Strong Approximations for Markovian Service Networks

Avi Mandelbaum [a] William A. Massey [b] Martin I. Reiman [c]

[a] *Davidson Faculty of Industrial Engineering and Management, Technion Institute, Haifa, 32000 ISRAEL,*
E-mail: avim@tx.technion.ac.il

[b] *Bell Laboratories, Lucent Technologies, Office 2C-320, Murray Hill, NJ 07974, U.S.A.,*
E-mail: will@research.bell-labs.com

[c] *Bell Laboratories, Lucent Technologies, Office 2C-315, Murray Hill, NJ 07974, U.S.A.,*
E-mail: marty@research.bell-labs.com

Inspired by service systems such as telephone call centers, we develop limit theorems for a large class of stochastic service network models. They are a special family of nonstationary Markov processes where parameters like arrival and service rates, routing topologies for the network, and the number of servers at a given node are all functions of time as well as the current state of the system. Included in our modeling framework are networks of $M_t/M_t/n_t$ queues with abandonment and retrials. The asymptotic limiting regime that we explore for these networks has a natural interpretation of scaling up the number of servers in response to a similar scaling up of the arrival rate for the customers. The individual service rates however, are not scaled.

We employ the theory of strong approximations to obtain functional strong laws of large numbers and functional central limit theorems for these networks. This gives us a tractable set of network fluid and diffusion approximations. A common theme for service network models with features like many servers, priorities, or abandonment is "non-smooth" state dependence that has not been covered systematically by previous work. We prove our central limit theorems in the presence of this non-smoothness by using a new notion of derivative.

Keywords: Strong Approximations, Fluid Approximations, Diffusion Approximations, Multiserver Queues, Queues with Abandonment, Queues with Retrials, Priority Queues, Queueing Networks, Jackson Networks, Nonstationary Queues.

AMS Subject classification: Primary: 60K25, 90B22, 60F17. Secondary: 60J60, 60J27, 60F15, 60F05.

Contents

1. Introduction and Summary

Motivated by the need to design and analyze Markovian service networks, we investigate fluid and diffusion limits for such systems. The main distinguishing feature of (most, but not all of) the systems we consider in this paper is that service is provided by a large supply of servers, and there is a corresponding large demand for this service. It is these large quantities that motivate the asymptotic regime we consider. Our methods allow us to consider networks with time dependent parameters, state dependent routing, abandonment, and retrials.

To make the description of our models and results easier to follow, we first consider a simple example (see Figure 1.1). The $M_t/M_t/n_t$ queue has a (time-inhomogeneous) Poisson arrival process with rate λ_t, a service rate (per server) of μ_t, and n_t servers, for all $t \geq 0$. We can construct the sample paths for the $M_t/M_t/n_t$ queue length process as the unique set of solutions to the functional equation

$$Q(t) = Q(0) + A_1\left(\int_0^t \lambda_s ds\right) - A_2\left(\int_0^t \mu_s \cdot (Q(s) \wedge n_s) ds\right), \qquad (1.1)$$

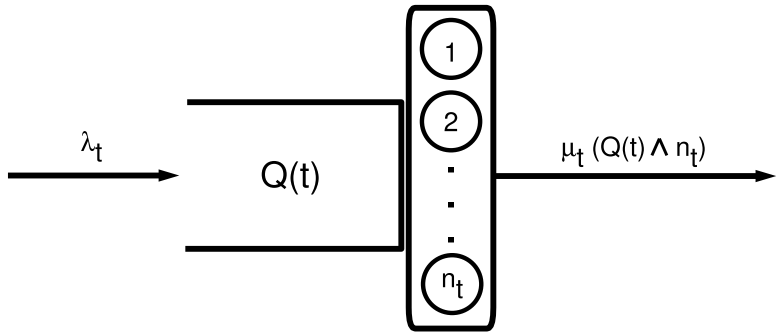

Figure 1.1. The $M_t/M_t/n_t$ queue.

where $A_1(\cdot)$ and $A_2(\cdot)$ are given independent, standard (rate 1) Poisson processes, and for all real x and y, $x \wedge y \equiv \min(x, y)$.

The asymptotic approach to the $M_t/M_t/1$ queue, as used in both Massey [15], [16] and Mandelbaum and Massey [11], was to create a family of associated $M_t/M_t/1$ queues where the queue indexed by $\eta > 0$ has arrival rate $\eta\lambda_t$ and service rate $\eta\mu_t$. We then determined the asymptotic behavior for the time evolution of this family of queues, when $\eta \to \infty$.

For the $M_t/M_t/n_t$ queue, we also create a family of associated processes. The key difference here is that, for the $M_t/M_t/n_t$ queue indexed by η, we want to have both the arrival rate and number of servers grow large, i.e. scaled up by η, but leave the service rate unscaled.

We are then interested in the asymptotic behavior of the processes

$$Q^\eta(t) = Q^\eta(0) + A_1\left(\int_0^t \eta\lambda_s ds\right) - A_2\left(\int_0^t \mu_s \cdot (Q^\eta(s) \wedge \eta n_s)\, ds\right) \tag{1.2}$$

$$= Q^\eta(0) + A_1\left(\int_0^t \eta\lambda_s ds\right) - A_2\left(\int_0^t \eta\mu_s \cdot \Big(\frac{1}{\eta} Q^\eta(s) \wedge n_s\Big) ds\right) \tag{1.3}$$

as $\eta \to \infty$. Note that (1.3) shows that the scaling we want is equivalent to the simultaneous scaling of λ_t and μ_t with multiplication by η, provided that we also divide Q^η by η (when $n_t \equiv 1$, this distinction does not matter since the $M_t/M_t/1$ service rate indexed by η is $\eta\mu_t 1_{\{Q_t^\eta > 0\}}$, where $1_{\{Q_t^\eta > 0\}}$ is the indicator function for the event $\{Q_t^\eta > 0\}$, which is the same as $\eta\mu_t 1_{\{Q_t^\eta/\eta > 0\}}$).

Equation (1.3) is a special case of equation (2.8), which is in turn a special case of equation (2.9). Equation (2.9) defines the processes of interest to us in the general service network setting. For systems with infinitesimal rates that are not state dependent, the scaling used in (1.3) and (2.8) is the same as *uniform* *acceleration*, as considered in Massey [15], [16] and Mandelbaum and Massey [11] for the $M_t/M_t/1$ queue, as well as Massey and Whitt [18] for the general case of finite state, time-inhomogeneous, continuous time Markov chains. In these articles a parameter ϵ is used and limits are taken such that $\epsilon \downarrow 0$. This parameterization can be reconciled with the notation in this paper by setting $\epsilon = 1/\eta$. We refer to the scaling in (1.3) and (2.8) as uniform acceleration also, even though this involves a slight abuse of terminology.

It seems appropriate to comment here on two distinguishing features of the above formulation that carry over to the general results of this paper: many 'unscaled' servers, and time-dependent parameters. Our original motivating examples were call centers, where service involves an interaction between either two people (the customer and server), or a person and a machine (the person is the customer). In either case, because a person is involved, it does not seem reasonable to scale the service rates with η. Thus, in order to accommodate the arrivals, whose rate is proportional to η, the number of servers must be scaled with η. Time dependent arrival rates should need no justification, since phenomena such as rush hours are quite common. Time dependent service rates can be used to model phenomena such as server fatigue or changes in the nature of services over the day. Finally, a time dependent number of servers arises with shift changes and in systems where the number of servers is varied to accommodate changes in the arrival rate.

Our "first-order" asymptotic result takes the form of a *functional strong law of large numbers* (FSLLN), and yields a fluid approximation for the original process. For the above $M_t/M_t/n_t$ example, the FSLLN states that

$$\lim_{\eta\to\infty} \frac{1}{\eta} Q^\eta(0) = Q^{(0)}(0) \quad \text{implies} \quad \lim_{\eta\to\infty} \frac{1}{\eta} Q^\eta(t) = Q^{(0)}(t) \ \text{ a.s.}, \tag{1.4}$$

uniformly on compact sets in t, where $Q^{(0)} = \left\{ Q^{(0)}(t) \mid t \geq 0 \right\}$ is the unique process that solves the integral equation

$$Q^{(0)}(t) = Q^{(0)}(0) + \int_0^t \left[\lambda_s - \mu_s \cdot (Q^{(0)}(s) \wedge n_s) \right] ds, \tag{1.5}$$

for all $t \geq 0$. The general version of our FSLLN is Theorem 2.2.

The above FSLLN can be refined with a *functional central limit theorem* (FCLT). A fundamental difficulty arises in attempting to apply prior results to obtain the FCLT, even for the $M_t/M_t/n_t$ queue. The resolution of this difficulty for general Markovian service networks is the purpose of this paper. Before stating the FCLT for the $M_t/M_t/n_t$ queue we first point out the essence of the difficulty.

Consider a sequence of real valued random variables $\{X_n,\ n \geq 1\}$ that correspond to partial sums of i.i.d. random variables (with finite means μ and variances σ^2). Letting $Y_n = X_n/n$ and

$$Z_n = \frac{\sqrt{n}}{\sigma}(Y_n - \mu) = \frac{X_n - n\mu}{\sigma\sqrt{n}}, \tag{1.6}$$

then the strong law of large numbers and central limit theorem yield

$$\lim_{n\to\infty} Y_n = \mu \text{ a.s.} \quad \text{and} \quad \lim_{n\to\infty} Z_n \stackrel{d}{=} Z, \tag{1.7}$$

where $\lim_{n\to\infty} Z_n \stackrel{d}{=} Z$ indicates convergence in distribution, and Z is a standard (mean 0, variance 1) normal random variable. Let $f : \mathbb{R} \to \mathbb{R}$ be differentiable in a neighborhood of μ, then (cf. [21])

$$\lim_{n\to\infty} f(Y_n) = f(\mu) \quad \text{a.s., and} \quad \lim_{n\to\infty} \frac{\sqrt{n}}{\sigma}\Big(f(Y_n) - f(\mu)\Big) \stackrel{d}{=} f'(\mu)Z. \tag{1.8}$$

What happens if f is continuous but not differentiable at μ? Continuity is sufficient to ensure that $\lim_{n\to\infty} f(Y_n) = f(\mu)$ a.s. If $f'(\mu+) \equiv \lim_{x\downarrow\mu} f'(x)$ and $f'(\mu-) \equiv \lim_{x\uparrow\mu} f'(x)$ both exist, then a more careful treatment can be used to show that

$$\lim_{n\to\infty} \frac{\sqrt{n}}{\sigma}\Big(f(Y_n) - f(\mu)\Big) \stackrel{d}{=} f'(\mu+)Z^+ - f'(\mu-)Z^-,$$

where for all real x and y, $x \vee y \equiv \max(x, y)$, $x^+ \equiv x \vee 0$, and $x^- \equiv (-x) \vee 0$.

Going back to the $M_t/M_t/n_t$ example, we cannot apply previous results, such as Kurtz [9] to obtain an FCLT for its queue length because the function $f_t(x) = x \wedge n_t$ is not differentiable with respect to x at $x = n_t$. To circumvent this difficulty we introduce a new notion of derivative (in the context of a multivariate function), which we call the *scalable Lipschitz derivative*. For example, if $\Lambda f_t(x; y)$ denotes the scalable Lipschitz derivative of f_t at x for any real y, then

$$\Lambda f_t(x; y) = y \cdot 1_{\{x<n_t\}} - y^- \cdot 1_{\{x=n_t\}} = y^+ \cdot 1_{\{x<n_t\}} - y^- \cdot 1_{\{x\leq n_t\}}. \tag{1.9}$$

Using this new notion of derivative we are able to obtain an FCLT for a wider class of stochastic models. For the $M_t/M_t/n_t$ queue, the FCLT that arises from uniform acceleration states that, if $\{ Q^\eta(0) \mid \eta > 0 \}$ is a family of random variables (see Section 2 for the independence assumption), then

$$\lim_{\eta \to \infty} \sqrt{\eta} \left[\frac{Q^\eta(0)}{\eta} - Q^{(0)}(0) \right] \stackrel{d}{=} Q^{(1)}(0) \tag{1.10}$$

implies

$$\lim_{\eta \to \infty} \sqrt{\eta} \left[\frac{Q^\eta(t)}{\eta} - Q^{(0)}(t) \right] \stackrel{d}{=} Q^{(1)}(t), \tag{1.11}$$

where $Q^{(0)}$ is defined in (1.5), $Q^{(1)} = \left\{ Q^{(1)}(t) \mid t \geq 0 \right\}$ is the unique stochastic process that solves the integral equation

$$\begin{aligned} Q^{(1)}(t) = Q^{(1)}(0) &- \int_0^t \mu_s 1_{\left\{Q^{(0)}(s) < n_s\right\}} Q^{(1)}(s)^+ ds + \int_0^t \mu_s 1_{\left\{Q^{(0)}(s) \leq n_s\right\}} Q^{(1)}(s)^- ds \\ &+ B_1 \left(\int_0^t \lambda_s ds \right) - B_2 \left(\int_0^t \mu_s \cdot \left(Q^{(0)}(s) \wedge n_s \right) ds \right), \end{aligned} \tag{1.12}$$

and $B_1(\cdot)$, $B_2(\cdot)$ are two independent, standard Brownian motions.

Although it seems clear that the FCLT for the $M_t/M_t/n_t$ queue could be proved on an ad-hoc basis without the scalable Lipschitz derivative, this notion is the key to proving the FCLT for more general systems, which is our Theorem 2.3. We are also able to obtain ordinary differential equations for the mean and covariance of the diffusion limit arising in the FCLT. These are given in Theorem 2.4.

We actually obtain a more refined FCLT which is motivated by the work of Halfin and Whitt [4]. They identify an important asymptotic regime that corresponds to parameter asymptotics of the form

$$\lambda_s^\eta = \eta \lambda_s + \sqrt{\eta} \ell_s + o(\sqrt{\eta}) \quad \text{and} \quad \mu_s^\eta = \eta \mu_s + \sqrt{\eta} m_s + o(\sqrt{\eta}). \tag{1.13}$$

The fluid limit is unchanged. The resulting refined diffusion limit for the $M_t/M_t/n_t$ queue is

$$\begin{aligned} Q^{(1)}(t) = Q^{(1)}(0) &- \int_0^t \mu_s 1_{\{Q^{(0)}(s) < n_s\}} Q^{(1)}(s)^+ ds \\ &+ \int_0^t \mu_s 1_{\{Q^{(0)}(s) \leq n_s\}} Q^{(1)}(s)^- ds + \int_0^t \left[\ell_s - m_s \cdot \left(Q^{(0)}(s) \wedge n_s \right) \right] ds \\ &+ B_1 \left(\int_0^t \lambda_s ds \right) - B_2 \left(\int_0^t \mu_s \cdot \left(Q^{(0)}(s) \wedge n_s \right) ds \right). \end{aligned} \tag{1.14}$$

If we set $\lambda_s = \lambda$, $\mu_s = \mu$, $n_s = n$, $m_s = 0$, and $\ell_s = -\mu\beta$ with $\lambda = \mu n$, and let $Q^\eta(0) = \eta n$, we recover the $M/M/n$ special case for the diffusion limit of [4].

The SLLN and the FCLT are proved in two steps. First, we prove a *strong approximation theorem*, which in the context of the $M_t/M_t/n_t$ queue states that, as $\eta \to \infty$

$$Q^\eta(t) = Q^\eta(0) + \int_0^t \left(\lambda_s^\eta - \mu_s^\eta \cdot \left(\frac{1}{\eta}Q^\eta(s) \wedge n_s\right)\right) ds \tag{1.15}$$
$$+B_1\left(\int_0^t \lambda_s^\eta ds\right) - B_2\left(\int_0^t \mu_s^\eta \cdot \left(\frac{1}{\eta}Q^\eta(s) \wedge n_s\right) ds\right) + O(\log \eta) \quad \text{a.s.,}$$

where $B_1(\cdot)$ and $B_2(\cdot)$ are as above and the convergence is uniform on compact t sets. The general version of this result is Theorem 2.1. The limit theorems then follow from a more detailed asymptotic analysis of this approximation theorem.

Although we leave the precise specification of our model and assumptions to Section 2, we describe here two more examples that illustrate the breath of our framework. One example (see Section 7) is a single node with several customer classes operating under the preemptive priority discipline (Figure 7.1), and the other (see Section 5) is a system with customer abandonment and retrials (Figure 5.1). More complicated examples, such as a Jackson network (Figure 4.1) and a network with state dependent routing (Figure 8.1) are treated in the body of the paper (Sections 4 and 8, respectively). All the network examples given in the paper have the features of time-varying rates and multiserver nodes.

The priority system we consider has c classes of customers and n_t servers. Customers of class i arrive as a Poisson process with rate λ_t^i and have service rate μ_t^i. (All the arrival and service processes are constructed from mutually independent Poisson processes.) Class i is given preemptive priority over any class j such that $i > j$, $1 \leq i, j \leq c$.

The system with abandonment and retrials has a single "service" node with n_t servers. New customers arrive to the service node in a Poisson process of rate λ_t. Customers arriving to find an idle server are taken into service with rate μ_t^1. Customers that find all servers busy join the queue, from which they are taken into service 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 the 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.

Systems with an infinite number of servers, or where the number of servers grows "fast enough" to be effectively infinite in the limit are also covered by our model and results. Examples of such results in the literature are Iglehart [6] and Whitt [23]. Although all of the examples that we consider in this paper correspond to systems with a large number of servers, this is not the only context in which our results are applicable. In particular it should be noted that our results can be applied to some closed queueing networks with a large number of customers. For discussion of such networks, we refer the reader to the references on finite population models found in the bibliography of Mandelbaum and Pats [14].

There has been a great deal of work on state dependent queues, time dependent queues, and related asymptotics. We make no attempt to survey this literature, focusing instead on four pieces of work related closely enough to ours to merit specific mention: [9], [13], [14], and [19]. The reader interested in more references on state dependent queues should consult [14] or [13]. References on time dependent queues are contained in Mandelbaum and Massey [11], and Massey and Whitt [17]. Motivated by population, epidemic, and chemical reaction models Kurtz [9] proves a FSLLN and a FCLT for systems with 'smooth' parameters. Our motivation is queueing systems that do not satisfy the smoothness required in [9]. We also generalize [9] in the sense that we allow time dependent rates, but this is mostly a notational issue. In Mandelbaum and Pats [14] limit theorems are proved for Markovian networks with state dependent rates. Systems whose limits may hit a boundary of the state space are allowed in [14], so that the issue of reflection must be dealt with. The limit processes that we obtain do not have the singular local time terms typically associated with reflection. Intuitively, this is because our limit processes do not hit any boundaries. The issue of piecewise continuous derivatives is treated in [13] (Theorem 4.3) for the one-dimensional case and is suggested as a subject for future research in [14]. Newell [19] considers approximations for the $G_t/G/n$ queue with large n. The approximations in [19] are of fluid and diffusion type, and are motivated by the strong law of large numbers and the central limit theorem, but no limit theorems are stated or proved in that work.

The rest of this paper is organized as follows. The model and main results are presented in Section 2. The some properties of scalable Lipschitz derivatives are described in Section 3. Some examples of Markovian service networks covered by our theorems are presented in Sections 4–8. In Section 4 we consider Jackson networks with many servers at each node. The system with abandonment and retrial discussed above is treated in Section 5. This is a special case of a Jackson network with abandonment, which is treated in Section 6. Section 7 deals with the priority system described above, and Section 8 covers Jackson networks with state dependent routing. The proofs of our main results are contained in Sections 9 and 10. The strong approximation theorem and the FSLLN are proved in Section 9. The FCLT is proved in Section 10. There are two appendices. The first appendix (Section 11) contains results on ordinary differential equations that we need, including a version of Gronwall's inequality and a uniqueness result for our limit processes. The second appendix (Section 12) contains proofs of some of the basic properties of Lipschitz derivatives.

2. The Model and Main Results

The primitives for our model are $\{\, A_i(\cdot) \mid i \in I \,\}$, a collection of mutually independent, standard (rate 1) Poisson processes, indexed by a set I which is at most countably infinite; a separable Banach space $\mathbb{V}$ with norm $|\cdot|$; a collection of "jump" vectors $\{\, \mathbf{v}_i \in \mathbb{V} \mid i \in I \,\}$ such that

$$\sum_{i \in I} |\mathbf{v}_i| < \infty; \tag{2.1}$$

a random initial state vector $\mathbf{Q}(0)$ in $\mathbb{V}$ that is always assumed to be independent of the collection of Poisson processes $\{\, A_i(\cdot) \mid i \in I \,\}$; and a collection of real-valued, non-negative Lipschitz "rate" functions on $\mathbb{V}$,

$$\{\, \alpha_t(\cdot, i) \mid t \geq 0,\, i \in I \,\}, \tag{2.2}$$

that jointly satisfy

$$\|\alpha_t(\cdot, i)\| \leq \beta_t \gamma^{(i)} \tag{2.3}$$

for some β_t, a locally integrable function, and $\left\{\, \gamma^{(i)} \,\middle|\, i \in I \,\right\}$, a summable sequence of real numbers; here $\|\cdot\|$ is the *Lipschitz norm* for real-valued functions on $\mathbb{V}$, namely

$$\|f\| \equiv \sup_{\mathbf{x},\mathbf{y} \in \mathbb{V},\, \mathbf{x} \neq \mathbf{y}} \frac{|f(\mathbf{x}) - f(\mathbf{y})|}{|\mathbf{x} - \mathbf{y}|} \vee |f(0)|. \tag{2.4}$$

It follows that for all $\mathbf{x}$ and $\mathbf{y}$ in $\mathbb{V}$, we have

$$|f(\mathbf{x}) - f(\mathbf{y})| \leq \|f\| \cdot |\mathbf{x} - \mathbf{y}| \tag{2.5}$$

and so f is a *Lipschitz function* whenever $\|f\| < \infty$. Moreover, for all $\mathbf{x} \in \mathbb{V}$,

$$|f(\mathbf{x})| \leq \|f\| \cdot (1 + |\mathbf{x}|). \tag{2.6}$$

For all of the examples that we consider in this paper, $\mathbb{V} = \mathbb{R}^N$ for some $1 \leq N < \infty$ and the number of elements in I is finite. Thus, although we prove the main results of the paper for a more general setting, any reader uncomfortable with the trappings of Banach spaces can replace $\mathbb{V}$ with $\mathbb{R}^N$ and still follow the examples we present. In that case $|\cdot|$ is the standard Euclidean norm on $\mathbb{R}^N$.

In terms of the primitives, we represent our *Markovian service network* to be the $\mathbb{V}$-valued stochastic process $\mathbf{Q} \equiv \{\, \mathbf{Q}(t) \mid t \geq 0 \,\}$, whose sample paths are uniquely determined by $\mathbf{Q}(0)$ and the functional equations

$$\mathbf{Q}(t) = \mathbf{Q}(0) + \sum_{i \in I} A_i \left(\int_0^t \alpha_s\Big(\mathbf{Q}(s), i\Big) ds \right) \mathbf{v}_i \tag{2.7}$$

for all $t \geq 0$ (for the $M_t/M_t/n_t$ example, $\mathbb{V} = \mathbb{R}$, $I = \{1, 2\}$, $\mathbf{v}_1 = 1$, $\mathbf{v}_2 = -1$, $\alpha_t(x, 1) = \lambda_t$, $\alpha_t(x, 2) = \mu_t \cdot (x \wedge n_t)$). Uniqueness of the solution to (2.7) is shown in Theorem 9.2. The special uniform acceleration that is used for the rate functions of the $M_t/M_t/n_t$ queue in (1.3) now generalizes to an asymptotic analysis of the processes $\{\, \mathbf{Q}^\eta \mid \eta > 0 \,\}$ as $\eta \to \infty$ where

$$\mathbf{Q}^\eta(t) = \mathbf{Q}^\eta(0) + \sum_{i \in I} A_i \left(\eta \int_0^t \alpha_s\Big(\frac{1}{\eta}\mathbf{Q}^\eta(s), i\Big) ds \right) \mathbf{v}_i. \tag{2.8}$$

Our goal is to characterize this asymptotic behavior as $\eta \uparrow \infty$ with a functional strong law of large numbers and a central limit theorem, but we do this with a more general type of asymptotic behavior for the rate functions.

The asymptotic analysis that we describe above was carried out by Kurtz [9] for the special case of rate functions having no explicit time dependence and state dependence that is continuously differentiable. In this paper, we extend his analysis to the following general class of processes:

$$\mathbf{Q}^\eta(t) = \mathbf{Q}^\eta(0) + \sum_{i \in I} A_i \left(\int_0^t \alpha_s^\eta\Big(\frac{1}{\eta}\mathbf{Q}^\eta(s), i\Big) ds \right) \mathbf{v}_i, \tag{2.9}$$

where

$$\|\alpha_t^\eta(\cdot, i)\| \leq \eta \beta_t \gamma^{(i)}. \tag{2.10}$$

In our extension, we allow the following

1. The rate functions $\alpha_t^\eta(\cdot, i)$ are functions of time as well as state.

2. The rate functions, which are indexed by the parameter η, are such that for each $i \in I$, $\alpha_t^\eta(\cdot, i)$ has the following asymptotic expansion as $\eta \to \infty$:

$$\alpha_t^\eta(\cdot, i) = \eta\alpha_t^{(0)}(\cdot, i) + \sqrt{\eta}\alpha_t^{(1)}(\cdot, i) + o(\sqrt{\eta}). \tag{2.11}$$

3. The rate functions, as a function of the state space $\mathbb{V}$, have a more general type of differentiability that include functions on the real line that are everywhere left and right differentiable.

The first condition is a minor extension of Kurtz but the latter two conditions are significant new extensions. The last condition is the most significant in that a new nonsmooth differential calculus must be developed to deal with these continuous but piecewise differentiable rate functions. These new conditions allow us to apply the limit theorems to a wider class of Markov processes that arise in the study of queueing networks with large numbers of servers.

Within the framework of strong approximations, we first approximate the sample-path representation (2.9) of the family $\{ \mathbf{Q}^\eta \mid \eta > 0 \}$ by the following theorem, which is proved in Section 9.

Theorem 2.1 (Strong Approximation). If (2.1) and (2.10) hold, then as $\eta \to \infty$ we have

$$\begin{aligned}\mathbf{Q}^\eta(t) = \mathbf{Q}^\eta(0) &+ \int_0^t \boldsymbol{\alpha}_s^\eta\Big(\frac{1}{\eta}\mathbf{Q}^\eta(s)\Big) ds \\ &+ \sum_{i\in I} B_i\left(\int_0^t \alpha_s^\eta\Big(\frac{1}{\eta}\mathbf{Q}^\eta(s), i\Big) ds\right) \mathbf{v}_i + O(\log\eta) \quad \text{a.s.}\end{aligned} \tag{2.12}$$

where the convergence is uniform on compact sets in t.

From this strong approximation, we deduce a FSLLN (Theorem 2.2), followed by a FCLT (Theorem 2.3). The limit theorems enable sample-path (2.17) and distributional approximations (2.30), which support computations and confidence intervals. The proof of the following functional strong law of large numbers is presented in Section 9.

Theorem 2.2 (FSLLN). Assume that (2.1) and (2.10) hold. Moreover, assume that

$$\lim_{\eta\to\infty} \sum_{i\in I} \int_0^t \left\| \frac{\alpha_s^\eta(\cdot, i)}{\eta} - \alpha_s^{(0)}(\cdot, i) \right\| ds = 0, \tag{2.13}$$

for all $t \geq 0$. If $\{\, \mathbf{Q}^{\eta}(0) \mid \eta > 0 \,\}$ is any family of random initial state vectors in $\mathbb{V}$, then

$$\lim_{\eta\to\infty} \frac{\mathbf{Q}^{\eta}(0)}{\eta} = \mathbf{Q}^{(0)}(0) \ \text{ a.s.} \quad \text{implies} \quad \lim_{\eta\to\infty} \frac{\mathbf{Q}^{\eta}(t)}{\eta} = \mathbf{Q}^{(0)}(t) \ \text{ a.s.} \tag{2.14}$$

where the convergence is uniform on compact sets in t, and $\mathbf{Q}^{(0)}$ is the unique deterministic process $\left\{\, \mathbf{Q}^{(0)}(t) \,\middle|\, t \geq 0 \,\right\}$ that solves the integral equation

$$\mathbf{Q}^{(0)}(t) = \mathbf{Q}^{(0)}(0) + \int_0^t \boldsymbol{\alpha}_s^{(0)}\Big(\mathbf{Q}^{(0)}(s)\Big)ds, \quad t \geq 0. \tag{2.15}$$

Here $\boldsymbol{\alpha}_t^{(0)}$, given by

$$\boldsymbol{\alpha}_t^{(0)}(\mathbf{x}) = \sum_{i\in I} \alpha_t^{(0)}(\mathbf{x}, i)\mathbf{v}_i, \quad \mathbf{x} \in \mathbb{V}, \tag{2.16}$$

is a Lipschitz mapping of $\mathbb{V}$ into itself and its Lipschitz norm $\left\|\boldsymbol{\alpha}_t^{(0)}\right\|$ is a locally integrable function of t.

We call $\mathbf{Q}^{(0)}$ the *fluid approximation* associated with the family $\{\, \mathbf{Q}^{\eta}(t) \mid t \geq 0 \,\}$. It gives rise to first-order "macroscopic" fluid approximations of the form

$$\mathbf{Q}^{\eta}(t, \omega) = \eta \mathbf{Q}^{(0)}(t) + o(\eta), \quad a.s., \quad t \geq 0. \tag{2.17}$$

In the development of a functional central limit theorem for our stochastic network, which refines the above fluid approximation, it is necessary to differentiate $\boldsymbol{\alpha}_t^{(0)}(\cdot)$ over the Banach space $\mathbb{V}$. There are specific examples of queueing systems that we analyze, like the $M_t/M_t/n_t$ queue, where the corresponding $\boldsymbol{\alpha}_t^{(0)}$ is piecewise differentiable but not everywhere differentiable. This poses a problem that is not easily ignored since these derivatives are evaluated at values for the fluid model $\mathbf{Q}^{(0)}(t)$. So even if $\boldsymbol{\alpha}_t^{(0)}(\cdot)$ has no derivative at only a finite number of points, the fluid process could spend all of its time at these points.

We resolve this issue by introducing a new type of differentiability. If $\mathbf{f}(\cdot)$ is a mapping from $\mathbb{V}_1$ into $\mathbb{V}_2$, we extend the Banach space norms $|\cdot|_1$ and $|\cdot|_2$ on $\mathbb{V}_1$ and $\mathbb{V}_2$ respectively, to define the following norm on $\mathbf{f}$:

$$\|\mathbf{f}\| \equiv \sup_{\mathbf{x},\mathbf{y}\in\mathbb{V}_1,\, \mathbf{x}\neq\mathbf{y}} \frac{|\mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{y})|_2}{|\mathbf{x} - \mathbf{y}|_1} \vee |\mathbf{f}(0)|_2, \tag{2.18}$$

and say that $\mathbf{f}$ is Lipschitz on $\mathbb{V}_1$ whenever $\|\mathbf{f}\| < \infty$. If $\mathcal{O}$ is an open subset of $\mathbb{V}_1$ and $\mathbf{x} \in \mathcal{O}$, we say that $\mathbf{f}$ is *locally Lipschitz* at x if

$$\|\mathbf{f}\|_{\mathcal{O}} \equiv \sup_{\mathbf{y},\mathbf{z}\in\mathcal{O},\ \mathbf{y}\neq\mathbf{z}} \frac{|\mathbf{f}(\mathbf{y}) - \mathbf{f}(\mathbf{z})|_2}{|\mathbf{y}-\mathbf{z}|_1} \vee |\mathbf{f}(0)|_2 < \infty. \tag{2.19}$$

Now we define $\mathbf{f}$ to have a *scalable Lipschitz derivative* at $\mathbf{x} \in \mathbb{V}_1$ if there exists another mapping from $\mathbb{V}_1$ into $\mathbb{V}_2$, denoted $\Lambda\mathbf{f}(\mathbf{x};\cdot)$, such that

$$\lim_{\mathbf{y}\to 0} \frac{|\mathbf{f}(\mathbf{x}+\mathbf{y}) - \mathbf{f}(\mathbf{x}) - \Lambda\mathbf{f}(\mathbf{x};\mathbf{y})|_2}{|\mathbf{y}|_1} = 0, \tag{2.20}$$

where the function $\Lambda\mathbf{f}(\mathbf{x};\cdot)$ is Lipschitz on $\mathbb{V}_1$ so that

$$\|\Lambda\mathbf{f}(\mathbf{x};\cdot)\| < \infty, \tag{2.21}$$

and for all real scalars with $\lambda \geq 0$,

$$\lambda\Lambda\mathbf{f}(\mathbf{x};\mathbf{y}) = \Lambda\mathbf{f}(\mathbf{x};\lambda\mathbf{y}). \tag{2.22}$$

Since all bounded linear mappings between $\mathbb{V}_1$ and $\mathbb{V}_2$ possess these last two properties, we see that differentiability is a special case of scalable Lipschitz differentiability. Non-smooth differentiation in the context of generalizing directional derivatives has been defined before, see Clarke [1] and Rockafellar [22] for details. Our definition (2.20) can be viewed as the analogue to the multivariate definition of differentiability or constructing the Jacobian.

We sometimes write this new derivative as

$$\Lambda\mathbf{f}_{\mathbf{x}}(\mathbf{y}) \equiv \Lambda\mathbf{f}(\mathbf{x};\mathbf{y}) \tag{2.23}$$

to emphasize that we should fix $\mathbf{x}$ and view the derivative as a function of $\mathbf{y}$.

We can now state the functional central limit theorem, whose proof is postponed to Section 10.

Theorem 2.3 (FCLT). Assume that (2.1) and (2.10) hold. Moreover assume that

$$\sum_{i\in I} \overline{\lim_{\eta\to\infty}} \int_0^t \left\| \sqrt{\eta}\left[\frac{\alpha_s^\eta(\cdot,i)}{\eta} - \alpha_s^{(0)}(\cdot,i)\right]\right\| ds < \infty \tag{2.24}$$

and

$$\lim_{\eta\to\infty} \sum_{i\in I} \int_0^t \left\| \sqrt{\eta}\left[\frac{\alpha_s^\eta(\cdot,i)}{\eta} - \alpha_s^{(0)}(\cdot,i)\right] - \alpha_s^{(1)}(\cdot,i)\right\| ds = 0. \tag{2.25}$$

It follows that $\boldsymbol{\alpha}_t^{(0)}$, given by (2.16), and $\boldsymbol{\alpha}_t^{(1)}$, given by

$$\boldsymbol{\alpha}_t^{(1)}(\mathbf{x}) = \sum_{i \in I} \alpha_t^{(1)}(\mathbf{x}, i)\mathbf{v}_i, \quad \mathbf{x} \in \mathbb{V}, \tag{2.26}$$

are both Lipschitz mappings of $\mathbb{V}$ into itself, and their Lipschitz norms are locally integrable functions of t,

Moreover, if we assume that $\boldsymbol{\alpha}_t^{(0)}(\cdot)$ has a scalable Lipschitz derivative $\Lambda\boldsymbol{\alpha}_t^{(0)}\Big(\mathbf{Q}^{(0)}(t);\cdot\Big)$ and we have a family of random initial state vectors $\{\, \mathbf{Q}^\eta(0) \mid \eta > 0 \,\}$ in $\mathbb{V}$, then for all random vectors $\mathbf{Q}^{(0)}(0)$ and $\mathbf{Q}^{(1)}(0)$ in $\mathbb{V}$, it follows that

$$\lim_{\eta\to\infty} \sqrt{\eta}\left[\frac{\mathbf{Q}^\eta(0)}{\eta} - \mathbf{Q}^{(0)}(0)\right] \stackrel{d}{=} \mathbf{Q}^{(1)}(0) \tag{2.27}$$

implies

$$\lim_{\eta\to\infty} \sqrt{\eta}\left[\frac{\mathbf{Q}^\eta(t)}{\eta} - \mathbf{Q}^{(0)}(t)\right] \stackrel{d}{=} \mathbf{Q}^{(1)}(t), \tag{2.28}$$

the convergence being weak-convergence in $D_{\mathbb{V}}[0,\infty)$, the space of $\mathbb{V}$-valued functions that are right-continuous with left-limits, equipped with the Skorohod J_1 topology.

Finally, the limit $\mathbf{Q}^{(1)} \equiv \Big\{\, \mathbf{Q}^{(1)}(t) \,\Big|\, t \geq 0 \,\Big\}$ is the unique stochastic process that solves the stochastic integral equation

$$\begin{aligned} \mathbf{Q}^{(1)}(t) = \mathbf{Q}^{(1)}(0) &+ \int_0^t \Big[\Lambda\boldsymbol{\alpha}_s^{(0)}\Big(\mathbf{Q}^{(0)}(s);\mathbf{Q}^{(1)}(s)\Big) + \boldsymbol{\alpha}_s^{(1)}\Big(\mathbf{Q}^{(0)}(s)\Big)\Big]\, ds \\ &+ \sum_{i\in I} B_i\left(\int_0^t \alpha_s^{(0)}\Big(\mathbf{Q}^{(0)}(s), i\Big) ds\right)\mathbf{v}_i, \quad t \geq 0, \end{aligned} \tag{2.29}$$

where the $\{\, B_i \mid i \in I \,\}$ are a family of mutually independent, standard Brownian motions.

We call $\mathbf{Q}^{(1)}$ the *diffusion approximation* associated with the family $\{\, \mathbf{Q}^\eta(t) \mid t \geq 0 \,\}$. It quantifies deviations from the fluid approximations, and it gives rise to second-order "mesoscopic" diffusion approximations of the form

$$\mathbf{Q}^\eta(t) \stackrel{d}{=} \eta\mathbf{Q}^{(0)}(t) + \sqrt{\eta}\mathbf{Q}^{(1)}(t) + o(\sqrt{\eta}) \tag{2.30}$$

as $\eta \to \infty$ for all $t \geq 0$, with the approximation being in distribution.

Although we state (and prove) Theorems 2.1, 2.2, and 2.3 for the setting of (2.9), all but one of the examples are presented in the more restrictive context of (2.8). This is done mainly to reduce the notational burden. The full generality of (2.9) is employed for the system with abandonment and retrials in Section 5. It should be clear from these results how to extend the other examples to the setting of (2.9).

Now consider the case of $\mathbb{V}$ being either a finite dimensional vector space or a Banach space that can be embedded into its own dual space (like a Hilbert space), so that we can define the notion of a transpose, denoted by a superscript "T" (for $\mathbb{V} = \mathbb{R}^N$, this corresponds to the standard transpose of a matrix). One consequence of the diffusion limit is an associated set of differential equations that become useful in the computation of its mean and covariance matrix. The proof is given at the end of Section 10.

Theorem 2.4. If conditions (2.1), (2.10), (2.24), and (2.25) all hold, then the mean vector and covariance matrix for $\mathbf{Q}^{(1)}(t)$ solve the following set of differential equations:

$$\frac{d}{dt}\mathsf{E}\left[\mathbf{Q}^{(1)}(t)\right] = \mathsf{E}\left[\Lambda\boldsymbol{\alpha}_t^{(0)}\left(\mathbf{Q}^{(0)}(t);\mathbf{Q}^{(1)}(t)\right)\right] + \boldsymbol{\alpha}_t^{(1)}\left(\mathbf{Q}^{(0)}(t)\right). \tag{2.31}$$

and

$$\begin{aligned}\frac{d}{dt}\mathsf{Cov}\left[\mathbf{Q}^{(1)}(t),\mathbf{Q}^{(1)}(t)\right] = &\left\{\mathsf{Cov}\left[\mathbf{Q}^{(1)}(t),\Lambda\boldsymbol{\alpha}_t^{(0)}\left(\mathbf{Q}^{(0)}(t);\mathbf{Q}^{(1)}(t)\right)\right]\right\}\\ &+\sum_{i\in I}\alpha_t^{(0)}\left(\mathbf{Q}^{(0)}(t),i\right)\mathbf{v}_i^{\mathsf{T}}\cdot\mathbf{v}_i\end{aligned} \tag{2.32}$$

for almost all t, where

$$\mathsf{Cov}\left[\mathbf{Q}^{(1)}(t),\mathbf{Q}^{(1)}(t)\right] \equiv \mathsf{E}\left[\mathbf{Q}^{(1)}(t)^{\mathsf{T}}\cdot\mathbf{Q}^{(1)}(t)\right] - \mathsf{E}\left[\mathbf{Q}^{(1)}(t)\right]^{\mathsf{T}}\cdot\mathsf{E}\left[\mathbf{Q}^{(1)}(t)\right] \tag{2.33}$$

and for all operators $\mathbf{A}$ on $\mathbb{V}$,

$$\{\mathbf{A}\} \equiv \mathbf{A} + \mathbf{A}^{\mathsf{T}}. \tag{2.34}$$

Moreover, if $\Lambda\boldsymbol{\alpha}_t^{(0)}\left(\mathbf{Q}^{(0)}(t);\cdot\right)$ is a linear operator for almost all t, then $\mathsf{E}\left[\mathbf{Q}^{(1)}(t)\right]$ is the unique solution for (2.31) and $\mathsf{Cov}\left[\mathbf{Q}^{(1)}(t),\mathbf{Q}^{(1)}(t)\right]$ is the unique solution for (2.32). Finally, for all $s < t$, $\mathsf{Cov}\left[\mathbf{Q}^{(1)}(s),\mathbf{Q}^{(1)}(t)\right]$ solves the same set of differential equations in t as does $\mathsf{E}\left[\mathbf{Q}^{(1)}(t)\right]$, but with a different set of initial conditions.

3. Calculus for Scalable Lipschitz Derivatives

Certain basic properties of the scalable Lipschitz derivative are useful in doing calculations for the diffusion limits of our service network processes. All of the theorems in this section are proved in Section 12. The first theorem states general properties for these functions.

Theorem 3.1. Scalable Lipschitz differentiability has the following properties:

1. If the function $\mathbf{f} : \mathbb{V}_1 \to \mathbb{V}_2$ is scalable Lipschitz differentiable at $\mathbf{x}$, then the resulting Lipschitz derivative function $\Lambda f_{\mathbf{x}} : \mathbb{V}_1 \to \mathbb{V}_2$ is unique.
2. If $\mathbf{f} : \mathbb{V}_1 \to \mathbb{V}_2$ and $\mathbf{g} : \mathbb{V}_2 \to \mathbb{V}_3$ are both scalable Lipschitz differentiable at $\mathbf{f}(\mathbf{x})$, then $\mathbf{g} \circ \mathbf{f} : \mathbb{V}_1 \to \mathbb{V}_3$ is scalable Lipschitz differentiable at $\mathbf{x}$, with
$$\Lambda(\mathbf{g} \circ \mathbf{f})_{\mathbf{x}}(\mathbf{y}) = \left(\Lambda \mathbf{g}_{\mathbf{f}(\mathbf{x})} \circ \Lambda \mathbf{f}_{\mathbf{x}}\right)(\mathbf{y}). \tag{3.1}$$
3. If $\mathbf{f} : \mathbb{V}_1 \to \mathbb{V}_2$ is locally Lipschitz, as defined in Section 2, in an open neighborhood $\mathcal{O} \subset \mathbb{V}_\infty$ of $\mathbf{x} \in \mathbb{V}_1$ and has a scalable Lipschitz derivative at $\mathbf{x}$, then
$$\|\Lambda \mathbf{f}_{\mathbf{x}}(\cdot)\| \leq \|\mathbf{f}\|_{\mathcal{O}}. \tag{3.2}$$

The next theorem is useful in the identification of scalable Lipschitz differentiable functions that act on finite dimensional vector spaces.

Theorem 3.2. The following results hold:

1. If $\mathbf{f} : \mathbb{R}^n \to \mathbb{R}^m$ is differentiable at $\mathbf{x} \in \mathbb{R}^n$ with Jacobian matrix $D\mathbf{f}(\mathbf{x})$, then it is scalable Lipschitz differentiable there and its scalable Lipschitz derivative is matrix multiplication by the Jacobian matrix so that
$$\Lambda \mathbf{f}_{\mathbf{x}}(\mathbf{y}) = \mathbf{y} \cdot D\mathbf{f}(\mathbf{x}) \tag{3.3}$$
for all $\mathbf{y} \in \mathbb{R}^m$, viewing $\mathbf{y}$ as a row vector.
2. If $\mathbf{f} : \mathbb{R}^n \to \mathbb{R}^m$ is locally Lipschitz at $\mathbf{x} \in \mathbb{R}^n$ and has all its radial derivatives at $\mathbf{x}$, then $\mathbf{f}$ has a scalable Lipschitz derivative at $\mathbf{x}$.

One simple consequence of the second statement of this theorem is that if $f : \mathbb{R} \to \mathbb{R}$ has left and right derivatives everywhere, then it is everywhere scalable Lipschitz differentiable and
$$\Lambda f_x(y) = f'(x+)y^+ - f'(x-)y^-, \tag{3.4}$$

for all real x and y.

For all $\mathbf{x}$ and $\mathbf{y}$ in $\mathbb{R}^m$, let $\mathbf{x} \wedge \mathbf{y}$ be the $\mathbb{R}^m$-vector whose i-th component equals $x_i \wedge y_i$ and define $\mathbf{x} \vee \mathbf{y}$ in a similar fashion. We can then define $\mathbf{x}^+ \equiv \mathbf{x} \vee \mathbf{0}$ and $\mathbf{x}^- \equiv (-\mathbf{x}) \vee \mathbf{0}$. Now let $\mathbf{f}, \mathbf{g} : \mathbb{V} \to \mathbb{R}^m$ and define $\mathbf{I}_{\{\mathbf{f}(\mathbf{x}) > \mathbf{g}(\mathbf{x})\}}$ to be the projection operator on $\mathbb{R}^m$ such that for any unit basis vector $\mathbf{e}_i$ for $i = 1, \ldots, m$ we have

$$\mathbf{e}_i \mathbf{I}_{\{\mathbf{f}(\mathbf{x}) > \mathbf{g}(\mathbf{x})\}} \equiv \begin{cases} \mathbf{e}_i \text{ if } f_i(\mathbf{x}) > g_i(\mathbf{x}), \\ 0 \text{ if } f_i(\mathbf{x}) \leq g_i(\mathbf{x}). \end{cases} \tag{3.5}$$

where $\mathbf{f}(\mathbf{x}) = (f_1(\mathbf{x}), \ldots, f_m(\mathbf{x}))$ and $\mathbf{g}(\mathbf{x}) = (g_1(\mathbf{x}), \ldots, g_m(\mathbf{x}))$. The projection operators $\mathbf{I}_{\{\mathbf{f}(\mathbf{x}) < \mathbf{g}(\mathbf{x})\}}$ and $\mathbf{I}_{\{\mathbf{f}(\mathbf{x}) = \mathbf{g}(\mathbf{x})\}}$ are defined similarly. We use these operators in the following theorem which gives us a non-smooth calculus for computing these scalable Lipschitz derivatives.

Theorem 3.3. The following operations preserve scalable Lipschitz differentiability:

1. If $\mathbf{f} : \mathbb{V}_1 \to \mathbb{V}_2$ and $\mathbf{g} : \mathbb{V}_1 \to \mathbb{V}_2$ are both scalable Lipschitz differentiable at $\mathbf{x}$, then $\mathbf{f} + \mathbf{g}$ is scalable Lipschitz differentiable at $\mathbf{x}$, where
$$\Lambda(\mathbf{f} + \mathbf{g})_{\mathbf{x}}(\mathbf{y}) = \Lambda \mathbf{f}_{\mathbf{x}}(\mathbf{y}) + \Lambda \mathbf{g}_{\mathbf{x}}(\mathbf{y}). \tag{3.6}$$
2. If $f : \mathbb{V} \to \mathbb{R}$ and $g : \mathbb{V} \to \mathbb{R}$ are both scalable Lipschitz differentiable at $\mathbf{x}$, then fg is scalable Lipschitz differentiable at $\mathbf{x}$, where
$$\Lambda(fg)_{\mathbf{x}}(\mathbf{y}) = f(\mathbf{x}) \Lambda g_{\mathbf{x}}(\mathbf{y}) + g(\mathbf{x}) \Lambda f_{\mathbf{x}}(\mathbf{y}). \tag{3.7}$$
3. If $\mathbf{f} : \mathbb{V} \to \mathbb{R}^m$ and $\mathbf{g} : \mathbb{V} \to \mathbb{R}^m$ are both scalable Lipschitz differentiable at $\mathbf{x}$, then $\mathbf{f} \vee \mathbf{g}$ and $\mathbf{f} \wedge \mathbf{g}$ are both scalable Lipschitz differentiable at $\mathbf{x}$, where
$$\begin{aligned} \Lambda(\mathbf{f} \vee \mathbf{g})_{\mathbf{x}}(\mathbf{y}) = {} & \Lambda \mathbf{f}_{\mathbf{x}}(\mathbf{y}) \mathbf{I}_{\{\mathbf{f}(\mathbf{x}) > \mathbf{g}(\mathbf{x})\}} + \Lambda \mathbf{g}_{\mathbf{x}}(\mathbf{y}) \mathbf{I}_{\{\mathbf{f}(\mathbf{x}) < \mathbf{g}(\mathbf{x})\}} \\ & + \Big(\Lambda \mathbf{f}_{\mathbf{x}}(\mathbf{y}) \vee \Lambda \mathbf{g}_{\mathbf{x}}(\mathbf{y}) \Big) \mathbf{I}_{\{\mathbf{f}(\mathbf{x}) = \mathbf{g}(\mathbf{x})\}} \end{aligned} \tag{3.8}$$
and
$$\begin{aligned} \Lambda(\mathbf{f} \wedge \mathbf{g})_{\mathbf{x}}(\mathbf{y}) = {} & \Lambda \mathbf{f}_{\mathbf{x}}(\mathbf{y}) \mathbf{I}_{\{\mathbf{f}(\mathbf{x}) < \mathbf{g}(\mathbf{x})\}} + \Lambda \mathbf{g}_{\mathbf{x}}(\mathbf{y}) \mathbf{I}_{\{\mathbf{f}(\mathbf{x}) > \mathbf{g}(\mathbf{x})\}} \\ & + \Big(\Lambda \mathbf{f}_{\mathbf{x}}(\mathbf{y}) \wedge \Lambda \mathbf{g}_{\mathbf{x}}(\mathbf{y}) \Big) \mathbf{I}_{\{\mathbf{f}(\mathbf{x}) = \mathbf{g}(\mathbf{x})\}}. \end{aligned} \tag{3.9}$$

Note that if $f : \mathbb{R}^n \to \mathbb{R}$, then we have $|f(\mathbf{x})| = f(\mathbf{x}) \vee \mathbf{0} + (-\mathbf{f}(\mathbf{x})) \vee \mathbf{0}$, and so

$$\Lambda |f|_{\mathbf{x}}(\mathbf{y}) = \Lambda f_{\mathbf{x}}(\mathbf{y}) 1_{\{f(\mathbf{x}) > 0\}} + |\Lambda f_{\mathbf{x}}(\mathbf{y})| 1_{\{f(\mathbf{x}) = 0\}} - \Lambda f_{\mathbf{x}}(\mathbf{y}) 1_{\{f(\mathbf{x}) < 0\}}. \tag{3.10}$$

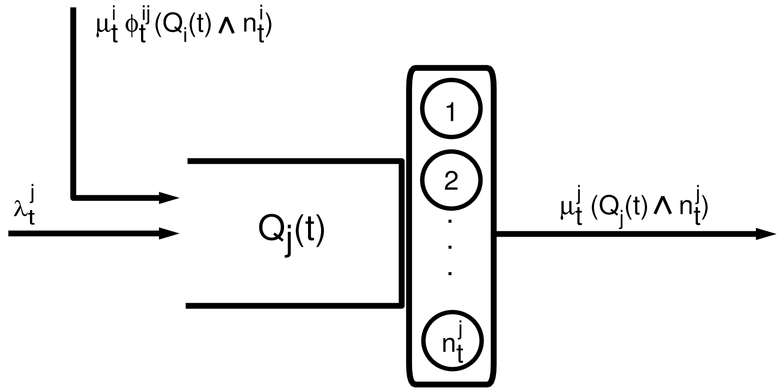

Figure 4.1. The Jackson network.

4. Classical Jackson Networks

We now consider the classical Jackson network but with the additional features of time varying rates and number of servers (see Figure 4.1). We extend Kendall notation and call it the $(M_t/M_t/n_t)^N$ network, where N denotes the number of nodes. We construct the $(M_t/M_t/n_t)^N$ network by first defining the following set of parameters:

$\lambda_t^i =$ external arrival rate to node i at time t,

$\mu_t^i =$ service rate for node i at time t,

$\phi_t^{ij} =$ service routing probability to node i from node j at time t,

$\phi_t^i =$ service departure probability from node i at time t,

$n_t^i =$ number of servers for node i at time t.

All these rate functions are assumed to be locally integrable functions of t and we require that

$$\phi_t^i + \sum_{j=1}^{N} \phi_t^{ij} = 1 \tag{4.1}$$

for all $t \geq 0$ and $i = 1, \ldots, N$.

We then set $\mathbb{V} = \mathbb{R}^N$ and define

$$\begin{aligned}\mathbf{Q}(t) = \mathbf{Q}(0) + \sum_{i=1}^{N} \Bigg[\sum_{j=1}^{N} A_{ij}^c \left(\int_0^t \Big(Q_i(s) \wedge n_s^i \Big) \mu_s^i \phi_s^{ij} ds \right) (\mathbf{e}_j - \mathbf{e}_i) \qquad (4.2)\\ + \left(A_i^a \left(\int_0^t \lambda_s^i ds \right) - A_i^b \left(\int_0^t \Big(Q_i(s) \wedge n_s^i \Big) \mu_s^i \phi_s^i ds \right) \right) \mathbf{e}_i \Bigg],\end{aligned}$$

where A_i^a, A_i^b, and A_{ij}^c for $i, j = 1, \ldots, N$ are mutually independent standard Poisson processes. For all $\mathbf{x} \in \mathbb{R}^N$, we define $\boldsymbol{\Delta}(\mathbf{x})$ to be the diagonal $N \times N$ matrix where the i-th diagonal entry is the i-th component of the vector $\mathbf{x}$. We define $\boldsymbol{\lambda}_t$, $\boldsymbol{\mu}_t$, and $\mathbf{n}_t$ to be row vectors where their i-th component equals λ_t^i, μ_t^i, and n_t^i respectively. We also define $\boldsymbol{\Phi}_t$ to be the $N \times N$ matrix whose (i, j) entry is ϕ_t^{ij}.

Theorem 4.1. Defining $\mathbf{Q}^\eta$ by uniform acceleration as in (2.8), the fluid limit for the $(M_t/M_t/n_t)^N$ network is the solution to the integral equation

$$\mathbf{Q}^{(0)}(t) = \mathbf{Q}^{(0)}(0) + \int_0^t \left[\boldsymbol{\lambda}_s + \left(\mathbf{Q}^{(0)}(s) \wedge \mathbf{n}_s\right)\boldsymbol{\Delta}(\boldsymbol{\mu}_s)(\Phi_s - \mathbf{I})\right] ds. \tag{4.3}$$

Moreover, the diffusion limit for the $(M_t/M_t/n_t)^N$ network is the unique solution to the integral equation

$$\begin{aligned}
\mathbf{Q}^{(1)}(t) = \mathbf{Q}^{(1)}(0)& \\
&+ \int_0^t \left(\mathbf{Q}^{(1)}(s)^+ \mathbf{I}_{\{\mathbf{Q}^{(0)}(s) < \mathbf{n}_s\}} - \mathbf{Q}^{(1)}(s)^- \mathbf{I}_{\{\mathbf{Q}^{(0)}(s) \leq \mathbf{n}_s\}}\right)\boldsymbol{\Delta}(\boldsymbol{\mu}_s)(\Phi_s - \mathbf{I}) ds \\
&+ \sum_{i=1}^N \left[\sum_{j=1}^N B_{ij}^c \left(\int_0^t \left(Q_i^{(0)}(s) \wedge n_s^i\right)\mu_s^i \phi_s^{ij} ds\right)(\mathbf{e}_j - \mathbf{e}_i)\right. \\
&\qquad \left. + \left(B_i^a \left(\int_0^t \lambda_s^i ds\right) - B_i^b \left(\int_0^t \left(Q_i^{(0)}(s) \wedge n_s^i\right)\mu_s^i \phi_s^i ds\right)\right)\mathbf{e}_i\right],
\end{aligned}$$

where B_i^a, B_i^b, and B_{ij}^c for $i, j = 1, \ldots, N$ are mutually independent standard Brownian motions.

Proof. From (4.2), it follows that

$$\begin{aligned}
\boldsymbol{\alpha}_t(\mathbf{x}) &= \sum_{i=1}^N \left[\lambda_t^i \mathbf{e}_i - (x_i \wedge n_t^i)\mu_t^i \phi_t^i \mathbf{e}_i + \sum_{j=1}^N (x_j \wedge n_t^j)\mu_t^j \phi_t^{ji}(\mathbf{e}_i - \mathbf{e}_j)\right] \\
&= \sum_{i=1}^N \left[\lambda_t^i + \sum_{j=1}^N (x_j \wedge n_t^j)\mu_t^j \phi_t^{ji} - (x_i \wedge n_t^i)\mu_t^i\right] \mathbf{e}_i \\
&= \boldsymbol{\lambda}_t + (\mathbf{x} \wedge \mathbf{n}_t)\boldsymbol{\Delta}(\boldsymbol{\mu}_t)\Phi_t - (\mathbf{x} \wedge \mathbf{n}_t)\boldsymbol{\Delta}(\boldsymbol{\mu}_t) \\
&= \boldsymbol{\lambda}_t + (\mathbf{x} \wedge \mathbf{n}_t)\boldsymbol{\Delta}(\boldsymbol{\mu}_t)(\Phi_t - \mathbf{I}).
\end{aligned} \tag{4.4}$$

The fluid limit now follows from applying Theorem 2.2.

If $\mathbf{f}(\mathbf{x}) \equiv \mathbf{x}$ and $\mathbf{g}(\mathbf{x}) = \mathbf{n}_t$ for all $\mathbf{x} \in \mathbb{R}^m$, then by Theorem 3.2

$$\Lambda\mathbf{f}(\mathbf{x};\mathbf{y}) = \mathbf{y} \quad \text{and} \quad \Lambda\mathbf{g}(\mathbf{x};\mathbf{y}) = \mathbf{0}. \tag{4.5}$$

Applying Theorem 3.3 to $\mathbf{f} \wedge \mathbf{g}(\mathbf{x}) = \mathbf{x} \wedge \mathbf{n}_t$ gives

$$\begin{aligned}\Lambda\Big(\mathbf{f} \wedge \mathbf{g}\Big)(\mathbf{x};\mathbf{y}) &= \mathbf{y}\,\mathbf{I}_{\{\mathbf{x}<\mathbf{n}_t\}} + (\mathbf{y} \wedge \mathbf{0})\,\mathbf{I}_{\{\mathbf{x}=\mathbf{n}_t\}} \\ &= (\mathbf{y}^+ - \mathbf{y}^-)\,\mathbf{I}_{\{\mathbf{x}<\mathbf{n}_t\}} - \mathbf{y}^-\,\mathbf{I}_{\{\mathbf{x}=\mathbf{n}_t\}} \\ &= \mathbf{y}^+\,\mathbf{I}_{\{\mathbf{x}<\mathbf{n}_t\}} - \mathbf{y}^-\,\mathbf{I}_{\{\mathbf{x}\leq\mathbf{n}_t\}}.\end{aligned} \tag{4.6}$$

Using (4.4) and (4.6), the scalable Lipschitz derivative of $\boldsymbol{\alpha}_t$ is

$$\Lambda\boldsymbol{\alpha}_t(\mathbf{x};\mathbf{y}) = \Big(\mathbf{y}^+\mathbf{I}_{\{\mathbf{x}<\mathbf{n}_t\}} - \mathbf{y}^-\mathbf{I}_{\{\mathbf{x}\leq\mathbf{n}_t\}}\Big)\boldsymbol{\Delta}(\boldsymbol{\mu}_t)(\Phi_t - \mathbf{I}) \tag{4.7}$$

and the diffusion limit follows from applying Theorem 2.3. □

The following result, which follows immediately from Theorem 2.4, provides ordinary differential equations for the mean vector and covariance matrix of $\mathbf{Q}^{(1)}$.

Theorem 4.2. The mean vector for the diffusion limit solves the differential equation

$$\begin{aligned}&\frac{d}{dt}\mathsf{E}\left[\mathbf{Q}^{(1)}(t)\right] \\ &\quad = \left(\mathsf{E}\left[\mathbf{Q}^{(1)}(t)^+\right]\mathbf{I}_{\{\mathbf{Q}^{(0)}(t)<\mathbf{n}_t\}} - \mathsf{E}\left[\mathbf{Q}^{(1)}(t)^-\right]\mathbf{I}_{\{\mathbf{Q}^{(0)}(t)\leq\mathbf{n}_t\}}\right)\boldsymbol{\Delta}(\boldsymbol{\mu}_t)(\Phi_t - \mathbf{I})\end{aligned}$$

and the covariance matrix for the diffusion limit solves the differential equation

$$\begin{aligned}&\frac{d}{dt}\mathsf{Cov}\left[\mathbf{Q}^{(1)}(t), \mathbf{Q}^{(1)}(t)\right] \\ &= \left\{\mathsf{Cov}\left[\mathbf{Q}^{(1)}(t), \mathbf{Q}^{(1)}(t)^+\mathbf{I}_{\{\mathbf{Q}^{(0)}(t)<\mathbf{n}_t\}} - \mathbf{Q}^{(1)}(t)^-\mathbf{I}_{\{\mathbf{Q}^{(0)}(t)\leq\mathbf{n}_t\}}\right]\boldsymbol{\Delta}(\boldsymbol{\mu}_t)(\Phi_t - \mathbf{I})\right\} \\ &\quad +\boldsymbol{\Delta}\left(\boldsymbol{\lambda}_t + \left(\mathbf{Q}^{(0)}(t) \wedge \mathbf{n}_t\right)\boldsymbol{\Delta}(\boldsymbol{\mu}_t)(\Phi_t + \mathbf{I})\right) - \left\{\boldsymbol{\Delta}\left(\mathbf{Q}^{(0)}(t) \wedge \mathbf{n}_t\right)\boldsymbol{\Delta}(\boldsymbol{\mu}_t)\Phi_t\right\}.\end{aligned}$$

Proof. Given (4.4) and (4.6), the proof is simply an application of Theorem 2.4. □

5. Queues with Abandonment and Retrials

We construct the multiserver queue with abandonment and retrials (see Figure 5.1) by first defining the following set of parameters:

$\lambda_t =$ external arrival rate to the service node at time t,

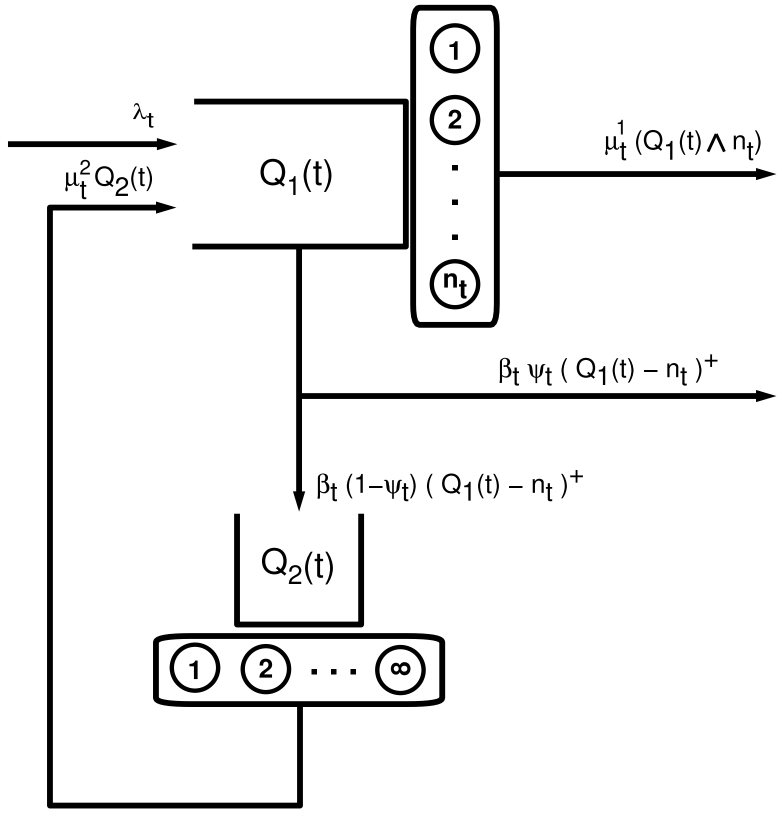

Figure 5.1. The abandonment queue with retrials.

$\beta_t =$ abandonment rate from the service node at time t,

$\mu_t^1 =$ service rate for the service node at time t,

$\mu_t^2 =$ service rate for the retry pool at time t,

$\psi_t =$ probability that a customer abandoning at time t does not retry,

$n_t =$ number of servers in service node at time t.

We then set $\mathbb{V} = \mathbb{R}^2$ and define $\mathbf{Q}(t) = (Q_1(t), Q_2(t))$, where

$$\begin{aligned}
Q_1(t) = Q_1(0) &+ A^a\left(\int_0^t \lambda_s ds\right) + A^c_{21}\left(\int_0^t Q_2(s)\mu_s^2 ds\right) \\
&- A^c\left(\int_0^t \Big(Q_1(s) \wedge n_s\Big)\mu_s^1 ds\right) - A^b\left(\int_0^t \Big(Q_1(s) - n_s\Big)^+ \beta_s\psi_s ds\right) \\
&- A^b_{12}\left(\int_0^t \Big(Q_1(s) - n_s\Big)^+ \beta_s(1-\psi_s) ds\right)
\end{aligned}$$

and

$$Q_2(t) = Q_2(0) + A_{12}^b\left(\int_0^t \Big(Q_1(s) - n_s\Big)^+ \beta_s(1-\psi_s)ds\right) - A_{21}^c\left(\int_0^t \Big(Q_2(s)\Big)\mu_s^2 ds\right).$$

Here we have a network with two nodes where the first one corresponds to the service node. The second node is the retrial pool and has an infinite number of servers to model retrial delay. Moreover, the act of abandoning the service queue due to impatience is modeled as abandonment routing where the customer enters the retrial pool with some probability or leaves the network entirely. Service routing instructs customers to leave the entire network after service completion at the first node and to enter the service queue after service completion at the retrial pool.

Theorem 5.1. Defining $\mathbf{Q}^\eta$ by uniform acceleration as in (2.8), the fluid limit for the multiserver queue with abandonment and retrials is the unique solution to the differential equations

$$\frac{d}{dt}Q_1^{(0)}(t) = \lambda_t + \mu_t^2 Q_2^{(0)}(t) - \mu_t^1\Big(Q_1^{(0)}(t)\wedge n_t\Big) - \beta_t\Big(Q_1^{(0)}(t) - n_t\Big)^+ \tag{5.1}$$

$$\frac{d}{dt}Q_2^{(0)}(t) = \beta_t(1-\psi_t)\Big(Q_1^{(0)}(t) - n_t\Big)^+ - \mu_t^2 Q_2^{(0)}(t). \tag{5.2}$$

Moreover, the diffusion limit for the multiserver queue with abandonment and retrials is the unique solution to the integral equations

$$\begin{aligned}
Q_1^{(1)}(t) = Q_1^{(1)}(0) + \int_0^t \Big[&\Big(\mu_s^1 1_{\{Q_1^{(0)}(s)\le n_s\}} + \beta_s 1_{\{Q_1^{(0)}(s)>n_s\}}\Big)Q_1^{(1)}(s)^- \\
&-\Big(\mu_s^1 1_{\{Q_1^{(0)}(s)<n_s\}} + \beta_s 1_{\{Q_1^{(0)}(s)\ge n_s\}}\Big)Q_1^{(1)}(s)^+ + \mu_s^2 Q_2^{(1)}(s)\Big]\,ds \\
&-B_{12}^b\left(\int_0^t \Big(Q_1^{(0)}(s) - n_s\Big)^+\beta_s(1-\psi_s)ds\right) - B_{21}^c\left(\int_0^t \Big(Q_2^{(0)}(s)\Big)\mu_s^2 ds\right) \\
&+B^a\left(\int_0^t \lambda_s ds\right) - B^b\left(\int_0^t \Big(Q_1^{(0)}(s) - n_s\Big)^+\beta_s\psi_s ds\right) \\
&-B^c\left(\int_0^t \Big(Q_1^{(0)}(s)\wedge n_s\Big)\mu_s^1 ds\right),
\end{aligned}$$

and

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

$$+\int_0^t \left[Q_1^{(1)}(s)^*\beta_s(1-\psi_s) - \mu_s^2 Q_2^{(1)}(s)\right] ds,$$

where

$$Q_1^{(1)}(t)^* = Q_1^{(1)}(t)^+ 1_{\{Q_1^{(0)}(t)\geq n_t\}} - Q_1^{(1)}(t)^- 1_{\{Q_1^{(0)}(t)>n_t\}}. \tag{5.3}$$

Proof. These results follow from Theorems 2.2 and 2.3. □

Theorem 5.2. The mean vector for the diffusion limit solves the set of differential equations

$$\begin{aligned}\frac{d}{dt}\mathsf{E}\left[Q_1^{(1)}(t)\right] &= \Big(\mu_t^1 1_{\{Q_1^{(0)}(t)\leq n_t\}} + \beta_t 1_{\{Q_1^{(0)}(t)>n_t\}}\Big)\mathsf{E}\left[Q_1^{(1)}(t)^-\right] \\ &\quad -\Big(\mu_t^1 1_{\{Q_1^{(0)}(t)<n_t\}} + \beta_t 1_{\{Q_1^{(0)}(t)\geq n_t\}}\Big)\mathsf{E}\left[Q_1^{(1)}(t)^+\right] + \mu_t^2\mathsf{E}\left[Q_2^{(1)}(t)\right]\end{aligned} \tag{5.4}$$

and

$$\begin{aligned}\frac{d}{dt}\mathsf{E}\left[Q_2^{(1)}(t)\right] &= \beta_t(1-\psi_t)\Big(\mathsf{E}\left[Q_1^{(1)}(t)^+\right] 1_{\{Q_1^{(0)}(t)\geq n_t\}} - \mathsf{E}\left[Q_1^{(1)}(t)^-\right] 1_{\{Q_1^{(0)}(t)>n_t\}}\Big) \\ &\quad -\mu_t^2\mathsf{E}\left[Q_2^{(1)}(t)\right].\end{aligned} \tag{5.5}$$

and the covariance matrix for the diffusion limit solves the differential equations

$$\begin{aligned}\frac{d}{dt}\mathsf{Var}\left[Q_1^{(1)}(t)\right] &= 2\left(\beta_t 1_{\{Q_1^{(0)}(t)>n_t\}} + \mu_t^1 1_{\{Q_1^{(0)}(t)\leq n_t\}}\right)\mathsf{Cov}\left[Q_1^{(1)}(t), Q_1^{(1)}(t)^-\right] \\ &\quad -2\left(\beta_t 1_{\{Q_1^{(0)}(t)\geq n_t\}} + \mu_t^1 1_{\{Q_1^{(0)}(t)<n_t\}}\right)\mathsf{Cov}\left[Q_1^{(1)}(t), Q_1^{(1)}(t)^+\right] \\ &\quad +\lambda_t + \beta_t\Big(Q_1^{(0)}(t)-n_t\Big)^+ + \mu_t^1\Big(Q_1^{(0)}(t)\wedge n_t\Big) + \mu_t^2 Q_2^{(0)}(t),\end{aligned}$$

$$\begin{aligned}&\frac{d}{dt}\mathsf{Var}\left[Q_2^{(1)}(t)\right] \\ &= -2\mu_t^2\mathsf{Var}\left[Q_2^{(1)}(t)\right] + \beta_t(1-\psi_t)\Big(Q_1^{(0)}(t)-n_t\Big)^+ + \mu_t^2 Q_2^{(0)}(t) + 2\beta_t(1-\psi_t) \\ &\quad\times\left(\mathsf{Cov}\left[Q_2^{(1)}(t), Q_1^{(1)}(t)^+\right] 1_{\{Q_1^{(0)}(t)\geq n_t\}} - \mathsf{Cov}\left[Q_2^{(1)}(t), Q_1^{(1)}(t)^-\right] 1_{\{Q_1^{(0)}(t)>n_t\}}\right),\end{aligned} \tag{5.6}$$

and

$$\begin{aligned}&\frac{d}{dt}\mathsf{Cov}\left[Q_1^{(1)}(t), Q_2^{(1)}(t)\right] \\ &= \left(\beta_t 1_{\{Q_1^{(0)}(t)>n_t\}} + \mu_t^1 1_{\{Q_1^{(0)}(t)\leq n_t\}}\right)\mathsf{Cov}\left[Q_2^{(1)}(t), Q_1^{(1)}(t)^-\right] \\ &\quad -\left(\beta_t 1_{\{Q_1^{(0)}(t)\geq n_t\}} + \mu_t^1 1_{\{Q_1^{(0)}(t)<n_t\}}\right)\mathsf{Cov}\left[Q_2^{(1)}(t), Q_1^{(1)}(t)^+\right]\end{aligned} \tag{5.7}$$

$$+\mu_t^2\left(\mathsf{Var}\left[Q_2^{(1)}(t)\right]-\mathsf{Cov}\left[Q_1^{(1)}(t),Q_2^{(1)}(t)\right]\right)$$
$$+\beta_t(1-\psi_t)\Big(Q_1^{(0)}(t)-n_t\Big)^+ + \mu_t^2 Q_2^{(0)}(t).$$

Proof. These results follow from Theorem 2.4. As we show in the next section, they are also a special case of Theorem 6.2 where the matrices that are specified by the given parameter rates are then

$$\mathbf{\Delta}(\boldsymbol{\beta}_t)=\begin{bmatrix}\beta_t & 0\\ 0 & 0\end{bmatrix},\ \mathbf{\Delta}(\boldsymbol{\mu}_t)=\begin{bmatrix}\mu_t^1 & 0\\ 0 & \mu_t^2\end{bmatrix},\ \mathbf{\Psi}_t=\begin{bmatrix}0 & 1-\psi_t\\ 0 & 0\end{bmatrix},\ \text{and}\ \ \mathbf{\Phi}_t=\begin{bmatrix}0 & 0\\ 1 & 0\end{bmatrix}. \tag{5.8}$$

Resulting products of these matrices are

$$\mathbf{\Delta}(\boldsymbol{\beta}_t)\Psi_t=\begin{bmatrix}0 & \beta_t(1-\psi_t)\\ 0 & 0\end{bmatrix}\ \text{ and }\ \mathbf{\Delta}(\boldsymbol{\mu}_t)\Phi_t=\begin{bmatrix}0 & 0\\ \mu_t^2 & 0\end{bmatrix}. \tag{5.9}$$

The special matrices that are functionals of the diffusion process are

$$\mathsf{Cov}\left[\mathbf{Q}^{(1)}(t),\mathbf{Q}^{(1)}(t)^*\right]=\begin{bmatrix}\mathsf{Cov}\left[Q_1^{(1)}(t),Q_1^{(1)}(t)^*\right] & 0\\ \mathsf{Cov}\left[Q_2^{(1)}(t),Q_1^{(1)}(t)^*\right] & 0\end{bmatrix} \tag{5.10}$$

and

$$\begin{aligned}&\mathsf{Cov}\left[\mathbf{Q}^{(1)}(t),\mathbf{Q}^{(1)}(t)-\mathbf{Q}^{(1)}(t)^*\right] \qquad (5.11)\\ &=\begin{bmatrix}\mathsf{Cov}\left[Q_1^{(1)}(t),Q_1^{(1)}(t)-Q_1^{(1)}(t)^*\right] & \mathsf{Cov}\left[Q_1^{(1)}(t),Q_2^{(1)}(t)\right]\\ \mathsf{Cov}\left[Q_2^{(1)}(t),Q_1^{(1)}(t)-Q_1^{(1)}(t)^*\right] & \mathsf{Var}\left[Q_2^{(1)}(t)\right]\end{bmatrix},\end{aligned}$$

where

$$\mathbf{Q}^{(1)}(t)^*=\mathbf{Q}^{(1)}(t)^+\mathbf{I}_{\left\{\mathbf{Q}^{(0)}(t)\geq\mathbf{n}_t\right\}}-\mathbf{Q}^{(1)}(t)^-\mathbf{I}_{\left\{\mathbf{Q}^{(0)}(t)>\mathbf{n}_t\right\}} \tag{5.12}$$

and

$$Q_1^{(1)}(t)^*=Q_1^{(1)}(t)^+1_{\{Q_1^{(0)}(t)\geq n_t\}}-Q_1^{(1)}(t)^-1_{\{Q_1^{(0)}(t)>n_t\}}. \tag{5.13}$$

The vector formulas of Theorem 6.2 reduce to

$$\begin{aligned}&\boldsymbol{\lambda}_t+\Big(\mathbf{Q}^{(0)}(t)-\mathbf{n}_t\Big)^+\mathbf{\Delta}(\boldsymbol{\beta}_t)(\Psi_t+\mathbf{I})+\Big(\mathbf{Q}^{(0)}(t)\wedge\mathbf{n}_t\Big)\mathbf{\Delta}(\boldsymbol{\mu}_t)(\Phi_t+\mathbf{I})\\ &=\begin{bmatrix}\lambda_t & 0\end{bmatrix}+\begin{bmatrix}\Big(Q_1^{(0)}(t)-n_t\Big)^+ & 0\end{bmatrix}\begin{bmatrix}\beta_t & 0\\ 0 & 0\end{bmatrix}\begin{bmatrix}1 & 1-\psi_t\\ 0 & 1\end{bmatrix}\\ &\quad+\begin{bmatrix}Q_1^{(0)}(t)\wedge n_t & Q_2^{(0)}(t)\end{bmatrix}\begin{bmatrix}\mu_t^1 & 0\\ 0 & \mu_t^2\end{bmatrix}\begin{bmatrix}1 & 0\\ 1 & 1\end{bmatrix}\end{aligned}$$

$$= \Big(\lambda_t + \beta_t\Big(Q_1^{(0)}(t) - n_t\Big)^+ + \mu_t^1\Big(Q_1^{(0)}(t) \wedge n_t\Big) + \mu_t^2 Q_2^{(0)}(t)\Big)\,\mathbf{e}_1$$
$$+ \Big(\beta_t(1-\psi_t)\Big(Q_1^{(0)}(t) - n_t\Big)^+ + \mu_t^2 Q_2^{(0)}(t)\Big)\,\mathbf{e}_2. \tag{5.14}$$

Combining all these identities gives us our answer. □

Finally, we explore the asymptotic regime suggested in Halfin and Whitt [4] by applying the full power of Theorems 2.2 and 2.3 to this multiserver queue with abandonment and retrials. First, we modify our rate functions so that

$$\lambda_t^\eta \equiv \eta\lambda_t + \sqrt{\eta}\,\ell_t, \tag{5.15}$$
$$(\mu_t^i)^\eta \equiv \mu_t^i + \frac{1}{\sqrt{\eta}} m_t^i \quad \text{for } i = 1, 2, \tag{5.16}$$
$$\beta_t^\eta \equiv \beta_t + \frac{1}{\sqrt{\eta}} b_t, \tag{5.17}$$
$$\psi_t^\eta \equiv \psi_t + \frac{1}{\sqrt{\eta}} p_t, \tag{5.18}$$
$$n_t^\eta \equiv \eta n_t, \tag{5.19}$$

where like λ_t, μ_t^i, β_t and ϕ_t, the functions ℓ_t, m_t^i, b_t, and p_t are locally integrable, but unlike them, not necessarily non-negative. By Theorems 2.2 and 2.3 we see that these additional terms of order $\sqrt{\eta}$ or $1/\sqrt{\eta}$ have no effect on the fluid approximation of $\mathbf{Q}^\eta$. However, the diffusion approximation is now the unique solution to the integral equation

$$\begin{aligned}
Q_1^{(1)}(t) = Q_1^{(1)}(0) &+ \int_0^t \Big[\Big(\mu_s^1 1_{\{Q_1^{(0)}(s) \le n_s\}} + \beta_s 1_{\{Q_1^{(0)}(s) > n_s\}}\Big) Q_1^{(1)}(s)^- \\
&- \Big(\mu_s^1 1_{\{Q_1^{(0)}(s) < n_s\}} + \beta_s 1_{\{Q_1^{(0)}(s) \ge n_s\}}\Big) Q_1^{(1)}(s)^+ + \mu_s^2 Q_2^{(1)}(s)\Big]\, ds \\
&+ \int_0^t \Big[\ell_s + m_s^2 Q_2^{(0)}(s) - m_s^1\Big(Q_1^{(0)}(s) \wedge n_s\Big) - b_s\Big(Q_1^{(0)}(s) - n_s\Big)^+\Big]\, ds \\
&- \Big[B_{12}^b\Big(\int_0^t \Big(Q_1^{(0)}(s) - n_s\Big)^+ \beta_s(1-\psi_s)ds\Big) + B_{21}^c\Big(\int_0^t \Big(Q_2^{(0)}(s)\Big)\mu_s^2 ds\Big)\Big] \\
&+ B^a\Big(\int_0^t \lambda_s ds\Big) - B^b\Big(\int_0^t \Big(Q_1^{(0)}(s) - n_s\Big)^+ \beta_s\psi_s ds\Big) \\
&- B^c\Big(\int_0^t \Big(Q_1^{(0)}(s) \wedge n_s\Big)\mu_s^1 ds\Big),
\end{aligned}$$

and

$Q_2^{(1)}(t)$

$$
\begin{aligned}
&= Q_2^{(1)}(0) + B_{12}^b\left(\int_0^t \left(Q_1^{(0)}(s) - n_s\right)^+ \beta_s(1-\psi_s)ds\right) + B_{21}^c\left(\int_0^t \left(Q_2^{(0)}(s)\right)\mu_s^2 ds\right) \\
&\quad + \int_0^t \left[\left(Q_1^{(1)}(s)^+ 1_{\{Q_1^{(0)}(s)\geq n_s\}} - Q_1^{(1)}(s)^- 1_{\{Q_1^{(0)}(s)>n_s\}}\right)\beta_s(1-\psi_s) - \mu_s^2 Q_2^{(1)}(s)\right] ds \\
&\quad + \int_0^t \left[\left(b_s(1-\psi_s) - \beta_s p_s\right)\left(Q_1^{(0)}(s) - n_s\right)^+ - m_s^2 Q_2^{(0)}(s)\right] ds.
\end{aligned}
$$

The differential equations for the covariance matrix of $\mathbf{Q}^{(1)}$ are unchanged but the equations for the mean vector are now

$$
\begin{aligned}
&\frac{d}{dt}\mathsf{E}\left[Q_1^{(1)}(t)\right] \\
&\quad = \left(\mu_t^1 1_{\{Q_1^{(0)}(t)\leq n_t\}} + \beta_t 1_{\{Q_1^{(0)}(t)>n_t\}}\right)\mathsf{E}\left[Q_1^{(1)}(t)^-\right] \\
&\quad\quad - \left(\mu_t^1 1_{\{Q_1^{(0)}(t)<n_t\}} + \beta_t 1_{\{Q_1^{(0)}(t)\geq n_t\}}\right)\mathsf{E}\left[Q_1^{(1)}(t)^+\right] + \mu_t^2\mathsf{E}\left[Q_2^{(1)}(t)\right] \\
&\quad\quad + \ell_t + m_t^2 Q_2^{(0)}(t) - m_t^1\left(Q_1^{(0)}(t)\wedge n_t\right) - b_t\left(Q_1^{(0)}(t) - n_t\right)^+
\end{aligned} \tag{5.20}
$$

and

$$
\begin{aligned}
&\frac{d}{dt}\mathsf{E}\left[Q_2^{(1)}(t)\right] \\
&\quad = \beta_t(1-\psi_t)\left(\mathsf{E}\left[Q_1^{(1)}(t)^+\right]1_{\{Q_1^{(0)}(t)\geq n_t\}} - \mathsf{E}\left[Q_1^{(1)}(t)^-\right]1_{\{Q_1^{(0)}(t)>n_t\}}\right) \\
&\quad\quad - \mu_t^2\mathsf{E}\left[Q_2^{(1)}(t)\right] + \left[b_t(1-\psi_t) - \beta_t p_t\right]\left(Q_1^{(0)}(t) - n_t\right)^+ - m_t^2 Q_2^{(0)}(t).
\end{aligned} \tag{5.21}
$$

6. Jackson Networks with Abandonment

The multiserver queue with abandonment and retrials is a special case of a more general network that we discuss in this section. Here, we construct a time varying analogue of the Jackson network that has the added feature of service abandonment (see Figure 6.1). Extending Kendall notation, we call it the $(M_t/M_t\backslash M_t/n_t)^N$ network for short. We construct it by first defining the following set of parameters:

$\lambda_t^i =$ external arrival rate to node i at time t,
$\beta_t^i =$ abandonment rate for node i at time t,
$\mu_t^i =$ service rate for node i at time t,
$\psi_t^{ij} =$ abandonment routing probability from node i to node j at time t,
$\phi_t^{ij} =$ service routing probability from node i to node j at time t,

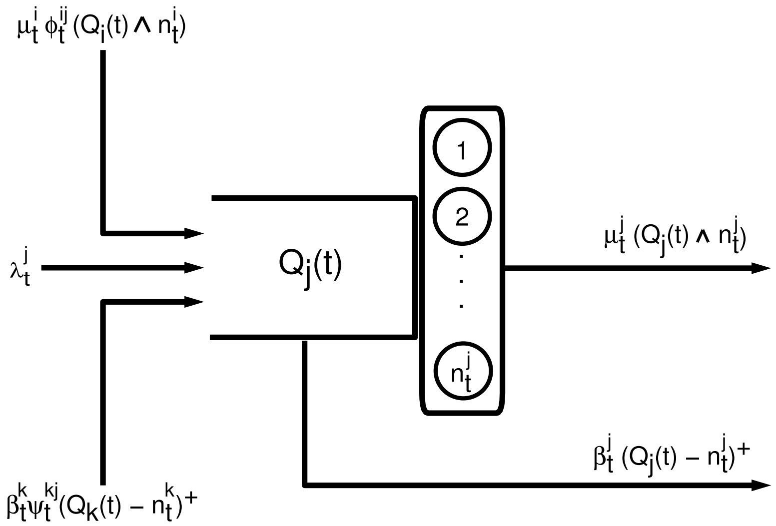

Figure 6.1. The Jackson network with abandonment.

$\psi_t^i =$ abandonment departure probability from node i at time t,

$\phi_t^i =$ service departure probability from node i at time t.

$n_t^i =$ number of servers for node i at time t,

where we require that

$$\psi_t^i + \sum_{j=1}^N \psi_t^{ij} = 1 \quad \text{and} \quad \phi_t^i + \sum_{j=1}^N \phi_t^{ij} = 1 \tag{6.1}$$

for all $t \geq 0$ and $i = 1, \ldots, N$.

We then set $\mathbb{V} = \mathbb{R}^N$ and define

$$\begin{aligned}\mathbf{Q}(t) = \mathbf{Q}(0) &+ \sum_{i=1}^N \sum_{j=1}^N A_{ij}^b \left(\int_0^t \left(Q_i(s) - n_s^i \right)^+ \beta_s^i \psi_s^{ij} ds \right) (\mathbf{e}_j - \mathbf{e}_i) \\ &+ \sum_{i=1}^N \sum_{j=1}^N A_{ij}^c \left(\int_0^t \left(Q_i(s) \wedge n_s^i \right) \mu_s^i \phi_s^{ij} ds \right) (\mathbf{e}_j - \mathbf{e}_i) \\ &+ \sum_{i=1}^N \left(A_i^a \left(\int_0^t \lambda_s^i ds \right) - A_i^b \left(\int_0^t \left(Q_i(s) - n_s^i \right)^+ \beta_s^i \psi_s^i ds \right) \right) \mathbf{e}_i \\ &- \sum_{i=1}^N A_i^c \left(\int_0^t \left(Q_i(s) \wedge n_s^i \right) \mu_s^i \phi_s^i ds \right) \mathbf{e}_i.\end{aligned} \tag{6.2}$$

Theorems 2.2 and 2.3 yield the following limiting results for these networks.

Theorem 6.1. Defining $\mathbf{Q}^\eta$ by uniform acceleration as in (2.8), the fluid limit for the $(M_t/M_t\backslash M_t/n_t)^N$ network is the unique solution to the integral equation

$$\mathbf{Q}^{(0)}(t) = \mathbf{Q}^{(0)}(0) + \int_0^t \Big[\boldsymbol{\lambda}_s + (\mathbf{Q}^{(0)}(s) - \mathbf{n}_s)^+ \boldsymbol{\Delta}(\boldsymbol{\beta}_s)(\Psi_s - \mathbf{I})\Big]\, ds \qquad (6.3)$$
$$+ \int_0^t (\mathbf{Q}^{(0)}(s) \wedge \mathbf{n}_s)\boldsymbol{\Delta}(\boldsymbol{\mu}_s)(\Phi_s - \mathbf{I})ds.$$

Moreover, the diffusion limit for the $(M_t/M_t\backslash M_t/n_t)^N$ network is the unique solution to the integral equation

$$\begin{aligned}
\mathbf{Q}^{(1)}(t) &= \mathbf{Q}^{(1)}(0) \\
&+ \int_0^t (\mathbf{Q}^{(1)}(s)^+ \mathbf{I}_{\{\mathbf{Q}^{(0)}(s)\geq \mathbf{n}_s\}} - \mathbf{Q}^{(1)}(s)^- \mathbf{I}_{\{\mathbf{Q}^{(0)}(s)> \mathbf{n}_s\}})\boldsymbol{\Delta}(\boldsymbol{\beta}_s)(\Psi_s - \mathbf{I})ds \\
&+ \int_0^t (\mathbf{Q}^{(1)}(s)^+ \mathbf{I}_{\{\mathbf{Q}^{(0)}(s)< \mathbf{n}_s\}} - \mathbf{Q}^{(1)}(s)^- \mathbf{I}_{\{\mathbf{Q}^{(0)}(s)\leq \mathbf{n}_s\}})\boldsymbol{\Delta}(\boldsymbol{\mu}_s)(\Phi_s - \mathbf{I})ds \\
&+ \sum_{i=1}^N \Bigg[\sum_{j=1}^N \Bigg(B^b_{ij} \left(\int_0^t \left(Q_i^{(0)}(s) - n_s^i\right)^+ \beta_s^i \psi_s^{ij} ds\right) \\
&\qquad + B^c_{ij} \left(\int_0^t \left(Q_i^{(0)}(s) \wedge n_s^i\right) \mu_s^i \phi_s^{ij} ds\right)\Bigg) (\mathbf{e}_j - \mathbf{e}_i) \\
&\qquad + \Bigg(B^a_i \left(\int_0^t \lambda_s^i ds\right) - B^b_i \left(\int_0^t \left(Q_i^{(0)}(s) - n_s^i\right)^+ \beta_s^i \psi_s^i ds\right) \\
&\qquad - B^c_i \left(\int_0^t \left(Q_i^{(0)}(s) \wedge n_s^i\right) \mu_s^i \phi_s^i ds\right)\Bigg) \mathbf{e}_i \Bigg].
\end{aligned} \qquad (6.4)$$

Proof. From (6.2), it follows that

$$\boldsymbol{\alpha}_t(\mathbf{x}) = \boldsymbol{\lambda}_t + (\mathbf{x} - \mathbf{n}_t)^+ \boldsymbol{\Delta}(\boldsymbol{\beta}_t)(\Psi_t - \mathbf{I}) + (\mathbf{x} \wedge \mathbf{n}_t)\boldsymbol{\Delta}(\boldsymbol{\mu}_t)(\Phi_t - \mathbf{I}). \qquad (6.5)$$

The fluid limit now follows from applying Theorem 2.2.

The scalable Lipschitz derivative of $\boldsymbol{\alpha}_t$ is

$$\Lambda\boldsymbol{\alpha}_t\Big(\mathbf{x}\Big)(\mathbf{y}) = (\mathbf{y}^+ \mathbf{I}_{\{\mathbf{x}\geq \mathbf{n}_t\}} - \mathbf{y}^- \mathbf{I}_{\{\mathbf{x}> \mathbf{n}_t\}})\boldsymbol{\Delta}(\boldsymbol{\beta}_t)(\Psi_t - \mathbf{I})$$
$$+ (\mathbf{y}^+ \mathbf{I}_{\{\mathbf{x}< \mathbf{n}_t\}} - \mathbf{y}^- \mathbf{I}_{\{\mathbf{x}\leq \mathbf{n}_t\}})\boldsymbol{\Delta}(\boldsymbol{\mu}_t)(\Phi_t - \mathbf{I}). \qquad (6.6)$$

The diffusion limit now follows from applying Theorem 2.3. □

Theorem 6.2. The mean vector for the diffusion limit solves the differential equation

$$\frac{d}{dt}\mathsf{E}\left[\mathbf{Q}^{(1)}(t)\right] = \mathsf{E}\left[\mathbf{Q}^{(1)}(t)^+ \mathbf{I}_{\{\mathbf{Q}^{(0)}(t)\geq \mathbf{n}_t\}} - \mathbf{Q}^{(1)}(t)^- \mathbf{I}_{\{\mathbf{Q}^{(0)}(t)> \mathbf{n}_t\}}\right] \boldsymbol{\Delta}(\boldsymbol{\beta}_t)(\Psi_t - \mathbf{I})$$

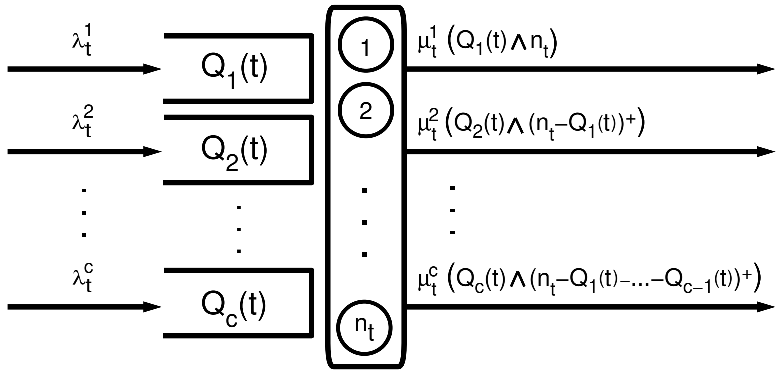

Figure 7.1. The preemptive priority queue.

$$+\mathsf{E}\left[\mathbf{Q}^{(1)}(t)^+\mathbf{I}_{\{\mathbf{Q}^{(0)}(t)<\mathbf{n}_t\}} - \mathbf{Q}^{(1)}(t)^-\mathbf{I}_{\{\mathbf{Q}^{(0)}(t)\le\mathbf{n}_t\}}\right]\boldsymbol{\Delta}(\boldsymbol{\mu}_t)(\Phi_t - \mathbf{I}),$$

and the covariance matrix for the diffusion limit solves the differential equation

$$\begin{aligned}
&\frac{d}{dt}\mathsf{Cov}\left[\mathbf{Q}^{(1)}(t),\mathbf{Q}^{(1)}(t)\right]\\
&= \left\{\mathsf{Cov}\left[\mathbf{Q}^{(1)}(t),\mathbf{Q}^{(1)}(t)^+\mathbf{I}_{\{\mathbf{Q}^{(0)}(t)\ge\mathbf{n}_t\}} - \mathbf{Q}^{(1)}(t)^-\mathbf{I}_{\{\mathbf{Q}^{(0)}(t)>\mathbf{n}_t\}}\right]\boldsymbol{\Delta}(\boldsymbol{\beta}_t)(\Psi_t - \mathbf{I})\right\}\\
&\quad + \left\{\mathsf{Cov}\left[\mathbf{Q}^{(1)}(t),\mathbf{Q}^{(1)}(t)^+\mathbf{I}_{\{\mathbf{Q}^{(0)}(t)<\mathbf{n}_t\}} - \mathbf{Q}^{(1)}(t)^-\mathbf{I}_{\{\mathbf{Q}^{(0)}(t)\le\mathbf{n}_t\}}\right]\boldsymbol{\Delta}(\boldsymbol{\mu}_t)(\Phi_t - \mathbf{I})\right\}\\
&\quad + \boldsymbol{\Delta}\left(\boldsymbol{\lambda}_t + \left(Q^{(0)}(t) - \mathbf{n}_t\right)^+\boldsymbol{\Delta}(\boldsymbol{\beta}_t)(\Psi_t + \mathbf{I}) + \left(Q^{(0)}(t)\wedge\mathbf{n}_t\right)\boldsymbol{\Delta}(\boldsymbol{\mu}_t)(\Phi_t + \mathbf{I})\right)\\
&\quad - \left\{\boldsymbol{\Delta}\left((Q^{(0)}(t) - \mathbf{n}_t)^+\right)\boldsymbol{\Delta}(\boldsymbol{\beta}_t)\Psi_t + \boldsymbol{\Delta}\left(Q^{(0)}(t)\wedge\mathbf{n}_t\right)\boldsymbol{\Delta}(\boldsymbol{\mu}_t)\Phi_t\right\}.
\end{aligned}$$

7. Priority Queues

A multiserver queue with preemptive priorities (see Figure 7.1) is defined using the following parameters:

$\lambda_t^i =$ arrival rate for class i customers at time t,

$\mu_t^i =$ service rate for class i customers at time t,

$n_t =$ number of servers at time t,

$c =$ number of customer classes.

We then set $\mathbb{V} = \mathbb{R}^c$ and define

$$\mathbf{Q}(t) = \mathbf{Q}(0) + \sum_{i=1}^{c}\left[A_i^a\left(\int_0^t \lambda_s^i ds\right) - A_i^b\left(\int_0^t \mu_s^i\, Q_i(s)\wedge\left(n_t - \sum_{j=1}^{i-1} Q_j(s)\right)^+ ds\right)\right]\mathbf{e}_i. \tag{7.1}$$

Theorem 7.1. Defining $\mathbf{Q}^\eta$ by uniform acceleration as in (2.8), the fluid limit for the priority queueing model is the solution to the integral equation

$$\mathbf{Q}^{(0)}(t) = \mathbf{Q}^{(0)}(0) + \int_0^t \left[\boldsymbol{\lambda}_s + \left(\mathbf{Q}^{(0)}(s) \wedge (n_s \mathbf{1} - \mathbf{Q}^{(0)}(s)\boldsymbol{\Theta})\right)^+ \boldsymbol{\Delta}(\boldsymbol{\mu}_s)\right] ds, \quad (7.2)$$

where $\boldsymbol{\Theta} = \{\, \theta_{ij} \mid 1 \leq i, j \leq c \,\}$ is the $c \times c$ matrix

$$\theta_{ij} = \begin{cases} 1 & \text{for } i < j, \\ 0 & \text{for } i \geq j. \end{cases} \quad (7.3)$$

Moreover, the diffusion limit is the solution to the integral equation

$$\begin{aligned} \mathbf{Q}^{(1)}(t) = \mathbf{Q}^{(1)}(0) &- \int_0^t \mathbf{Q}^{(1)}(s)\mathbf{I}_{\{\mathbf{Q}^{(0)}(s) < (n_t\mathbf{1} - \mathbf{Q}^{(0)}(s)\boldsymbol{\Theta})^+\}} \boldsymbol{\Delta}(\boldsymbol{\mu}_s) ds \\ &- \int_0^t (\mathbf{Q}^{(1)}(s)\boldsymbol{\Theta})^+ \mathbf{I}_{\{\mathbf{Q}^{(0)}(s) > (n_t\mathbf{1} - \mathbf{Q}^{(0)}(s)\boldsymbol{\Theta})^+,\ n_t\mathbf{1} \leq \mathbf{Q}^{(0)}(s)\boldsymbol{\Theta}\}} \boldsymbol{\Delta}(\boldsymbol{\mu}_s) ds \\ &+ \int_0^t (\mathbf{Q}^{(1)}(s)\boldsymbol{\Theta})^- \mathbf{I}_{\{\mathbf{Q}^{(0)}(s) > (n_t\mathbf{1} - \mathbf{Q}^{(0)}(s)\boldsymbol{\Theta})^+,\ n_t\mathbf{1} < \mathbf{Q}^{(0)}(s)\boldsymbol{\Theta}\}} \boldsymbol{\Delta}(\boldsymbol{\mu}_s) ds \\ &- \int_0^t \left(\mathbf{Q}^{(1)}(s) \wedge \left(-(\mathbf{Q}^{(1)}(s)\boldsymbol{\Theta})^*\right)\right) \cdot \mathbf{I}_{\{\mathbf{Q}^{(0)}(s) = (n_t\mathbf{1} - \mathbf{Q}^{(0)}(s)\boldsymbol{\Theta})^+\}} \boldsymbol{\Delta}(\boldsymbol{\mu}_s) ds \\ &+ \sum_{i=1}^c \left[B_i^a \left(\int_0^t \lambda_s^i ds\right) - B_i^b \left(\int_0^t \mu_s^i\, Q_i(s) \wedge \left(n_t - \sum_{j=1}^{i-1} Q_j(s)\right) ds\right)\right] \mathbf{e}_i, \end{aligned}$$

where

$$(\mathbf{Q}^{(1)}(t)\boldsymbol{\Theta})^* = (\mathbf{Q}^{(1)}(t)\boldsymbol{\Theta})^+ \mathbf{I}_{\{n_t\mathbf{1} < \mathbf{Q}^{(0)}(t)\boldsymbol{\Theta}\}} - (\mathbf{Q}^{(1)}(t)\boldsymbol{\Theta})^- \mathbf{I}_{\{n_t\mathbf{1} \leq \mathbf{Q}^{(0)}(t)\boldsymbol{\Theta}\}}. \quad (7.4)$$

Proof. From (7.1), it follows that

$$\boldsymbol{\alpha}_t(\mathbf{x}) = \boldsymbol{\lambda}_t - (\mathbf{x} \wedge (n_t\mathbf{1} - \mathbf{x}\boldsymbol{\Theta})^+)\boldsymbol{\Delta}(\boldsymbol{\mu}_t). \quad (7.5)$$

The fluid limit now follows from applying Theorem 2.2.

Applying (3.8) and (3.9), the scalable Lipschitz derivative of $\boldsymbol{\alpha}_t$ is

$$\begin{aligned} &\Lambda\boldsymbol{\alpha}_t(\mathbf{x}; \mathbf{y}) \\ &= \left(-(\mathbf{y}\boldsymbol{\Theta})^+ \mathbf{I}_{\{\mathbf{x} > (n_t\mathbf{1} - \mathbf{x}\boldsymbol{\Theta})^+,\ n_t\mathbf{1} \leq \mathbf{x}\boldsymbol{\Theta}\}} + (\mathbf{y}\boldsymbol{\Theta})^- \mathbf{I}_{\{\mathbf{x} > (n_t\mathbf{1} - \mathbf{x}\boldsymbol{\Theta})^+,\ n_t\mathbf{1} < \mathbf{x}\boldsymbol{\Theta}\}}\right) \boldsymbol{\Delta}(\boldsymbol{\mu}_t) \\ &\quad - \left(\mathbf{y} \wedge \left(-(\mathbf{y}\boldsymbol{\Theta})^+ \mathbf{I}_{\{n_t\mathbf{1} < \mathbf{x}\boldsymbol{\Theta}\}} + (\mathbf{y}\boldsymbol{\Theta})^- \mathbf{I}_{\{n_t\mathbf{1} \leq \mathbf{x}\boldsymbol{\Theta}\}}\right)\right) \mathbf{I}_{\{\mathbf{x} = (n_t\mathbf{1} - \mathbf{x}\boldsymbol{\Theta})^+\}} \boldsymbol{\Delta}(\boldsymbol{\mu}_t) \\ &\quad - \mathbf{y}\mathbf{I}_{\{\mathbf{x} < (n_t\mathbf{1} - \mathbf{x}\boldsymbol{\Theta})^+\}} \boldsymbol{\Delta}(\boldsymbol{\mu}_t). \end{aligned}$$

The diffusion limit now follows from applying Theorem 2.3. $\square$

Theorem 7.2. The mean vector for the diffusion limit solves the differential equation

$$
\begin{aligned}
\frac{d}{dt}\mathsf{E}\left[\mathbf{Q}^{(1)}(t)\right] = {} & \mathsf{E}\left[\mathbf{Q}^{(1)}(t)\right]\mathbf{I}_{\{\mathbf{Q}^{(0)}(t)<(n_t\mathbf{1}-\mathbf{x}\mathbf{\Theta})^+\}} \\
& +\mathsf{E}\left[\left(\mathbf{Q}^{(1)}(t)\mathbf{\Theta}\right)^+\right]\mathbf{I}_{\{\mathbf{Q}^{(0)}(t)>(n_t\mathbf{1}-\mathbf{Q}^{(0)}(t)\mathbf{\Theta})^+,\ n_t\mathbf{1}\leq\mathbf{Q}^{(0)}(t)\mathbf{\Theta}\}} \\
& -\mathsf{E}\left[\left(\mathbf{Q}^{(1)}(t)\mathbf{\Theta}\right)^-\right]\mathbf{I}_{\{\mathbf{Q}^{(0)}(t)>(n_t\mathbf{1}-\mathbf{Q}^{(0)}(t)\mathbf{\Theta})^+,\ n_t\mathbf{1}<\mathbf{Q}^{(0)}(t)\mathbf{\Theta}\}} \\
& +\mathsf{E}\left[\mathbf{Q}^{(1)}(t)\wedge\left(-(\mathbf{Q}^{(1)}(t)\mathbf{\Theta})^*\right)\right]\cdot\mathbf{I}_{\{\mathbf{Q}^{(0)}(t)=(n_t\mathbf{1}-\mathbf{Q}^{(0)}(t)\mathbf{\Theta})^+\}}
\end{aligned}
$$

and the covariance matrix for the diffusion limit solves the differential equation

$$
\begin{aligned}
& \frac{d}{dt}\mathsf{Cov}\left[\mathbf{Q}^{(1)}(t),\mathbf{Q}^{(1)}(t)\right] \\
& \quad = \left\{\mathsf{Cov}\left[\mathbf{Q}^{(1)}(t),\mathbf{Q}^{(1)}(t)\right]\mathbf{I}_{\{\mathbf{Q}^{(0)}(t)<(n_t\mathbf{1}-\mathbf{x}\mathbf{\Theta})^+\}}\right\} \\
& \quad + \left\{\mathsf{Cov}\left[\mathbf{Q}^{(1)}(t),\left(\mathbf{Q}^{(1)}(t)\mathbf{\Theta}\right)^+\right]\mathbf{I}_{\{\mathbf{Q}^{(0)}(t)>(n_t\mathbf{1}-\mathbf{Q}^{(0)}(t)\mathbf{\Theta})^+,\ n_t\mathbf{1}\leq\mathbf{Q}^{(0)}(t)\mathbf{\Theta}\}}\right\} \\
& \quad - \left\{\mathsf{Cov}\left[\mathbf{Q}^{(1)}(t),\left(\mathbf{Q}^{(1)}(t)\mathbf{\Theta}\right)^-\right]\mathbf{I}_{\{\mathbf{Q}^{(0)}(t)>(n_t\mathbf{1}-\mathbf{Q}^{(0)}(t)\mathbf{\Theta})^+,\ n_t\mathbf{1}<\mathbf{Q}^{(0)}(t)\mathbf{\Theta}\}}\right\} \\
& \quad + \left\{\mathsf{Cov}\left[\mathbf{Q}^{(1)}(t),\mathbf{Q}^{(1)}(t)\wedge\left(-(\mathbf{Q}^{(1)}(t)\mathbf{\Theta})^*\right)\right]\cdot\mathbf{I}_{\{\mathbf{Q}^{(0)}(t)=(n_t\mathbf{1}-\mathbf{Q}^{(0)}(t)\mathbf{\Theta})^+\}}\right\} \\
& \quad +\mathbf{\Delta}(\boldsymbol{\lambda}_t)-\mathbf{\Delta}\left(\mathbf{Q}^{(0)}(t)\wedge\left(n_t\mathbf{1}-\mathbf{Q}^{(0)}(t)\mathbf{\Theta}\right)^+\right)\cdot\mathbf{\Delta}(\boldsymbol{\mu}_t).
\end{aligned}
$$

where $(\mathbf{Q}^{(1)}(t)\mathbf{\Theta})^*$ is given by (7.4).

8. Jackson Networks with State Dependent Routing

We now consider another generalization of the classical Jackson network where the arrival rate, service rate, and routing probabilities are all functions of the state of the joint queue length vector $\mathbf{Q}(t)$ (see Figure 8.1). We extend Kendall notation and call it the $(M_t(Q)/M_t(Q)/n_t)^N$ network. The examples considered in Sections 4 and 7 are special cases of this network. We construct the $(M_t(Q)/M_t(Q)/n_t)^N$ network by first defining the following set of parameters:

$$
\begin{aligned}
\lambda_t^i\big(\mathbf{Q}(t)\big) &= \text{ external arrival rate to node } i \text{ at time } t, \\
\mu_t^i\big(\mathbf{Q}(t)\big) &= \text{ service rate for node } i \text{ at time } t, \\
\phi_t^{ij}\big(\mathbf{Q}(t)\big) &= \text{ service routing probability to node } i \text{ from node } j \text{ at time } t,
\end{aligned}
$$

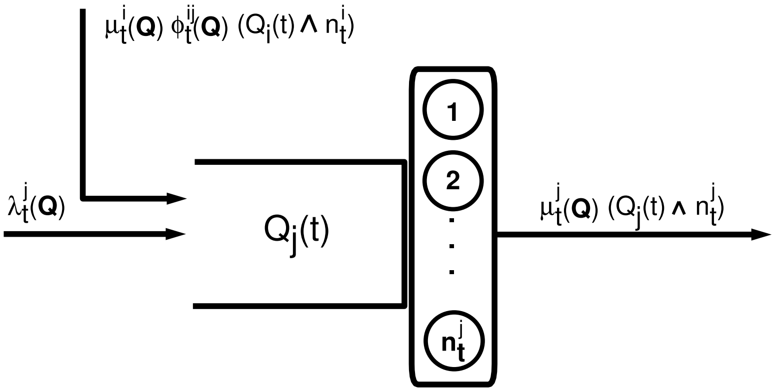

Figure 8.1. The Jackson network with state dependent routing.

$\phi_t^i\big(\mathbf{Q}(t)\big)$ = service departure probability from node i at time t.
n_t^i = number of servers for node i at time t,

where $\lambda_t^i(\cdot)$, $\mu_t^i(\cdot)$, $\phi_t^{ij}(\cdot)$, and $\phi_t^i(\cdot)$ are all Lipschitz functions with scalable Lipschitz derivatives and we require that

$$\psi_t^i(\mathbf{x}) + \sum_{j=1}^N \psi_t^{ij}(\mathbf{x}) = 1 \quad \text{and} \quad \phi_t^i(\mathbf{x}) + \sum_{j=1}^N \phi_t^{ij}(\mathbf{x}) = 1 \tag{8.1}$$

for all $t \geq 0$, all $\mathbf{x} \in \mathbb{V} = \mathbb{R}^N$, and $i = 1, \ldots, N$.

We then define

$$\begin{aligned}\mathbf{Q}(t) = \mathbf{Q}(0) &+ \sum_{i=1}^N A_i^a \left(\int_0^t \lambda_s^i\big(\mathbf{Q}(s)\big)\, ds \right) \mathbf{e}_i \\ &- \sum_{i=1}^N A_i^b \left(\int_0^t \big(Q_i(s) \wedge n_s^i\big) \mu_s^i\big(\mathbf{Q}(s)\big) \phi_s^i\big(\mathbf{Q}(s)\big)\, ds \right) \mathbf{e}_i \\ &+ \sum_{i=1}^N \sum_{j=1}^N A_{ij}^c \left(\int_0^t \big(Q_i(s) \wedge n_s^i\big) \mu_s^i\big(\mathbf{Q}(s)\big) \phi_s^{ij}\big(\mathbf{Q}(s)\big)\, ds \right) (\mathbf{e}_j - \mathbf{e}_i).\end{aligned} \tag{8.2}$$

Theorem 8.1. Defining $\mathbf{Q}^\eta$ by uniform acceleration as in (2.8), the fluid limit for the $(M_t(Q)/M_t(Q)/n_t)^N$ network is the unique solution to the integral equation

$$\begin{aligned}\mathbf{Q}^{(0)}(t) = \mathbf{Q}^{(0)}(0) &+ \int_0^t \boldsymbol{\lambda}_s\big(\mathbf{Q}^{(0)}(s)\big)\, ds \\ &+ \int_0^t \big(\mathbf{Q}^{(0)}(s) \wedge \mathbf{n}_s\big) \boldsymbol{\Delta}\left(\boldsymbol{\mu}_s\big(\mathbf{Q}^{(0)}(s)\big)\right) \left(\boldsymbol{\Phi}_s\big(\mathbf{Q}^{(0)}(s)\big) - \mathbf{I}\right) ds.\end{aligned} \tag{8.3}$$

Moreover, the diffusion limit for the $(M_t(Q)/M_t(Q)/n_t)^N$ network is the unique solution to the integral equation

$$
\begin{aligned}
\mathbf{Q}^{(1)}(t) = \mathbf{Q}^{(1)}(0) &+ \int_0^t \Lambda\boldsymbol{\lambda}_t\Big(\mathbf{Q}^{(0)}(s);\mathbf{Q}^{(1)}(s)\Big)\,ds \\
&+ \int_0^t \mathbf{Q}^{(1)}(s)^*\boldsymbol{\Delta}\left(\boldsymbol{\mu}_s\Big(\mathbf{Q}^{(0)}(s)\Big)\right)\left(\Phi_s\Big(\mathbf{Q}^{(0)}(s)\Big)-\mathbf{I}\right)ds \\
&+ \int_0^t \Big(\mathbf{Q}^{(0)}(s)\wedge\mathbf{n}_s\Big)\left(\boldsymbol{\Delta}\Big(\Lambda\boldsymbol{\mu}_t(\mathbf{Q}^{(0)}(s);\mathbf{Q}^{(1)}(s))\Big)\Big(\Phi_t(\mathbf{Q}^{(0)}(s))-\mathbf{I}\Big)\right)ds \\
&+ \int_0^t \Big(\mathbf{Q}^{(0)}(s)\wedge\mathbf{n}_s\Big)\left(\boldsymbol{\Delta}\left(\boldsymbol{\mu}_t\Big(\mathbf{Q}^{(0)}(s)\Big)\right)\Lambda\Phi_t\left(\mathbf{Q}^{(0)}(s);\mathbf{Q}^{(1)}(s)\right)\right)ds \\
&+ \sum_{i=1}^N\sum_{j=1}^N B_{ij}^c\left(\int_0^t\Big(Q_i^{(0)}(s)\wedge n_s^i\Big)\mu_s^i\Big(\mathbf{Q}^{(0)}(s)\Big)\phi_s^{ij}\Big(\mathbf{Q}^{(0)}(s)\Big)ds\right)(\mathbf{e}_j-\mathbf{e}_i) \\
&- \sum_{i=1}^N B_i^c\left(\int_0^t\Big(Q_i^{(0)}(s)\wedge n_s^i\Big)\mu_s^i\Big(\mathbf{Q}^{(0)}(s)\Big)\phi_s^{i}\Big(\mathbf{Q}^{(0)}(s)\Big)ds\right)\mathbf{e}_i \\
&+ \sum_{i=1}^N B_i^a\left(\int_0^t\lambda_s^i\Big(\mathbf{Q}^{(0)}(s)\Big)ds\right)\mathbf{e}_i,
\end{aligned}
$$

where

$$
\mathbf{Q}^{(1)}(t)^* = \mathbf{Q}^{(1)}(t)^+\mathbf{I}_{\{\mathbf{Q}^{(0)}(t)<\mathbf{n}_t\}} - \mathbf{Q}^{(1)}(t)^-\mathbf{I}_{\{\mathbf{Q}^{(0)}(t)\le\mathbf{n}_t\}}. \tag{8.4}
$$

Proof. From (8.2), it follows that

$$
\boldsymbol{\alpha}_t(\mathbf{x}) = \boldsymbol{\lambda}_t(\mathbf{x}) + (\mathbf{x}\wedge\mathbf{n}_t)\boldsymbol{\Delta}\Big(\boldsymbol{\mu}_t(\mathbf{x})\Big)\Big(\Phi_t(\mathbf{x})-\mathbf{I}\Big). \tag{8.5}
$$

The fluid limit now follows from applying Theorem 2.2.

The scalable Lipschitz derivative of $\boldsymbol{\alpha}_t$ is

$$
\begin{aligned}
\Lambda\boldsymbol{\alpha}_t(\mathbf{x};\mathbf{y}) = \Lambda\boldsymbol{\lambda}_t(\mathbf{x};\mathbf{y}) &+ \Big(\mathbf{y}^+\mathbf{I}_{\{\mathbf{x}<\mathbf{n}_t\}} - \mathbf{y}^-\mathbf{I}_{\{\mathbf{x}\le\mathbf{n}_t\}}\Big)\boldsymbol{\Delta}\Big(\boldsymbol{\mu}_t(\mathbf{x})\Big)\Big(\Phi_t(\mathbf{x})-\mathbf{I}\Big) \\
&+ (\mathbf{x}\wedge\mathbf{n}_t)\left(\boldsymbol{\Delta}\Big(\Lambda\boldsymbol{\mu}_t(\mathbf{x};\mathbf{y})\Big)\Big(\Phi_t(\mathbf{x})-\mathbf{I}\Big) + \boldsymbol{\Delta}(\boldsymbol{\mu}_t(\mathbf{x}))\Lambda\Phi_t(\mathbf{x};\mathbf{y})\right).
\end{aligned} \tag{8.6}
$$

The diffusion limit now follows from applying Theorem 2.3. □

Theorem 8.2. The mean vector for the diffusion limit solves the differential equation

$$
\begin{aligned}
\frac{d}{dt}\mathsf{E}\Big[\mathbf{Q}^{(1)}(t)\Big] = {} & \mathsf{E}\Big[\Lambda\boldsymbol{\lambda}_t\Big(\mathbf{Q}^{(0)}(t);\mathbf{Q}^{(1)}(t)\Big)\Big] \\
&+ \mathsf{E}\Big[\mathbf{Q}^{(1)}(t)^+\Big]\,\mathbf{I}_{\{\mathbf{Q}^{(0)}(t)<\mathbf{n}_t\}}\cdot\boldsymbol{\Delta}\left(\boldsymbol{\mu}_t\Big(\mathbf{Q}^{(0)}(t)\Big)\right)\left(\Phi_t\Big(\mathbf{Q}^{(0)}(t)\Big)-\mathbf{I}\right)
\end{aligned} \tag{8.7}
$$

$$
\begin{aligned}
&-\mathsf{E}\left[\mathbf{Q}^{(1)}(t)^-\right] \mathbf{I}_{\{\mathbf{Q}^{(0)}(t)\leq \mathbf{n}_t\}} \cdot \mathbf{\Delta}\left(\boldsymbol{\mu}_t\Big(\mathbf{Q}^{(0)}(t)\Big)\right)\left(\Phi_t\Big(\mathbf{Q}^{(0)}(t)\Big) - \mathbf{I}\right) \\
&+(\mathbf{Q}^{(0)}(t)\wedge \mathbf{n}_t)\mathbf{\Delta}\left(\mathsf{E}\left[\Lambda\boldsymbol{\mu}_t\Big(\mathbf{Q}^{(0)}(t);\mathbf{Q}^{(1)}(t)\Big)\right]\right)\left(\Phi_t\Big(\mathbf{Q}^{(0)}(t)\Big) - \mathbf{I}\right) \\
&+(\mathbf{Q}^{(0)}(t)\wedge \mathbf{n}_t)\mathbf{\Delta}\left(\boldsymbol{\mu}_t\Big(\mathbf{Q}^{(0)}(t)\Big)\right)\mathsf{E}\left[\Lambda\Phi_t\left(\mathbf{Q}^{(0)}(t);\mathbf{Q}^{(1)}(t)\right)\right]
\end{aligned}
$$

and the covariance matrix for the diffusion limit solves the differential equation

$$
\begin{aligned}
&\frac{d}{dt}\mathsf{Cov}\left[\mathbf{Q}^{(1)}(t),\mathbf{Q}^{(1)}(t)\right] = \qquad\qquad (8.8)\\
&\left\{\mathsf{Cov}\left[\mathbf{Q}^{(1)}(t),\Lambda\boldsymbol{\lambda}_t\Big(\mathbf{Q}^{(0)}(t);\mathbf{Q}^{(1)}(t)\Big)\right]\right\} \\
&+\left\{\mathsf{Cov}\left[\mathbf{Q}^{(1)}(t),\mathbf{Q}^{(1)}(t)^+\right]\mathbf{I}_{\{\mathbf{Q}^{(0)}(t)<\mathbf{n}_t\}}\cdot\mathbf{\Delta}\left(\boldsymbol{\mu}_t\Big(\mathbf{Q}^{(0)}(t)\Big)\right)\cdot\left(\Phi_t\Big(\mathbf{Q}^{(0)}(t)\Big)-\mathbf{I}\right)\right\} \\
&-\left\{\mathsf{Cov}\left[\mathbf{Q}^{(1)}(t),\mathbf{Q}^{(1)}(t)^-\right]\mathbf{I}_{\{\mathbf{Q}^{(0)}(t)\leq\mathbf{n}_t\}}\cdot\mathbf{\Delta}\left(\boldsymbol{\mu}_t\Big(\mathbf{Q}^{(0)}(t)\Big)\right)\cdot\left(\Phi_t\Big(\mathbf{Q}^{(0)}(t)\Big)-\mathbf{I}\right)\right\} \\
&+\left\{\mathsf{Cov}\left[\mathbf{Q}^{(1)}(t),(\mathbf{Q}^{(0)}(t)\wedge\mathbf{n}_t)\mathbf{\Delta}\left(\Lambda\boldsymbol{\mu}_t\Big(\mathbf{Q}^{(0)}(t);\mathbf{Q}^{(1)}(t)\Big)\right)\left(\Phi_t\Big(\mathbf{Q}^{(0)}(t)\Big)-\mathbf{I}\right)\right]\right\} \\
&+\left\{\mathsf{Cov}\left[\mathbf{Q}^{(1)}(t),(\mathbf{Q}^{(0)}(t)\wedge\mathbf{n}_t)\mathbf{\Delta}\left(\boldsymbol{\mu}_t\Big(\mathbf{Q}^{(0)}(t)\Big)\right)\Lambda\Phi_t\left(\mathbf{Q}^{(0)}(t);\mathbf{Q}^{(1)}(t)\right)\right]\right\} \\
&+\mathbf{\Delta}\Big(\boldsymbol{\lambda}_t+\Big(\mathbf{Q}^{(0)}(t)\wedge\mathbf{n}_t\Big)\mathbf{\Delta}\Big(\boldsymbol{\mu}_t(\mathbf{Q}^{(0)}(t))\Big)\Big(\Phi_t(\mathbf{Q}^{(0)}(t))+\mathbf{I}\Big)\Big) \\
&-\left\{\mathbf{\Delta}\Big(\mathbf{Q}^{(0)}(t)\wedge\mathbf{n}_t\Big)\mathbf{\Delta}\Big(\boldsymbol{\mu}_t(\mathbf{Q}^{(0)}(t))\Big)\Phi_t\Big(\mathbf{Q}^{(0)}(t)\Big)\right\}.
\end{aligned}
$$

9. Proofs of the Strong Limit Theorems

In this section, we prove the strong approximation and strong law of large number theorems stated in Section 2. As preparation, we first show existence and uniqueness for the process $\mathbf{Q} = \{\, \mathbf{Q}(t) \mid t \geq 0 \,\}$. In defining this process, we also construct a process $Z = \{\, Z(t) \mid t \geq 0 \,\}$ that we use as a bound on its growth. In a fashion similar to results found in Kurtz [9], the Z process plays the key role in a stochastic analogue to Gronwall's inequality.

Recall, from Section 2, that for all the results in this section we make the following set of assumptions:

1. The family of Lipschitz rate functions $\{\, \alpha_t(\cdot, i) \mid i \in I \,\}$ has the growth condition

$$\|\alpha_t(\cdot, i)\| \leq \beta_t \gamma^{(i)}, \tag{9.1}$$

for all $i \in I$, where β_t is a positive, locally integrable function and

$\left\{ \gamma^{(i)} \,\middle|\, i \in I \right\}$ is an absolutely summable sequence. Similarly, the family of Lipschitz rate functions $\{ \alpha_t^\eta(\cdot, i) \mid i \in I, \eta > 0 \}$ has the property that

$$\|\alpha_t^\eta(\cdot, i)\| \leq \eta \beta_t \gamma^{(i)}. \tag{9.2}$$

2. The family of transition vectors $\{ \mathbf{v}_i \mid i \in I \}$ has the property that

$$\sum_{i \in I} |\mathbf{v}_i| < \infty. \tag{9.3}$$

It should be noted that the last condition is not as limiting as it seems. If $\mathbb{V}$ is the Banach space ℓ_1 of absolutely summable sequences and $\{ \mathbf{v}_i \mid i \in I \}$ is the set of unit basis vectors, merely redefine the norm to give each basis vector a weight where all of the weights are summable.

Lemma 9.1. There exists a positive, increasing process $Z \equiv \{ Z(t) \mid t \geq 0 \}$ that is the unique solution to the equation

$$Z(t) \equiv X\left(\int_0^t \beta_s Z(s) ds \right), \tag{9.4}$$

for all $t \geq 0$, where the process $X \equiv \{ X(t) \mid t \geq 0 \}$ is defined by a random variable $X(0) > 0$ that is independent of the collection of Poisson processes $\{ A_i \mid i \in I \}$ and

$$X(t) \equiv X(0) + \sum_{i \in I} A_i(\gamma^{(i)} t) |\mathbf{v}_i|, \tag{9.5}$$

is an increasing pure jump process with no explosions.

Moreover, the process $M \equiv \{ M(t) \mid t \geq 0 \}$ defined by

$$M(t) \equiv Z(t) \exp\left(- \sum_{i \in I} \gamma^{(i)} |\mathbf{v}_i| \cdot \int_0^t \beta_s ds \right), \tag{9.6}$$

is a martingale.

Proof. Since $\sum_{i \in I} \gamma^{(i)} < \infty$, it follows that the process $A \equiv \{ A(t) \mid t \geq 0 \}$, where

$$A(t) = \sum_{i \in I} A_i(\gamma^{(i)} t) \tag{9.7}$$

is Poisson with mean rate $\sum_{i\in I}\gamma^{(i)}$. Given that $\sum_{i\in I}|\mathbf{v}_i|<\infty$ also, we then have for all $t>0$

$$\mathsf{E}[X(t)] = \mathsf{E}[X(0)] + \sum_{i\in I}\gamma^{(i)}|\mathbf{v}_i|t < \mathsf{E}[X(0)] + \sum_{i\in I}\gamma^{(i)}\cdot\sum_{i\in I}|\mathbf{v}_i|t < \infty. \tag{9.8}$$

Hence $X(t)<\infty$ a.s. for all t and its jump times are given by the Poisson process A.

Now let $b(t)\equiv\int_0^t\beta_s ds$. If β is a strictly positive function then b^{-1}, the inverse function for b, is well defined. For all $t\geq 0$, define the random process $\{\,\tau(t)\,|\,t\geq 0\,\}$ such that

$$\tau^{-1}(t) = b^{-1}\left(\int_0^t \frac{ds}{X(s)}\right). \tag{9.9}$$

This is well defined since $X(t)\geq X(0)>0$ for all $t\geq 0$. Hence, we can define the process $\{\,Z(t)\,|\,t\geq 0\,\}$ to be

$$Z(t)\equiv\frac{\tau'(t)}{\beta_t}. \tag{9.10}$$

Since $\tau(t)\equiv\int_0^t\beta_s Z(s)ds$, we have uniqueness.

Now consider the process $M_*\equiv\{\,M_*(t)\,|\,t\geq 0\,\}$, where

$$M_*(t)\equiv Z(t)-\sum_{i\in I}\gamma^{(i)}|\mathbf{v}_i|\cdot\int_0^t\beta_s Z(s)ds. \tag{9.11}$$

This process is a martingale, since $\left\{\,X(t)-\sum_{i\in I}\gamma^{(i)}|\mathbf{v}_i|\,t\,\middle|\,t\geq 0\right\}$ is one, and by (9.9) we see that $\tau(t)=\int_0^t\beta_s Z(s)ds$ is a stopping time with respect to the filtration generated by the process X for all $t\geq 0$. Finally, M is a martingale since M_* is one of bounded variation and

$$\begin{aligned} M(t) &= Z(t)\,\exp\left(-\sum_{i\in I}\gamma^{(i)}|\mathbf{v}_i|\cdot\int_0^t\beta_s ds\right)\\ &= Z(0)+\int_0^t\exp\left(-\sum_{i\in I}\gamma^{(i)}|\mathbf{v}_i|\cdot\int_0^s\beta_r dr\right)dM_*(s), \end{aligned} \tag{9.12}$$

which completes the proof. □

Theorem 9.2. Given the rate functions $\{\,\alpha_t(\cdot,i)\,|\,t\geq 0,\,i\in I\,\}$ and the initial state vector $\mathbf{Q}(0)$ is independent of the collection of Poisson processes

$\{ A_i \mid i \in I \}$, we can construct a unique stochastic process $\mathbf{Q} \equiv \{ \mathbf{Q}(t) \mid t \geq 0 \}$ such that

$$\mathbf{Q}(t) = \mathbf{Q}(0) + \sum_{i \in I} A_i \left(\int_0^t \alpha_s \Big(\mathbf{Q}(s), i \Big) ds \right) \mathbf{v}_i. \tag{9.13}$$

Moreover, we have for all $t \geq 0$,

$$1 + \sup_{0 \leq s \leq t} |\mathbf{Q}(s)| \leq Z(t), \tag{9.14}$$

where the process $Z \equiv \{ Z(t) \mid t \geq 0 \}$ is uniquely defined by (9.4) and (9.5) with $X(0) = 1 + |\mathbf{Q}(0)|$.

Proof. Define the following sequence $\mathbf{Q}_n \equiv \{ \mathbf{Q}_n(t) \mid t \geq 0 \}$, where $\mathbf{Q}_0(t) \equiv \mathbf{Q}(0)$ for all $t \geq 0$ and for all positive integers n we have

$$\mathbf{Q}_n(t) \equiv \mathbf{Q}(0) + \sum_{i \in I} A_i \left(\int_0^{t \wedge T_n} \alpha_s \Big(\mathbf{Q}_{n-1}(s), i \Big) ds \right) \mathbf{v}_i, \tag{9.15}$$

where

$$T_n \equiv \inf \left\{ t \,\middle|\, \sum_{i \in I} A_i \left(\int_0^t \alpha_s \Big(\mathbf{Q}_{n-1}(s), i \Big) ds \right) = n \right\}. \tag{9.16}$$

We are done once we prove the following two statements:

1. $\mathbf{Q}_n(t) = \mathbf{Q}_{n-1}(t)$ for all $0 \leq t < T_n$.
2. $\lim_{n \to \infty} T_n = \infty$ a.s.

We can then construct the desired process $\mathbf{Q} \equiv \{ \mathbf{Q}(t) \mid t \geq 0 \}$ by defining for all $n \geq 1$,

$$\mathbf{Q}(t) = \mathbf{Q}_{n-1}(t) \quad \text{for all} \quad 0 \leq t < T_n. \tag{9.17}$$

Uniqueness follows by using induction on n. Using (9.15) shows that the uniqueness of $\mathbf{Q}_n$ implies the uniqueness of $\mathbf{Q}_{n+1}$.

The first statement is proved by using induction on n. The result holds for $n = 1$, since $t < T_1$ implies that $\mathbf{Q}_1(t) = \mathbf{Q}_0(t) = \mathbf{Q}(0)$, since $t < T_1$ means that

$$\sum_{i \in I} A_i \left(\int_0^t \alpha_s \Big(\mathbf{Q}_0(s), i \Big) ds \right) = 0, \tag{9.18}$$

and for all $i \in I$, we have

$$A_i \left(\int_0^t \alpha_s \Big(\mathbf{Q}_0(s), i \Big) ds \right) = 0. \tag{9.19}$$

If we assume that $\mathbf{Q}_{n-1}(t) = \mathbf{Q}_n(t)$ for all $0 \le t < T_n$, then it follows that

$$\int_0^t \alpha_s\Big(\mathbf{Q}_n(s), i\Big)ds = \int_0^t \alpha_s\Big(\mathbf{Q}_{n-1}(s), i\Big)ds \tag{9.20}$$

for all $i \in I$ and $0 \le t \le T_n$. We then must have

$$\mathbf{Q}_{n+1}(t) = \mathbf{Q}_n(t) \;\text{ for all }\; 0 \le t < T_n. \tag{9.21}$$

Now consider the case of $T_n \le t < T_{n+1}$. By definition of the $\mathbf{Q}_n$'s we have

$$\mathbf{Q}_n(t) = \mathbf{Q}_n(T_n) \tag{9.22}$$

and

$$\mathbf{Q}_{n+1}(t) = \mathbf{Q}(0) + \sum_{i\in I} A_i\left(\int_0^t \alpha_s\Big(\mathbf{Q}_n(s), i\Big)ds\right)\mathbf{v}_i. \tag{9.23}$$

However, by the definition of T_n and T_{n+1}, we have

$$n \le \sum_{i\in I} A_i\left(\int_0^t \alpha_s\Big(\mathbf{Q}_n(s), i\Big)ds\right) < n+1. \tag{9.24}$$

This follows from the fact that the A_i's are increasing processes. Combining this with the fact that the processes are also integer valued not only shows that the sum in (9.24) equals n, but that for all $i \in I$,

$$\begin{aligned} A_i\left(\int_0^t \alpha_s\Big(\mathbf{Q}_n(s), i\Big)ds\right) &= A_i\left(\int_0^{T_n} \alpha_s\Big(\mathbf{Q}_n(s), i\Big)ds\right) \\ &= A_i\left(\int_0^{T_n} \alpha_s\Big(\mathbf{Q}_{n-1}(s), i\Big)ds\right) \end{aligned}$$

where the last equality follows from (9.20). Finally, this gives us

$$\mathbf{Q}_{n+1}(t) = \mathbf{Q}_n(t) = \mathbf{Q}_n(T_n) \;\text{ for all }\; T_n \le t < T_{n+1}, \tag{9.25}$$

which means that

$$\mathbf{Q}_{n+1}(t) = \mathbf{Q}_n(t) \;\text{ for all }\; 0 \le t < T_{n+1}, \tag{9.26}$$

completing the induction argument.

To prove the second statement, we observe that

$$|\mathbf{Q}_n(t)| \le |\mathbf{Q}(0)| + \sum_{i\in I} A_i\left(\int_0^t \Big\|\alpha_s\Big(\cdot, i\Big)\Big\|\,(1 + |\mathbf{Q}_{n-1}(s)|)\,ds\right)|\mathbf{v}_i| \tag{9.27}$$

$$\le |\mathbf{Q}(0)| + \sum_{i\in I} A_i\left(\gamma^{(i)}\int_0^t \beta_s\,(1 + |\mathbf{Q}_{n-1}(s)|)\,ds\right)|\mathbf{v}_i|. \tag{9.28}$$

It follows by induction that for all $n \geq 0$ and $t \geq 0$ we have

$$1 + |\mathbf{Q}_n(t)| \leq Z(t), \tag{9.29}$$

where $Z(t)$ is the process defined in (9.4), with $X(0) = 1 + |\mathbf{Q}(0)|$. Consequently,

$$\sum_{i \in I} A_i \left(\int_0^t \alpha_s \Big(\mathbf{Q}_n(s), i \Big) ds \right) \leq \sum_{i \in I} A_i \left(\gamma^{(i)} \int_0^t \beta_s Z(s) ds \right). \tag{9.30}$$

If we set $\gamma \equiv \sum_{i \in I} \gamma^{(i)}$, then $\sum_{i \in I} A_i(\gamma^{(i)} t)$ is a Poisson process with rate γ. Let T_n^γ be the time of the n-th jump for this Poisson process. It now follows from (9.16) and (9.30) that for all $n \geq 1$,

$$S_n^\gamma \leq T_n, \tag{9.31}$$

where

$$T_n^\gamma = \int_0^{S_n^\gamma} \beta_s Z(s) ds. \tag{9.32}$$

Since $\lim_{n \to \infty} T_n^\gamma = \infty$ a.s., then $\lim_{n \to \infty} S_n^\gamma = \infty$ a.s., and so we have $\lim_{n \to \infty} T_n = \infty$ a.s.

Finally, (9.14) follows from the fact that $1 + |\mathbf{Q}(t)| \leq Z(t)$ for all t and Z is a non-decreasing process. □

Now consider the family of uniformly accelerated processes $\{ \mathbf{Q}^\eta \mid \eta > 0 \}$ as defined in (2.9). We will always assume that every element of $\{ \mathbf{Q}^\eta(0) \mid \eta > 0 \}$ is a random vector in $\mathbb{V}$ that is independent of the collection of Poisson processes $\{ A_i \mid i \in I \}$.

Lemma 9.3. If $\{ \mathbf{Q}^\eta(0) \mid \eta > 0 \}$ is a family of random vectors independent of the Poisson processes $\{ A_i \mid i \in I \}$, then

$$\overline{\lim_{\eta \to \infty}} \frac{|\mathbf{Q}^\eta(0)|}{\eta} < \infty \quad \text{a.s.} \quad \text{implies} \quad \overline{\lim_{\eta \to \infty}} Z^\eta(s) < \infty \quad \text{a.s.}, \tag{9.33}$$

which implies

$$\overline{\lim_{\eta \to \infty}} \sup_{0 \leq s \leq t} \frac{|\mathbf{Q}^\eta(s)|}{\eta} < \infty \quad \text{a.s.} \tag{9.34}$$

Proof. Let $\{ Z^\eta(t) \mid t \geq 0 \}$ be the unique process such that

$$Z^\eta(t) = X^\eta \left(\int_0^t \beta_s Z^\eta(s) ds \right) \tag{9.35}$$

and

$$X^{\eta}(t) \equiv 1 + \frac{1}{\eta} |\mathbf{Q}^{\eta}(0)| + \sum_{i \in I} \frac{1}{\eta} A_i \left(\gamma^{(i)} \eta t \right) |\mathbf{v}_i|. \tag{9.36}$$

A simple modification of the proof for Theorem 9.2 gives us

$$1 + \frac{1}{\eta} \sup_{0 \le s \le t} |\mathbf{Q}^{\eta}(s)| \le Z^{\eta}(t). \tag{9.37}$$

A similar modification of the proof for Lemma 9.1 shows that

$$\left\{ Z^{\eta}(t) \exp\Big(- \sum_{i \in I} \gamma^{(i)} |\mathbf{v}_i| \cdot \int_0^t \beta_s ds \Big) \,\Bigg|\, t \ge 0 \right\} \tag{9.38}$$

is a martingale, and so for all $t \ge 0$.

$$\mathsf{E}\left[Z^{\eta}(t) \right] = \left(1 + \frac{1}{\eta} |\mathbf{Q}^{\eta}(0)| \right) \exp\left(\gamma \int_0^t \beta_s ds \right). \tag{9.39}$$

Using Chebyshev's inequality, we see that for all $t > 0$, the set $\{ Z^{\eta}(t) \mid \eta > 0 \}$ is a tight family of random variables and so

$$\overline{\lim_{\eta \to \infty}} Z^{\eta}(t) < \infty \quad \text{a.s.} \tag{9.40}$$

which combined with (9.37), completes the proof. □

Our fundamental results, stated in Section 2, are proved within the framework of strong approximations. The framework is based on a pathwise approximation of the Poisson process, articulated in the following lemma.

Lemma 9.4 (Kurtz [9], Lemma 3.1). A standard (rate 1) Poisson process $\{ A(t) \mid t \ge 0 \}$ can be realized on the same probability space as a standard Brownian motion $\{ B(t) \mid t \ge 0 \}$ in such a way that the positive random variable X, given by

$$X \equiv \sup_{t \ge 0} \frac{|A(t) - t - B(t)|}{\log(2 \vee t)} < \infty, \tag{9.41}$$

has a finite moment generating function in a neighborhood of the origin. In particular it has a finite mean.

Using Lemma 9.4, we associate with our stochastic primitives $\{ A_i \mid i \in I \}$, namely the family of mutually independent standard Poisson processes, another family $\{ B_i \mid i \in I \}$ of mutually independent standard Brownian motions, such that

$$X_i \equiv \sup_{t \geq 0} \frac{|A_i(t) - t - B_i(t)|}{\log(2 \vee t)} < \infty \quad \text{a.s.} \tag{9.42}$$

Moreover, the random variables X_i's can be taken i.i.d. with a finite mean. Note that both families, as well as the X_i's, must be realized on a common probability space. In the sequel, we write $\omega \in \Omega$ for elementary outcomes in this common probability space.

Now we give the proof for our strong approximation theorem.

Proof of Theorem 2.1. Using Lemma 9.4, we have

$$\sup_{0 \leq s \leq t} \left| \mathbf{Q}^\eta(s) - \mathbf{Q}^\eta(0) - \int_0^s \boldsymbol{\alpha}_r^\eta\Big(\frac{1}{\eta}\mathbf{Q}^\eta(r)\Big) dr - \sum_{i \in I} B_i \left(\int_0^s \alpha_r^\eta\Big(\frac{1}{\eta}\mathbf{Q}^\eta(r), i\Big) dr \right) \mathbf{v}_i \right|$$

$$\leq \sum_{i \in I} X_i \log \left(2 \vee \int_0^t \alpha_s^\eta\Big(\frac{1}{\eta}\mathbf{Q}^\eta(s), i\Big) ds \right) |\mathbf{v}_i| \tag{9.43}$$

$$\leq \sum_{i \in I} X_i \log \left(2 \vee \left(\int_0^t \|\alpha_s^\eta(\cdot, i)\| \left(1 + \frac{1}{\eta}|\mathbf{Q}^\eta(s)|\right) ds \right) \right) |\mathbf{v}_i| \tag{9.44}$$

$$\leq \sum_{i \in I} X_i \log \left(2 \vee \left(\eta \gamma^{(i)} Z^\eta(t) \int_0^t \beta_s ds \right) \right) |\mathbf{v}_i| \tag{9.45}$$

$$\leq \sum_{i \in I} X_i |\mathbf{v}_i| \log \left(2 \vee \left(\eta \gamma^* Z^\eta(t) \int_0^t \beta_s ds \right) \right), \tag{9.46}$$

where $\gamma^* \equiv \sup_{i \in I} \gamma_i$. The first step follows from (9.42). The second step follows from (2.6). The third step uses (9.2) and (9.14). Since the X_i are nonnegative, using (2.10) and the fact that the X_i's are i.i.d. with the same finite mean implies that

$$\mathsf{E}\left[\sum_{i \in I} X_i |\mathbf{v}_i| \right] < \infty, \tag{9.47}$$

which gives us the desired result. □

Now we give the proof for our functional strong law of large numbers.

Proof of Theorem 2.2. First observe that $\boldsymbol{\alpha}_t^{(0)}$ is well defined for almost all t as given by (2.16) since for all $t \geq 0$,

$$\int_0^t \left\| \boldsymbol{\alpha}_s^{(0)} \right\| ds \leq \int_0^t \left\| \frac{\boldsymbol{\alpha}_s^\eta}{\eta} \right\| ds + \int_0^t \left\| \frac{\boldsymbol{\alpha}_s^\eta}{\eta} - \boldsymbol{\alpha}_s^{(0)} \right\| ds \tag{9.48}$$

$$\leq \sup_{i \in I} |\mathbf{v}_i| \cdot \sum_{i \in I} \left[\gamma_i \int_0^t \beta_s ds + \int_0^t \left\| \frac{\alpha_s^\eta(\cdot, i)}{\eta} - \alpha_s^{(0)}(\cdot, i) \right\| ds \right]$$

$$< \infty$$

Moreover, we have for all $t \geq 0$,

$$\int_0^t \left\| \frac{\boldsymbol{\alpha}_s^\eta}{\eta} - \boldsymbol{\alpha}_s^{(0)} \right\| ds \leq \sup_{i \in I} |\mathbf{v}_i| \cdot \sum_{i \in I} \int_0^t \left\| \frac{\alpha_s^\eta(\cdot, i)}{\eta} - \alpha_s^{(0)}(\cdot, i) \right\| ds < \infty. \tag{9.49}$$

Applying condition (9.3) then gives us

$$\lim_{\eta \to \infty} \int_0^t \left\| \frac{\boldsymbol{\alpha}_s^\eta}{\eta} - \boldsymbol{\alpha}_s^{(0)} \right\| ds = 0 \tag{9.50}$$

Combining (2.12) with (2.15), the integral equation for $\mathbf{Q}^{(0)}$, we obtain

$$\begin{aligned} \frac{\mathbf{Q}^\eta(t)}{\eta} - \mathbf{Q}^{(0)}(t) = & \int_0^t \left[\frac{1}{\eta} \boldsymbol{\alpha}_s^\eta \Big(\frac{1}{\eta} \mathbf{Q}^\eta(s) \Big) - \boldsymbol{\alpha}_s^{(0)} \Big(\mathbf{Q}^{(0)}(s) \Big) \right] ds \\ & + \sum_{i \in I} \frac{1}{\eta} B_i \left(\int_0^t \alpha_s^\eta \Big(\frac{1}{\eta} \mathbf{Q}^\eta(s), i \Big) ds \right) \mathbf{v}_i + O(\frac{\log \eta}{\eta}), \end{aligned} \tag{9.51}$$

where $O(\frac{\log \eta}{\eta})$ holds uniformly on compact sets in t.

Taking the supremum of (9.51), we obtain

$$\begin{aligned} & \sup_{0 \leq s \leq t} \left| \frac{\mathbf{Q}^\eta(s)}{\eta} - \mathbf{Q}^{(0)}(s) \right| \\ & \quad \leq \int_0^t \left| \frac{1}{\eta} \boldsymbol{\alpha}_s^\eta \Big(\frac{1}{\eta} \mathbf{Q}^\eta(s) \Big) - \boldsymbol{\alpha}_s^{(0)} \Big(\frac{1}{\eta} \mathbf{Q}^\eta(s) \Big) \right| ds \\ & \quad + \int_0^t \left| \boldsymbol{\alpha}_s^{(0)} \Big(\frac{1}{\eta} \mathbf{Q}^\eta(s) \Big) - \boldsymbol{\alpha}_s^{(0)} \Big(\mathbf{Q}^{(0)}(s) \Big) \right| ds \\ & \quad + \sum_{i \in I} \sup_{0 \leq s \leq t} \frac{1}{\eta} \left| B_i \left(\int_0^s \alpha_r^\eta \Big(\frac{1}{\eta} \mathbf{Q}^\eta(r), i \Big) dr \right) \right| |\mathbf{v}_i| + O(\frac{\log \eta}{\eta}) \\ & \quad \leq \int_0^t \left\| \frac{\boldsymbol{\alpha}_s^\eta}{\eta} - \boldsymbol{\alpha}_s^{(0)} \right\| ds \cdot Z^\eta(t) + \int_0^t \left\| \boldsymbol{\alpha}_s^{(0)} \right\| \sup_{0 \leq r \leq s} \left| \frac{\mathbf{Q}^\eta(r)}{\eta} - \mathbf{Q}^{(0)}(r) \right| ds \\ & \quad + \sum_{i \in I} \sup_{0 \leq s \leq t} \frac{1}{\eta} \left| B_i \left(\eta \gamma^{(i)} \int_0^s \beta_r dr \cdot Z^\eta(s) \right) \right| |\mathbf{v}_i| + O(\frac{\log \eta}{\eta}). \end{aligned} \tag{9.52}$$

Since (9.48) holds, we can apply Gronwall's inequality (Lemma 11.1, Corollary 11.2, or Hale [3] page 36) and obtain

$$\begin{aligned} & \sup_{0 \leq s \leq t} \left| \frac{\mathbf{Q}^\eta(s)}{\eta} - \mathbf{Q}^{(0)}(s) \right| \\ & \quad \leq \exp\Big(\int_0^t \left\| \boldsymbol{\alpha}_s^{(0)} \right\| ds \Big) \left(\int_0^t \left\| \frac{\boldsymbol{\alpha}_s^\eta}{\eta} - \boldsymbol{\alpha}_s^{(0)} \right\| ds \cdot Z^\eta(t) \right. \end{aligned} \tag{9.53}$$

$$+\sum_{i\in I}\sup_{0\le s\le t}\frac{1}{\eta}\left|B_i\left(\eta\gamma^{(i)}\int_0^s\beta_r\,dr\cdot Z^\eta(s)\right)\right||\mathbf{v}_i|+O(\frac{\log\eta}{\eta})\Bigg).$$

Using (9.50) and Lemma 9.3, taking the limit yields

$$\lim_{\eta\to\infty}\sup_{0\le s\le t}\left|\frac{\mathbf{Q}^\eta(s)}{\eta}-\mathbf{Q}^{(0)}(s)\right|=0, \tag{9.54}$$

which completes the proof. □

10. Proof of the Central Limit Theorem

In this section, we prove results related to convergence in distribution, so it is useful to define partial ordering between two real-valued random variables in distribution. For any two real-valued random variables X and Y, we define the relation $X \le_{st} Y$ to mean that there exists some $\hat{X}$ and $\hat{Y}$ with $X \stackrel{d}{=} \hat{X}$ and $Y \stackrel{d}{=} \hat{Y}$ such that $\mathsf{P}(\hat{X} \le \hat{Y}) = 1$. We use $<_{st}$ to denote a strict inequality.

Before we prove Theorem 2.3, we show that the limit supremum of $\left|\mathbf{Q}^\eta(t) - \eta\mathbf{Q}^{(0)}(t)\right|/\sqrt{\eta}$ as $\eta \to \infty$ is always finite in distribution.

Theorem 10.1. For all $t \ge 0$, if $\boldsymbol{\alpha}_t(\cdot)$ satisfies the same hypothesis as in Theorem 2.1 and

$$\sum_{i\in I}\overline{\lim_{\eta\to\infty}}\int_0^t\left\|\frac{\alpha_s^\eta(\cdot,i)-\eta\alpha_s^{(0)}(\cdot,i)}{\sqrt{\eta}}\right\|ds<\infty, \tag{10.1}$$

we then have

$$\overline{\lim_{\eta\to\infty}}\sup_{0\le s\le t}\frac{\left|\mathbf{Q}^\eta(s)-\eta\mathbf{Q}^{(0)}(s)\right|}{\sqrt{\eta}}<_{st}\infty. \tag{10.2}$$

Proof. By Theorem 2.2, we know that $\boldsymbol{\alpha}_t^{(0)}$ is a Lipschitz mapping of $\mathbb{V}$ into itself where $\left\|\boldsymbol{\alpha}_t^{(0)}\right\|$ is locally integrable over t, since dividing the sum in condition (10.1) by $\sqrt{\eta}$ and taking the limit as $\eta \to \infty$ gives us condition (2.13). Moreover, multiplying the inequality (9.49) by $\sqrt{\eta}$ gives us

$$\int_0^t\left\|\frac{\boldsymbol{\alpha}_s^\eta-\eta\boldsymbol{\alpha}_s^{(0)}}{\sqrt{\eta}}\right\|ds<\infty \tag{10.3}$$

for all $t \ge 0$.

Using Theorem 2.1 and the self similarity of standard Brownian motion, we have

$$\mathbf{Q}^\eta(t) \stackrel{d}{=} \mathbf{Q}^\eta(0) + \int_0^t \boldsymbol{\alpha}_s^\eta\Big(\frac{1}{\eta}\mathbf{Q}^\eta(s)\Big)ds \qquad (10.4)$$

$$+ \sum_{i\in I}\sqrt{\eta}B_i^*\left(\frac{1}{\eta}\int_0^t \alpha_s^\eta\Big(\frac{1}{\eta}\mathbf{Q}^\eta(s), i\Big)ds\right)\mathbf{v}_i + O(\log\eta),$$

which holds uniformly on compact sets of t. The collection of independent Brownian motions used here are denoted B_i^* for all $i \in I$ to distinguish them from the Brownian motions B_i used in Section 9. This gives us

$$\frac{\mathbf{Q}^\eta(t) - \eta\mathbf{Q}^{(0)}(t)}{\sqrt{\eta}} \stackrel{d}{=} \frac{1}{\sqrt{\eta}}\int_0^t \boldsymbol{\alpha}_s^\eta\Big(\frac{1}{\eta}\mathbf{Q}^\eta(s)\Big) - \eta\boldsymbol{\alpha}_s^{(0)}\Big(\mathbf{Q}^{(0)}(s)\Big)ds \qquad (10.5)$$

$$+ \sum_{i\in I} B_i^*\left(\frac{1}{\eta}\int_0^t \alpha_s^\eta\Big(\frac{1}{\eta}\mathbf{Q}^\eta(s), i\Big)ds\right)\mathbf{v}_i + O(\frac{\log\eta}{\sqrt{\eta}}),$$

which holds uniformly on compact sets of t. Applying the Lipschitz property of $\boldsymbol{\alpha}_s^\eta$, we have

$$\sup_{0\le s\le t}\frac{\Big|\mathbf{Q}^\eta(s) - \eta\mathbf{Q}^{(0)}(s)\Big|}{\sqrt{\eta}} \le_{st} \frac{1}{\sqrt{\eta}}\int_0^t \left|\boldsymbol{\alpha}_s^\eta\Big(\frac{1}{\eta}\mathbf{Q}^\eta(s)\Big) - \eta\boldsymbol{\alpha}_s^{(0)}\Big(\frac{1}{\eta}\mathbf{Q}^\eta(s)\Big)\right| ds$$

$$+ \frac{1}{\sqrt{\eta}}\int_0^t \left|\eta\boldsymbol{\alpha}_s^{(0)}\Big(\frac{1}{\eta}\mathbf{Q}^\eta(s)\Big) - \eta\boldsymbol{\alpha}_s^{(0)}\Big(\mathbf{Q}^{(0)}(s)\Big)\right| ds$$

$$+ \sum_{i\in I}\sup_{0\le s\le t}\left|B_i^*\left(\frac{1}{\eta}\int_0^s \alpha_r^\eta\Big(\frac{1}{\eta}\mathbf{Q}^\eta(r)\Big)dr\right)\right| |\mathbf{v}_i| + O(\frac{\log\eta}{\sqrt{\eta}})$$

$$\le_{st} \int_0^t \frac{1}{\sqrt{\eta}}\left\|\boldsymbol{\alpha}_s^\eta - \eta\boldsymbol{\alpha}_s^{(0)}\right\| ds \cdot Z^\eta(t)$$

$$+ \int_0^t \left\|\boldsymbol{\alpha}_s^{(0)}\right\| \frac{1}{\sqrt{\eta}}\sup_{0\le r\le s}\Big|\mathbf{Q}^\eta(r) - \eta\mathbf{Q}^{(0)}(r)\Big| ds$$

$$+ \sum_{i\in I}\sup_{0\le s\le t}\left|B_i^*\left(\gamma^{(i)}\int_0^s \beta_r dr \cdot Z^\eta(s)\right)\right| |\mathbf{v}_i| + O(\frac{\log\eta}{\sqrt{\eta}}).$$

Now we apply Gronwall's inequality, which yields

$$\sup_{0\le s\le t}\frac{\Big|\mathbf{Q}^\eta(s) - \eta\mathbf{Q}^{(0)}(s)\Big|}{\sqrt{\eta}}$$

$$\le_{st} \exp\Big(\int_0^t \left\|\boldsymbol{\alpha}_s^{(0)}\right\| ds\Big) \cdot \Bigg(\int_0^t \frac{1}{\sqrt{\eta}}\left\|\boldsymbol{\alpha}_s^\eta - \eta\boldsymbol{\alpha}_s^{(0)}\right\| ds \cdot Z^\eta(t)$$

$$+ \sum_{i \in I} \sup_{0 \le s \le t} \left| B_i^* \left(\gamma^{(i)} \int_0^s \beta_r dr \cdot Z^\eta(s) \right) \right| |\mathbf{v}_i| + O(\frac{\log \eta}{\sqrt{\eta}}) \Big) .$$

Taking limits on both sides we obtain

$$\overline{\lim_{\eta \to \infty}} \sup_{0 \le s \le t} \frac{\left| \mathbf{Q}^\eta(s) - \eta \mathbf{Q}^{(0)}(s) \right|}{\sqrt{\eta}}$$
$$\le_{st} \exp\Big(\int_0^t \|\boldsymbol{\alpha}_s\| \, ds \Big) \cdot \sum_{i \in I} \sup_{0 \le s \le t} \left| B_i^* \left(\gamma^{(i)} \int_0^s \beta_r dr \cdot Z(s) \right) \right| |\mathbf{v}_i|,$$

and this completes the proof. □

Proof of Theorem 2.3. By applying arguments similar to those in the proof of Theorem 2.2 combined with conditions (2.24) and (2.25), we can show that both that $\boldsymbol{\alpha}_t^{(0)}$ and $\boldsymbol{\alpha}_t^{(1)}$ are well defined Lipschitz functions mapping $\mathbb{V}$ into itself for almost all t, where $\left\|\boldsymbol{\alpha}_t^{(0)}\right\|$ and $\left\|\boldsymbol{\alpha}_t^{(1)}\right\|$ are locally integrable functions of t. Moreover by similar arguments we can show that conditions (2.24) and (2.25) imply

$$\lim_{\eta \to \infty} \int_0^t \left\| \frac{\boldsymbol{\alpha}_s^\eta}{\eta} - \boldsymbol{\alpha}_s^{(0)} \right\| ds = \lim_{\eta \to \infty} \int_0^t \left\| \frac{\boldsymbol{\alpha}_s^\eta - \eta \boldsymbol{\alpha}_s^{(0)}}{\sqrt{\eta}} - \boldsymbol{\alpha}_s^{(1)} \right\| ds = 0. \qquad (10.6)$$

Now combine the integral equations for $\mathbf{Q}^\eta$, $\mathbf{Q}^{(0)}$, and $\mathbf{Q}^{(1)}$ to obtain

$$\frac{\mathbf{Q}^\eta(t) - \eta \mathbf{Q}^{(0)}(t)}{\sqrt{\eta}} - \mathbf{Q}^{(1)}(t)$$
$$\stackrel{d}{=} \int_0^t \frac{1}{\sqrt{\eta}} \Big(\boldsymbol{\alpha}_s^\eta \Big(\frac{1}{\eta} \mathbf{Q}^\eta(s) \Big) - \eta \boldsymbol{\alpha}_s^{(0)} \Big(\mathbf{Q}^{(0)}(s) \Big) \Big) ds$$
$$- \int_0^t \Lambda \boldsymbol{\alpha}_s^{(0)} \Big(\mathbf{Q}^{(0)}(s); \mathbf{Q}^{(1)}(s) \Big) + \boldsymbol{\alpha}_s^{(1)} \Big(\mathbf{Q}^{(0)}(s) \Big) ds$$
$$+ \sum_{i \in I} \left[B_i^* \Big(\frac{1}{\eta} \int_0^t \alpha_s^\eta \Big(\frac{1}{\eta} \mathbf{Q}^\eta(s), i \Big) ds \Big) - B_i^* \Big(\int_0^t \alpha_s^{(0)} \Big(\mathbf{Q}^{(0)}(s), i \Big) ds \Big) \right] \mathbf{v}_i + O(\frac{\log \eta}{\sqrt{\eta}})$$
$$\stackrel{d}{=} \int_0^t \frac{1}{\sqrt{\eta}} \Big(\boldsymbol{\alpha}_s^\eta \Big(\frac{1}{\eta} \mathbf{Q}^\eta(s) \Big) - \eta \boldsymbol{\alpha}_s^{(0)} \Big(\frac{1}{\eta} \mathbf{Q}^\eta(s) \Big) \Big) - \boldsymbol{\alpha}_s^{(1)} \Big(\frac{1}{\eta} \mathbf{Q}^\eta(s) \Big) ds$$
$$+ \int_0^t \frac{1}{\sqrt{\eta}} \Big(\eta \boldsymbol{\alpha}_s^{(0)} \Big(\frac{1}{\eta} \mathbf{Q}^\eta(s) \Big) - \eta \boldsymbol{\alpha}_s^{(0)} \Big(\mathbf{Q}^{(0)}(s) \Big) \Big) - \Lambda \boldsymbol{\alpha}_s^{(0)} \Big(\mathbf{Q}^{(0)}(s); \mathbf{Q}^{(1)}(s) \Big) ds$$
$$+ \int_0^t \boldsymbol{\alpha}_s^{(1)} \Big(\frac{1}{\eta} \mathbf{Q}^\eta(s) \Big) - \boldsymbol{\alpha}_s^{(1)} \Big(\mathbf{Q}^{(0)}(s) \Big) ds$$
$$+ \sum_{i \in I} \left[B_i^* \Big(\frac{1}{\eta} \int_0^t \alpha_s^\eta \Big(\frac{1}{\eta} \mathbf{Q}^\eta(s), i \Big) ds \Big) - B_i^* \Big(\int_0^t \alpha_s^{(0)} \Big(\mathbf{Q}^{(0)}(s), i \Big) ds \Big) \right] \mathbf{v}_i + O(\frac{\log \eta}{\sqrt{\eta}})$$

$$
\begin{aligned}
&\stackrel{d}{=} \int_0^t \left(\frac{\boldsymbol{\alpha}_s^\eta - \eta \boldsymbol{\alpha}_s^{(0)}}{\sqrt{\eta}} - \boldsymbol{\alpha}_s^{(1)} \right) \left(\frac{1}{\eta} \mathbf{Q}^\eta(s) \right) ds && (10.7) \\
&\quad + \int_0^t \sqrt{\eta} \Big(\boldsymbol{\alpha}_s^{(0)} \Big(\mathbf{Q}^{(0)}(s) + \frac{1}{\eta} \mathbf{Q}^\eta(s) - \mathbf{Q}^{(0)}(s) \Big) && (10.8) \\
&\qquad\qquad - \boldsymbol{\alpha}_s^{(0)} \Big(\mathbf{Q}^{(0)}(s) \Big) - \Lambda \boldsymbol{\alpha}_s^{(0)} \Big(\mathbf{Q}^{(0)}(s); \frac{1}{\eta} \mathbf{Q}^\eta(s) - \mathbf{Q}^{(0)}(s) \Big) \Big) ds \\
&\quad + \int_0^t \sqrt{\eta} \Lambda \boldsymbol{\alpha}_s^{(0)} \Big(\mathbf{Q}^{(0)}(s); \frac{1}{\eta} \mathbf{Q}^\eta(s) - \mathbf{Q}^{(0)}(s) \Big) - \Lambda \boldsymbol{\alpha}_s^{(0)} \Big(\mathbf{Q}^{(0)}(s); \mathbf{Q}^{(1)}(s) \Big) ds && (10.9) \\
&\quad + \int_0^t \boldsymbol{\alpha}_s^{(1)} \Big(\frac{1}{\eta} \mathbf{Q}^\eta(s) \Big) - \boldsymbol{\alpha}_s^{(1)} \Big(\mathbf{Q}^{(0)}(s) \Big) ds && (10.10) \\
&\quad + \sum_{i \in I} \left[B_i^* \Big(\frac{1}{\eta} \int_0^t \alpha_s^\eta \Big(\frac{1}{\eta} \mathbf{Q}^\eta(s), i \Big) ds \Big) - B_i^* \Big(\int_0^t \alpha_s^{(0)} \Big(\mathbf{Q}^{(0)}(s), i \Big) ds \Big) \right] \mathbf{v}_i && (10.11) \\
&\quad + O(\frac{\log \eta}{\sqrt{\eta}}).
\end{aligned}
$$

Now we show that summands (10.7), (10.8), (10.10), and (10.11) all converge to zero.

For the first summand (10.7), we have

$$
\left| \int_0^t \left(\frac{\boldsymbol{\alpha}_s^\eta - \eta \boldsymbol{\alpha}_s^{(0)}}{\sqrt{\eta}} - \boldsymbol{\alpha}_s^{(1)} \right) \left(\frac{1}{\eta} \mathbf{Q}^\eta(s) \right) ds \right| \leq \int_0^t \left\| \frac{\boldsymbol{\alpha}_s^\eta - \eta \boldsymbol{\alpha}_s^{(0)}}{\sqrt{\eta}} - \boldsymbol{\alpha}_s^{(1)} \right\| ds \cdot Z^\eta(t). \tag{10.12}
$$

For the second summand (10.8) when $\mathbf{Q}^\eta(s) \neq \eta \mathbf{Q}^{(0)}(s)$, we have

$$
\begin{aligned}
&\sqrt{\eta} \left| \boldsymbol{\alpha}_s^{(0)} \Big(\frac{1}{\eta} \mathbf{Q}^\eta(s) \Big) - \boldsymbol{\alpha}_s^{(0)} \Big(\mathbf{Q}^{(0)}(s) \Big) - \Lambda \boldsymbol{\alpha}_s^{(0)} \Big(\frac{1}{\eta} \mathbf{Q}^\eta(s) - \mathbf{Q}^{(0)}(s) \Big) \right| && (10.13) \\
&\qquad = \sqrt{\eta} \left| \frac{1}{\eta} \mathbf{Q}^\eta(s) - \mathbf{Q}^{(0)}(s) \right| \\
&\qquad\quad \times \frac{\left| \boldsymbol{\alpha}_s^{(0)} \Big(\frac{1}{\eta} \mathbf{Q}^\eta(s) \Big) - \boldsymbol{\alpha}_s^{(0)} \Big(\mathbf{Q}^{(0)}(s) \Big) - \Lambda \boldsymbol{\alpha}_s^{(0)} \Big(\frac{1}{\eta} \mathbf{Q}^\eta(s) - \mathbf{Q}^{(0)}(s) \Big) \right|}{\left| \frac{1}{\eta} \mathbf{Q}^\eta(s) - \mathbf{Q}^{(0)}(s) \right|},
\end{aligned}
$$

so combining the fact that $\boldsymbol{\alpha}_t^{(0)}(\cdot)$ is scalable Lipschitz differentiable with Theorem 2.2 we obtain

$$
\lim_{\eta \to \infty} \sqrt{\eta} \left| \boldsymbol{\alpha}_s \Big(\frac{1}{\eta} \mathbf{Q}^\eta(s) \Big) - \boldsymbol{\alpha}_s \Big(\mathbf{Q}^{(0)}(s) \Big) - \Lambda \boldsymbol{\alpha}_s \Big(\frac{1}{\eta} \mathbf{Q}^\eta(s) - \mathbf{Q}^{(0)}(s) \Big) \right| = 0. \tag{10.14}
$$

For the fourth summand (10.10), we have

$$\left|\int_0^t \boldsymbol{\alpha}_s^{(1)}\Big(\frac{1}{\eta}\mathbf{Q}^\eta(s)\Big) - \boldsymbol{\alpha}_s^{(1)}\Big(\mathbf{Q}^{(0)}(s)\Big)ds\right| \leq \int_0^t \left\|\boldsymbol{\alpha}_s^{(1)}\right\| ds \cdot \sup_{0\leq s\leq t}\left|\frac{1}{\eta}\mathbf{Q}^\eta(s) - \mathbf{Q}^{(0)}(s)\right|. \tag{10.15}$$

Finally for the fifth summand (10.11), observe that

$$\begin{aligned}
&\sup_{0\leq s\leq t}\left|\frac{1}{\eta}\int_0^s \alpha_r^\eta\Big(\frac{1}{\eta}\mathbf{Q}^\eta(r),i\Big)dr - \int_0^s \alpha_r^{(0)}\Big(\mathbf{Q}^{(0)}(r),i\Big)dr\right| \\
&\quad\leq \int_0^t \left|\frac{1}{\eta}\alpha_s^\eta\Big(\frac{1}{\eta}\mathbf{Q}^\eta(s),i\Big) - \alpha_s^{(0)}\Big(\frac{1}{\eta}\mathbf{Q}^\eta(s),i\Big)\right| ds \\
&\qquad + \int_0^t \left|\alpha_s^{(0)}\Big(\frac{1}{\eta}\mathbf{Q}^\eta(s),i\Big) - \alpha_s^{(0)}\Big(\frac{1}{\eta}\mathbf{Q}^\eta(s),i\Big)\right| ds \\
&\quad\leq \int_0^t \left\|\frac{1}{\eta}\alpha_s^\eta - \alpha_s^{(0)}\right\| ds \cdot Z^\eta(t) + \int_0^t \left\|\alpha_s^{(0)}\right\| ds \cdot \sup_{0\leq s\leq t}\left|\frac{1}{\eta}\mathbf{Q}^\eta(s) - \mathbf{Q}^{(0)}(s)\right|
\end{aligned}$$

so by Theorem 2.2, we have

$$\lim_{\eta\to\infty}\sup_{0\leq s\leq t}\left|\int_0^s \alpha_r\Big(\frac{\mathbf{Q}^\eta(r)}{\eta},i\Big)dr - \int_0^s \alpha_r\Big(\mathbf{Q}^{(0)}(r),i\Big)dr\right| = 0. \tag{10.16}$$

From this it follows that

$$\lim_{\eta\to\infty}\sup_{0\leq s\leq t}\left|\sum_{i\in I}\left[B_i^*\Big(\frac{1}{\eta}\int_0^s \alpha_r^\eta\Big(\frac{\mathbf{Q}^\eta(r)}{\eta},i\Big)dr\Big) - B_i^*\Big(\int_0^s \alpha_r^{(0)}\Big(\mathbf{Q}^{(0)}(r),i\Big)dr\Big)\right]\mathbf{v}_i\right| = 0, \tag{10.17}$$

by dominated convergence and the fact that continuous functions are uniformly continuous on compact sets.

If we make the observation that

$$\left\|\Lambda\boldsymbol{\alpha}_t^{(0)}\Big(\mathbf{Q}^{(0)}(t),\cdot\Big)\right\| \leq \left\|\boldsymbol{\alpha}_t^{(0)}\right\|, \tag{10.18}$$

then we have

$$\begin{aligned}
&\overline{\lim_{\eta\to\infty}}\sup_{0\leq s\leq t}\left|\frac{\mathbf{Q}^\eta(s)-\eta\mathbf{Q}^{(0)}(s)}{\sqrt{\eta}} - \mathbf{Q}^{(1)}(s)\right| \qquad (10.19)\\
&\qquad\leq \int_0^t \left\|\boldsymbol{\alpha}_s^{(0)}\right\| \cdot \overline{\lim_{\eta\to\infty}}\sup_{0\leq r\leq s}\left|\frac{\mathbf{Q}^\eta(r)-\eta\mathbf{Q}^{(0)}(r)}{\sqrt{\eta}} - \mathbf{Q}^{(1)}(r)\right| ds.
\end{aligned}$$

Now we apply Gronwall's inequality and obtain

$$\lim_{\eta\to\infty}\sup_{0\leq s\leq t}\left|\frac{\mathbf{Q}^\eta(s)-\eta\mathbf{Q}^{(0)}(s)}{\sqrt{\eta}} - \mathbf{Q}^{(1)}(s)\right| = 0, \tag{10.20}$$

which completes the proof. □

Proof of Theorem 2.4. Given the integral equation (2.29) that $\mathbf{Q}^{(1)}(t)$ solves, we immediately have for the mean vector $\mathsf{E}\left[\mathbf{Q}^{(1)}(t)\right]$

$$\mathsf{E}\left[\mathbf{Q}^{(1)}(t)\right] = \mathsf{E}\left[\mathbf{Q}^{(1)}(0)\right] + \int_0^t \mathsf{E}\left[\Lambda\boldsymbol{\alpha}_s^{(0)}\Big(\mathbf{Q}^{(0)}(s);\mathbf{Q}^{(1)}(s)\Big)\right]ds + \int_0^t \boldsymbol{\alpha}_s^{(1)}\Big(\mathbf{Q}^{(0)}(s)\Big)ds. \tag{10.21}$$

Differentiating (10.21) gives us (2.31).

The solution to the integral equation (2.29) also solves the stochastic differential equation

$$\begin{aligned} d\mathbf{Q}^{(1)}(t) = {} & \Big(\Lambda\boldsymbol{\alpha}_t^{(0)}\Big(\mathbf{Q}^{(0)}(t);\mathbf{Q}^{(1)}(t)\Big) + \boldsymbol{\alpha}_t^{(1)}\Big(\mathbf{Q}^{(0)}(t)\Big)\Big)\,dt \\ & + \sum_{i\in I}\sqrt{\alpha_t^{(0)}\Big(\mathbf{Q}^{(0)}(t),i\Big)}\,\mathbf{v}_i\,dB_i^*(t), \end{aligned} \tag{10.22}$$

see Karatzas and Shreve [8] for more details on stochastic calculus. Using Ito's formula ([8] page 149) we can rigorously show the following result obtained through formal manipulations

$$\begin{aligned} & d\left(\mathbf{Q}^{(1)}(t)^{\mathsf{T}}\cdot\mathbf{Q}^{(1)}(t)\right) \\ & = d\mathbf{Q}^{(1)}(t)^{\mathsf{T}}\cdot\mathbf{Q}^{(1)}(t) + \mathbf{Q}^{(1)}(t)^{\mathsf{T}}\cdot d\mathbf{Q}^{(1)}(t) + d\mathbf{Q}^{(1)}(t)^{\mathsf{T}}\cdot d\mathbf{Q}^{(1)}(t) \\ & = \Big(\Lambda\boldsymbol{\alpha}_t^{(0)}\Big(\mathbf{Q}^{(0)}(t);\mathbf{Q}^{(1)}(t)\Big) + \boldsymbol{\alpha}_t^{(1)}\Big(\mathbf{Q}^{(0)}(t)\Big)\Big)^{\mathsf{T}}\cdot\mathbf{Q}^{(1)}(t)dt \\ & \quad + \sum_{i\in I}\sqrt{\alpha_t^{(0)}\Big(\mathbf{Q}^{(0)}(t),i\Big)}\,\mathbf{v}_i^{\mathsf{T}}\cdot\mathbf{Q}^{(1)}(t)dB_i^*(t) \\ & \quad + \mathbf{Q}^{(1)}(t)^{\mathsf{T}}\cdot\Big(\Lambda\boldsymbol{\alpha}_t^{(0)}\Big(\mathbf{Q}^{(0)}(t);\mathbf{Q}^{(1)}(t)\Big) + \boldsymbol{\alpha}_t^{(1)}\Big(\mathbf{Q}^{(0)}(t)\Big)\Big)\,dt \\ & \quad + \sum_{i\in I}\sqrt{\alpha_t^{(0)}\Big(\mathbf{Q}^{(0)}(t),i\Big)}\,\mathbf{Q}^{(1)}(t)^{\mathsf{T}}\cdot\mathbf{v}_i dB_i^*(t) + \sum_{i\in I}\alpha_t^{(0)}\Big(\mathbf{Q}^{(0)}(t),i\Big)\mathbf{v}_i^{\mathsf{T}}\cdot\mathbf{v}_i dt. \end{aligned}$$

Taking expectations, the Brownian motion differential terms disappear and we get

$$\begin{aligned} & \frac{d}{dt}\mathsf{E}\left[\mathbf{Q}^{(1)}(t)^{\mathsf{T}}\cdot\mathbf{Q}^{(1)}(t)\right] \\ & \qquad = \mathsf{E}\left[\Big(\Lambda\boldsymbol{\alpha}_t^{(0)}\Big(\mathbf{Q}^{(0)}(t);\mathbf{Q}^{(1)}(t)\Big) + \boldsymbol{\alpha}_t^{(1)}\Big(\mathbf{Q}^{(0)}(t)\Big)\Big)^{\mathsf{T}}\cdot\mathbf{Q}^{(1)}(t)\right] \\ & \qquad\quad + \mathsf{E}\left[\mathbf{Q}^{(1)}(t)^{\mathsf{T}}\cdot\Big(\Lambda\boldsymbol{\alpha}_t^{(0)}\Big(\mathbf{Q}^{(0)}(t);\mathbf{Q}^{(1)}(t)\Big) + \boldsymbol{\alpha}_t^{(1)}\Big(\mathbf{Q}^{(0)}(t)\Big)\Big)\right] \\ & \qquad\quad + \sum_{i\in I}\alpha_t^{(0)}\Big(\mathbf{Q}^{(0)}(t),i\Big)\mathbf{v}_i^{\mathsf{T}}\cdot\mathbf{v}_i, \end{aligned} \tag{10.23}$$

for almost all t. Using the derivative of (10.21), we obtain

$$\begin{aligned}
&\frac{d}{dt}\mathsf{E}\left[\mathbf{Q}^{(1)}(t)\right]^{\mathsf{T}}\cdot\mathsf{E}\left[\mathbf{Q}^{(1)}(t)\right] && (10.24)\\
&\qquad= \left(\mathsf{E}\left[\Lambda\boldsymbol{\alpha}_t^{(0)}\left(\mathbf{Q}^{(0)}(t);\mathbf{Q}^{(1)}(t)\right)\right]+\boldsymbol{\alpha}_t^{(1)}\left(\mathbf{Q}^{(0)}(t)\right)\right)^{\mathsf{T}}\cdot\mathsf{E}\left[\mathbf{Q}^{(1)}(t)\right]\\
&\qquad\quad+\mathsf{E}\left[\mathbf{Q}^{(1)}(t)\right]^{\mathsf{T}}\cdot\mathsf{E}\left[\Lambda\boldsymbol{\alpha}_t^{(0)}\left(\mathbf{Q}^{(0)}(t);\mathbf{Q}^{(1)}(t)\right)+\boldsymbol{\alpha}_t^{(1)}\left(\mathbf{Q}^{(0)}(t)\right)\right].
\end{aligned}$$

Subtracting (10.24) from (10.23) gives us (2.32).

If $\Lambda\boldsymbol{\alpha}_t^{(0)}\left(\mathbf{Q}^{(0)}(t);\cdot\right)$ is a linear operator for almost all t, then let $\mathbf{A}_t$ be the matrix that represents its action on $\mathbb{V}$ or

$$\Lambda\boldsymbol{\alpha}_t^{(0)}\left(\mathbf{Q}^{(0)}(t);\mathbf{Q}^{(1)}(t)\right) = \mathbf{Q}^{(1)}(t)\mathbf{A}_t. \tag{10.25}$$

If $|\mathbf{A}_t|$ is the bounded linear operator norm of $\mathbf{A}_t$, then by (2.18) and Theorem 3.1 we have

$$|\mathbf{A}_t| = \left\|\Lambda\boldsymbol{\alpha}_t^{(0)}\left(\mathbf{Q}^{(0)}(t);\cdot\right)\right\| \leq \left\|\boldsymbol{\alpha}_t^{(0)}\right\|, \tag{10.26}$$

and so by Theorem 2.2, $|\mathbf{A}_t|$ is a locally integrable function of t.

We can then rewrite (10.21) as

$$\mathsf{E}\left[\mathbf{Q}^{(1)}(t)\right] = \mathsf{E}\left[\mathbf{Q}^{(1)}(0)\right] + \int_0^t \mathsf{E}\left[\mathbf{Q}^{(1)}(s)\right]\mathbf{A}_s ds + \int_0^t \boldsymbol{\alpha}_s^{(1)}\left(\mathbf{Q}^{(0)}(s)\right)ds \tag{10.27}$$

and observe that by (2.6)

$$\left|\boldsymbol{\alpha}_t^{(1)}\left(\mathbf{Q}^{(0)}(t)\right)\right| \leq \left\|\boldsymbol{\alpha}_t^{(1)}\right\|(1+|\mathbf{Q}^{(0)}(t)|). \tag{10.28}$$

Since we know that $\left\|\boldsymbol{\alpha}_t^{(1)}\right\|$ is a locally integrable function of t by Theorem 2.3, then we need only prove that $\sup_{0\leq s\leq t}\left|\mathbf{Q}^{(0)}(s)\right|$ is finite for all t to show that $\left|\boldsymbol{\alpha}_t^{(1)}\left(\mathbf{Q}^{(0)}(t)\right)\right|$ is a locally integrable function of t. From this and Theorem 11.4, the uniqueness of the solution to the integral equation (10.27) immediately follows.

Using the integral equation for the fluid approximation of Theorem 2.2, we have

$$\begin{aligned}
1+\left|\mathbf{Q}^{(0)}(t)\right| &= 1+\left|\mathbf{Q}^{(0)}(0)+\int_0^t \boldsymbol{\alpha}_s^{(0)}\left(\mathbf{Q}^{(0)}(s)\right)ds\right| && (10.29)\\
&\leq 1+\left|\mathbf{Q}^{(0)}(0)\right|+\int_0^t\left|\boldsymbol{\alpha}_s^{(0)}\left(\mathbf{Q}^{(0)}(s)\right)\right|ds && (10.30)\\
&\leq 1+\left|\mathbf{Q}^{(0)}(0)\right|+\int_0^t\left\|\boldsymbol{\alpha}_s^{(0)}\right\|\left(1+\left|\mathbf{Q}^{(0)}(s)\right|\right)ds, && (10.31)
\end{aligned}$$

where the second step is an application of (2.6). Applying Gronwall's inequality, we have

$$1+\sup_{0\le s\le t}\left|\mathbf{Q}^{(0)}(s)\right| \le \left(1+\left|\mathbf{Q}^{(0)}(0)\right|\right)\cdot\exp\left(\int_0^t \left\|\boldsymbol{\alpha}_s^{(0)}\right\| ds\right) \tag{10.32}$$

and so (10.28) for all t, gives us

$$\begin{aligned}\int_0^t \left|\boldsymbol{\alpha}_s^{(1)}\left(\mathbf{Q}^{(0)}(s)\right)\right| ds &\le \int_0^t \left|\boldsymbol{\alpha}_s^{(1)}\right|\cdot\left(1+\left|\mathbf{Q}^{(0)}(s)\right|\right) ds \\ &\le \left(1+\sup_{0\le s\le t}\left|\mathbf{Q}^{(0)}(s)\right|\right)\cdot\int_0^t \left\|\boldsymbol{\alpha}_s^{(1)}\right\| ds \\ &\le \left(1+\left|\mathbf{Q}^{(0)}(0)\right|\right)\cdot\exp\left(\int_0^t \left\|\boldsymbol{\alpha}_s^{(0)}\right\| ds\right)\cdot\int_0^t \left\|\boldsymbol{\alpha}_s^{(1)}\right\| ds \\ &< \infty.\end{aligned}$$

Finally, observe that

$$\mathsf{Cov}\left[\mathbf{Q}^{(1)}(t), \Lambda\boldsymbol{\alpha}_t^{(1)}\left(\mathbf{Q}^{(0)}(t);\mathbf{Q}^{(1)}(t)\right)\right] = \mathsf{Cov}\left[\mathbf{Q}^{(1)}(t),\mathbf{Q}^{(1)}(t)\right]\mathbf{A}_t \tag{10.33}$$

for almost all t, and so the integral equation for the covariance matrix is

$$\begin{aligned}\mathsf{Cov}\left[\mathbf{Q}^{(1)}(t),\mathbf{Q}^{(1)}(t)\right] = \mathsf{Cov}\left[\mathbf{Q}^{(1)}(0),\mathbf{Q}^{(1)}(0)\right] &+ \int_0^t \left\{\mathsf{Cov}\left[\mathbf{Q}^{(1)}(s),\mathbf{Q}^{(1)}(s)\right]\mathbf{A}_s\right\} ds \\ &+ \int_0^t \sum_{i\in I} \alpha_s^{(0)}\left(\mathbf{Q}^{(0)}(s), i\right)\mathbf{v}_i^{\mathsf{T}}\cdot\mathbf{v}_i ds,\end{aligned} \tag{10.34}$$

Uniqueness then follows by a similar argument as for the mean vector. □

11. Appendix: Ordinary Differential Equations

The existence and uniqueness of the fluid and diffusion approximations for our service network processes rely heavily on the theory of non-linear ordinary differential equations. In this section, we provide a self-contained summary of these results. More details can be found in books by Hale [3] or Hochstadt [5].

Lemma 11.1 (Gronwall, c.f. [2], [3], or [8]). Let x, y, and z be measurable, non-negative functions on the reals. If y is bounded and z is integrable on $[0,T]$ and for all $0\le t\le T$,

$$x(t)\le z(t)+\int_0^t x(s)y(s)\,ds, \tag{11.1}$$

then

$$x(t) \leq z(t) + \int_0^t z(s)y(s)\exp\Bigl(\int_s^t y(r)\,dr\Bigr)ds. \tag{11.2}$$

Proof. If we multiply both sides of (11.1) by y, then

$$x(t)y(t) - \left(\int_0^t x(s)y(s)ds\right) y(t) \leq z(t)y(t). \tag{11.3}$$

Now if we multiply both sides by $\exp\left(-\int_0^t y(s)ds\right)$ we obtain

$$\begin{aligned} x(t)y(t)\exp\left(-\int_0^t y(s)ds\right) - \left(\int_0^t x(s)y(s)ds\right) y(t)\exp\left(-\int_0^t y(s)ds\right) \\ \leq z(t)y(t)\exp\left(-\int_0^t y(s)ds\right). \end{aligned} \tag{11.4}$$

Simplifying the left side of the inequality as the derivative of a product of absolutely continuous functions, we have

$$\frac{d}{dt}\left[\left(\int_0^t x(s)y(s)ds\right)\exp\left(-\int_0^t y(s)ds\right)\right] \leq z(t)y(t)\exp\left(-\int_0^t y(s)ds\right). \tag{11.5}$$

Integrating both sides yields

$$\left(\int_0^t x(s)y(s)ds\right)\exp\left(-\int_0^t y(s)ds\right) \leq \int_0^t z(s)y(s)\exp\left(-\int_0^s y(r)dr\right)ds \tag{11.6}$$

which is equivalent to

$$\int_0^t x(s)y(s)ds \leq \int_0^t z(s)y(s)\exp\left(\int_s^t y(r)dr\right)ds. \tag{11.7}$$

Combining (11.1) to (11.7) gives us (11.2). □

Corollary 11.2. If x, y, and z satisfy the same hypotheses as above, then

$$\sup_{0\leq t\leq T} x(t) \leq \sup_{0\leq t\leq T} z(t)\cdot\exp\Bigl(\int_0^T y(t)\,dt\Bigr). \tag{11.8}$$

Proof. Using (11.2), we have for all $0 \leq t \leq T$,

$$x(t) \leq \sup_{0\leq t\leq T} z(t) + \int_0^T z(t)y(t)\exp\Bigl(\int_t^T y(s)\,ds\Bigr)dt, \tag{11.9}$$

which in turn gives us

$$\sup_{0\leq t\leq T} x(t) \leq \sup_{0\leq t\leq T} z(t) + \int_0^T z(t)y(t)\exp\Bigl(\int_t^T y(s)\,ds\Bigr)dt \tag{11.10}$$

$$\leq \sup_{0\leq t\leq T} z(t) + \sup_{0\leq t\leq T} z(t) \int_0^T y(t) \exp\Big(\int_t^T y(s)\, ds\Big) dt \quad (11.11)$$

$$\leq \sup_{0\leq t\leq T} z(t) + \sup_{0\leq t\leq T} z(t) \left(\exp\Big(\int_0^T y(t)\, dt\Big) - 1\right) \quad (11.12)$$

$$\leq \sup_{0\leq t\leq T} z(t) \exp\Big(\int_0^T y(t)\, dt\Big), \quad (11.13)$$

and this completes the proof. □

Lemma 11.3. Let $\left\{ x^{(n)} \,\middle|\, n \geq 0 \right\}$ be a sequence of bounded, non-negative functions on the interval $[0,T]$ and let y be a non-negative, integrable function on $[0,T]$. If we have for all $n \geq 0$ and $0 \leq t \leq T$

$$x^{(n+1)}(t) \leq \int_0^t x^{(n)}(s) y(s)\, ds, \quad (11.14)$$

then we have for all $0 \leq t \leq T$,

$$x^{(n)}(t) \leq \frac{1}{n!} \left(\int_0^t y(s)\, ds\right)^n \sup_{0\leq s\leq t} x^{(0)}(s). \quad (11.15)$$

Proof. Since $x^{(0)}(t) \leq \sup_{0\leq s\leq t} x^{(0)}(s)$, then (11.15) follows from induction on n. □

Now we state and prove the result that is referred to in Hochstadt (see [5] page 204) as the fundamental existence and uniqueness theorem for nonlinear ordinary differential equations.

Theorem 11.4. If for all t in $[0,T]$, the function $\mathbf{f}_t : \mathbb{V} \to \mathbb{V}$ is Lipschitz such that $\|\mathbf{f}_t\|$ is integrable over $[0,T]$, then the integral equation

$$\mathbf{X}(t) = \int_0^t \mathbf{f}_s\Big(\mathbf{X}(s)\Big) ds \quad (11.16)$$

has a unique bounded solution $\mathbf{X} : [0,T] \to \mathbb{V}$.

Proof. For all $n \geq 0$, let

$$\mathbf{X}^{(n+1)}(t) \equiv \int_0^t \mathbf{f}_s\Big(\mathbf{X}^{(n)}(s)\Big) ds \quad (11.17)$$

and

$$x^{(n)}(t) \equiv \sup_{0\leq s\leq t} \left| \mathbf{X}^{(n+1)}(s) - \mathbf{X}^{(n)}(s) \right|. \quad (11.18)$$

This gives us

$$
\begin{aligned}
x^{(n+1)}(t) &\leq \sup_{0\leq s\leq t} \left|\mathbf{X}^{(n+2)}(s) - \mathbf{X}^{(n+1)}(s)\right| \\
&\leq \sup_{0\leq s\leq t} \left|\int_0^s \left[\mathbf{f}_r\left(\mathbf{X}^{(n+1)}(r)\right) - \mathbf{f}_r\left(\mathbf{X}^{(n)}(r)\right)\right] dr\right| \\
&\leq \int_0^t \left|\mathbf{f}_s\left(\mathbf{X}^{(n+1)}(s)\right) - \mathbf{f}_s\left(\mathbf{X}^{(n)}(s)\right)\right| ds \\
&\leq \int_0^t \|\mathbf{f}_s\| \left|\mathbf{X}^{(n+1)}(s) - \mathbf{X}^{(n)}(s)\right| ds \\
&\leq \int_0^t \|\mathbf{f}_s\| \, x^{(n)}(s) ds
\end{aligned}
$$

which by the previous lemma means that

$$
x^{(n)}(t) \leq \frac{1}{n!}\left(\int_0^t \|\mathbf{f}_s\| \, ds\right)^n \sup_{0\leq s\leq t} x^{(0)}(s). \tag{11.19}
$$

For all $m > n$, we then have

$$
\begin{aligned}
&\sup_{0\leq t\leq T} \left|\mathbf{X}^{(m)}(t) - \mathbf{X}^{(n)}(t)\right| \\
&\quad\leq \sum_{i=n}^{m-1} \sup_{0\leq t\leq T} \left|\mathbf{X}^{(i+1)}(t) - \mathbf{X}^{(i)}(t)\right| \qquad (11.20) \\
&\quad\leq \sum_{i=n}^{m-1} \frac{1}{i!}\left(\int_0^T \|\mathbf{f}_t\| \, dt\right)^i \cdot \sup_{0\leq t\leq T} \left|\mathbf{X}^{(1)}(t) - \mathbf{X}^{(0)}(t)\right| \qquad (11.21) \\
&\quad= \left(\sum_{i=n}^{\infty} \frac{1}{i!}\left(\int_0^T \|\mathbf{f}_t\| \, dt\right)^i - \sum_{i=m}^{\infty} \frac{1}{i!}\left(\int_0^T \|\mathbf{f}_t\| \, dt\right)^i\right) \cdot \sup_{0\leq t\leq T} \left|\mathbf{X}^{(1)}(t) - \mathbf{X}^{(0)}(t)\right|,
\end{aligned}
$$

which gives us

$$
\lim_{m,n\to\infty} \sup_{0\leq t\leq T} \left|\mathbf{X}^{(m)}(t) - \mathbf{X}^{(n)}(t)\right| = 0, \tag{11.22}
$$

and makes $\left\{ \mathbf{X}^{(n)}(\cdot) \,\middle|\, n \geq 0 \right\}$ a Cauchy sequence.

Now let $\mathbf{X}(\cdot) \equiv \lim_{n\to\infty} \mathbf{X}^{(n)}(\cdot)$. It follows that $\mathbf{X}$ is a bounded solution to the integral equation (11.16). Moreover, it is unique. If $\hat{\mathbf{X}}(\cdot)$ is some other bounded solution to (11.16), we then have

$$
\left|\mathbf{X}(t) - \hat{\mathbf{X}}(t)\right| = \left|\int_0^t \mathbf{f}_s\left(\mathbf{X}(s)\right) - \mathbf{f}_s\left(\hat{\mathbf{X}}(s)\right) ds\right| \leq \int_0^t \|\mathbf{f}_s\| \left|\mathbf{X}(s) - \hat{\mathbf{X}}(s)\right| ds \tag{11.23}
$$

and by Gronwall's inequality, $\mathbf{X} = \hat{\mathbf{X}}$. □

12. Appendix: Scalable Lipschitz Derivatives

In this section, we provide all the proofs for the theorems about scalable Lipschitz differentiability stated in Section 3.

Proof of Theorem 3.1. Let $\hat{\Lambda}\mathbf{f_x}$ be another scalable Lipschitz function such that

$$\lim_{\mathbf{y}\to 0} \frac{\left|\mathbf{f}(\mathbf{x}+\mathbf{y}) - \mathbf{f}(\mathbf{x}) - \hat{\Lambda}\mathbf{f_x}(\mathbf{y})\right|}{|\mathbf{y}|} = 0. \tag{12.1}$$

We then have for all $\mathbf{y}$ that

$$\begin{aligned}
\left|\Lambda\mathbf{f_x}(\mathbf{y}) - \hat{\Lambda}\mathbf{f_x}(\mathbf{y})\right| &= \lim_{\lambda\downarrow 0} \frac{\left|\Lambda\mathbf{f_x}(\lambda\mathbf{y}) - \hat{\Lambda}\mathbf{f_x}(\lambda\mathbf{y})\right|}{\lambda} \\
&\leq \lim_{\lambda\downarrow 0} \frac{|\mathbf{f}(\mathbf{x}+\lambda\mathbf{y}) - \mathbf{f}(\mathbf{x}) - \Lambda\mathbf{f_x}(\lambda\mathbf{y})|}{\lambda} \\
&\quad + \lim_{\lambda\downarrow 0} \frac{\left|\mathbf{f}(\mathbf{x}+\lambda\mathbf{y}) - \mathbf{f}(\mathbf{x}) - \hat{\Lambda}\mathbf{f_x}(\lambda\mathbf{y})\right|}{\lambda} \\
&= 0,
\end{aligned}$$

and so $\Lambda\mathbf{f_x} = \hat{\Lambda}\mathbf{f_x}$.

To show that the property of scalable Lipschitz differentiability is closed under composition, we must show that

$$\lim_{\mathbf{y}\to 0} \frac{\left|\mathbf{g}\circ\mathbf{f}(\mathbf{x}+\mathbf{y}) - \mathbf{g}\circ\mathbf{f}(\mathbf{x}) - \left(\Lambda\mathbf{g}_{\mathbf{f}(\mathbf{x})}\circ\Lambda\mathbf{f_x}\right)(\mathbf{y})\right|}{|\mathbf{y}|} = 0 \tag{12.2}$$

We have

$$\begin{aligned}
&\frac{\left|\mathbf{g}\circ\mathbf{f}(\mathbf{x}+\mathbf{y}) - \mathbf{g}\circ\mathbf{f}(\mathbf{x}) - \left(\Lambda\mathbf{g}_{\mathbf{f}(\mathbf{x})}\circ\Lambda\mathbf{f_x}\right)(\mathbf{y})\right|}{|\mathbf{y}|} \\
&\leq \frac{\left|\mathbf{g}\Big(\mathbf{f}(\mathbf{x}+\mathbf{y}) - \mathbf{f}(\mathbf{x}) + \mathbf{f}(\mathbf{x})\Big) - \mathbf{g}\Big(\mathbf{f}(\mathbf{x})\Big) - \Lambda\mathbf{g}_{\mathbf{f}(\mathbf{x})}\Big(\mathbf{f}(\mathbf{x}+\mathbf{y}) - \mathbf{f}(\mathbf{x})\Big)\right|}{|\mathbf{y}|} \\
&\quad + \frac{\left|\left(\Lambda\mathbf{g}_{\mathbf{f}(\mathbf{x})}\circ\Lambda\mathbf{f_x}\right)(\mathbf{y}) - \Lambda\mathbf{g}_{\mathbf{f}(\mathbf{x})}\Big(\mathbf{f}(\mathbf{x}+\mathbf{y}) - \mathbf{f}(\mathbf{x})\Big)\right|}{|\mathbf{y}|} && (12.3) \\
&\leq 1_{\{\mathbf{f}(\mathbf{x}+\mathbf{y})\neq\mathbf{f}(\mathbf{x})\}} \frac{\left|\mathbf{g}\Big(\mathbf{f}(\mathbf{x}+\mathbf{y}) - \mathbf{f}(\mathbf{x}) + \mathbf{f}(\mathbf{x})\Big) - \mathbf{g}\Big(\mathbf{f}(\mathbf{x})\Big) - \Lambda\mathbf{g}_{\mathbf{f}(\mathbf{x})}\Big(\mathbf{f}(\mathbf{x}+\mathbf{y}) - \mathbf{f}(\mathbf{x})\Big)\right|}{|\mathbf{y}|} \\
&\quad + \left\|\Lambda\mathbf{g}_{\mathbf{f}(\mathbf{x})}\right\| \frac{|\mathbf{f}(\mathbf{x}+\mathbf{y}) - \mathbf{f}(\mathbf{x}) - \Lambda\mathbf{f_x}(\mathbf{y})|}{|\mathbf{y}|} && (12.4)
\end{aligned}$$

$$
\begin{aligned}
&\le 1_{\{\mathbf{f}(\mathbf{x}+\mathbf{y})\neq\mathbf{f}(\mathbf{x})\}} \frac{\left|\mathbf{g}\Big(\mathbf{f}(\mathbf{x}+\mathbf{y})-\mathbf{f}(\mathbf{x})+\mathbf{f}(\mathbf{x})\Big)-\mathbf{g}\Big(\mathbf{f}(\mathbf{x})\Big)-\Lambda\mathbf{g}_{\mathbf{f}(\mathbf{x})}\Big(\mathbf{f}(\mathbf{x}+\mathbf{y})-\mathbf{f}(\mathbf{x})\Big)\right|}{|\mathbf{f}(\mathbf{x}+\mathbf{y})-\mathbf{f}(\mathbf{x})|} \\
&\quad\times \frac{|\mathbf{f}(\mathbf{x}+\mathbf{y})-\mathbf{f}(\mathbf{x})|}{|\mathbf{y}|} + \left\|\Lambda\mathbf{g}_{\mathbf{f}(\mathbf{x})}\right\| \frac{|\mathbf{f}(\mathbf{x}+\mathbf{y})-\mathbf{f}(\mathbf{x})-\Lambda\mathbf{f}_{\mathbf{x}}(\mathbf{y})|}{|\mathbf{y}|} \qquad (12.5)\\
&\le 1_{\{\mathbf{f}(\mathbf{x}+\mathbf{y})\neq\mathbf{f}(\mathbf{x})\}} \frac{\left|\mathbf{g}\Big(\mathbf{f}(\mathbf{x}+\mathbf{y})-\mathbf{f}(\mathbf{x})+\mathbf{f}(\mathbf{x})\Big)-\mathbf{g}\Big(\mathbf{f}(\mathbf{x})\Big)-\Lambda\mathbf{g}_{\mathbf{f}(\mathbf{x})}\Big(\mathbf{f}(\mathbf{x}+\mathbf{y})-\mathbf{f}(\mathbf{x})\Big)\right|}{|\mathbf{f}(\mathbf{x}+\mathbf{y})-\mathbf{f}(\mathbf{x})|} \\
&\quad\times \frac{|\mathbf{f}(\mathbf{x}+\mathbf{y})-\mathbf{f}(\mathbf{x})-\Lambda\mathbf{f}_{\mathbf{x}}(\mathbf{y})|}{|\mathbf{y}|} \\
&\quad+1_{\{\mathbf{f}(\mathbf{x}+\mathbf{y})\neq\mathbf{f}(\mathbf{x})\}} \frac{|\Lambda\mathbf{f}_{\mathbf{x}}(\mathbf{y})|}{|\mathbf{y}|} \\
&\quad\times \frac{\left|\mathbf{g}\Big(\mathbf{f}(\mathbf{x}+\mathbf{y})-\mathbf{f}(\mathbf{x})+\mathbf{f}(\mathbf{x})\Big)-\mathbf{g}\Big(\mathbf{f}(\mathbf{x})\Big)-\Lambda\mathbf{g}_{\mathbf{f}(\mathbf{x})}\Big(\mathbf{f}(\mathbf{x}+\mathbf{y})-\mathbf{f}(\mathbf{x})\Big)\right|}{|\mathbf{f}(\mathbf{x}+\mathbf{y})-\mathbf{f}(\mathbf{x})|} \\
&\quad+ \left\|\Lambda\mathbf{g}_{\mathbf{f}(\mathbf{x})}\right\| \frac{|\mathbf{f}(\mathbf{x}+\mathbf{y})-\mathbf{f}(\mathbf{x})-\Lambda\mathbf{f}_{\mathbf{x}}(\mathbf{y})|}{|\mathbf{y}|} \qquad (12.6)\\
&\le 1_{\{\mathbf{f}(\mathbf{x}+\mathbf{y})\neq\mathbf{f}(\mathbf{x})\}} \frac{\left|\mathbf{g}\Big(\mathbf{f}(\mathbf{x}+\mathbf{y})-\mathbf{f}(\mathbf{x})+\mathbf{f}(\mathbf{x})\Big)-\mathbf{g}\Big(\mathbf{f}(\mathbf{x})\Big)-\Lambda\mathbf{g}_{\mathbf{f}(\mathbf{x})}\Big(\mathbf{f}(\mathbf{x}+\mathbf{y})-\mathbf{f}(\mathbf{x})\Big)\right|}{|\mathbf{f}(\mathbf{x}+\mathbf{y})-\mathbf{f}(\mathbf{x})|} \\
&\quad\times \frac{|\mathbf{f}(\mathbf{x}+\mathbf{y})-\mathbf{f}(\mathbf{x})-\Lambda\mathbf{f}_{\mathbf{x}}(\mathbf{y})|}{|\mathbf{y}|} \\
&\quad+1_{\{\mathbf{f}(\mathbf{x}+\mathbf{y})\neq\mathbf{f}(\mathbf{x})\}} \|\Lambda\mathbf{f}_{\mathbf{x}}\| \\
&\quad\times \frac{\left|\mathbf{g}\Big(\mathbf{f}(\mathbf{x}+\mathbf{y})-\mathbf{f}(\mathbf{x})+\mathbf{f}(\mathbf{x})\Big)-\mathbf{g}\Big(\mathbf{f}(\mathbf{x})\Big)-\Lambda\mathbf{g}_{\mathbf{f}(\mathbf{x})}\Big(\mathbf{f}(\mathbf{x}+\mathbf{y})-\mathbf{f}(\mathbf{x})\Big)\right|}{|\mathbf{f}(\mathbf{x}+\mathbf{y})-\mathbf{f}(\mathbf{x})|} \\
&\quad+ \left\|\Lambda\mathbf{g}_{\mathbf{f}(\mathbf{x})}\right\| \frac{|\mathbf{f}(\mathbf{x}+\mathbf{y})-\mathbf{f}(\mathbf{x})-\Lambda\mathbf{f}_{\mathbf{x}}(\mathbf{y})|}{|\mathbf{y}|}, \qquad (12.7)
\end{aligned}
$$

and this proves (12.2).

Finally, if $\mathcal{O}$ is an open subset of $\mathbb{V}_1$ and $\mathbf{x} \in \mathcal{O}$, then for all $\mathbf{y}_1$ and $\mathbf{y}_2$ in $\mathbb{V}_1$, there exists a positive scalar λ_0 such that

$$
\mathbf{x} + \lambda\mathbf{y}_1 \in \mathcal{O} \quad \text{and} \quad \mathbf{x} + \lambda\mathbf{y}_2 \in \mathcal{O} \qquad (12.8)
$$

for all $0 \le \lambda \le \lambda_0$. Since it follows from the definition of scalable Lipschitz differentiability at $\mathbf{x}$ that for all $\mathbf{y} \in \mathbb{V}_1$

$$
\lim_{\lambda\downarrow 0} \frac{\mathbf{f}(\mathbf{x}+\lambda\mathbf{y})-\mathbf{f}(\mathbf{x})}{\lambda} = \Lambda_{\mathbf{x}}(\mathbf{y}), \qquad (12.9)
$$

we then have

$$|\Lambda \mathbf{f}_{\mathbf{x}}(\mathbf{y}_1) - \Lambda \mathbf{f}_{\mathbf{x}}(\mathbf{y}_2)| = \lim_{\lambda \downarrow 0} \frac{|\mathbf{f}(\mathbf{x}+\lambda \mathbf{y}_1) - \mathbf{f}(\mathbf{x}+\lambda \mathbf{y}_2)|}{\lambda} \leq \|\mathbf{f}\|_{\mathcal{O}} \cdot |\mathbf{y}_1 - \mathbf{y}_2|, \quad (12.10)$$

which completes the proof. □

Proof of Theorem 3.2. Combining the uniqueness result (statement 1) of Theorem 3.1 with the fact that all linear operators are scalable and Lipschitz, we see that any differentiable function with Jacobian matrix $D\mathbf{f}(\mathbf{x})$ is scalable Lipschitz differentiable with

$$\Lambda \mathbf{f}_{\mathbf{x}}(\mathbf{y}) = \mathbf{y} \cdot D\mathbf{f}(\mathbf{x}). \quad (12.11)$$

Define $D\mathbf{f}(\mathbf{x};\mathbf{y})$ to be the *radial derivative* of $\mathbf{f}$ at $\mathbf{x}$ with respect to $\mathbf{y}$, namely

$$D\mathbf{f}(\mathbf{x};\mathbf{y}) = \lim_{\lambda \downarrow 0} \frac{\mathbf{f}(\mathbf{x}+\lambda \mathbf{y}) - \mathbf{f}(\mathbf{x})}{\lambda}. \quad (12.12)$$

It follows immediately that $D\mathbf{f}(\mathbf{x};\lambda\mathbf{y}) = \lambda D\mathbf{f}(\mathbf{x};\mathbf{y})$ for all $\lambda \geq 0$. To show that $D\mathbf{f}(\mathbf{x};\cdot)$ is Lipschitz, we use the same proof as for Theorem 3.1 to derive

$$|D\mathbf{f}(\mathbf{x};\mathbf{y}_1) - D\mathbf{f}(\mathbf{x};\mathbf{y}_2)| \leq \|\mathbf{f}\|_{\mathcal{O}} \cdot |\mathbf{y}_1 - \mathbf{y}_2| \quad (12.13)$$

Finally, we show that $D\mathbf{f}(\mathbf{x};\mathbf{y})$ satisfies (2.20), hence it is the scalable Lipschitz derivative of $\mathbf{f}$ at $\mathbf{x}$. We do this by observing that the surface of the unit ball defined by some norm $|\cdot|$ for $\mathbb{R}^m$ is compact. It follows that any sequence $\{\, \mathbf{y}_n \mid n \geq 0 \,\}$ in $\mathbb{R}^m$ converging to zero must have some subsequence $\{\, \mathbf{z}_n \mid n \geq 0 \,\}$ converging to zero with

$$\lim_{n \to \infty} \frac{\mathbf{z}_n}{|\mathbf{z}_n|} = \hat{\mathbf{z}}, \quad (12.14)$$

where $|\hat{\mathbf{z}}| = 1$. Assuming that $\mathbf{z}_n \neq 0$ for all $n \geq 0$, if we define $\hat{\mathbf{z}}_n \equiv \mathbf{z}_n / |\mathbf{z}_n|$, then

$$\begin{aligned}
&\lim_{n \to \infty} \frac{|\mathbf{f}(\mathbf{x}+\mathbf{z}_n) - \mathbf{f}(\mathbf{x}) - D\mathbf{f}(\mathbf{x};\mathbf{z}_n)|}{|\mathbf{z}_n|} \\
&\quad \leq \lim_{n \to \infty} \frac{|\mathbf{f}(\mathbf{x}+\mathbf{z}_n) - \mathbf{f}(\mathbf{x}+|\mathbf{z}_n|\hat{\mathbf{z}})|}{|\mathbf{z}_n|} + \lim_{n \to \infty} \frac{|D\mathbf{f}(\mathbf{x};\mathbf{z}_n) - D\mathbf{f}(\mathbf{x};|\mathbf{z}_n|\hat{\mathbf{z}})|}{|\mathbf{z}_n|} \\
&\qquad + \lim_{n \to \infty} \frac{|\mathbf{f}(\mathbf{x}+|\mathbf{z}_n|\hat{\mathbf{z}}) - \mathbf{f}(\mathbf{x}) - D\mathbf{f}(\mathbf{x};|\mathbf{z}_n|\hat{\mathbf{z}})|}{|\mathbf{z}_n|} \\
&\quad \leq 2\,\|\mathbf{f}\|_{\mathcal{O}} \cdot \lim_{n \to \infty} |\hat{\mathbf{z}}_n - \hat{\mathbf{z}}| + \lim_{n \to \infty} \left| \frac{\mathbf{f}(\mathbf{x}+|\mathbf{z}_n|\hat{\mathbf{z}}) - \mathbf{f}(\mathbf{x})}{|\mathbf{z}_n|} - D\mathbf{f}(\mathbf{x};\hat{\mathbf{z}}) \right| \\
&\quad = 0.
\end{aligned}$$

Since $\hat{\mathbf{z}}$ is arbitrary, then (2.20), with $D\mathbf{f}(\mathbf{x};\mathbf{y})$ replacing $\Lambda\mathbf{f}(\mathbf{x};\mathbf{y})$, must hold for all sequences that converge to zero, which proves statement 2 of the theorem. □

Proof of Theorem 3.3. We write out the proof for the case for the maximum of two functions. All of the proofs for the other statements in this theorem are done in a similar manner, applying the composition formula for scalable Lipschitz derivatives in statement 2 of Theorem 3.1, as well as statement 1 of Theorem 3.2 when it applies.

Let $m : \mathbb{R}^2 \to \mathbb{R}$ equal the maximum function, so that $m(\mathbf{x}) = x_1 \vee x_2$ where $\mathbf{x} = (x_1, x_2)$. We then have

$$\begin{aligned}
&\left|m(\mathbf{x}+\mathbf{y}) - m(\mathbf{x}) - 1_{\{x_1>x_2\}}y_1 - 1_{\{x_1<x_2\}}y_2 - 1_{\{x_1=x_2\}}y_1 \vee y_2\right| \\
&= \left|(x_1+y_1)\vee(x_2+y_2) - x_1 \vee x_2 - 1_{\{x_1>x_2\}}y_1 - 1_{\{x_1<x_2\}}y_2 - 1_{\{x_1=x_2\}}y_1 \vee y_2\right| \\
&= 1_{\{x_1>x_2\}}\left|(x_1+y_1)\vee(x_2+y_2) - (x_1+y_1)\right| \\
&\quad +1_{\{x_1<x_2\}}\left|(x_1+y_1)\vee(x_2+y_2) - (x_2+y_2)\right| \\
&\quad +1_{\{x_1=x_2\}}\left|(x_1+y_1)\vee(x_1+y_2) - \Big(x_1 + (y_1 \vee y_2)\Big)\right| \\
&= 1_{\{x_1>x_2\}}\left|(x_1+y_1)\vee(x_2+y_2) - (x_1+y_1)\right| \\
&\quad +1_{\{x_1<x_2\}}\left|(x_1+y_1)\vee(x_2+y_2) - (x_2+y_2)\right|.
\end{aligned}$$

If $x_1 > x_2$, then we can find sufficiently small y_1 and y_2 such that $x_1+y_1 > x_2+y_2$. Since a similar argument can be made for $x_1 < x_2$, we see that

$$\lim_{y_1,y_2\to 0}\left|m(\mathbf{x}+\mathbf{y}) - m(\mathbf{x}) - y_1\,1_{\{x_1>x_2\}} - y_2\,1_{\{x_1<x_2\}} - y_1 \vee y_2\,1_{\{x_1=x_2\}}\right| = 0 \tag{12.15}$$

and this expression equals zero when y_1 and y_2 are small but non-zero. This means that

$$\lim_{|\mathbf{y}|\to\mathbf{0}} \frac{\left|m(\mathbf{x}+\mathbf{y}) - m(\mathbf{x}) - y_1\,1_{\{x_1>x_2\}} - y_2\,1_{\{x_1<x_2\}} - y_1 \vee y_2\,1_{\{x_1=x_2\}}\right|}{|\mathbf{y}|} = 0, \tag{12.16}$$

which makes m scalable Lipschitz differentiable. By composition, $f \vee g$ is scalable Lipschitz differentiable at $\mathbf{x}$ if f and g are and the composition formula gives us the desired formula. □

Acknowledgements

We would like to thank the two referees for their many suggestions that led to significant improvements in the current version of this paper.

References

[1] F.H. Clarke. *Optimization and Nonsmooth Analysis.* SIAM, 1990 (originally published 1983).

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

[3] J. Hale. *Ordinary Differential Equations.* J. Wiley & Sons/Interscience New York, 1969.

[4] S. Halfin and W. Whitt. Heavy-traffic limits theorem for queues with many exponential servers. *Operations Research*, 29:567–588, 1981.

[5] Hochstadt, H. *Differential Equations: A Modern Approach.* Dover Publications, Inc. New York, 1963.

[6] D. L. Iglehart. Limit diffusion approximations for the many server queue and the repairman problem. *Journal of Applied Probability*, 2:429–441, 1965.

[7] J.R. Jackson. Jobshop-like queueing systems. *Management Science*, 10(1):131–142, 1963.

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

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

[10] T. G. Kurtz. Representation of Markov processes as multiparameter time changes. *The Annals of Probability*, 8(4):682–715, 1980.

[11] A. Mandelbaum and W. A. Massey. Strong approximations for time–dependent queues. *Mathematics of Operations Research*, 20(1):33–64, 1995.

[12] A. Mandelbaum, W. A. Massey, G. Pats and K. Ramanan. Approximations for time–dependent networks. In preparation.

[13] A. Mandelbaum and G. Pats. State–dependent queues: Approximation and applications. In F. Kelly and R.J. Williams, editors, *IMA Volumes in Mathematics and Its Applications*, volume 71, pages 239–282. Springer-Verlag, 1995.

[14] A. Mandelbaum and G. Pats. State-dependent stochastic networks, Part I: Approximations and applications with continuous diffusion limits. *Annals of Applied Probability*, 8(2):569–646, 1998.

[15] W.A. Massey. Asymptotic analysis of the time dependent M/M/1 queue, *Math. of Op. Res.,* **10**, 305–327, 1985.

[16] W.A. Massey. *Nonstationary Queues*, Ph. D. Thesis, Stanford University, 1981.

[17] W. A. Massey and W. Whitt. Networks of infinite–server queues with non-stationary Poisson input. *Queueing Systems*, 13:183–250, 1993.

[18] W. A. Massey and W. Whitt. Uniform acceleration expansions for Markov chains with time-varying rates. *Annals of Applied Probability*, to appear in 1998.

[19] G. F. Newell. *Approximate Stochastic Behavior of n-Server Service Systems with Large n.* Springer-Verlag, Berlin, 1973.

[20] G. Pats. *State–Dependent Queueing Networks: Approximations and Applications.* PhD thesis, Faculty of Industrial Engineering and Management, Technion, 1994.

[21] C. R. Rao. *Linear Statistical Inference and its Applications.* J. Wiley, New York, second edition, 1973.

[22] R.T. Rockafellar. *Convex Analysis.* Princeton University Press, 1970 (paperback edition 1997).

[23] W. Whitt. On the heavy-traffic limit theorem for $GI/G/\infty$ queues. *Advances in Applied Probability.* Vol. 14, No. 1, pp. 171–190, 1982.