L^*(t) \le x_1$. Then, both sides of (114) are necessarily positive because the right side is positive if $L^*(t) < x_1$ and the left side is positive if $L^*(t) = x_1$ because $\xi \in \mathbf{M}_1$. Rearranging (114) gives $s(L^*(t)-) = t$. By (20), $s(s_r^{-1}(t)) \ge t$. Hence, by monotonicity of $s(\cdot)$, $s(x) = t$ for all $x \in [s_r^{-1}(t), L^*(t))$, which contradicts (20). Thus, $s_r^{-1}(t) = L^*(t)$ and (110) holds.

Case 2(b). Assume that $t \in [t_1, \infty)$ and $\rho < 1$. Then, $t_1 < \infty$ and because $\rho < 1$, (C2) implies that $\mathcal{Z}^*(t) = \mathbf{0}$ and (109) holds. Moreover, $L^*(t) = \infty$. Because $\rho < 1$ and $l_\xi < x_2 = x_1$, by (21), $s_r^{-1}(t) = x_1 = \infty$. Hence, (110) holds.

Case 2(c). Assume that $t \in [t_1, \infty)$, $\rho_{x_1} \le 1$, and $\rho \ge 1$. If $t_1 > 0$, then, $\xi \in \mathbf{M}_1$. If $t \in (t_1, \infty)$, then, by (21) and (45), $s_\ell^{-1}(t) = s_r^{-1}(t) = x_1$ and thus, by (108), (109) holds. Because $s_r^{-1}(t_1) = x_1$, right continuity of $\mathcal{Z}^*(\cdot)$ and (109) on (t_1, ∞) together imply (109) for $t = t_1$. If $t_1 = 0$, then $\xi \notin \mathbf{M}_1$ and $t \in (0, \infty)$. Hence, $s_r^{-1}(t) = x_2 \wedge l_\xi = l_\xi$ and by Lemma 5.12(ii) and (85), (109) holds. Thus, in Case 2(c), (109) holds.

Next, we verify (110). By (109), $L^*(t) \ge s_r^{-1}(t)$. Suppose that $L^*(t) > s_r^{-1}(t) = x_1 \vee l_\xi \ge l_\xi$. Then, we may replace $s_\ell^{-1}(t)$ with $L^*(t)$ in the upper bound in (108), let $g \nearrow \chi$, invoke the monotone convergence theorem, apply (C2) and the inequality $\rho \ge 1$, and rearrange terms to obtain

$$\langle \chi 1_{[0, L^*(t))}, \xi + \alpha t \nu\rangle \le t.$$

Then, because $L^*(t) > l_\xi$, $\langle \chi 1_{[0, L^*(t))}, \xi\rangle > 0$. Hence,

$$\alpha \langle \chi 1_{[0, L^*(t))}, \nu\rangle < 1.$$

Then $L^*(t) \le x_1$, which is a contradiction. Hence, (110) holds.

Case 3. Assume that $l_\xi < x_2$, $t \in [t_1, \infty)$, and $\rho_{x_1} > 1$. Then $x_1 = x_2$, $\xi \in \mathbf{M}_1$, $\rho > 1$, and $t_1 \in (0, \infty)$.

Case 3(a). Assume that $t = t_1$. By Case 2(a), continuity of $\mathcal{Z}^*(\cdot)$, and the fact that $\lim_{t \nearrow t_1} s_r^{-1}(t) = x_0 \le x_1 = x_2$ (see Lemma 4.2(iii)), it follows that for all $g \in C_b^+(\mathbb{R}_+)$,

$$\langle g 1_{(x_0,\infty)}, \xi + \alpha t_1 \nu\rangle \le \langle g, \mathcal{Z}^*(t_1)\rangle \le \langle g 1_{[x_0,\infty)}, \xi + \alpha t_1 \nu\rangle.$$

Hence, $L^*(t_1) \ge x_0$. In addition, by the same reasoning used in Case 1 to argue from (106), $L^*(t_1) \le x_2$. If $x_0 = x_2$, (C3) follows at time t_1. Otherwise, $x_0 < x_2 = x_1$. Because $s(x) = t_1$ for all $x \in [x_0, x_1)$, it follows that the union of the supports of ν and ξ does not intersect (x_0, x_1). Hence, for some $a \ge 0$ and for all $g \in C_b^+(\mathbb{R}_+)$,

$$\langle g, \mathcal{Z}^*(t_1)\rangle = a g(x_0) + \langle g 1_{[x_2,\infty)}, \xi + \alpha t_1 \nu\rangle.$$

Because $\xi \in \mathbf{M}_1$, then $x_0 \ne 0$. Letting $g \nearrow \chi$, applying the monotone convergence theorem, using (C2), and the fact that ν and ξ do not charge (x_0, x_1) implies that $a = (\langle \chi 1_{[0, x_0]}, \xi + \alpha t_1 \nu\rangle - t_1)/x_0$. Hence, by (21), (45), and Lemma 4.3(vi), $a = 0$. Then, $L^*(t_1) = x_2$ and (C3) follows at time t_1.

Case 3(b): Assume that $t \in (t_1, \infty)$. Then consider the time-shifted fluid limit point $\mathcal{Z}_{t_1}^*(\cdot) = \mathcal{Z}^*(\cdot + t_1)$. By Case 3(a), $L_{t_1}^*(0) = L^*(t_1) = x_2$. Since $x_2 > 0$, $\mathcal{Z}_{t_1}^*(0)$ does not charge the origin; thus, by the commentary on time-shifted fluid limit points preceding the statement of Lemma 5.14 and Case 1, (C3) follows. □

Acknowledgments. The research of Douglas G. Down is supported in part by the Natural Sciences and Engineering Research Council of Canada. The research of H. Christian Gromoll is supported in part by National Science Foundation (NSF) Mathematical Sciences Postdoctoral Research Fellowship and NSF Grant DMS 0707111. The research of Amber L. Puha is supported in part by NSF Grant 03-05272 and a National Security Agency Young Investigators Award. The authors gratefully acknowledge the stimulating environment at EURANDOM, where this work was begun.

References

[1] Bansal, N., M. Harchol-Balter. 2001. Analysis of SRPT scheduling: Investigating unfairness. *Proc. 2001 ACM Sigmetrics Internat. Conf. Measurement Modeling Comput. Systems*, SIGMETRICS, Cambridge, MA.
[2] Billingsley, P. 1968. *Convergence of Probability Measures*. John Wiley & Sons, Inc., New York.
[3] Down, D. G., R. Wu. 2006. Multi-layered round robin routing for parallel servers. *Queueing Systems* **53** 177–188.
[4] Down, D., H. C. Gromoll, A. L. Puha. 2009. State-dependent response times via fluid limits for shortest remaining processing time queues. *ACM Sigmetrics Performance Evaluation Review* (poster session). Forthcoming.
[5] Doytchinov, B., J. L. Lehoczky, S. E. Shreve. 2001. Real-time queues in heavy traffic with earliest-deadline-first queue discipline. *Ann. Appl. Probab.* **11** 332–378.
[6] Ethier, S., T. Kurtz. 1986. *Markov Processes: Characterization and Convergence*. John Wiley & Sons, Inc., New York.
[7] Gromoll, H. C., R. J. Williams. 2009. Fluid limits for networks with bandwidth sharing and general document size distributions. *Ann. Appl. Probab.* **19** 243–280.
[8] Gromoll, H. C., A. L. Puha, R. J. Williams. 2002. The fluid limit for a processor sharing queue. *Ann. Appl. Probab.* **12** 797–859.
[9] Kallenberg, O. 1986. *Random Measures*. Academic Press, New York.
[10] Kruk, Ł., J. L. Lehoczky, S. E. Shreve, S. N. Yeung. 2004. Earliest-deadline-first service in heavy traffic acyclic networks. *Ann. Appl. Probab.* **14** 1306–1352.
[11] Núñez Queija, R. 2002. Queues with equally heavy sojourn time and service requirement distributions. *Ann. Oper. Res.* **113** 101–117.
[12] Nuyens, M., B. Zwart. 2006. A large deviations analysis of the GI/GI/1 SRPT queue. *Queueing Systems* **54** 85–97.
[13] Perera, R. 1993. The variance of delay time in queueing system M/G/1 with optimal strategy SRPT. *Archiv Elektronik Übertragungstechnik* **47** 110–114.
[14] Prohorov, Yu. V. 1956. Convergence of random processes and limit theorems in probability theory. *Theory Probab. Appl.* **1** 157–214.
[15] Puha, A. L., A. L. Stoylar, R. J. Williams. 2006. The fluid limit of an overloaded processor sharing queue. *Math. Oper. Res.* **31** 316–350.
[16] Schassberger, R. 1990. The steady-state appearance of the M/G/1 queue under the discipline of shortest remaining processing time. *Adv. Appl. Probab.* **22** 456–479.
[17] Schrage, L. 1968. A proof of the optimality of the shortest remaining processing time discipline. *Oper. Res.* **16** 687–690.
[18] Schrage, L. E., L. W. Miller. 1966. The queue M/G/1 with the shortest remaining processing time discipline. *Oper. Res.* **14** 670-684.
[19] Schreiber, F. 1993. Properties and applications of the optimal queueing strategy SRPT: A survey. *Archiv Elektronik Übertragungstechnik* **47** 372–378.
[20] Smith, D. R. 1976. A new proof of the optimality of the shortest remaining processing time discipline. *Oper. Res.* **26** 197–199.
[21] Wierman, A., M. Harchol-Balter. 2003. Classifying scheduling policies with respect to unfairness in an M/G/1. *Proc. 2003 ACM Sigmetrics Internat. Conf. Measurement Modeling Comput. Systems*, SIGMETRICS, New York.