Waiting Time Asymptotics for Time Varying Multiserver Queues with Abandonment and Retrials

A. Mandelbaum
Technion Institute
Haifa, 32000
ISRAEL
avim@tx.technion.ac.il

W. A. Massey
Bell Labs, Lucent Technologies
Murray Hill, NJ 07974
U.S.A.
will@research.bell-labs.com

M. I. Reiman
Bell Labs, Lucent Technologies
Murray Hill, NJ 07974
U.S.A.
marty@research.bell-labs.com

A. L. Stolyar
Bell Labs, Lucent Technologies
Murray Hill, NJ 07974
U.S.A.
stolyar@research.bell-labs.com

Abstract

We consider a nonstationary Markov multiserver queueing model where waiting customers may abandon and subsequently retry. In this paper we derive fluid and diffusion approximations for the associated *waiting time* process. The fluid and diffusion approximations for the corresponding *queue length* process were obtained in [4] (see also [5]).

1 Introduction

The model we consider in this paper is a multi-server queue with time-varying parameters, in which customers are impatient and hence abandon after (subjectively) excessive wait. Moreover, obtaining service is important enough for some customers that they return and seek service (retry) after a "time-out". Formally, our model is depicted in Figure 1: there is a single "service" node with n_t, $t \geq 0$, servers. New customers arrive to the service node following a Poisson process of rate λ_t. Customers arriving to find an idle server are taken into service that has rate μ_t^1. Customers that find all servers busy join a queue, from which they are served in a FCFS manner. Each customer waiting in the queue abandons at rate β_t. An abandoning customer leaves the system with probability ψ_t or joins a retrial pool with probability $1 - \psi_t$. Each customer in the retrial pool leaves to enter the service node at rate μ_t^2. Upon entry to the service node, these customers are treated the same as new customers. The behavior of the system is described by the two-dimensional, continuous time Markov chain $\mathbf{Q}(t) = \big(Q_1(t), Q_2(t)\big)$ where $Q_1(t)$ equals the number of customers residing in the service node (waiting or being served) and $Q_2(t)$ equals the number of customers in the retrial pool.

Our work is motivated by the need to develop analytical tools that support performance analysis of *large* telecommunication systems, such as telephone *call centers*, where

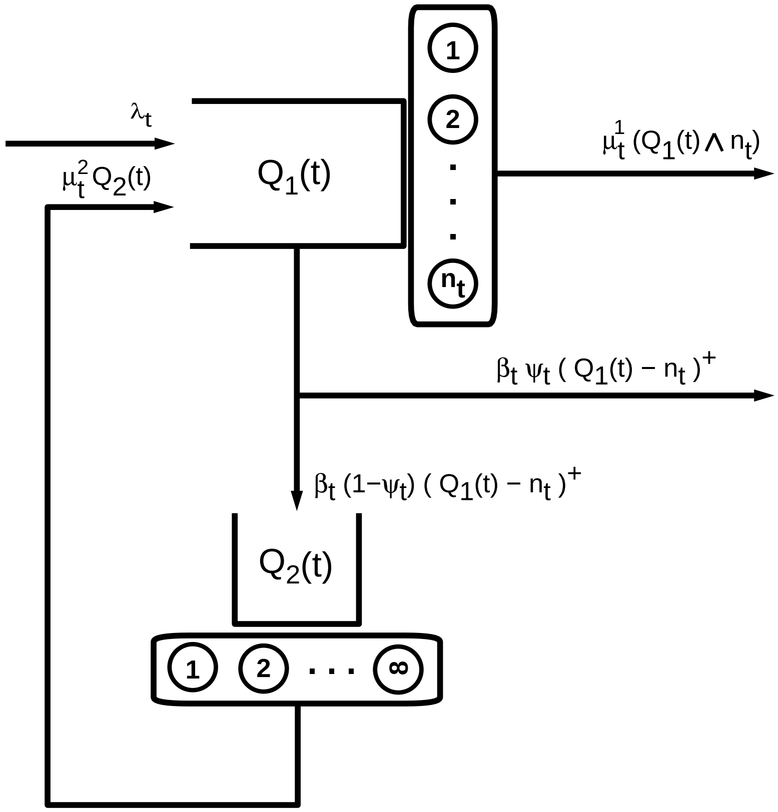

Figure 1: The abandonment queue with retrials.

abandonments and retrials arise naturally. Call centers are constantly subject to time-varying conditions, and waiting customers in phone queues are unable to observe the state of the system. It follows that time-dependent modeling (as opposed to also state-dependent) is natural for call centers. Finally, we point out that the analysis of *waiting times* is (typically) analytically more challenging than that of the queue lengths, and in many applications (like call centers) is probably more important. For more discussion and related references on these issues, see [5].

The remainder of this paper is organized as follows. The asymptotic regime we consider is introduced in Section 2. In Section 3 we provide fluid and diffusion limits for the virtual waiting time at a fixed time τ. Process level fluid and diffusion limits for the virtual waiting time are presented in Section 4.

2 Asymptotic Regime

As mentioned above we are interested in the behavior of a system with large number of servers and large input rate. Thus, we consider the asymptotic regime where we scale up the number of servers in response to a similar scaling up of the arrival rate by customers.

More precisely, the asymptotic regime is as follows. Assume that $\lambda_t, \beta_t, \mu_t^1, \mu_t^2, \psi_t, n_t$ are fixed functions of time t. We consider a *sequence* of systems indexed by scaling parameter $\eta = \eta_1, \eta_2, \ldots, \eta_k \to \infty$ as $k \to \infty$. (To avoid cumbersome notation, in what follows, we index a system by η, and when we write $\eta \to \infty$, we mean that η goes to infinity by taking values from the sequence $\eta_1, \eta_2, \ldots.$) In a system with index η, the arrival rate (i.e., the intensity of the Poisson arrival process) is $\eta\lambda_t$ and the number of servers is ηn_t. (Actually, the latter should be, for example, the integer part of ηn_t, but again, to avoid trivial complications and simplify notation, we assume it's just ηn_t.) We also make the following additional

Assumption 2.1 The function n_t is *continuously differentiable* in $[0, \infty)$.

Sample paths of the family of queue length processes $\mathbf{Q}^\eta(t) = \left(Q_1^\eta(t), Q_2^\eta(t)\right)$, indexed by the scaling parameter η, are determined by the following equations:

$$\begin{aligned} Q_1^\eta(t) &= Q_1^\eta(0) + \Pi_{21}^c\left(\int_0^t Q_2^\eta(s)\mu_s^2 ds\right) - \Pi_{12}^b\left(\int_0^t \left(Q_1^\eta(s) - \eta n_s\right)^+ \beta_s(1-\psi_s)ds\right) \qquad (2.1) \\ &\quad + \Pi^a\left(\int_0^t \eta\lambda_s ds\right) - \Pi^b\left(\int_0^t \left(Q_1^\eta(s) - \eta n_s\right)^+ \beta_s\psi_s ds\right) - \Pi^c\left(\int_0^t \left(Q_1^\eta(s) \wedge \eta n_s\right)\mu_s^1 ds\right) \end{aligned}$$

and

$$Q_2^\eta(t) = Q_2^\eta(0) + \Pi_{12}^b\left(\int_0^t \left(Q_1^\eta(s) - \eta n_s\right)^+ \beta_s(1-\psi_s)ds\right) - \Pi_{21}^c\left(\int_0^t \left(Q_2^\eta(s)\right)\mu_s^2 ds\right) \ , \quad (2.2)$$

where $\Pi^a, \Pi^b, \Pi^c, \Pi_{12}^b, \Pi_{21}^c$, are independent standard (rate 1) Poisson processes. In this paper we use the notation $x \wedge y = \min(x, y)$ and $x^+ = \max(x, 0)$ for all real x and y.

Throughout this paper we assume that the following *initial conditions* hold:

$$\frac{1}{\eta}\mathbf{Q}^\eta(0) \to \mathbf{Q}^{(0)}(0) \ , \qquad (2.3)$$

$$\eta^{-1/2}[\mathbf{Q}^\eta(0) - \eta\mathbf{Q}^{(0)}(0)] \to \mathbf{Q}^{(1)}(0) \ , \qquad (2.4)$$

where $\mathbf{Q}^{(0)}(0)$ and $\mathbf{Q}^{(1)}(0)$ are fixed vectors, and

$$Q_1^{(0)}(0) > 0 \ . \qquad (2.5)$$

In the rest of the paper we also use the following notation. Let E be a complete separable metric space, and a be a real number. Then we denote by $\mathcal{D}(E, a)$ the Skorohod space of E-valued functions defined in the interval $[a, \infty)$ which are right continuous and have left limits. The space $\mathcal{D}(E, a)$ is endowed with Skorohod J_1-metric and the corresponding topology.

3 Waiting Time in Node 1: Marginal Distribution at a Given Time.

Suppose that we are interested in the waiting time of a "virtual" customer arriving at station 1 at a *fixed* time $\tau \geq 0$. Since we have a system with abandonment, a convenient way to approach this problem is to consider the system that is obtained from the original one by the following modification. *Suppose, that after time τ, there are no new exogenous arrivals into the system, and any customer departing any station i leaves the system.* In other words, starting time τ, each station i has no new arrivals, and it just serves the customers which were at the station at time τ. Theorem 5.1 in [4] still applies to the modified system; the only difference is that the terms in the equations, corresponding to the arrivals after time τ, should be "zeroed out". Namely, the following results follow directly from Theorem 5.1 (and its proof) in [4].

Denote the arrival and departure processes for station 1 by

$$A^\eta = \{\, A^\eta(t) \mid t \geq 0 \,\} \quad \text{and} \quad \Delta^\eta = \{\, \Delta^\eta(t) \mid t \geq 0 \,\}$$

respectively. Let, by convention, the arrival process include the customers in node 1 at time 0, so $A^\eta(0) = \mathbf{Q}_1^\eta(0)$, $\Delta^\eta(0) = 0$, and $A^\eta(t) - \Delta^\eta(t) = Q_1^\eta(t),\ t \geq 0$.

Then we obtain the following *fluid* limit result.

Theorem 3.1 *With probability* 1*, the following convergence holds uniformly on compact sets (u.o.c.) of* t*:*

$$\frac{1}{\eta}(\mathbf{Q}^\eta, A^\eta, \Delta^\eta) \to (\mathbf{Q}^{(0)}, A^{(0)}, \Delta^{(0)}) \tag{3.1}$$

where $\mathbf{Q}^\eta = (Q_1^\eta, Q_2^\eta)$, $\mathbf{Q}^{(0)} = (Q_1^{(0)}, Q_2^{(0)})$*, the fluid limit* $Q^{(0)}$ *satisfies the following equations*

$$Q_1^{(0)}(t) = Q_1^{(0)}(0) + \int_0^t \left[\lambda_s + \mu_s^2 Q_2^{(0)}(s)\right] 1_{\{s \leq \tau\}} - \mu_s^1\Big(Q_1^{(0)}(s) \wedge n_s\Big) - \beta_s\Big(Q_1^{(0)}(s) - n_s\Big)^+ ds \tag{3.2}$$

and

$$Q_2^{(0)}(t) = Q_2^{(0)}(0) + \int_0^{t\wedge\tau} \beta_s(1-\psi_s)\Big(Q_1^{(0)}(s) - n_s\Big)^+ ds - \int_0^t \mu_s^2 Q_2^{(0)}(s) ds\ . \tag{3.3}$$

Moreover, $A^{(0)}$ *and* $\Delta^{(0)}$ *are equal to*

$$A^{(0)}(t) = Q_1^{(0)}(0) + \int_0^{t\wedge\tau} \left[\lambda_s + \mu_s^2 Q_2^{(0)}(s)\right] ds \tag{3.4}$$

and

$$\Delta^{(0)}(t) = \int_0^t \left[\mu_s^1\Big(Q_1^{(0)}(s) \wedge n_s\Big) + \beta_s\Big(Q_1^{(0)}(s) - n_s\Big)^+\right] ds\ , \tag{3.5}$$

where $\Delta^{(0)}$ *is a continuously differentiable non-decreasing function in* $[0,\infty)$.

We also obtain the following *diffusion* limit.

Theorem 3.2 *The following weak convergence holds (in the space being the direct product of corresponding Skorohod spaces* $\mathcal{D}(\mathbb{R},0)$*) :*

$$\sqrt{\eta}\left(\frac{1}{\eta}\mathbf{Q}^\eta - \mathbf{Q}^{(0)}, \frac{1}{\eta}A^\eta - A^{(0)}, \frac{1}{\eta}\Delta^\eta - \Delta^{(0)}\right) \xrightarrow{d} (\mathbf{Q}^{(1)}, A^{(1)}, \Delta^{(1)}), \tag{3.6}$$

where $\mathbf{Q}^{(1)} = (Q_1^{(1)}, Q_2^{(1)})$ *is the unique continuous solution to the stochastic differential equations*

$$\begin{aligned} Q_1^{(1)}(t) \;=\; & Q_1^{(1)}(0) + \int_0^t \left[\mu_s^1\Big(Q_1^{(1)}(s) - Q_1^{(1)}(s)^*\Big) + \beta_s Q_1^{(1)}(s)^*\right] ds \\ & + \int_0^{t\wedge\tau} \mu_s^2 Q_2^{(1)}(s) ds - B_{21}^c\left(\int_0^{t\wedge\tau}\Big(Q_2^{(0)}(s)\Big)\mu_s^2 ds\right) + B^a\left(\int_0^{t\wedge\tau}\lambda_s ds\right) \\ & - B_{12}^b\left(\int_0^t \Big(Q_1^{(0)}(s) - n_s\Big)^+ \beta_s(1-\psi_s) ds\right) - B^b\left(\int_0^t \Big(Q_1^{(0)}(s) - n_s\Big)^+ \beta_s\psi_s ds\right) \\ & - B^c\left(\int_0^t \Big(Q_1^{(0)}(s) \wedge n_s\Big)\mu_s^1 ds\right) \end{aligned} \tag{3.7}$$

and

$$\begin{aligned} Q_2^{(1)}(t) \;=\; & Q_2^{(1)}(0) + \int_0^{t\wedge\tau} Q_1^{(1)}(s)^* \beta_s(1-\psi_s) ds - \int_0^t \mu_s^2 Q_2^{(1)}(s) ds \\ & + B_{21}^c\left(\int_0^{t\wedge\tau}\Big(Q_2^{(0)}(s)\Big)\mu_s^2 ds\right) + B_{12}^b\left(\int_0^{t\wedge\tau}\Big(Q_1^{(0)}(s) - n_s\Big)^+ \beta_s(1-\psi_s) ds\right), \end{aligned} \tag{3.8}$$

with

$$Q_1^{(1)}(t)^* = Q_1^{(1)}(t)^+ 1_{\{Q_1^{(0)}(t)\geq n_t\}} - Q_1^{(1)}(t)^- 1_{\{Q_1^{(0)}(t)>n_t\}}, \tag{3.9}$$

where $A^{(1)}$ *and* $\Delta^{(1)}$ *are defined as*

$$A^{(1)}(t) = Q_1^{(1)}(0) + \int_0^{t\wedge\tau} \mu_s^2 Q_2^{(1)}(s)ds - B_{21}^c\left(\int_0^{t\wedge\tau}\left(Q_2^{(0)}(s)\right)\mu_s^2 ds\right) + B^a\left(\int_0^{t\wedge\tau}\lambda_s ds\right) \tag{3.10}$$

and

$$\Delta^{(1)}(t) = \int_0^t \left[\mu_s^1\left(Q_1^{(1)}(s) - Q_1^{(1)}(s)^*\right) + \beta_s Q_1^{(1)}(s)^*\right] ds + B^c\left(\int_0^t \left(Q_1^{(0)}(s)\wedge n_s\right)\mu_s^1 ds\right) \tag{3.11}$$
$$+B_{12}^b\left(\int_0^t\left(Q_1^{(0)}(s) - n_s\right)^+ \beta_s(1-\psi_s)ds\right) + B^b\left(\int_0^t\left(Q_1^{(0)}(s)-n_s\right)^+\beta_s\psi_s ds\right).$$

Clearly,

$$Q_1^{(1)}(t) = A^{(1)}(t) - \Delta^{(1)}(t)\ . \tag{3.12}$$

Now, let us define the "potential service initiation" process D^η for node 1 by

$$D^\eta(t) = \Delta^\eta(t) + \eta n_t,\ \ t\geq 0\ .$$

Note that if $Q_1^\eta(t) < \eta n_t$, then $A^\eta(t) < D^\eta(t)$; so the potential service can be "ahead" of arrivals.

Obviously, we have the (probability 1, u.o.c.) convergence:

$$\frac{1}{\eta}D^\eta(t) \to D^{(0)}(t),\ \ t\geq 0,$$

where $D^{(0)}(t) = \Delta^{(0)}(t) + n_t, t\geq 0$. Since n_t is continuously differentiable by assumption and we know that $\Delta^{(0)}(t)$ is continuously differentiable, $D^{(0)}(t)$ is also continuously differentiable and we denote its derivative by $d^{(0)}(t)$. Now we will make an important (but not very restrictive in majority of applications) additional assumption.

Assumption 3.1. The function $D^{(0)}$ (of t) is continuously differentiable with *strictly positive derivative*, and

$$\lim_{t\to\infty} D^{(0)}(t) > A^{(0)}(\tau)\ . \tag{3.13}$$

(Note, that according to our definitions, both $A^\eta(\cdot)$ and $A^{(0)}(\cdot)$ are constant in the interval $[\tau,\infty)$.)

Also, it will be convenient to adopt a convention that all the processes we consider are defined in the interval $[-T,\infty)$, with

$$T = n_0/d^{(0)}(0)\ .$$

We make this extension by assuming that nothing is happening in the interval $[-T,0)$ (no arrivals or departures) except the number of servers is increasing linearly from 0 to ηn_0 (for the unscaled process with index η).

We then can rewrite (3.1) and (3.6) as follows (with all the functions being now defined for $t\geq -T$):

$$\frac{1}{\eta}(\mathbf{Q}^\eta, A^\eta, D^\eta) \to (\mathbf{Q}^{(0)}, A^{(0)}, D^{(0)}) \tag{3.14}$$

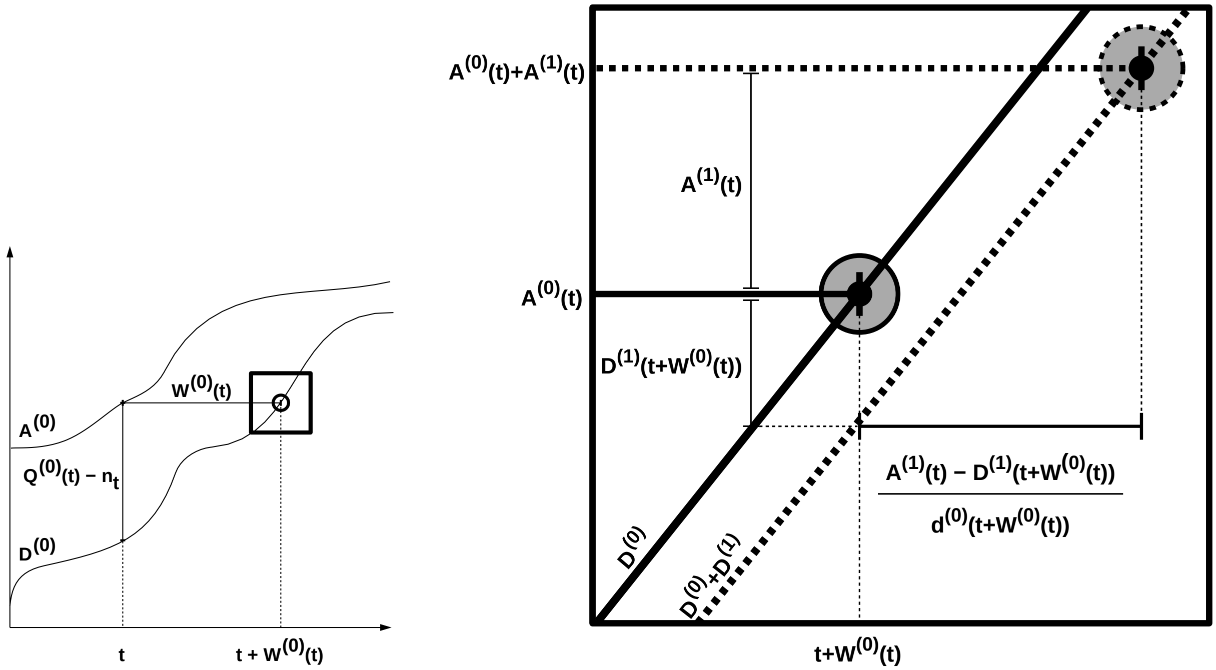

Figure 2: The diffusion term for the attainment waiting time

and

$$\sqrt{\eta}\Big(\frac{1}{\eta}\mathbf{Q}^{\eta} - \mathbf{Q}^{(0)}, \frac{1}{\eta}A^{\eta} - A^{(0)}, \frac{1}{\eta}D^{\eta} - D^{(0)}\Big) \xrightarrow{d} (\mathbf{Q}^{(1)}, A^{(1)}, D^{(1)}) \ , \tag{3.15}$$

where

$$D^{(1)} = \Delta^{(1)} \ . \tag{3.16}$$

Note that processes $A^{(0)}, D^{(0)}, A^{(1)}, D^{(1)}$ are continuous and $D^{(0)}(-T) = D^{(1)}(-T) = 0$.

Our conventions together with the Assumption 3.1 make the following processes well defined and finite with probability 1 for all sufficiently large η. Let us define, for all $t \geq -T$, the *first attainment* processes

$$S^{\eta}(t) = \inf\{s \geq -T : \ D^{\eta}(s) > A^{\eta}(t)\}$$

and

$$S^{(0)}(t) = \inf\{s \geq -T : \ D^{(0)}(s) > A^{(0)}(t)\}, \tag{3.17}$$

and the *attainment* waiting time processes

$$W^{\eta}(t) = S^{\eta}(t) - t$$

and

$$W^{(0)}(t) = S^{(0)}(t) - t \ . \tag{3.18}$$

Denote by $\hat{W}^{\eta}(\tau)$ the *virtual* waiting time at τ, i.e. the time a "test" customer (in the original non-modified system) arriving in node 1 at time τ would have to wait until its service starts, assuming this customer *does not abandon* while waiting. Then the relation between the virtual waiting time $\hat{W}^{\eta}(\tau)$ and the attainment waiting time $W^{\eta}(\tau)$ is simply

$$\hat{W}^{\eta}(\tau) = W^{\eta}(\tau)^{+} \ . \tag{3.19}$$

Indeed, note that $W^\eta(\tau)$ (and $W^{(0)}(\tau)$) may be negative. All this means is that $Q_1^\eta(\tau) < \eta n_\tau$, and therefore in this case $\hat{W}^\eta(\tau) = 0$. If $W^\eta(\tau)$ is non-negative, then its value is exactly equal to the virtual waiting time.

It follows directly from Theorem and Corollary in [7] that (3.14), (3.15), and Assumption 3.1, imply the following convergences.

With probability 1, u.o.c.,

$$(\frac{1}{\eta}\mathbf{Q}^\eta, \frac{1}{\eta}A^\eta, \frac{1}{\eta}D^\eta, W^\eta) \to (\mathbf{Q}^{(0)}, A^{(0)}, D^{(0)}, W^{(0)}) \ . \tag{3.20}$$

In distribution,

$$\sqrt{\eta}(\frac{1}{\eta}\mathbf{Q}^\eta - Q^{(0)}, \frac{1}{\eta}A^\eta - A^{(0)}, \frac{1}{\eta}D^\eta - D^{(0)}, W^\eta - W^{(0)}) \xrightarrow{d} (Q^{(1)}, A^{(1)}, D^{(1)}, W^{(1)}) \ , \tag{3.21}$$

where

$$W^{(1)}(t) = \frac{A^{(1)}(t) - D^{(1)}(S^{(0)}(t))}{d^{(0)}(S^{(0)}(t))} \ .$$

Since the processes $A^{(1)}, D^{(1)}, Q^{(1)}, W^{(1)}$ are continuous with probability 1, we automatically obtain the weak convergence of finite dimensional distributions.

In particular, consider the non-trivial case $S^{(0)}(\tau) \geq \tau$ (which is equivalent to $Q_1^{(0)}(\tau) \geq n_\tau$). We obtain

$$W^\eta(\tau) \to W^{(0)}(\tau)$$

and

$$\sqrt{\eta}(W^\eta(\tau) - W^{(0)}(\tau)) \xrightarrow{d} W^{(1)}(\tau) = \frac{Q_1^{(1)}(S^{(0)}(\tau))}{d^{(0)}(S^{(0)}(\tau))} \ .$$

Solving equation (3.2) for $Q_1^{(0)}(\cdot)$ in the interval $[\tau, \infty)$, we obtain

$$Q_1^{(0)}(t) = Q_1^{(0)}(\tau) \exp\left(-\int_\tau^t \beta_s ds\right) + \int_\tau^t \exp\left(-\int_s^t \beta_r dr\right)(\beta_s - \mu_s^1) n_s ds \ , \ t \geq \tau.$$

We can find $S^{(0)}(\tau)$ from

$$S^{(0)}(\tau) = \min\{t \geq \tau \mid Q_1^{(0)}(t) = n_t\} \ .$$

Solving a stochastic differential equation for $Q_1^{(1)}(\cdot)$ in the interval $[\tau, S^{(0)}(\tau)]$, we obtain (cf. [2]

$$Q_1^{(1)}(S^{(0)}(\tau)) \stackrel{d}{=} Q_1^{(1)}(\tau) \exp\left(-\int_\tau^{S^{(0)}(\tau)} \beta_s ds\right) + \int_\tau^{S^{(0)}(\tau)} \exp\left(-\int_s^t \beta_r dr\right) f_s dB(s-\tau) \ ,$$

where

$$f_t^2 = (Q_1^{(0)}(t) - n_t)\beta_t + n_t \mu_t^1 \ ,$$

and B is a standard Brownian motion process. In particular,

$$\mathsf{E}[Q_1^{(1)}(S^{(0)}(\tau))] = \mathsf{E}[Q_1^{(1)}(\tau)] \exp\left(-\int_\tau^{S^{(0)}(\tau)} \beta_s ds\right)$$

and

$$\mathsf{Var}[Q_1^{(1)}(S^{(0)}(\tau))] = \mathsf{Var}[Q_1^{(1)}(\tau)] \exp\left(-\int_\tau^{S^{(0)}(\tau)} 2\beta_r dr\right) + \int_\tau^{S^{(0)}(\tau)} \exp\left(-\int_s^{S^{(0)}(\tau)} 2\beta_r dr\right) f_s^2 ds \ .$$

Note that in case $Q_1^{(0)}(\tau) = n_\tau$, we obtain

$$S^{(0)}(\tau) = \tau,\ W^{(0)}(\tau) = 0,\ d^{(0)}(\tau) = \mu_\tau^1 n_\tau + n'_\tau,$$

and, therefore,

$$\sqrt{\eta} W^\eta(\tau) \xrightarrow{d} W^{(1)}(\tau) = \frac{Q_1^{(1)}(\tau)}{\mu_\tau^1 n_\tau + n'_\tau}\ .$$

Recalling (3.19), we obtain the following diffusion limit for the virtual waiting time in this case

$$\sqrt{\eta}\hat{W}^\eta(\tau) \xrightarrow{d} \frac{Q_1^{(1)}(\tau)^+}{\mu_\tau^1 n_\tau + n'_\tau},\ \ \text{if}\ \ Q_1^{(0)}(\tau) = n_\tau\ ,$$

which is what we intuitively expected.

We checked the accuracy of the fluid approximation for the virtual waiting time via simulation. The system we considered has all parameters constant except for λ_t. In particular we considered $n_t = 50$, $\mu_t^1 = 1$, $\mu_t^2 = 0.2$, $\beta_t = 0.25$, and $\psi_t = 0.5$, with $\lambda_t = 10 + 20t - t^2$, $0 \le t \le 20$. The results are shown in Figure 3. The graph on top compares the fluid and simulation results for the queue length, and the graph on the bottom compares the fluid and simulation results for the virtual waiting time. (The simulation results depicted are an average of 5000 independent replications. More details on the simulation method are contained in [5].)

4 Waiting Time in Node 1: A Process

In the previous section we derived fluid and diffusion approximations of the marginal distribution of the attainment waiting time, which uniquely determines those for the virtual waiting time, in node 1 *at a given time* $\tau \ge 0$. A natural conjecture is that one can obtain similar asymptotics for the attainment waiting time as a *random process* defined for $\tau \in [0, \infty)$. In this section we present results showing that the above conjecture is indeed true.

We need more definitions. First, in this section, unless otherwise explicitly stated, we view all the processes as random processes of two time variables, $t \in [-T, \infty)$ and $\tau \in [0, \infty)$. (In the previous section τ was a fixed parameter.) More precisely, we view them as random elements $X = ((X(t,\tau),\ t \in [-T,\infty)),\ \tau \ge 0)$ (X can be Q_i^η or A^η or $Q_i^{(j)}$, etc.) taking values in the space $\mathcal{D}(\mathcal{D}(\mathbb{R}, -T), 0)$.

Note that for each fixed τ all processes of interest are well defined in the previous section, and the convergences (3.20) and (3.21) do hold for any fixed τ.

Assumption 4.1 Assumption 3.1 holds for any $\tau \ge 0$.

A generalization of the argument used in the proofs in [4] (roughly, making all estimates in the convergence proofs "uniform on τ"), and a generalization of the results in [7], lead to the following results which are extensions of (3.20) and (3.21). The details are contained in [6]. First, we state our functional strong law of large numbers result.

Theorem 4.1 *With probability* 1*, uniformly on compact sets of* (t, τ)*,*

$$(\frac{1}{\eta}\mathbf{Q}^\eta, \frac{1}{\eta}A^\eta, \frac{1}{\eta}D^\eta, S^\eta, W^\eta) \to (\mathbf{Q}^{(0)}, A^{(0)}, D^{(0)}, S^{(0)}, W^{(0)})\ , \tag{4.1}$$

where all functions $\mathbf{Q}^{(0)}$*,* $A^{(0)}$*,* $D^{(0)}$*,* $S^{(0)}$*,* $W^{(0)}$*, are continuous jointly on* τ *and* t*, and for each fixed* τ *they (as functions of* t*) satisfy the ODE (3.2), (3.3), and equations (3.4), (3.5), (3.17), (3.18). Moreover,* $d^{(0)}(t,\tau) \equiv (\partial/\partial t)D^{(0)}(t,\tau)$ *is strictly positive.*

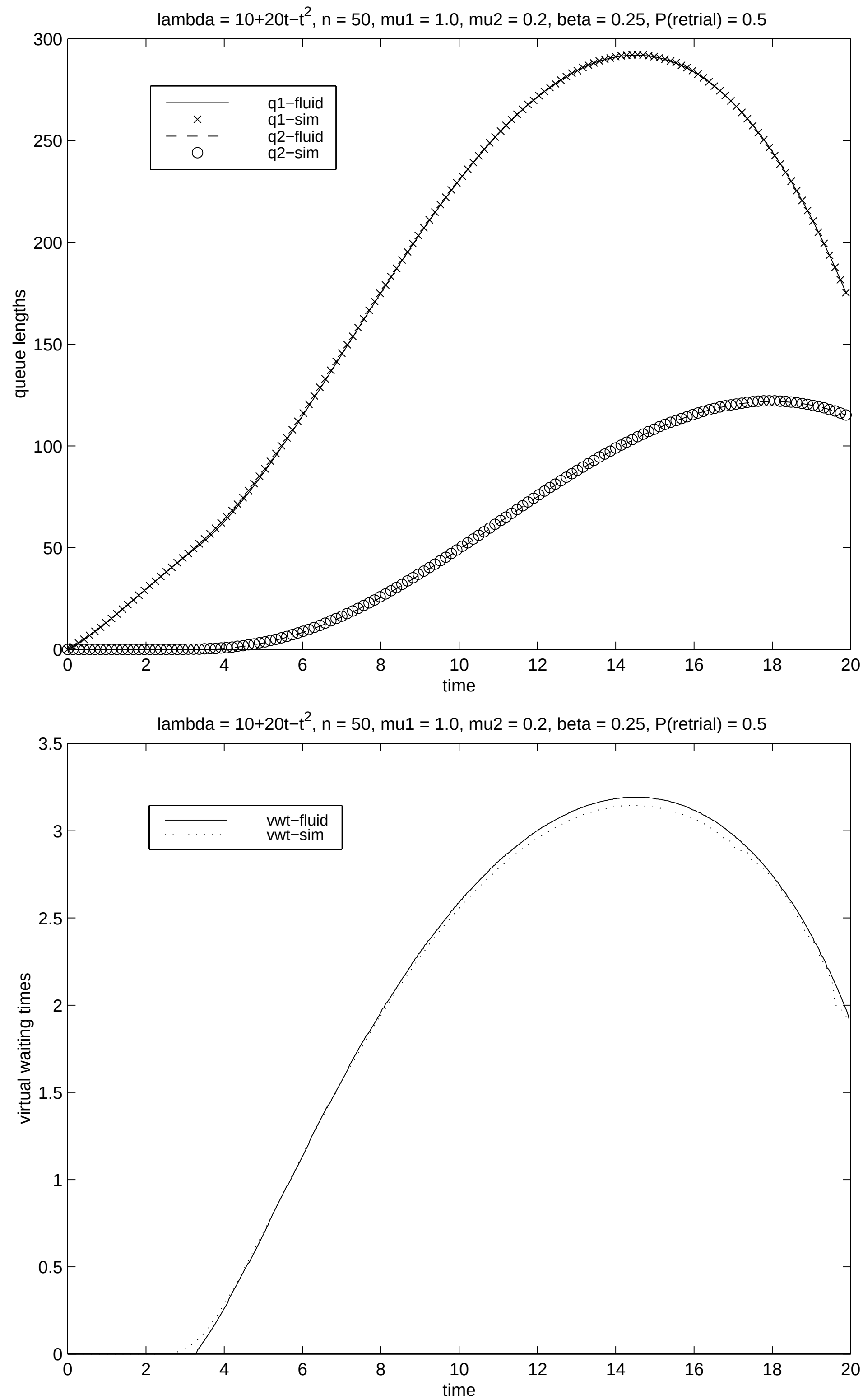

Figure 3: The fluid approximations for the queue lengths and virtual waiting time

Now we state our functional central limit theorem.

Theorem 4.2 *The following weak convergence holds:*

$$\sqrt{\eta}(\frac{1}{\eta}\mathbf{Q}^{\eta} - \mathbf{Q}^{(0)}, \frac{1}{\eta}A^{\eta} - A^{(0)}, \frac{1}{\eta}D^{\eta} - D^{(0)}, (W^{\eta}(\tau,\tau) - W^{(0)}(\tau,\tau),\ \tau \geq 0)) \ \stackrel{d}{\to}$$
$$(\mathbf{Q}^{(1)}, A^{(1)}, D^{(1)}, (W^{(1)}(\tau,\tau), \tau \geq 0)) \ , \qquad (4.2)$$

where $\mathbf{Q}^{(1)}$ is a jointly continuous on τ and t random process, which (as a function of t, with τ fixed) is the unique solution to the stochastic differential equations (3.7) and (3.8); $A^{(1)}$ and $D^{(1)}$ (as functions of t) satisfy (3.10), (3.11), (3.16), and are jointly continuous on τ and t; and

$$W^{(1)}(\tau,\tau) = \frac{A^{(1)}(\tau,\tau) - D^{(1)}(S^{(0)}(\tau,\tau),\tau)}{d^{(0)}(S^{(0)}(\tau,\tau),\tau)}$$

is continuous on τ.

5 Acknowledgements

The authors thank Brian Rider of Courant Institute for his work on the numerical simulation, computation, and plots of the graphs shown in Figure 3.

References

[1] S. N. Ethier and T. G. Kurtz. *Markov Process: Characterization and Convergence.* John Wiley and Sons, New York, 1986.

[2] I. Karatzas and S. E. Shreve. *Brownian Motion and Stochastic Calculus (Second Edition).* Springer-Verlag, New York, 1991.

[3] T. G. Kurtz. Strong approximation theorems for density dependent Markov chains. *Stochastic Processes and Their Applications*, 6:223–240, 1978.

[4] A. Mandelbaum, W. A. Massey, M. I. Reiman. Strong Approximations for Markovian Service Networks. *Queueing Systems*, (1998).

[5] A. Mandelbaum, W. A. Massey, M. I. Reiman, B. Rider. Time Varying Multiserver Queues with Abandonment and Retrials. *ITC-16*, Edinburgh, Scotland, (1999).

[6] A. Mandelbaum, W. A. Massey, M. I. Reiman, A. L. Stolyar. Waiting Time Asymptotics for Multiserver, Nonstationary Jackson Networks with Abandonment. *In preparation.*

[7] A. Puhalskii. On the Invariance Principle for the First Passage Time. *Mathematics of Operatioins Research*, Vol. 19, (1994), pp. 946-954.