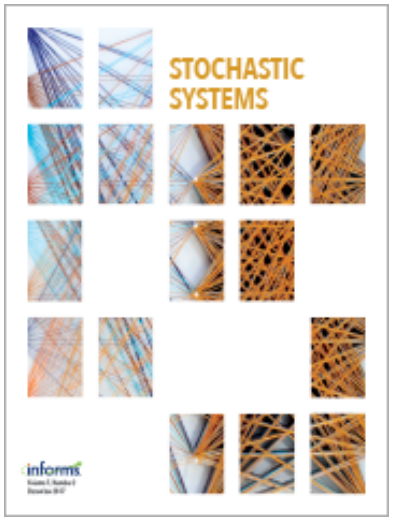
This article was downloaded by: [216.73.217.10] On: 05 November 2026, At: 13:43
Publisher: Institute for Operations Research and the Management Sciences (INFORMS)
INFORMS is located in Maryland, USA



Stochastic Systems

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

Heavy Traffic Approximation for the Stationary Distribution of a Generalized Jackson Network: The BAR Approach

Anton Braverman, J. G. Dai, Masakiyo Miyazawa

To cite this article:
Anton Braverman, J. G. Dai, Masakiyo Miyazawa (2017) Heavy Traffic Approximation for the Stationary Distribution of a Generalized Jackson Network: The BAR Approach. Stochastic Systems 7(1):143-196. https://doi.org/10.1287/15-SSY199

This work is licensed under a Creative Commons Attribution 4.0 International License. You are free to copy, distribute, transmit and adapt this work, but you must attribute this work as "*Stochastic Systems*. Copyright 2017 The Author(s). https://doi.org/10.1287/15-SSY199, used under a Creative Commons Attribution License: http://creativecommons.org/licenses/by/4.0/."

Copyright © 2017, The author(s)

Please scroll down for article—it is on subsequent pages

With 12,500 members from nearly 90 countries, INFORMS is the largest international association of operations research (O.R.) and analytics professionals and students. INFORMS provides unique networking and learning opportunities for individual professionals, and organizations of all types and sizes, to better understand and use O.R. and analytics tools and methods to transform strategic visions and achieve better outcomes.
For more information on INFORMS, its publications, membership, or meetings visit http://www.informs.org

Stochastic Systems
2017, Vol. 7, No. 1, 143–196
DOI: 10.1214/15-SSY199

Downloaded from informs.org by [216.73.217.10] on 05 November 2026, at 13:43 . For personal use only, all rights reserved.

HEAVY TRAFFIC APPROXIMATION FOR THE STATIONARY DISTRIBUTION OF A GENERALIZED JACKSON NETWORK: THE BAR APPROACH

By Anton Braverman*, J. G. Dai*, and Masakiyo Miyazawa†

*Cornell University** *and Tokyo University of Science*†

In the seminal paper of Gamarnik and Zeevi [17], the authors justify the steady-state diffusion approximation of a generalized Jackson network (GJN) in heavy traffic. Their approach involves the so-called limit interchange argument, which has since become a popular tool employed by many others who study diffusion approximations. In this paper we illustrate a novel approach by using it to justify the steady-state approximation of a GJN in heavy traffic. Our approach involves working directly with the basic adjoint relationship (BAR), an integral equation that characterizes the stationary distribution of a Markov process. As we will show, the BAR approach is a more natural choice than the limit interchange approach for justifying steady-state approximations, and can potentially be applied to the study of other stochastic processing networks such as multiclass queueing networks.

1. Introduction. This paper considers open single-class queueing networks that have d service stations. Each station has a single server operating under the first-in-first-out (FIFO) service discipline. Upon completing service at a particular station, customers are either routed to another station, or exit the network. There is a single class of customers at each station, meaning that all customers are homogenous in terms of service times and routing. A customer entering the network will exit in finite time with probability one, hence the term open network. For each station, the external interarrival times (possibly null), service times, and routing decisions are assumed to follow three separate i.i.d. sequences of random variables; the three sequences are assumed to be independent. Furthermore, the interarrival times, service times and routing decisions are assumed to be independent between different stations. Such a network is hereafter referred to as a generalized Jackson network (GJN).

Received August 2015.

Keywords and phrases: Stochastic processing networks, single class networks, multiclass networks, stationary distributions, heavy traffic approximation, interchange of limits, reflecting Brownian motions, SRBM.

Downloaded from informs.org by [216.73.217.10] on 05 November 2026, at 13:43 . For personal use only, all rights reserved.

In a seminal paper, Gamarnik and Zeevi [17] proved that for a sequence of GJNs indexed by $n = 1, 2, \ldots$,

$$r_n L^{(n)}(\infty) \Rightarrow Z(\infty) \quad \text{as } n \to \infty, \tag{1.1}$$

where the symbol $\Rightarrow$ denotes convergence in distribution, $\{r_n\}$ is a sequence of positive numbers that converge to zero, $L^{(n)}(\infty)$ is a random vector whose ith component represents the steady-state number of customers at station i in the nth network, and $Z(\infty)$ is a random vector that has the stationary distribution of a certain d-dimensional semimartingale reflecting Brownian motion (SRBM) $Z = \{Z(t), t \geq 0\}$ that was first defined in [21]. Readers are referred to the introduction of [17] for the importat motivation of this problem, and a review of then recent literature. Gamarnik and Zeevi [17] proved (1.1) under two key conditions: (a) the heavy traffic condition, and (b) the exponential moment condition.

Condition (a) is standard and can be expressed in terms of a d-dimensional vector $\rho^{(n)}$, where $\rho_i^{(n)}$ is the traffic intensity at station i in the nth network. This condition requires that $\rho_i^{(n)} < 1$ at each station i, and $\rho_i^{(n)} \to 1$ as $n \to \infty$. The scaling parameter r_n in (1.1) is closely tied to this heavy traffic condition and describes how quickly each $\rho_i^{(n)}$ converges to one. In particular, r_n goes to zero at the same rate as $1-\rho_i^{(n)}$. Namely, $\lim_{n\to\infty}(1-\rho_i^{(n)})/r_n > 0$ exists for each station i. Condition (b) requires that interarrival and service times have finite exponential moments; such a condition is unnecessarily strong. In a follow up work by Budhiraja and Lee [7], this condition is relaxed to a new moment condition: (b') interarrival and service times have finite second moments, and the sequences (indexed by n) of square interarrival and square service times are uniformly integrable. Conditions (a) and (b') in [7] represent the weakest possible conditions for (1.1) to hold.

In this paper, we prove (1.1) under conditions (a) and (b'); see Theorem 2.1 in Section 2.4. Our proof of (1.1) uses a novel approach, and is drastically different from the ones in [7] and [17]. This approach was used in [28] to study the steady-state approximation of a single server queue in heavy traffic. However, ours is the first paper to apply it to the network setting. In addition to proving Theorem 2.1, the ideas laid out in this paper can be applied to study steady-state diffusion approximations of other systems of interest. One promising direction for future work is to generalize this approach to study multiclass queuing networks, such as those studied by Bramson and Dai [4]. We now outline the approach.

For $\theta \in \mathbb{R}^d$ with $\theta \leq 0$, let $\varphi^{(n)}(\theta)$ be the moment generating function (MGF) of $Z^{(n)}(\infty) = r_n L^{(n)}(\infty)$, defined in (4.1). To prove (1.1), we show

Downloaded from informs.org by [216.73.217.10] on 05 November 2026, at 13:43 . For personal use only, all rights reserved.

that $\varphi^{(n)}(\theta)$ converges to $\varphi(\theta)$, the MGF of $Z(\infty)$. To do so, it suffices to prove that the pointwise limit of any convergent subsequence $\{\varphi^{(n_k)},\ k \geq 1\}$ must be $\varphi(\theta)$, i.e.

$$\lim_{k\to\infty} \varphi^{(n_k)}(\theta) = \varphi(\theta) \text{ for each } \theta \in \mathbb{R}^d \text{ with } \theta \leq 0. \tag{1.2}$$

Also associated with $Z(\infty)$ are boundary MGFs $\varphi_j(\theta)$ $(j = 1, \ldots, d)$, defined in (2.26). By a uniqueness result in [13], we know that $\varphi(\theta)$ and its boundary counterparts are characterized by a basic adjoint relationship (BAR) given in (2.30). We know that $\varphi^{(n_k)}(\theta)$ also has associated boundary MGFs $\varphi_j^{(n_k)}(\theta)$ $(j = 1, \ldots, d)$ that are defined in (4.1). To prove (1.2), we show in Proposition 4.1 that $\varphi^{(n_k)}(\theta)$ and its boundary counterparts satisfy BAR (2.30) asymptotically. Namely, the limits $\varphi^*(\theta) = \lim_{k\to\infty} \varphi^{(n_k)}(\theta)$, along with $\varphi_j^*(\theta) = \lim_{k\to\infty} \varphi_j^{(n_k)}(\theta)$, satisfy BAR (2.30) exactly. On its own, this result does not yet imply that $\varphi^*(\theta) = \varphi(\theta)$.

Apriori, we cannot exclude the possibility that $\varphi^*(\theta)$ may be degenerate, i.e. be the MGF of some nonnegative measure on $\mathbb{R}^d$ that has total mass strictly less than 1 (or possibly 0). To invoke the uniquness result in [13] and conclude that $\varphi^*(\theta) = \varphi(\theta)$, we must also prove that $\varphi^*(\theta)$ is the MGF of *some* probability measure, i.e. that it is not degenerate. For example, $\varphi^*(\theta) \equiv 0$ and $\varphi_j^*(\theta) \equiv 0$ clearly satisfy BAR (2.30), but $\varphi(\theta) \neq 0$. To this point, we show that

$$\lim_{\theta\uparrow 0} \varphi^*(\theta) = 1, \tag{1.3}$$

which implies that the sequence of probability measures corresponding to $\{\varphi^{(n_k)},\ k \geq 1\}$ is tight; see Lemma 6.1. It turns out that condition (1.3) can be verified algebraically from the fact that $\varphi^*(\theta)$ and $\varphi_j^*(\theta)$ $(j = 1, \ldots, d)$ satisfy BAR (2.30). Most of the steps in this algebraic procedure are carried out in Proposition 5.1. Once we have this proposition, showing (1.3) becomes straightforward; see for instance, the proof of (6.3). The proof of Proposition 5.1 requires that the reflection matrix of the SRBM be an $\mathcal{M}$-matrix. An $\mathcal{M}$-matrix is an invertible square matrix whose diagonal entries are non-negative, and off diagonal entries are non-positive [2, Chapter 6]. This condition is always satisfied in the GJN setting, because the reflection matrix of the SRBM has the form $(I - P^{\mathrm{T}})$, where P routing matrix of the GJN.

The procedure of considering a sequence $\varphi^{(n_k)}(\theta) \to \varphi^*(\theta)$ (and $\varphi_j^{(n_k)}(\theta) \to \varphi_j^*(\theta)$), and then verifying (1.3) looks like an application of Lévy's convergence theorem [37, Section 18.1], with one key difference. Lévy's result says

Downloaded from informs.org by [216.73.217.10] on 05 November 2026, at 13:43 . For personal use only, all rights reserved.

that if a sequence of characteristic functions (CF) converges, and the limit is continuous at zero, then the corresponding sequence of probability measures converges weakly to a limiting probability measure. Our use of MGFs in this paper, instead of CFs, is not accidental. Applying Lévy's theorem requires knowing apriori that the sequence (or at least a subsequence) of CFs converges, and for a complicated model like a GJN this is nigh impossible to verify. In contrast, MGFs are monotone functions, and any sequence of MGFs always has a convergent subsequence due to Helly's selection principle [10, Theorem 4.3.3].

To prove Proposition 4.1 (that $\varphi^{(n_k)}(\theta)$ and its boundary counterparts satisfy BAR (2.30) asymptotically), we work with a continuous time Markov process $X^{(n)} = \{X^{(n)}(t), t \geq 0\}$ that describes the dynamics of the nth GJN. In addition to the queue length process $L^{(n)} = \{L^{(n)}(t), t \geq 0\}$, this Markov process also keeps track of the remaining interarrival times and remaining service times at all stations. Although $L^{(n)}$ is a jump process taking values ℓ in $\mathbb{Z}_+^d$, the other component of $X^{(n)}$ is a piecewise deterministic process taking values (u, v) in $\mathbb{R}_+^{2d}$. Nevertheless, the stationary distribution of $X^{(n)}$ satisfies a BAR (3.15) for all "good" test functions $f(\ell, u, v)$. The BAR for $X^{(n)}$ has two components. The first component involves the deterministic process and the second involves the jump process; the latter is generally difficult to analyze. To handle this difficulty, we choose $f(\ell, u, v)$ to be an exponential function of the state variable (ℓ, u, v) of the form

$$f(\ell, u, v) = e^{\langle\theta,\ell\rangle+\langle\eta,u\rangle+\langle\zeta,v\rangle}, \tag{1.4}$$

where $(\theta, \eta, \zeta) \in \mathbb{R}^{3d}$ are parameters and $\langle\cdot,\cdot\rangle$ is the Euclidean inner product (we actually use truncated versions of these functions to accommodate general interarrival and service time distributions, which may not have exponential moments, see (3.21)). By judiciously choosing $\eta = \eta(\theta)$ and $\zeta = \zeta(\theta)$ as functions of $\theta \in \mathbb{R}^d$, we eliminate the jump term of the BAR (3.15), leaving us with the jump free BAR (3.26). To obtain Proposition 4.1 from this jumpless BAR (3.26), we perform Taylor expansion on $\eta(\theta)$ and $\zeta(\theta)$ to obtain their quadratic approximations (Lemma 4.1), and establish corresponding error bounds (Lemma 4.2).

Both [7] and [17] focused on proving the tightness of $\{r_n L^{(n)}(\infty), n \geq 1\}$ by ingenuously constructing appropriate Lyapunov functions. Both papers rely on the Lipschitz continuity of the Skorohod map corresponding to the SRBM for their tightness argument. Such a map does not exist in the multiclass queueing network setting, which makes the generalization of results from [7] and [17] to the multiclass setting difficult. Gurvich [19] is the only paper that provides a sufficient condition for proving (1.1) in the multiclass

Downloaded from informs.org by [216.73.217.10] on 05 November 2026, at 13:43 . For personal use only, all rights reserved.

setting. In addition to the strong exponential moment assumption, [19] assumes a strong state space collapse (SSC) condition: (c) fluid solutions of a critically loaded fluid model converge to their equilibria at a "uniform linear rate" in finite time. Gurvich [19] focused on generalizing the approach in [17]. In particular, he retained the strong exponential moment assumption. In a recent paper, Ye and Yao [38] focused on generalizing the approach in [7] to resource-sharing networks that lie outside of the multiclass queueing network setting. Using a multiclass queueing network example, in Section 5 of [38], the authors outline the steps needed for generalizing their approach to the multiclass queueing network setting. The authors are able to keep the weak moment condition (b'), relaxing condition (c) to condition (c'): fluid solutions converge to their equilibria "uniformly fast", but not necessarily in finite time. However, they imposed a strong "bounded workload condition", which is difficult to check in general.

The approach of Gamarnik and Zeevi [17] is known as the limit interchange argument, and has since been used by others to study steady-state approximations of various queueing systems. In the single server setting, Katsuda [25] studied a multiclass single server queue with feedback, and Zhang and Zwart [39] studied a limited processor sharing queue. In the many-server setting, Tezcan [36] considered a parallel-server system with multiple server pools and no customer abandonment, Gamarnik and Stolyar [16] examined a multiclass, many-server queue with abandonment, where customer service and patience times are exponentially distributed with means varying between different customer classes, and Dai et al. [12] considered a many-server queue with abandonment, where service times follow a phase-type distribution. In recent years, several papers [5, 6, 18, 20, 23] have gone beyond limit theorems, and establish rates of convergence to the approximating distribution. The framework underlying those papers (except for [20]) is known as Stein's method [35, 9, 32].

The limit interchange and Stein method frameworks represent the two general approaches used to establish convergence of steady-state distributions. Our paper adds a third approach to this set. Each approach has its own pros and cons. The Stein approach is able to provide rates of convergence, which is a step beyond just convergence, but this comes at a cost. Successfully applying it requires deeper knowledge about the underlying system than either the limit interchange approach, or the one presented in this paper. In particular, Stein's method has not been applied to queueing networks and has so far been limited to systems with a single station. The limit interchange approach has been the prominent method in the past decade. At its core, it requires the use of a Lyapunov function to prove tightness. How-

Downloaded from informs.org by [216.73.217.10] on 05 November 2026, at 13:43 . For personal use only, all rights reserved.

ever, each stochastic system requires a separate Lyapunov function. Finding one is typically very difficult and can be considered an art. In contrast, the method in this paper is algorithmic in nature, and requires no guesswork to find any Lyapunov function. For instance, there is little wiggle room in choosing the exponential test function in (1.4), and our tightness argument in Section 5 is algebraic and "procedural". In terms of generality of our method, multiclass queueing networks do add an extra layer of difficulty to our approach. Namely, the presence of SSC in those networks and the fact that the reflection matrix of the SRBM no longer has to be an $\mathcal{M}$-matrix. It is the subject of ongoing research to extend our approach to the multiclass setting.

The rest of the paper is structured as follows. In Section 2, we introduce the sequence of GJNs, the heavy traffic condition, describe the approximating SRBM, and state our main results. In Section 3, we derive the BAR (3.15) for each GJN in the sequence, and introduce conditions on test functions under which the jump term there disappears. Section 4 is devoted to proving Proposition 4.1, which states that the MGFs of the queue lengths of the GJN approximately satisfy the BAR of the SRBM. In Section 5 we present Proposition 5.1, which we use together with Proposition 4.1 to prove our main result, Theorem 2.1, in Section 6. We defer proofs of certain technical lemmas to the Appendix.

1.1. *Notation.* All random variables and stochastic processes are defined on a common probability space $(\Omega, \mathcal{F}, \mathbb{P})$, and all stochastic processes $X = \{X(t), t \geq 0\}$ are assumed to be right continuous on $[0, \infty)$, and having left limits on $(0, \infty)$. For a sequence of random variables $\{Y_n,\ n \geq 1\}$ and a random variable Y, we write $Y_n \Rightarrow Y$ if Y_n converge in distribution to Y. For an integer $d \geq 1$, $\mathbb{R}^d$ denotes the d-dimensional Euclidean space, and $\mathbb{Z}_+^d$ and $\mathbb{R}_+^d$ denote the spaces of d-dimensional vectors whose elements are non-negative integers and non-negative real numbers, respectively. For vectors $x, y \in \mathbb{R}^d$, we write x_i to denote the ith component of x, $1 \leq i \leq d$. Furthermore, we write $x \leq y$ if $x_i \leq y_i$ for all $i = 1, ..., d$ and we let $\langle x, y \rangle$ be their Euclidean inner product. All vectors are understood to be column vectors. A function $f : \mathbb{R}^d \to \mathbb{R}$ is said to be non-decreasing if $x \leq y$ implies $f(x) \leq f(y)$. For a vector $x \in \mathbb{R}^d$, define the sup-norm $||x||_\infty = \sup_i |x_i|$. For integers a, b with $a > b$, we define $\sum_{i=a}^b = 0$. For integers i, j, we let δ_{ij} be the Kronecker delta; i.e. $\delta_{ij} = 1$ if $i = j$, and zero otherwise. We let x^{T} and A^{T} denote the transpose of a vector x and matrix A, respectively. We reserve I for the identity matrix, e for the vector of all ones and $e^{(i)}$ for the vector that has a one in the ith element and zeroes elsewhere; the

dimensions of these vectors will be clear from the context.

2. Heavy traffic approximation. In this section, we introduce the generalized Jackson network and state the main results of this paper.

2.1. *Network description.* To be able to state our main results, we first introduce a generalized Jackson network, and define a Markov process that describes it. This network has d stations, numbered $1, 2, \ldots, d$. Let $\mathcal{J} = \{1, 2, \ldots, d\}$. Each station has a single server that serves customers in the first-in-first-out (FIFO) manner. A station may have customers arriving from outside the network; we refer to such arrivals as external arrivals. A customer who completes service at a station either goes to another station or leaves the network. The following is a mathematical description of a GJN.

Let $\mathcal{E}$ be the subset of $\mathcal{J}$ whose members are the stations that have external arrivals. External arrivals at station $i \in \mathcal{E}$ follow a renewal process with i.i.d. interarrival times

$$\{T_{e,i}(m),\ m = 1, 2, \ldots, \}, \tag{2.1}$$

and we let $T_{e,i}$ be a nonnegative random variable having the distribution of the interarrival times. We assume that $T_{e,i}$ has finite variance and a non-zero mean. External arrivals at different stations are independent.

Service times at station i are i.i.d. random variables

$$\{T_{s,i}(m),\ m = 1, 2, \ldots, \}, \tag{2.2}$$

and we let $T_{s,i}$ be a nonnegative random variable having the distribution of the service times. We assume that $T_{s,i}$ has finite variance and a non-zero mean. Service times at different stations are independent. Service time sequences and external interarrival time sequences are assumed to be independent.

A customer that completes service at station $i \in \mathcal{J}$ goes to station $j \in \mathcal{J}$ with probability p_{ij} or exits the network with probability $p_{i0} = 1 - \sum_{j \in \mathcal{J}} p_{ij}$, independently of everything else. Let P be the $d \times d$ square matrix whose (i, j)th entry is p_{ij} for $i, j \in \mathcal{J}$.

This queueing network is referred to as a generalized Jackson network (GJN). We now introduce a Markov process for describing the GJN. For time $t \geq 0$, denote the number of customers, the residual external arrival time, and the residual service time at station $i \in \mathcal{J}$ by $L_i(t)$, $R_{e,i}(t)$ and $R_{s,i}(t)$, respectively. We set $R_{e,i}(t) = 0$ for $i \in \mathcal{J} \setminus \mathcal{E}$, and $R_{s,i}(t) = T_{s,i}(m)$ if no customer is in service at station i at time t, and the service time of the next customer at station i is $T_{s,i}(m)$. In Section 2.1 of [11] and Section 2.1.1

Downloaded from informs.org by [216.73.217.10] on 05 November 2026, at 13:43 . For personal use only, all rights reserved.

Downloaded from informs.org by [216.73.217.10] on 05 November 2026, at 13:43 . For personal use only, all rights reserved.

of [7], $R_{s,i}(t)$ is defined to be zero if no customer is in service at time t. Our definition will make it slightly easier to derive condition (3.19) in Section 3.1 to annihilate the jump term in (3.15) there.

Denote the vectors whose entries are $L_i(t)$, $R_{e,i}(t)$ and $R_{s,i}(t)$ by $L(t)$, $R_e(t)$ and $R_s(t)$, respectively. Throughout the paper we refer to $\{L(t), t \geq 0\}$ as the queue length process (even though it includes customers currently in service as well). Let $X(t) = (L(t), R_e(t), R_s(t))$, then $X = \{X(t), t \geq 0\}$ is a Markov process with state space $\mathbb{Z}_+^d \times \mathbb{R}_+^{2d}$.

The GJN is a natural generalization of the Jackson network, but its stationary distribution is hard to get. This motivates the study of heavy traffic approximations in [30] and [24]. To introduce the notation of heavy traffic, we introduce a sequence of generalized Jackson networks in the next section.

2.2. *A sequence of networks and their assumptions.* Consider a sequence of GJNs indexed by $n = 1, 2, \ldots$. We denote the nth GJN by superscript $^{(n)}$. For example, $X(t), L(t), R_e(t), R_s(t)$ are denoted by $X^{(n)}(t)$, $L^{(n)}(t)$, $R_e^{(n)}(t)$, $R_s^{(n)}(t)$, respectively. We assume the routing matrix $P = (p_{ij})$ is independent of n. The networks are assumed to be open, i.e., the matrix $(I - P)$ is invertible; the inverse is given by

$$(I - P)^{-1} = I + P + P^2 + \ldots. \tag{2.3}$$

For $i \in \mathcal{E}$, we denote the mean and variance of $T_{e,i}^{(n)}$ by $1/\lambda_{e,i}^{(n)}$ and $(\sigma_{e,i}^{(n)})^2$, respectively. For notational simplicity, we adopt the conventions that $\lambda_{e,i} = 0$ and $(\sigma_{e,i}^{(n)})^2 = 0$ for $i \in \mathcal{J} \setminus \mathcal{E}$. Similarly, for $j \in \mathcal{J}$, we denote the mean and variance of $T_{s,j}^{(n)}$ by $1/\lambda_{s,j}^{(n)}$ and $(\sigma_{s,j}^{(n)})^2$, respectively. Let $\lambda_{a,i}^{(n)}$ for $i \in \mathcal{J}$ be the solution of the traffic equation:

$$\lambda_{a,i}^{(n)} = \lambda_{e,i}^{(n)} + \sum_{j \in \mathcal{J}} \lambda_{a,j}^{(n)} p_{ji}, \qquad i \in \mathcal{J}.$$

Then $\lambda_{a,i}^{(n)}$ can be interpreted as the total arrival rate at station i. The traffic equation can be written as the vector valued equation:

$$\lambda_a^{(n)} = \lambda_e^{(n)} + P^{\mathrm{T}} \lambda_a^{(n)}, \tag{2.4}$$

where P^{T} is the transpose of P. Equation (2.4) has a unique solution given by $\lambda_a^{(n)} = (I - P^{\mathrm{T}})^{-1} \lambda_e^{(n)}$.

We assume the following heavy traffic conditions: there exists a positive vector $b \in \mathbb{R}_+^d$, and a sequence of positive numbers r_n such that

$$\lambda_s^{(n)} - \lambda_a^{(n)} = b\, r_n, \qquad n \geq 1, \tag{2.5}$$

$$\lim_{n\to\infty} r_n = 0. \tag{2.6}$$

It will be convenient to express condition (2.5) in terms of the primitive data $\lambda_e^{(n)}$ and $\lambda_s^{(n)}$. For this, we substitute $\lambda_a^{(n)} = \lambda_s^{(n)} - b\, r_n$ into both sides of (2.4) to get

$$\lambda_s^{(n)} - \lambda_e^{(n)} - P^{\mathrm{T}}\lambda_s^{(n)} = r_n R b, \tag{2.7}$$

where

$$R = I - P^{\mathrm{T}}. \tag{2.8}$$

Note that r_n is chosen to be $1/\sqrt{n}$ in [30] and much of the literature as well, but it is intuitively clear that this is not essential as long as (2.5) holds and r_n converges to zero as $n \to \infty$. For example, some authors take $r_n = 1/n$ (see, e.g., [8]). In this paper, we do not make any specific choice for r_n; this conveys the same spirit of Kingman [26]'s heavy traffic approximation (see [28] for details). We make the following moment assumptions on the sequence of networks:

$$\sigma_{e,i}^{(n)} \to \sigma_{e,i} < \infty, \quad \lambda_{e,i}^{(n)} \to \lambda_{e,i} > 0 \quad \text{for each } i \in \mathcal{E}, \tag{2.9}$$

$$\sigma_{s,j}^{(n)} \to \sigma_{s,j} < \infty \quad \text{for each } j \in \mathcal{J}. \tag{2.10}$$

In addition, we assume

$$\left\{\left(T_{e,i}^{(n)}\right)^2, n \geq 1\right\} \text{ is uniformly integrable for each } \quad i \in \mathcal{E}, \tag{2.11}$$

$$\left\{\left(T_{s,j}^{(n)}\right)^2, n \geq 1\right\} \text{ is uniformly integrable for each } \quad j \in \mathcal{J}. \tag{2.12}$$

Following traffic equation (2.4), conditions (2.5), (2.6), and (2.9) imply that

$$\lambda_{a,j}^{(n)} \to \lambda_{a,j}, \quad \text{and} \quad \lambda_{s,j}^{(n)} \to \lambda_{s,j} = \lambda_{a,j} \quad \text{for } j \in \mathcal{J}, \tag{2.13}$$

where $\lambda_a = (I - P^{\mathrm{T}})^{-1}\lambda_e$.

The diffusion approximation focuses on the sequence $L^{(n)} = \{L^{(n)}(t), t \geq 0\}$, which is the first component of $X^{(n)}$. Clearly, $L^{(n)}$ is not a Markov process in general, but the standard approach in the literature (e.g. [24, 30]) shows that the diffusion-scaled process $Z^{(n)} = \{Z^{(n)}(t), t \geq 0\}$, defined as

$$Z^{(n)}(t) = r_n L^{(n)}\left(t/r_n^2\right), \quad t \geq 0, \tag{2.14}$$

converges in distribution to a semimartingale reflecting Brownian motion (SRBM) $Z = \{Z(t), t \geq 0\}$ (to be defined in Section 2.3). The heavy traffic

Downloaded from informs.org by [216.73.217.10] on 05 November 2026, at 13:43 . For personal use only, all rights reserved.

assumptions (2.5) and (2.6) are crucial for the time and space scalings in (2.14) to be correct.

We also assume that for each n,

$$X^{(n)} = \{X^{(n)}(t), t \geq 0\} \text{ is positive Harris recurrent.} \tag{2.15}$$

This assumption is satisfied under heavy traffic condition (2.5) and some additional regularity assumptions on the interarrival time distributions; see, for example, Theorem 3.8 of Down and Meyn [15] and Theorem 5.1 of Dai [11]. Since $X^{(n)}$ is positive Harris recurrent, it has a unique stationary distribution. We let $X^{(n)}(\infty)$ be the vector having that stationary distribution.

Since $X^{(n)}$ is assumed to have a stationary distribution, $L^{(n)}$ has a stationary distribution. We use $L^{(n)}(\infty)$ and $Z^{(n)}(\infty)$ to denote the random vectors having the stationary distributions of $L^{(n)}$ and $Z^{(n)}$, respectively. Note that the stationary distribution of $Z^{(n)}$ is independent of the time scaling in (2.14) for each fixed n. Furthermore, it is clear that $Z^{(n)}(\infty) \stackrel{d}{=} r_n L^{(n)}(\infty)$. For future reference, we use $\pi^{(n)}$ to denote the stationary distribution of $Z^{(n)}$. As stated in (1.1), the primary result of this paper is to prove that $Z^{(n)}(\infty) \Rightarrow Z(\infty)$, where $Z(\infty)$ has the stationary distribution of the SRBM Z, which we now define.

2.3. *Semimartingale reflecting Brownian motions and BAR.* Recall the matrix R, defined in (2.8), and set

$$\mu = -Rb \in \mathbb{R}^d, \tag{2.16}$$

where b is given in (2.5). Let $\Sigma = (\Sigma_{ij})$ be a $d \times d$ symmetric matrix given by

$$\begin{aligned}\Sigma_{ij} = \sum_{k \in \mathcal{J}} \lambda_{s,k} \big[& p_{ki}(\delta_{ij} - p_{kj}) \\ & + \lambda_{s,k}^2 \sigma_{s,k}^2 (p_{ki} - \delta_{ki})(p_{kj} - \delta_{kj}) \big] + \lambda_{e,i}^3 \sigma_{e,i}^2 \delta_{ij},\end{aligned} \tag{2.17}$$

where δ_{ij} is the Kronecker delta, p_{ij} are the routing probabilities in the GJN, and the quantities $\lambda_{e,i}$, $\sigma_{e,i}$, $\sigma_{s,i}$, and $\lambda_{s,i}$ are given in (2.9), (2.10), and (2.13), respectively. The matrix Σ is always non-negative definite. Throughout the document, we assume that

$$\Sigma \text{ is positive definite,} \tag{2.18}$$

which is a standard assumption in both [17, 7]. Associated with the data (μ, Σ, R) is $Z = \{Z(t), t \geq 0\}$, a semimartingale reflecting Brownian motion (SRBM) that satisfies

$$Z(t) = \xi(t) + RY(t), \quad t \geq 0, \tag{2.19}$$

Downloaded from informs.org by [216.73.217.10] on 05 November 2026, at 13:43 . For personal use only, all rights reserved.

Downloaded from informs.org by [216.73.217.10] on 05 November 2026, at 13:43 . For personal use only, all rights reserved.

(2.20) ξ is a d-dimensional Brownian motion with drift μ and covariance matrix Σ,

(2.21) $Y(0) = 0$, each component of Y is non-decreasing, and

(2.22) $$\int_0^\infty Z_j(t)dY_j(t) = 0, \quad j \in \mathcal{J}.$$

The matrix R is called the reflection matrix of Z. Recall that an $\mathcal{M}$-matrix is an invertible square matrix whose diagonal entries are non-negative, and off diagonal entries are non-positive [2, Chapter 6]. Since R in (2.8) is an $\mathcal{M}$-matrix, it follows from [21] that the SRBM Z exists and is unique as a strong solution to (2.19)–(2.22).

Again, because R is an $\mathcal{M}$-matrix and condition

(2.23) $$R^{-1}\mu = -b < 0$$

is satisfied, it follows from [22] that the SRBM Z has a unique stationary distribution π on $(\mathbb{R}^d_+, \mathcal{B}(\mathbb{R}^d_+))$, where $\mathcal{B}(\mathbb{R}^d_+)$ is the Borel field of $\mathbb{R}^d_+$. We now discuss the characterization of π.

Let $\mathbb{E}_\pi$ denote the expectation when $Z(0)$ has distribution π. For $j \in \mathcal{J}$, define the boundary probability measure π_j by

$$\pi_j(B) = \frac{1}{\mathbb{E}_\pi[Y_j(1)]}\mathbb{E}_\pi\left[\int_0^1 1(Z(t) \in B)dY_j(t)\right], \qquad B \in \mathcal{B}(\mathbb{R}^d_+).$$

We know that $\mathbb{E}_\pi[Y_j(1)] < \infty$ by [22, Theorem 1]. Note that $Y_j(t)$ increases only on the face

$$F_j = \{x \in \mathbb{R}^d_+ : x_j = 0\}, \qquad j \in \mathcal{J}.$$

Therefore, π_j concentrates on F_j, but we define it on $\mathbb{R}^d_+$ for notational simplicity.

The following lemma gives a characterization of the stationary distribution π and its associated boundary measures $\pi_1, \ldots, \pi_d$.

LEMMA 2.1. *(a) The stationary distribution π and its associated boundary measures $\pi_1, \ldots, \pi_d$ must satisfy the following basic adjoint relationship (BAR)*

(2.24) $$\int_{\mathbb{R}^d_+} Gf(x)\pi(dx) + \sum_{j\in\mathcal{J}} \mathbb{E}_\pi[Y_j(1)] \int_{\mathbb{R}^d_+} \langle \nabla f(x), R^{(j)}\rangle \pi_j(dx) = 0,$$
for each $f \in C_b^2(\mathbb{R}^d_+)$,

where $R^{(j)}$ is the jth column of R, $C_b^2(\mathbb{R}_+^d)$ is the set of functions f on $\mathbb{R}_+^d$ such that f, its first order derivatives, and its second order derivatives are bounded and continuous, and

$$Gf(x) = \frac{1}{2}\sum_{i,j=1}^{d} \Sigma_{ij}\frac{\partial^2 f}{\partial x_i \partial x_j}(x) + \sum_{j=1}^{d} \mu_j \frac{\partial f}{\partial x_j}(x). \tag{2.25}$$

(b) Conversely, assume that π, π_1, …, π_d are probability measures on $\mathbb{R}_+^d$ satisfying BAR (2.24). Then π must be the stationary distribution of the (μ, Σ, R)-SRBM, and π_1, …, π_d the corresponding boundary measures.

Proof. Part (a) follows from [22]. Part (b) follows from [13]. □

We denote the moment generating functions (MGFs) of π and π_j by φ and φ_j, respectively. Namely, for $\theta \in \mathbb{R}^d$ with $\theta \le 0$,

$$\begin{aligned} \varphi(\theta) &= \mathbb{E}_\pi[e^{\langle \theta, Z(0)\rangle}], \\ \varphi_j(\theta) &= \frac{1}{\mathbb{E}_\pi[Y_j(1)]}\mathbb{E}_\pi\left[\int_0^1 e^{\langle \theta, Z(t)\rangle} dY_j(t)\right], \quad j \in \mathcal{J}. \end{aligned} \tag{2.26}$$

We also define

$$\begin{aligned} \gamma(\theta) &= \frac{1}{2}\langle \theta, \Sigma\theta\rangle + \langle -Rb, \theta\rangle, \\ \gamma_j(\theta) &= \langle R^{(j)}, \theta\rangle, \qquad \theta \in \mathbb{R}^d,\ j \in \mathcal{J}. \end{aligned} \tag{2.27}$$

Plugging $f(x) = e^{\langle \theta, x\rangle}$ into (2.24), the BAR (2.24) becomes

$$\gamma(\theta)\varphi(\theta) + \sum_{j=1}^{d} \mathbb{E}_\pi[Y_j(1)]\gamma_j(\theta)\varphi_j\big(\theta\big) = 0, \quad \text{for each } \theta \in \mathbb{R}^d \text{ with } \theta \le 0, \tag{2.28}$$

and we now use it to show that

$$\mathbb{E}_\pi[Y_j(1)] = b_j, \quad j \in \mathcal{J}. \tag{2.29}$$

To do so, we let $\theta = \alpha e^{(k)}$, where $\alpha \le 0$ is a real number and $k \in \mathcal{J}$. Dividing both sides of (2.28) by α and taking $\alpha \uparrow 0$, we obtain

$$-\langle Rb, e^{(k)}\rangle + \sum_{j=1}^{d} \mathbb{E}_\pi[Y_j(1)]\langle R^{(j)}, e^{(k)}\rangle = \sum_{j=1}^{d} \big(\mathbb{E}_\pi[Y_j(1)] - b_j\big) r_{kj} = 0$$

Downloaded from informs.org by [216.73.217.10] on 05 November 2026, at 13:43 . For personal use only, all rights reserved.

Downloaded from informs.org by [216.73.217.10] on 05 November 2026, at 13:43 . For personal use only, all rights reserved.

for all $k \in \mathcal{J}$, where r_{kj} is the (k,j)th entry of R. In vector form, this is equivalent to

$$R\big(\mathbb{E}_\pi[Y(1)] - b\big) = 0,$$

and since R is invertible, (2.29) follows. Therefore, (2.28) can be rewritten in the more practical form

$$\gamma(\theta)\varphi(\theta) + \sum_{j=1}^{d} b_j \gamma_j(\theta)\varphi_j\big(\theta\big) = 0, \qquad \text{for each } \theta \in \mathbb{R}^d \text{ with } \theta \le 0. \tag{2.30}$$

We refer to this as the MGF version of the BAR.

It is shown in the appendix of the arXiv version of [14] that the MGF version of the BAR (2.30) and the standard version the BAR (2.24) are equivalent. Thus, it follows from the characterization obtained in [13] that $\varphi(\theta)$ and $\varphi_j\big(\theta\big)$ are the unique moment generating functions that satisfy (2.30).

2.4. *Main results.* Recall that $\pi^{(n)}$ is the stationary distribution of $Z^{(n)}$ and for $j \in \mathcal{J}$, define $\pi_j^{(n)}$ to be the distribution of $\big[Z^{(n)}(\infty)\big|Z_j^{(n)}(\infty) = 0\big]$. We are now ready to present our main results.

THEOREM 2.1. *Consider the sequence of GJNs indexed by n. Assume the heavy traffic conditions (2.5) and (2.6), positive recurrence condition (2.15), and moment conditions (2.9)–(2.12) hold. Then,*

(a) As $n \to \infty$, the stationary distribution $\pi^{(n)}$ of $Z^{(n)}$ converges weakly to π, the stationary distribution of the SRBM Z.

(b) For each $j \in \mathcal{J}$, $\pi_j^{(n)}$ converges weakly to π_j, the corresponding boundary measure on F_j.

Part (b) appears to be new. As mentioned in Section 2.2, Part (a) of the theorem is known, but we prove it using a new approach. To elaborate on this approach, we first discuss the existing approach that uses process limit and tightness.

Assuming the distribution of $Z^{(n)}(0)$ converges, Reiman [30] proves that the scaled queue length process $\{Z^{(n)}(t), t \ge 0\}$ converges in distribution to the SRBM Z; we refer to this as the process limit. Namely,

$$Z^{(n)}(\cdot) \Rightarrow Z(\cdot) \quad \text{as } n \to \infty, \tag{2.31}$$

Also, it is proved in [22] that

$$Z(t) \Rightarrow Z(\infty) \quad \text{as } t \to \infty, \tag{2.32}$$

Downloaded from informs.org by [216.73.217.10] on 05 November 2026, at 13:43 . For personal use only, all rights reserved.

where $Z(\infty)$ is a random vector having distribution π. From (2.31) and (2.32), it is not surprising that $Z^{(n)}(\infty) \Rightarrow Z(\infty)$, which is the content of part (a) of Theorem 2.1. Here, $Z^{(n)}(\infty)$ is the random vector having the stationary distribution of $\{Z^{(n)}(t), t \geq 0\}$. As we mentioned earlier, part (a) of Theorem 2.1 was already proved by Gamarnik and Zeevi [17], and Budhiraja and Lee [7]. Both sets of authors rely on the process limit result (2.31), and reduce the problem to showing tightness of the stationary distributions $\{\pi^{(n)} : n = 1, 2, \ldots\}$. Together with the process limit, this tightness implies (1.1). Our proof of Theorem 2.1 will not use the process limit (2.31). Instead, we work directly with the BAR associated with each network in the sequence, which we now derive in Section 3.

3. Network dynamics and basic adjoint relationship. In Section 3.1 we derive the BAR for the GJN and describe a special class of test functions for which the BAR reduces to a very simple expression. In Section 3.2, we focus on a family of exponential test functions that belong to this special class. This lays down the foundation for Section 4, where we compare the BAR of the GJN to that of the approximating SRBM.

3.1. *Network dynamics and tractable BAR.* In this section, we derive the BAR (3.15) for a GJN using sample path dynamics of the network. The main result of this section is Lemma 3.1 below. It describes a special class of functions for which the BAR has a simple form. This lemma can be compared to the BAR in Miyazawa [29], which studies a heterogeneous multiserver queue. However, our approach is different and we make detailed comments on this difference at the end of this section.

We first describe the network dynamics in terms of network primitives. We fix a network in the sequence and temporarily drop the index n for the remainder of the section. For station $j \in \mathcal{J}$, we introduce the sequence of i.i.d. random vectors

$$\{\phi^{(j)}(m) \in \mathbb{Z}_+^d,\ m = 1, 2, \ldots\} \tag{3.1}$$

with

$$\mathbb{P}(\phi^{(j)}(1) = e^{(k)}) = p_{jk}, \quad k \in \mathcal{J}$$
$$\mathbb{P}(\phi^{(j)}(1) = 0) = 1 - \sum_{k \in \mathcal{J}} p_{jk} = p_{j0},$$

where p_{jk} is the probability that a customer goes to station k after completing service at station j. These random vectors represent the routing of customers after completion of service at station j.

Downloaded from informs.org by [216.73.217.10] on 05 November 2026, at 13:43 . For personal use only, all rights reserved.

Recall the definition of $X = \{X(t), t \geq 0\}$ from Section 2.1, and let

$$X(0) = (L(0), R_e(0), R_s(0)) \in \mathbb{Z}_+^d \times \mathbb{R}_+^{2d}$$

be the initial state of the network. We recall the sequences of interarrival and service times from (2.1) and (2.2), and for each integer $q \geq 1$ we define the primitive processes

$$U_i(q) = R_{e,i}(0) + \sum_{m=1}^{q-1} T_{e,i}(m), \quad i \in \mathcal{E}, \tag{3.2}$$

$$V_j(q) = R_{s,j}(0) + \sum_{m=1}^{q-1} T_{s,j}(m), \quad j \in \mathcal{J}, \tag{3.3}$$

$$\Phi^{(j)}(q) = \sum_{m=1}^{q} \phi^{(j)}(m), \quad j \in \mathcal{J}. \tag{3.4}$$

For station $j \in \mathcal{J}$, and time $t \geq 0$, let $E_j(t)$ and $D_j(t)$ be the total number of external arrivals and total number of departures, respectively, on the interval $[0, t]$. Let

$$B_j(t) = \int_0^t 1(L_j(s) \geq 1) ds, \quad j \in \mathcal{J} \tag{3.5}$$

be the cumulative busy time of the server at station j. We can then define

$$S_j(q) = \inf\{t > 0 : \ B_j(t) = V_j(q)\}, \quad j \in \mathcal{J}, q \geq 1, \tag{3.6}$$

to be the time of the qth service completion at station j. Recalling that only stations in $\mathcal{E}$ have external arrivals, we see that for every $t \geq 0$ and on every sample path,

$$E_i(t) = \max\left\{q \in \mathbb{Z}_+ : U_i(q) \leq t\right\}, \quad i \in \mathcal{E}, \tag{3.7}$$

$$E_i(t) \equiv 0, \quad i \in \mathcal{J} \setminus \mathcal{E},$$

$$D_j(t) = \max\left\{q \in \mathbb{Z}_+ : S_j(q) \leq t\right\}, \quad j \in \mathcal{J}. \tag{3.8}$$

Furthermore, one may check that the queue lengths, residual interarrival times and residual service times satisfy

$$L_j(t) = L_j(0) + E_j(t) - D_j(t) + \sum_{k \in \mathcal{J}} \Phi_j^{(k)}(D_k(t)), \quad j \in \mathcal{J}, \tag{3.9}$$

$$R_{e,i}(t) = U_i(E_i(t) + 1) - t, \quad i \in \mathcal{E},$$

$$R_{e,i}(t) \equiv 0, \quad i \in \mathcal{J} \setminus \mathcal{E},$$
$$R_{s,j}(t) = V_j(D_j(t)+1) - B_j(t), \quad j \in \mathcal{J}.$$

In the last equation, we have adopted the convention that when station j is empty at time t, the remaining service time $R_{s,j}(t)$ is set to be the service time of the next customer at this station. Clearly,

$$\begin{aligned} \frac{\partial}{\partial t} R_{e,i}(t) &= -1, \text{ for } i \in \mathcal{E} \\ \frac{\partial}{\partial t} R_{s,j}(t) &= -1(L_j(t) > 0), \quad \text{for } j \in \mathcal{J}, \end{aligned} \tag{3.10}$$

where the derivatives at a jump time t are interpreted as the right derivatives. With the network dynamics rigorously defined, we proceed to derive a BAR for the network.

Let $\mathcal{D}$ be the space of all bounded functions $f(\ell, y) : \mathbb{Z}_+^d \times \mathbb{R}_+^{2d} \to \mathbb{R}$ defined as follows. For any $i = 1, \ldots, 2d$, fix $(x, (y_k)_{k \neq i}) \in \mathbb{Z}_+^d \times \mathbb{R}_+^{2d-1}$ and view $f(x, y)$ as a one dimensional function in the y_i component. We require that this one dimensional function be continuously differentiable at all but finitely many points, and have bounded derivatives whose bound is independent of the point $(x, (y_k)_{k \neq i})$. For instance, $\mathcal{D}$ contains the space of all bounded functions $f : \mathbb{Z}_+^d \times \mathbb{R}_+^{2d} \to \mathbb{R}$, such that for any $\ell \in \mathbb{Z}_+^d$, the function $f(\ell, \cdot) : \mathbb{R}_+^{2d} \to \mathbb{R}$ is continuously differentiable with a bounded derivative, whose bound is independent of ℓ. The reason for enforcing the component-wise conditions on the set $\mathcal{D}$ is that the test functions we use in this paper always have some form of truncation, which prevents them from being everywhere differentiable in the variable belonging to $\mathbb{R}_+^{2d}$.

Now for any $f \in \mathcal{D}$, and any interval $[t, t+h] \subset \mathbb{R}_+$ free of jumps, we can use the fundamental theorem of calculus together with (3.10) to see that

$$f(X(t+h)) - f(X(t)) = \int_t^{t+h} \mathcal{A}f(X(s))\, ds,$$

where

$$\mathcal{A}f(x) = -\sum_{i \in \mathcal{E}} \frac{\partial f}{\partial u_i}(x) - \sum_{j \in \mathcal{J}} \frac{\partial f}{\partial v_j}(x) 1(\ell_j > 0), \tag{3.11}$$
$$x = (\ell, u, v) \in \mathbb{Z}_+^d \times \mathbb{R}_+^d \times \mathbb{R}_+^d.$$

For the remainder of this section, we let ν denote the stationary distribution of X, and let $\mathbb{P}_\nu$ be the probability measure conditioned on $X(0)$ having the distribution ν. To deal with the jumps of X, we introduce some notation.

Downloaded from informs.org by [216.73.217.10] on 05 November 2026, at 13:43 . For personal use only, all rights reserved.

Downloaded from informs.org by [216.73.217.10] on 05 November 2026, at 13:43 . For personal use only, all rights reserved.

We know that the jumps of the process X correspond to external arrivals and departures at various stations. We use the term event to refer to a single external arrival or departure. At time instance s, there may be simultaneous events. A jump of X at time s constitutes all the events that occur at s. Since we assumed that $\mathbb{E}T_{e,i} > 0$ and $\mathbb{E}T_{s,j} > 0$ for all $i \in \mathcal{E}$ and $j \in \mathcal{J}$, it follows from basic renewal theory (see for instance [31, Theorem 3.3.1]) that

$$\sum_{i\in\mathcal{E}} \mathbb{E}_\nu\big[E_i(t) - E_i(0)\big] + \sum_{j\in\mathcal{J}} \mathbb{E}_\nu\big[D_j(t) - D_j(0)\big] < \infty. \tag{3.12}$$

The finiteness of $\mathbb{E}_\nu\big[D_j(t)-D_j(0)\big]$ comes from the fact that $\{D_j(t),\ t \geq 0\}$ is dominated by the renewal process corresponding to the times $\{V_j(q),\ q \geq 1\}$. Therefore, for every $t > 0$ we know that $\mathbb{P}_\nu$-almost surely, the process X has finitely many events on the interval $(0, t]$. It follows that $\mathbb{P}_\nu$-almost surely,

$$\begin{aligned}&\text{the process } X \text{ has countably many jumps on the interval } (0,\infty),\\ &\text{and every jump instant on this interval has finitely many events.}\end{aligned} \tag{3.13}$$

In what follows, we deal only with the sample paths where (3.13) holds. Suppose now that k events occur simultaneously at time s. We can order them in an arbitrary manner, provided that we do not violate the network dynamics. For example, if station s is empty at time $s-$, and experiences both an arrival and departure at time s, then the arrival must happen first. The particular order assigned to the simultaneous events does not matter, because (3.13) implies there are always finitely many events at a time instance. We can therefore order all of the events that occur on the interval $(0,\infty)$, and represent them as $\delta_1 < \delta_2 < \ldots$, where δ_m represents the mth event to occur after time 0. We let $T(\delta_m)$ represent the time at which the mth event happens.

Now for integer $m \geq 1$, we write X_{δ_m} to represent the value of the process immediately after the mth event has been applied to it. Our use of X_{δ_m} as opposed to $X(T(\delta_m))$ is intentional. If δ_{m_1} and δ_{m_2} represent two simultaneous arrivals to some station, then $X(T(\delta_{m_1})) = X(T(\delta_{m_2}))$, but $X_{\delta_{m_1}} \neq X_{\delta_{m_2}}$. We also write X_{δ_m-} to represent the value of the process X right before the mth event is applied to it, but after the first $m-1$ events have been applied.

From (3.12), we see that

$$\begin{aligned}&\mathbb{E}_\nu\left[\sum_{m=1}^{\infty} 1(0 < T(\delta_m) \leq t)\right]\\ &\quad= \sum_{i\in\mathcal{E}} \mathbb{E}_\nu\big[E_i(t) - E_i(0)\big] + \sum_{j\in\mathcal{J}} \mathbb{E}_\nu\big[D_j(t) - D_j(0)\big] < \infty.\end{aligned} \tag{3.14}$$

Downloaded from informs.org by [216.73.217.10] on 05 November 2026, at 13:43 . For personal use only, all rights reserved.

By isolating times when jumps occur, one can verify that for any $t > 0$ and $f \in \mathcal{D}$,

$$f(X(t)) = f(X(0)) + \int_0^t \mathcal{A}f(X(s))ds + \sum_{m=1}^{\infty} \big(f(X_{\delta_m}) - f(X_{\delta_m -})\big)1(0 < T(\delta_m) \le t), \quad \mathbb{P}_\nu\text{-almost surely.}$$

Since $f \in \mathcal{D}$ is bounded, we can take the expectation under ν to see that

$$\mathbb{E}_\nu\left[\int_0^t \mathcal{A}f(X(s))ds\right] + \mathbb{E}_\nu\left[\sum_{m=1}^{\infty} \big(f(X_{\delta_m}) - f(X_{\delta_m -})\big)1(0 < T(\delta_m) \le t)\right] = 0.$$

Furthermore, we know that $\mathcal{A}f(x)$ is bounded for all $x \in \mathbb{Z}_+^d \times \mathbb{R}_+^{2d}$, meaning we can apply the Fubini-Tonelli theorem to interchange the expectation with the integral in the first term. Stationarity implies that

$$\mathbb{E}_\nu\big[\mathcal{A}f(X(s))\big] = \mathbb{E}_\nu\big[\mathcal{A}f(X(0))\big], \quad \text{for all } s \in [0, t].$$

We therefore have the intermediate result

$$\mathbb{E}_\nu\big[\mathcal{A}f(X(0))\big] + \mathbb{E}_\nu\left[\frac{1}{t}\sum_{m=1}^{\infty} \big(f(X_{\delta_m}) - f(X_{\delta_m -})\big)1(0 < T(\delta_m) \le t)\right] = 0.$$

We now use (3.14) and the boundedness of $f \in \mathcal{D}$ to apply the Fubini-Tonelli theorem again and arrive at the BAR for the GJN:

$$\mathbb{E}_\nu\big[\mathcal{A}f(X(0))\big] + \frac{1}{t}\sum_{m=1}^{\infty} \mathbb{E}_\nu\Big[\big(f(X_{\delta_m}) - f(X_{\delta_m -})\big)1(0 < T(\delta_m) \le t)\Big] = 0, \quad f \in \mathcal{D}. \tag{3.15}$$

Before proceeding further, let us pause and discuss the implications of (3.15). The equation above encodes information about the distribution ν. To access this information, we choose various test functions $f \in \mathcal{D}$ and plug them into (3.15). However, (3.15) is of limited practical use for most test functions because the jump term is hard to handle analytically. To get around this difficulty, we now describe how to engineer $f \in \mathcal{D}$ so that the jump terms above vanish. Roughly speaking, the idea is to ensure that

$$\mathbb{E}[f(X(s))|X(s-) = x] = f(x)$$

Downloaded from informs.org by [216.73.217.10] on 05 November 2026, at 13:43 . For personal use only, all rights reserved.

at each jump time s and for each state x, i.e. the test function remains unchanged in expectation after a jump. In Section 3.2, we describe a family of exponential test functions satisfying the above property, which is also rich enough to asymptotically characterize the distribution of the queue lengths.

It may be helpful for the reader to compare the arguments that follow to the proof of Lemma 2.3 in [29, Appendix A.2], where the author also seeks to find test functions for which the jump terms vanish. In that paper, simultaneous events are handled by an approach that is slightly different from ours. To make the jump terms vanish, we now analyze the terms $f(X_{\delta_m}) - f(X_{\delta_m-})$. If the event δ_m corresponds to the qth external arrival to station $i \in \mathcal{E}$, then

$$X_{\delta_m} = X_{\delta_m-} + (e^{(i)},\ T_{e,i}(q)e^{(i)},\ 0),$$

where $(e^{(i)},\ T_{e,i}(q)e^{(i)},\ 0) \in \mathbb{Z}_+^d \times \mathbb{R}^d \times \mathbb{R}^d$. Since $T_{e,i}(q)$ is independent of X_{δ_m-} and $U_i(q)$, we have

$$\mathbb{E}_\nu\Big[\big(f(X_{\delta_m}) - f(X_{\delta_m-})\big)1(0 < U_i(q) \le t)\Big] = 0 \tag{3.16}$$

if

$$\mathbb{E}\Big[f(\ell + e^i, u + T_{e,i}(q)e^{(i)}, v)\Big] = f(\ell, u, v), \quad i \in \mathcal{E}, \tag{3.17}$$

for every feasible state $(\ell, u, v) \in \mathbb{Z}_+^d \times \mathbb{R}_+^{2d}$ with $u_i = 0$. Similarly, if δ_m corresponds to the qth departure from station $j \in \mathcal{J}$, then

$$X_{\delta_m} = X_{\delta_m-} + (-e^{(j)} + \phi^{(j)}(q),\ 0,\ T_{s,j}(q)e^{(j)}).$$

Since $T_{s,j}(q)$ and $\phi^{(j)}(q)$ are independent of X_{δ_m-} and $S_j(q)$, we have

$$\mathbb{E}_\nu\Big[\big(f(X_{\delta_m}) - f(X_{\delta_m-})\big)1(0 < S_j(q) \le t)\Big] = 0 \tag{3.18}$$

if

$$\mathbb{E}\Big[f(\ell - e^j + \phi^{(j)}(q), u, v + T_{s,j}(q)e^{(j)})\Big] = f(\ell, u, v), \quad j \in \mathcal{J}, \tag{3.19}$$

for every feasible state $(\ell, u, v) \in \mathbb{Z}_+^d \times \mathbb{R}_+^{2d}$ with $\ell_j > 0$ and $v_j = 0$.

We summarize our analysis in the following lemma.

LEMMA 3.1. *Assume that X is positive Harris recurrent and $X(0)$ follows the stationary distribution of X. For any function $f \in \mathcal{D}$ satisfying conditions (3.17) and (3.19), the basic adjoint relationship* (3.15) *reduces to*

$$\sum_{i\in\mathcal{E}} \mathbb{E}\Big[\frac{\partial f}{\partial u_i}(X(0))\Big] + \sum_{j\in\mathcal{J}} \mathbb{E}\Big[\frac{\partial f}{\partial v_j}(X(0))\Big] \tag{3.20}$$

$$-\sum_{j\in\mathcal{J}}\mathbb{E}\Big[\frac{\partial f}{\partial v_j}(X(0))1(L_j(0)=0)\Big]=0.$$

PROOF. Apply (3.16), (3.18), and the definition of $\mathcal{A}$ in (3.11) to (3.15). □

REMARK 3.1. *The idea of engineering a test function to kill the jump terms in* (3.15) *first appeared in the work of Miyazawa [28] for a single server queue, which is further studied for a multiserver queue with heterogeneous servers in [29]. The novel feature of the present paper is the careful consideration of tightness in Section 5. This issue was not present in [28, 29] because those papers deal with a single queue as opposed to a queueing network.*

In the next section, we describe a useful family of exponential test functions that satisfies (3.17) and (3.19).

3.2. *Exponential test functions for a sequence of GJNs.* We consider the sequence of Markov processes $\{X^{(n)},\ n\geq 1\}$ from Section 2.2. Recall our goal, which is to show that the MGF of $Z^{(n)}(\infty)=r_nL^{(n)}(\infty)$ asymptotically satisfies the BAR (2.30) of the approximating SRBM. As a first step, we have already derived a BAR for $X^{(n)}$. We now describe a family of exponential test functions that satisfy conditions (3.17) and (3.19).

For all $n\in\mathbb{Z}_+$, let us first define the truncation function $g^{(n)}:\mathbb{R}\to\mathbb{R}$ as

$$g^{(n)}(y)=\min(y,1/r_n),\quad y\in\mathbb{R}.$$

Now for each $\theta\in\mathbb{R}^d$, $\eta\in\mathbb{R}^d$ and $\zeta\in\mathbb{R}^d$, define $f^{(n)}_{\theta,\eta,\zeta}:\mathbb{Z}^d_+\times\mathbb{R}^{2d}_+\to\mathbb{R}_+$ as

$$f^{(n)}_{\theta,\eta,\zeta}(x)=e^{\langle\theta,\ell\rangle+\sum_{i\in\mathcal{E}}\eta_ig^{(n)}(u_i)+\sum_{j\in\mathcal{J}}\zeta_jg^{(n)}(v_j)},\quad \text{for } x=(\ell,u,v)\in\mathbb{Z}^d_+\times\mathbb{R}^{2d}_+. \tag{3.21}$$

It is not hard to verify that $f^{(n)}_{\theta,\eta,\zeta}\in\mathcal{D}$. Our in (3.21) is meant to resemble a moment generating function. Indeed, if η and ζ were allowed to be independent of θ and if we chose $g^{(n)}(y)=y$, then this family of test functions would characterize the stationary distribution of $X^{(n)}$ via its BAR (3.15). However, applying (3.15) to $f^{(n)}_{\theta,\eta,\zeta}$ is of little practical use, because the jump terms become too complicated to work with.

Instead, we want to choose $f^{(n)}_{\theta,\eta,\zeta}$ to satisfy the conditions of Lemma 3.1, so that we can use (3.20) instead. To do so, we must choose η and ζ as functions of θ (i.e. $\eta=\eta^{(n)}(\theta)$ and $\zeta=\zeta^{(n)}(\theta)$), significantly reducing the

Downloaded from informs.org by [216.73.217.10] on 05 November 2026, at 13:43 . For personal use only, all rights reserved.

size of this family of functions. Following the logic behind (3.17) and (3.19), we choose $\eta^{(n)}(\theta)$ and $\zeta^{(n)}(\theta)$ to satisfy

$$e^{\theta_i}\mathbb{E}\big(e^{\eta_i^{(n)}(\theta_i)g^{(n)}(T_{e,i}^{(n)})}\big) = 1, \qquad i \in \mathcal{E}, \tag{3.22}$$

$$t_j(\theta)\mathbb{E}\big(e^{\zeta_j^{(n)}(\theta)g^{(n)}(T_{s,j}^{(n)})}\big) = 1, \qquad j \in \mathcal{J}, \tag{3.23}$$

where

$$t_j(\theta) = e^{-\theta_j}\Big(\sum_{k\in\mathcal{J}} p_{jk}e^{\theta_k} + p_{j0}\Big).$$

We have rewritten $\eta_i^{(n)}(\theta)$ as $\eta_i^{(n)}(\theta_i)$ because it is independent of θ_k for $k \neq i$. For the remainder of the paper, we write

$$f_\theta^{(n)}(x) = f^{(n)}_{\theta,\eta^{(n)}(\theta),\zeta^{(n)}(\theta)}(x).$$

This reduced family of test functions can only characterize the distribution of $Z^{(n)}(\infty)$ asymptotically as $n \to \infty$, which is enough for our purposes. The following lemma is similar to Lemma 2.3 of [29], and says that $f_\theta^{(n)}$ satisfies the conditions of Lemma 3.1.

LEMMA 3.2. *There exists $M > 0$ such that for all $\theta \in \mathbb{R}^d$ with $\|\theta\| \le M$, the solutions $\eta_i^{(n)}(\theta_i)$ and $\zeta_j^{(n)}(\theta)$ to (3.22)–(3.23) are well defined and finite for all $n \ge 1$, $i \in \mathcal{E}$ and $j \in \mathcal{J}$. Furthermore, conditions (3.17) and (3.19) are satisfied for $f = f_\theta^{(n)}$.*

PROOF. Once we assume that $\eta_i^{(n)}(\theta_i)$ and $\zeta_j^{(n)}(\theta)$ are well defined and finite, the second claim of the lemma becomes trivial to verify by (3.22) and (3.23). We now check that these functions are well defined for all $\theta \in \mathbb{R}^d$ with $\|\theta\| \le M$. The argument to show that $\eta_i^{(n)}(\theta_i)$ is well defined is simple, and is repeated here from Section 2.3 of [29]. For $i \in \mathcal{E}$, let

$$G_{e,i}^{(n)}(y) = \mathbb{E}\big(e^{yg^{(n)}(T_{e,i}^{(n)})}\big), \qquad y \in \mathbb{R},$$

be the moment generating function of $g^{(n)}(T_{e,i}^{(n)})$. Then $G_{e,i}^{(n)}(y)$ exists for all $y \in \mathbb{R}$ because $g^{(n)}(T_{e,i}^{(n)})$ is a bounded random variable for every n. Furthermore, the inverse $(G_{e,i}^{(n)})^{-1}$ exists as $G_{e,i}^{(n)}(y)$ is strictly increasing. Observe however that this inverse is only defined on the interval

$$\Big(\lim_{y\to-\infty} G_{e,i}^{(n)}(y), \lim_{y\to\infty} G_{e,i}^{(n)}(y)\Big) = \big(\mathbb{P}(T_{e,i}^{(n)} = 0), \infty\big).$$

Downloaded from informs.org by [216.73.217.10] on 05 November 2026, at 13:43 . For personal use only, all rights reserved.

Downloaded from informs.org by [216.73.217.10] on 05 November 2026, at 13:43 . For personal use only, all rights reserved.

Hence, (3.22) yields

$$\eta_i^{(n)}(\theta_i) = (G_{e,i}^{(n)})^{-1}(e^{-\theta_i}), \qquad \theta_i \in \big(-\infty, -\log(\mathbb{P}(T_{e,i}^{(n)} = 0))\big),$$

with the convention that $-\log(0) = \infty$. We know that there exists some $M_{e,i} > 0$ such that

$$0 < M_{e,i} < -\log(\mathbb{P}(T_{e,i}^{(n)} = 0)), \quad n \geq 1$$

because $\mathbb{P}(T_{e,i}^{(n)} = 0) \neq 1$ for any $n \geq 1$, and $\lim_{n\to\infty} \mathbb{P}(T_{e,i}^{(n)} = 0) \neq 1$. The former is true because $1/\mathbb{E}(T_{e,i}^{(n)}) = \lambda_{e,i}^{(n)} < \infty$ for each n, and the latter follows from (2.9) and the fact that $\lambda_{e,i} < \infty$ there. Hence, $\eta_i^{(n)}(\theta_i)$ is well defined and finite for all $\theta_i \in (-\infty, M_{e,i})$. We now derive a similar expression for $\zeta_j^{(n)}(\theta)$. For $j \in \mathcal{J}$ and $y \in \mathbb{R}$, define

$$G_{s,j}^{(n)}(y) = \mathbb{E}\big(e^{yg^{(n)}(T_{s,j}^{(n)})}\big),$$

and observe that $(G_{s,j}^{(n)})^{-1}$ exists because $G_{s,j}^{(n)}(y)$ is an increasing function. Define $\chi_j^{(n)}(y)$ as

$$\chi_j^{(n)}(y) = (G_{s,j}^{(n)})^{-1}(e^{-y}), \qquad y \in \big(-\infty, -\log(\mathbb{P}(T_{s,j}^{(n)} = 0))\big), \tag{3.24}$$

and let $M_{s,j} > 0$ be such that

$$0 < M_{s,j} < -\log(\mathbb{P}(T_{s,j}^{(n)} = 0), \quad n \geq 1.$$

We conclude that $\zeta_j^{(n)}(\theta)$ is well defined and satisfies

$$\zeta_j^{(n)}(\theta) = \chi_j^{(n)}(\log t_j(\theta)) = (G_{s,j}^{(n)})^{-1}\Big(\frac{1}{t_j(\theta)}\Big), \quad \text{for } \theta \text{ such that } \log t_j(\theta) < M_{s,j}, \tag{3.25}$$

where $t_j(\theta)$ is defined in (3.23). □

We now present the BAR for this special family of exponential test functions.

LEMMA 3.3. *Assume that $X^{(n)}$ is positive Harris recurrent, and let*

$$X^{(n)}(\infty) = (L^{(n)}(\infty), R_e^{(n)}(\infty), R_s^{(n)}(\infty))$$

Downloaded from informs.org by [216.73.217.10] on 05 November 2026, at 13:43 . For personal use only, all rights reserved.

have the stationary distribution of $X^{(n)}$. *Then*

$$\begin{aligned}(3.26)\quad &\sum_{i\in\mathcal{E}}\eta_i^{(n)}(\theta_i)\mathbb{E}\big[1(R_{e,i}^{(n)}(\infty)<1/r_n)f_\theta^{(n)}(X^{(n)}(\infty))\big]\\ &\quad+\sum_{j\in\mathcal{J}}\zeta_j^{(n)}(\theta)\mathbb{E}\big[1(R_{s,j}^{(n)}(\infty)<1/r_n)f_\theta^{(n)}(X^{(n)}(\infty))\big]\\ &\quad-\sum_{j\in\mathcal{J}}\zeta_j^{(n)}(\theta)\mathbb{E}\big[1(R_{s,j}^{(n)}(\infty)<1/r_n, L_j^{(n)}(\infty)=0)f_\theta^{(n)}(X^{(n)}(\infty))\big]\\ &=0\end{aligned}$$

for all $\|\theta\|\le M$, *where* $M>0$ *is as in Lemma 3.2.*

PROOF. This lemma follows immediately from applying Lemmas 3.1 and 3.2 to the test function $f_\theta^{(n)}(x)$ from (3.21), because its right side partial derivatives are

$$\begin{aligned}\frac{\partial}{\partial u_i}f_\theta^{(n)}(x)&=\eta_i^{(n)}(\theta_i)1(u_i<1/r_n)f_\theta^{(n)}(x),\quad i\in\mathcal{E},\\ \frac{\partial}{\partial v_j}f_\theta^{(n)}(x)&=\zeta_j^{(n)}(\theta)1(v_j<1/r_n)f_\theta^{(n)}(x),\quad j\in\mathcal{J}.\end{aligned}$$

□

In order for Lemma 3.3 to be of practical use, we need to know the behavior of $\eta_i^{(n)}(\theta_i)$ and $\zeta_j^{(n)}(\theta)$; however, these functions are defined implicitly. In the next section, we use quadratic approximations of $\eta_i^{(n)}(\theta_i)$ and $\zeta_j^{(n)}(\theta)$ to convert (3.26) into a more convenient expression that resembles (2.30).

4. Approximate BAR. This section is devoted to proving Proposition 4.1, which we now state.

PROPOSITION 4.1. *Assume that all conditions stated in Theorem 2.1 are satisfied. Recall the definitions of* $\gamma(\theta)$ *and* $\gamma_j(\theta)$ *from (2.27). For* $j\in\mathcal{J}$ *and* $\theta\le 0$, *define*

$$\begin{aligned}(4.1)\qquad \varphi^{(n)}(\theta)&=\mathbb{E}[e^{\langle\theta,Z^{(n)}(\infty)\rangle}],\\ \varphi_j^{(n)}(\theta)&=\mathbb{E}[e^{\langle\theta,Z^{(n)}(\infty)\rangle}|Z_j^{(n)}(\infty)=0].\end{aligned}$$

Let

$$(4.2)\qquad \epsilon^{(n)}(\theta)=\gamma(\theta)\varphi^{(n)}(\theta)+\sum_{j\in\mathcal{J}}b_j\gamma_j(\theta)\varphi_j^{(n)}(\theta)\quad n\ge 1,\ \theta\le 0,$$

then

$$\lim_{n\to\infty} \sup_{\substack{\theta<0 \\ 0<\|\theta\|\le M}} \frac{|\epsilon^{(n)}(\theta)|}{\|\theta\|} = 0,$$

where $M > 0$ is as in Lemma 3.2.

This result states that the steady state distributions of the queue lengths asymptotically satisfy (2.30) as $n \to \infty$. It plays an essential role in the proof of Theorem 2.1 in Section 6.

The idea behind the proof is to use (3.26) as a starting point, and use quadratic approximations of $\eta_i^{(n)}(\theta_i)$ and $\zeta_j^{(n)}(\theta)$ to arrive at (4.2). In Section 4.1 we use Taylor expansion to obtain the quadratic approximations of $\eta_i^{(n)}(\theta_i)$ and $\zeta_j^{(n)}(\theta)$. We then show that under heavy traffic scaling, the associated approximation error vanishes in an appropriate fashion. We prove Lemma 4.1 in Section 4.3.

4.1. *Taylor expansions.* We begin with a general lemma, which describes the behavior of functions that are implicitly defined in a manner similar to $\eta_i^{(n)}(\theta_i)$ in (3.22). The lemma is similar to Lemma 2.4 of [29].

LEMMA 4.1. *Let H be a bounded, non-negative random variable with $\mathbb{E}H > 0$ and set*

$$\lambda_H = \frac{1}{\mathbb{E}H}, \quad \sigma_H^2 = \mathrm{Var}(H).$$

Then the function $f(x) : \mathbb{R} \to \mathbb{R}$ satisfying

$$\mathbb{E}(e^{f(x)H}) = e^{-x}, \quad x \in \big(-\infty, -\log(\mathbb{P}(H = 0))\big),$$

is well defined and finite, with the convention that $-\log(0) = \infty$. Furthermore,

(a) $f(x)$ is infinitely differentiable.
(b) $f(x)$ is decreasing and concave.
(c) For any $K \in (0, -\log(\mathbb{P}(H = 0))$, $f(x)$ satisfies

$$\Big|f(x) + \lambda_H x + \frac{1}{2}\lambda_H^3 \sigma_H^2 x^2\Big| \le c_H(x), \tag{4.3}$$

$$|f(x)| \le \max\{\lambda_H, |f(K)|/K\}|x|, \quad |x| \le K, \tag{4.4}$$

where

$$c_H(x) = \frac{x^2}{2} \sup_{|y|<|x|} |f''(y) - f''(0)|. \tag{4.5}$$

Downloaded from informs.org by [216.73.217.10] on 05 November 2026, at 13:43 . For personal use only, all rights reserved.

Downloaded from informs.org by [216.73.217.10] on 05 November 2026, at 13:43 . For personal use only, all rights reserved.

The proof of this lemma is deferred to Section A.1. We now apply it to obtain expansions of $\eta_i^{(n)}(\theta_i)$ and $\zeta_j^{(n)}(\theta)$. Let $M > 0$ be as in Lemma 3.2. Set

$$\frac{1}{\widetilde{\lambda}_{e,i}^{(n)}} = \mathbb{E}(g^{(n)}(T_{e,i}^{(n)})), \quad \big(\widetilde{\sigma}_{e,i}^{(n)}\big)^2 = \mathrm{Var}(g^{(n)}(T_{e,i}^{(n)})), \quad i \in \mathcal{E},$$
$$\frac{1}{\widetilde{\lambda}_{s,j}^{(n)}} = \mathbb{E}(g^{(n)}(T_{s,j}^{(n)})), \quad \big(\widetilde{\sigma}_{s,j}^{(n)}\big)^2 = \mathrm{Var}(g^{(n)}(T_{s,j}^{(n)})), \quad j \in \mathcal{J}.$$

Observe that the uniform integrability assumptions (2.11) and (2.12), together with the definition of $g^{(n)}(x)$, imply that for all $i \in \mathcal{E}$ and $j \in \mathcal{J}$,

$$\widetilde{\lambda}_{e,i}^{(n)} \to \lambda_{e,i}, \quad \widetilde{\lambda}_{s,j}^{(n)} \to \lambda_{s,j}, \quad \widetilde{\sigma}_{e,i}^{(n)} \to \sigma_{e,i}, \quad \text{and} \quad \widetilde{\sigma}_{s,j}^{(n)} \to \sigma_{s,j} \tag{4.6}$$

(cf. (2.9), (2.10) and (2.13)). Lemma 4.1 now trivially applies to $\eta_i^{(n)}(\theta_i)$ to obtain

$$\Big|\eta_i^{(n)}(\theta_i) + \widetilde{\lambda}_{e,i}^{(n)}\theta_i + \frac{1}{2}\big(\widetilde{\lambda}_{e,i}^{(n)}\big)^3\big(\widetilde{\sigma}_{e,i}^{(n)}\big)^2\theta_i^2\Big| \le c_{e,i}^{(n)}(\theta_i), \tag{4.7}$$

$$|\eta_i^{(n)}(\theta_i)| \le \hat{c}_{e,i}^{(n)}(M)|\theta_i|, \quad |\theta_i| \le M, \tag{4.8}$$

where $c_{e,i}^{(n)}$ corresponds to c_H in (4.5) and for any $K \in (0, M]$, we define $\hat{c}_{e,i}^{(n)}(K)$ as

$$\hat{c}_{e,i}^{(n)}(K) = \max\big\{\widetilde{\lambda}_{e,i}^{(n)}, |\eta_i^{(n)}(K)|/K\big\}.$$

Recall that $\zeta_j^{(n)}(\theta) = \chi_j^{(n)}(\log t_j(\theta))$, where $\chi_j^{(n)}$ is defined in (3.24) and

$$\log t_j(\theta) = -\theta_j + \log\Big(\sum_{k\in\mathcal{J}} p_{jk}e^{\theta_k} + p_{j0}\Big) \in C^\infty(\mathbb{R}^d).$$

By Taylor expansion, one can verify that $\log t_j(\theta)$ satisfies

$$\begin{aligned}\Big|\log t_j(\theta) - \Big(-\theta_j + \sum_{k\in\mathcal{J}} p_{jk}\theta_k\Big) \\ + \frac{1}{2}\Big(\sum_{k\in\mathcal{J}}\theta_k^2 p_{jk} - \big(\sum_{k\in\mathcal{J}}\theta_k p_{jk}\big)^2\Big)\Big| \le c_{1,j}(\theta),\end{aligned} \tag{4.9}$$

$$\Big|\log^2 t_j(\theta) - \Big(-\theta_j + \sum_{k\in\mathcal{J}} p_{jk}\theta_k\Big)^2\Big| \le c_{2,j}(\theta),$$

where

$$c_{1,j}(\theta) = \frac{1}{6} \sup_{\|y\| \le \|\theta\|} \left| \sum_{k,l,m \in \mathcal{J}} \theta_k \theta_l \theta_m \frac{\partial^3 \log t_j}{\partial \theta_k \partial \theta_l \partial \theta_m}(y) \right|, \tag{4.10}$$

$$c_{2,j}(\theta) = \frac{1}{6} \sup_{\|y\| \le \|\theta\|} \left| \sum_{k,l,m \in \mathcal{J}} \theta_k \theta_l \theta_m \frac{\partial^3 \log^2 t_j}{\partial \theta_k \partial \theta_l \partial \theta_m}(y) \right|.$$

Note that both $c_{1,j}(\theta)$ and $c_{2,j}(\theta)$ are finite for each θ and $j \in \mathcal{J}$ because $\log t_j(\theta)$ belongs to $C^\infty(\mathbb{R}^d)$.

Applying Lemma 4.1 to $\chi_j^{(n)}$, together with the Taylor expansions above, we have the following result about $\zeta_j^{(n)}(\theta)$. For each $j \in \mathcal{J}$ and $n \ge 1$,

$$\begin{aligned} \Big| \zeta_j^{(n)}(\theta) + \widetilde{\lambda}_{s,j}^{(n)} \Big(-\theta_j + \sum_{k \in \mathcal{J}} p_{jk} \theta_k \Big) \\ + \frac{1}{2} \widetilde{\lambda}_{s,j}^{(n)} \Big(\sum_{k \in \mathcal{J}} \theta_k^2 p_{jk} - \big(\sum_{k \in \mathcal{J}} \theta_k p_{jk} \big)^2 \Big) \\ + \frac{1}{2} \big(\widetilde{\lambda}_{s,j}^{(n)} \big)^3 \big(\widetilde{\sigma}_{s,j}^{(n)} \big)^2 \Big(-\theta_j + \sum_{k \in \mathcal{J}} p_{jk} \theta_k \Big)^2 \Big| \le c_{s,j}^{(n)}(\theta), \end{aligned} \tag{4.11}$$

$$|\zeta_j^{(n)}(\theta)| \le \hat{c}_{s,j}^{(n)}(M) \|\theta\|, \quad \|\theta\| \le M, \tag{4.12}$$

where

$$\begin{aligned} c_{s,j}^{(n)}(\theta) = \sup_{|y| \le |\log t_j(\theta)|} \frac{1}{2} \log^2 t_j(\theta) \Big| \big(\chi_j^{(n)} \big)''(y) - \big(\chi_j^{(n)} \big)''(0) \Big| \\ + \widetilde{\lambda}_{s,j}^{(n)} c_{1,j}(\theta) + \frac{1}{2} \big(\widetilde{\lambda}_{s,j}^{(n)} \big)^3 \big(\widetilde{\sigma}_{s,j}^{(n)} \big)^2 c_{2,j}(\theta), \end{aligned} \tag{4.13}$$

$c_{1,j}(\theta)$ and $c_{2,j}(\theta)$ are as in (4.10), and for any $K \in (0, M]$,

$$\begin{aligned} \hat{c}_{s,j}^{(n)}(K) &= c_{Lip,j}(K) \max \big\{ \widetilde{\lambda}_{s,j}^{(n)}, |\chi_j^{(n)}(\widehat{K})| / \widehat{K} \big\}, \\ \widehat{K} &= \sup_{\|\theta\| \le K} |\log t_j(\theta)|, \quad \text{and} \quad c_{Lip,j}(K) = \sup_{0 < \|\theta\| \le K} \frac{|\log t_j(\theta)|}{\|\theta\|}. \end{aligned} \tag{4.14}$$

We know $c_{Lip,j}(K) < \infty$ because $\log t_j(\theta) \in C^\infty(\mathbb{R}^d)$ and is therefore is locally Lipschitz.

4.2. *Error bounds.* In order for the quadratic approximations of $\eta_i^{(n)}(\theta_i)$ and $\zeta_j^{(n)}(\theta)$ to be useful, we need the error bounds $c_{e,i}^{(n)}(\theta)$ and $c_{s,j}^{(n)}(\theta)$ to be

Downloaded from informs.org by [216.73.217.10] on 05 November 2026, at 13:43 . For personal use only, all rights reserved.

Downloaded from informs.org by [216.73.217.10] on 05 November 2026, at 13:43 . For personal use only, all rights reserved.

small. Recall that (3.26) is a statement about the unscaled vector $X^{(n)}(\infty)$, but (4.2) deals with the scaled queue length vector $Z^{(n)}(\infty) = r_n L^{(n)}(\infty)$. From the form of the test function in (3.21), we see that by replacing θ by $r_n\theta$, (3.26) becomes a statement about the scaled queue length $Z^{(n)}(\infty)$. Under the heavy traffic scaling, the errors from the quadratic approximations vanish asymptotically in neighborhoods of the origin. The following lemma presents this statement formally.

LEMMA 4.2. *Let $M > 0$ be as in Lemma 3.2. For any $K > 0$ such that $r_n K \le M$,*

$$\lim_{n\to\infty} \sup_{0<|\theta_i|<K} \frac{c_{e,i}^{(n)}(r_n\theta_i)}{r_n^2\theta_i^2} = 0, \quad i \in \mathcal{E}, \tag{4.15}$$

and

$$\lim_{n\to\infty} \sup_{0<\|\theta\|<K} \frac{c_{s,j}^{(n)}(r_n\theta)}{r_n^2\|\theta\|^2} = 0, \quad j \in \mathcal{J}. \tag{4.16}$$

This lemma is proved in Section A.2. Our next result states that the functions $\eta_i^{(n)}(\theta_i)$ and $\zeta_j^{(n)}(\theta)$ are locally Lipschitz in small neighborhoods of the origin, with Lipschitz constants that do not depend on n. Its proof is postponed to Section A.3.

LEMMA 4.3. *Let $M > 0$ be as in Lemma 3.2. For $i \in \mathcal{E}$ and $j \in \mathcal{J}$, and $K \in (0, r_n M]$,*

$$\sup_{n\ge1} \hat{c}_{e,i}^{(n)}(r_n K) < \infty \quad \text{and} \quad \sup_{n\ge1} \hat{c}_{s,j}^{(n)}(r_n K) < \infty.$$

In the rest of this section, we will frequently use the following bound. Recall the definition of $f_{r_n\theta}^{(n)}(x)$ from (3.21), where $x = (\ell, u, v) \in \mathbb{Z}_+^d \times \mathbb{R}_+^{2d}$. Using (4.8), (4.12), and Lemma 4.3, it follows that for any $K \in (0, M]$, we can define $c_f(K)$ to satisfy

$$\begin{aligned} &\sup_{n\ge1} \sup_{\substack{\theta<0\\ \|\theta\|\le K}} f_{r_n\theta}^{(n)}(x) \\ &\quad = \sup_{n\ge1} \sup_{\substack{\theta<0\\ \|\theta\|\le K}} e^{\langle r_n\theta,\ell\rangle + \sum_{i\in\mathcal{E}} \eta_i^{(n)}(r_n\theta_i) g^{(n)}(u_i) + \sum_{j\in\mathcal{J}} \zeta_j^{(n)}(r_n\theta) g^{(n)}(v_j)} \\ &\quad \le c_f(K) \equiv \sup_{n\ge1} e^{\sum_{i\in\mathcal{E}} \hat{c}_{e,i}^{(n)}(r_n K)K + \sum_{j\in\mathcal{J}} \hat{c}_{s,j}^{(n)}(r_n K)K} < \infty. \end{aligned} \tag{4.17}$$

Downloaded from informs.org by [216.73.217.10] on 05 November 2026, at 13:43 . For personal use only, all rights reserved.

This bound holds for all $x \in \mathbb{Z}_+^d \times \mathbb{R}_+^{2d}$. We now state several lemmas that we will use to prove Proposition 4.1.

LEMMA 4.4. *Recall the heavy traffic condition* (2.5). *For any station* $j \in \mathcal{J}$,

$$\mathbb{P}(L_j^{(n)}(\infty) = 0) = 1 - \lambda_{a,j}^{(n)}/\lambda_{s,j}^{(n)} = r_n b_j/\lambda_{s,j}^{(n)}.$$

LEMMA 4.5. *The sequences* $\{R_{e,i}^{(n)}(\infty), n \geq 1\}$ *and* $\{R_{s,j}^{(n)}(\infty), n \geq 1\}$ *are uniformly integrable for all* $i \in \mathcal{E}$ *and* $j \in \mathcal{J}$.

We would like to point out that Lemmas 4.4 and 4.5 are not novel, and can be proved using the well developed theory of Palm calculus (see for example [1, Chapter 1] or [3, Chapter 4]). However, to keep this paper self-contained, we avoid using Palm calculus and prove these lemmas in Section A.4 using the BAR (3.15).

LEMMA 4.6. *Let* $M > 0$ *be as in Lemma 3.2, and* $K \in (0, r_n M]$. *The following statements are true:*

$$\lim_{n\to\infty} \sup_{\substack{\theta<0 \\ 0<\|\theta\|\leq K}} \frac{1}{\|\theta\|} \Big| - \frac{1}{r_n^2} \Big(\sum_{i\in\mathcal{E}} \eta_i^{(n)}(r_n\theta_i) + \sum_{j\in\mathcal{J}} \zeta_j^{(n)}(r_n\theta) \Big) - \gamma(\theta) \Big| = 0, \tag{4.18}$$

$$\lim_{n\to\infty} \sup_{\substack{\theta<0 \\ \|\theta\|\leq K}} \Big| \mathbb{E}\big[f_{r_n\theta}^{(n)}(X^{(n)}(\infty))\big] - \varphi^{(n)}(\theta) \Big| = 0, \tag{4.19}$$

$$\lim_{n\to\infty} \sup_{\substack{\theta<0 \\ \|\theta\|\leq K}} \frac{1}{r_n} \Big| \mathbb{E}\big[1(L_j^{(n)}(\infty) = 0)\big(f_{r_n\theta}^{(n)}(X^{(n)}(\infty)) - e^{\langle\theta, Z^{(n)}(\infty)\rangle}\big)\big] \Big| = 0, \qquad j \in \mathcal{J}. \tag{4.20}$$

The proof of this lemma is postponed until Section A.5.

4.3. *Proof of Lemma 4.1.* We now prove Proposition 4.1 using the auxiliary lemmas stated in the previous section. As a starting point, we multiply (3.26) by $-1/r_n^2$ to see that

$$\begin{aligned} 0 = &-\frac{1}{r_n^2}\mathbb{E}\big[f_{r_n\theta}^{(n)}(X^{(n)}(\infty))\big] \Big(\sum_{i\in\mathcal{E}} \eta_i^{(n)}(r_n\theta_i) + \sum_{j\in\mathcal{J}} \zeta_j^{(n)}(r_n\theta) \Big) \\ &+ \frac{1}{r_n^2} \sum_{j\in\mathcal{J}} \zeta_j^{(n)}(r_n\theta)\mathbb{E}\big[1(L_j^{(n)}(\infty) = 0) f_{r_n\theta}^{(n)}(X^{(n)}(\infty))\big] \end{aligned} \tag{4.21}$$

$$
\begin{aligned}
&- \frac{1}{r_n^2} \sum_{j \in \mathcal{J}} \zeta_j^{(n)}(r_n\theta) \\
&\qquad \times \mathbb{E}\big[1(R_{s,j}^{(n)}(\infty) \geq 1/r_n, L_j^{(n)}(\infty) = 0) f_{r_n\theta}^{(n)}(X^{(n)}(\infty))\big] \\
&+ \frac{1}{r_n^2} \sum_{i \in \mathcal{E}} \eta_i^{(n)}(r_n\theta_i) \mathbb{E}\big[1(R_{e,i}^{(n)}(\infty) \geq 1/r_n) f_{r_n\theta}^{(n)}(X^{(n)}(\infty))\big] \\
&+ \frac{1}{r_n^2} \sum_{j \in \mathcal{J}} \zeta_j^{(n)}(r_n\theta) \mathbb{E}\big[1(R_{s,j}^{(n)}(\infty) \geq 1/r_n) f_{r_n\theta}^{(n)}(X^{(n)}(\infty))\big].
\end{aligned}
$$

We claim that the last three lines are negligible. For this, we wish to show that

$$
\begin{aligned}
&\lim_{n\to\infty} \sup_{\substack{\theta<0 \\ 0<\|\theta\|\leq M}} \frac{1}{\|\theta\|} \frac{1}{r_n^2} \Big| \sum_{j \in \mathcal{J}} \zeta_j^{(n)}(r_n\theta) \mathbb{E}\big[1(R_{s,j}^{(n)}(\infty) \geq 1/r_n) f_{r_n\theta}^{(n)}(X^{(n)}(\infty))\big] \Big| \\
&\quad = 0,
\end{aligned}
$$

then similar statements hold for the remaining two lines. Observe that for each $j \in \mathcal{J}$,

$$
\begin{aligned}
&\sup_{\substack{\theta<0 \\ 0<\|\theta\|\leq M}} \frac{1}{r_n^2\|\theta\|} \Big| \zeta_j^{(n)}(r_n\theta) \mathbb{E}\big[1(R_{s,j}^{(n)}(\infty) \geq 1/r_n) f_{r_n\theta}^{(n)}(X^{(n)}(\infty))\big] \Big| \\
&\qquad \leq \hat{c}_{s,j}^{(n)}(r_nM) c_f(M) \frac{1}{r_n} \mathbb{P}(R_{s,j}^{(n)}(\infty) \geq 1/r_n)
\end{aligned}
$$

by (4.12) and (4.17). By Lemma 4.3 and Lemma 4.5, this upper bound vanishes as $n \to \infty$. Thus, we have succeeded in proving that

$$
\begin{aligned}
&\lim_{n\to\infty} \sup_{\substack{\theta<0 \\ 0<\|\theta\|\leq M}} \frac{1}{\|\theta\|} \Big| \sum_{j \in \mathcal{J}} \frac{1}{r_n^2} \zeta_j^{(n)}(r_n\theta) \mathbb{E}\big[f_{r_n\theta}^{(n)}(X^{(n)}(\infty)) 1(L_j^{(n)}(\infty) = 0)\big] \\
&\qquad - \frac{1}{r_n^2} \Big(\sum_{i \in \mathcal{E}} \eta_i^{(n)}(r_n\theta_i) + \sum_{j \in \mathcal{J}} \zeta_j^{(n)}(r_n\theta) \Big) \mathbb{E}\big[f_{r_n\theta}^{(n)}(X^{(n)}(\infty))\big] \Big| \\
&\quad = 0.
\end{aligned}
$$

For the next step, we apply (4.17) and Lemma 4.6 to see that

$$
\begin{aligned}
&\lim_{n\to\infty} \sup_{\substack{\theta<0 \\ 0<\|\theta\|\leq M}} \frac{1}{\|\theta\|} \Big| - \frac{1}{r_n^2} \Big(\sum_{i \in \mathcal{E}} \eta_i^{(n)}(r_n\theta_i) + \sum_{j \in \mathcal{J}} \zeta_j^{(n)}(r_n\theta) \Big) \\
&\qquad \times \mathbb{E}\big[f_{r_n\theta}^{(n)}(X^{(n)}(\infty))\big] - \gamma(\theta)\varphi^{(n)}(\theta) \Big|
\end{aligned}
$$

Downloaded from informs.org by [216.73.217.10] on 05 November 2026, at 13:43 . For personal use only, all rights reserved.

$$\leq \lim_{n\to\infty} \sup_{\substack{\theta<0\\ 0<\|\theta\|\leq M}} \frac{1}{\|\theta\|}\Big| - \frac{1}{r_n^2}\Big(\sum_{i\in\mathcal{E}} \eta_i^{(n)}(r_n\theta_i) + \sum_{j\in\mathcal{J}} \zeta_j^{(n)}(r_n\theta)\Big) - \gamma(\theta)\Big| \times \mathbb{E}\big[f^{(n)}_{r_n\theta}(X^{(n)}(\infty))\big]$$
$$+ \lim_{n\to\infty} \sup_{\substack{\theta<0\\ 0<\|\theta\|\leq M}} \frac{|\gamma(\theta)|}{\|\theta\|}\Big|\mathbb{E}\big[f^{(n)}_{r_n\theta}(X^{(n)}(\infty))\big] - \varphi^{(n)}(\theta)\Big| = 0.$$

We arrive at the intermediate result

$$\lim_{n\to\infty} \sup_{\substack{\theta<0\\ 0<\|\theta\|\leq M}} \frac{1}{\|\theta\|}\Big|\gamma(\theta)\varphi^{(n)}(\theta) + \sum_{j\in\mathcal{J}} \frac{1}{r_n^2}\zeta_j^{(n)}(r_n\theta)\mathbb{E}\big[f^{(n)}_{r_n\theta}(X^{(n)}(\infty))1(L_j^{(n)}(\infty)=0)\big]\Big| = 0. \tag{4.22}$$

Recall the definition of $\varphi_j^{(n)}(\theta)$ from (4.1) and use the telescoping sum

$$\lim_{n\to\infty} \sup_{\substack{\theta<0\\ 0<\|\theta\|\leq M}} \frac{1}{\|\theta\|}\Big|\frac{1}{r_n^2}\zeta_j^{(n)}(r_n\theta)\mathbb{E}\big[f^{(n)}_{r_n\theta}(X^{(n)}(\infty))1(L_j^{(n)}(\infty)=0)\big] - b_j\gamma_j(\theta)\varphi_j^{(n)}(\theta)\Big|$$
$$\leq \lim_{n\to\infty} \sup_{\substack{\theta<0\\ 0<\|\theta\|\leq M}} \frac{1}{\|\theta\|}\Big|\frac{1}{r_n^2}\zeta_j^{(n)}(r_n\theta) - \frac{b_j}{P(L_j^{(n)}(\infty)=0)}\gamma_j(\theta)\Big| \times \mathbb{E}\big[f^{(n)}_{r_n\theta}(X^{(n)}(\infty))1(L_j^{(n)}(\infty)=0)\big]$$
$$+ \lim_{n\to\infty} \sup_{\substack{\theta<0\\ 0<\|\theta\|\leq M}} \frac{1}{\|\theta\|}\frac{b_j}{P(L_j^{(n)}(\infty)=0)}\gamma_j(\theta) \times \Big|\mathbb{E}\Big[\big[f^{(n)}_{r_n\theta}(X^{(n)}(\infty)) - e^{\langle\theta,Z^{(n)}(\infty)\rangle}\big]1(L_j^{(n)}(\infty)=0)\Big]\Big|.$$

Using (4.22) we see that to complete the proof of Proposition 4.1, all we need to do is show the upper bound above equals zero. To show that the first term is zero, we recall from Lemma 4.4 that

$$\mathbb{P}(L_j^{(n)}(\infty)=0) = 1 - \lambda_{a,j}^{(n)}/\lambda_{s,j}^{(n)} = r_n b_j/\lambda_{s,j}^{(n)}, \quad j\in\mathcal{J}.$$

Recalling the form of $\gamma_j(\theta)$ from (2.27) and the bound $c_f(M)$ from (4.17), we see that

$$\lim_{n\to\infty} \sup_{\substack{\theta<0\\ 0<\|\theta\|\leq M}} \frac{1}{\|\theta\|}\Big|\frac{1}{r_n^2}\zeta_j^{(n)}(r_n\theta) - \frac{b_j}{\mathbb{P}(L_j^{(n)}(\infty)=0)}\gamma_j(\theta)\Big|$$

Downloaded from informs.org by [216.73.217.10] on 05 November 2026, at 13:43 . For personal use only, all rights reserved.

$$\times \mathbb{E}\big[f^{(n)}_{r_n\theta}(X^{(n)}(\infty))1(L^{(n)}_j(\infty)=0)\big]$$

$$= \lim_{n\to\infty} \sup_{\substack{\theta<0\\ 0<\|\theta\|\le M}} \frac{1}{\|\theta\|}\Big|\frac{\mathbb{P}(L^{(n)}_j(\infty)=0)}{r_n^2}\zeta^{(n)}_j(r_n\theta) - b_j\gamma_j(\theta)\Big|$$

$$\times \mathbb{E}\big[f^{(n)}_{r_n\theta}(X^{(n)}(\infty))\big|L^{(n)}_j(\infty)=0\big]$$

$$\le \lim_{n\to\infty} \sup_{\substack{\theta<0\\ 0<\|\theta\|\le M}} \frac{1}{\|\theta\|}\Big|\frac{b_j}{r_n\lambda^{(n)}_{s,j}}\zeta^{(n)}_j(r_n\theta) - b_j\langle R^{(j)},\theta\rangle\Big| c_f(M).$$

Furthermore,

$$\lim_{n\to\infty} \sup_{\substack{\theta<0\\ 0<\|\theta\|\le M}} \frac{1}{\|\theta\|}\Big|\frac{b_j}{r_n\lambda^{(n)}_{s,j}}\zeta^{(n)}_j(r_n\theta) - b_j\langle R^{(j)},\theta\rangle\Big| c_f(M)$$

$$= \lim_{n\to\infty} \sup_{\substack{\theta<0\\ 0<\|\theta\|\le M}} \frac{1}{\|\theta\|}\Big|\frac{\widetilde{\lambda}^{(n)}_{s,j}}{\lambda^{(n)}_{s,j}} b_j\langle R^{(j)},\theta\rangle - b_j\langle R^{(j)},\theta\rangle\Big| c_f(M) = 0,$$

where the first equality is justified by (2.13) together with the approximation of $\zeta^{(n)}_j(r_n\theta)$ from (4.11) and Lemma 4.2, and the second equality follows from (2.13) and (4.6). Now for the second term, Lemmas 4.4 and 4.6 tell us that for $j \in \mathcal{J}$,

$$\lim_{n\to\infty} \sup_{\substack{\theta<0\\ \|\theta\|\le M}} \frac{\Big|\mathbb{E}\big[1(L^{(n)}_j(\infty)=0)\big(f^{(n)}_{r_n\theta}(X^{(n)}(\infty)) - e^{\langle\theta,Z^{(n)}(\infty)\rangle}\big)\big]\Big|}{\mathbb{P}(L^{(n)}_j(\infty)=0)} = 0.$$

This concludes the proof of Proposition 4.1.

5. Tightness of stationary distributions: An essential Proposition. This section is centered around the statement and proof of Proposition 5.1, which is critical to proving that the sequence $\{Z^{(n)}(\infty),\ n \ge 1\}$ is tight. The tightness argument itself is provided as a part the proof of Theorem 2.1 in Section 6, and relies on both Propositions 4.1 and 5.1. We now motivate the proposition and introduce some notation needed to state it.

Let $\mathcal{C}$ be the class of functions f that satisfy

(a) $f : \{\theta \le 0 : \theta \in \mathbb{R}^d\} \to [0,1]$
(b) f is continuous on $\{\theta < 0 : \theta \in \mathbb{R}^d\}$
(c) $\theta_1 \le \theta_2 \Rightarrow f(\theta_1) \le f(\theta_2)$,
(d) $f(0) = 1$.

Downloaded from informs.org by [216.73.217.10] on 05 November 2026, at 13:43 . For personal use only, all rights reserved.

Downloaded from informs.org by [216.73.217.10] on 05 November 2026, at 13:43 . For personal use only, all rights reserved.

Clearly, the MGF of any probability measure on $\mathbb{R}^d_+$ belongs to $\mathcal{C}$. Suppose that f is a pointwise limit of a sequence of MGFs of probability measures. Then, $f \in \mathcal{C}$. Such a pointwise limit is not necessarily the MGF of a probability measure; for example, this happens when the sequence of measures is not tight. One can prove that the pointwise limit f is an MGF of a probability measure if and only if f is left continuous at 0; see Lemma 6.1 in Section 6. By left continuity, we mean

$$\lim_{\theta \uparrow 0} f(\theta) = 1,$$

where $\theta \uparrow 0$ means that $\theta \in \mathbb{R}^d$ approaches 0 from left in arbitrary way.

In the tightness argument of Section 6, we deal with the sequence of MGFs corresponding to $\{Z^{(n)}(\infty),\ n \geq 1\}$. Loosely speaking, Proposition 4.1 tells us that the pointwise limit of every convergent subsequence of the MGFs satisfies the BAR of the SRBM (2.30). Proposition 5.1 then leverages the structure of this BAR to prove that the pointwise limit of the MGFs is left continuous at the origin, thereby proving tightness of the corresponding sequence of probability measures. This argument is made precise in Section 6. Having sufficiently motivated the necessity of Proposition 5.1, we now continue with its setup.

For our purposes, it will be beneficial to take limits as $\theta \uparrow 0$ along rays stemming from the origin. Each ray corresponds to some direction vector $\boldsymbol{c}_A = (c_1, \ldots, c_d)^{\mathrm{T}}$, where $A \subset \mathcal{J}$ and $c_i = 0$ when $i \notin A$. We write $\boldsymbol{c}_A > 0$ if $c_i > 0$ when $i \in A$. For a fixed ray $\boldsymbol{c}_A > 0$, we consider the limits

$$\lim_{\alpha \uparrow 0} f(\alpha \boldsymbol{c}_A), \quad \alpha \in \mathbb{R},$$

which always exist by the monotonicity of f. We note the following fact:

LEMMA 5.1. *Fix $A \subset \mathcal{J}$ and consider any two direction vectors $\boldsymbol{c}_A > 0$ and $\tilde{\boldsymbol{c}}_A > 0$. For any function $f \in \mathcal{C}$,*

$$\lim_{\alpha \uparrow 0} f(\alpha \boldsymbol{c}_A) = \lim_{\alpha \uparrow 0} f(\alpha \tilde{\boldsymbol{c}}_A).$$

PROOF. Since $\boldsymbol{c}_A > 0$ and $\tilde{\boldsymbol{c}}_A > 0$, there exist constants $m, M > 0$ such that

$$m\boldsymbol{c}_A \leq \tilde{\boldsymbol{c}}_A \leq M\boldsymbol{c}_A.$$

Using the monotonicity of $f(\theta)$,

$$\lim_{\alpha \uparrow 0} f(\alpha \boldsymbol{c}_A) = \lim_{\alpha \uparrow 0} f(\alpha M \boldsymbol{c}_A) \leq \lim_{\alpha \uparrow 0} f(\alpha \tilde{\boldsymbol{c}}_A) \leq \lim_{\alpha \uparrow 0} f(\alpha m \boldsymbol{c}_A) = \lim_{\alpha \uparrow 0} f(\alpha \boldsymbol{c}_A). \quad \square$$

Downloaded from informs.org by [216.73.217.10] on 05 November 2026, at 13:43 . For personal use only, all rights reserved.

In view of this lemma, we write

$$f(0_A-) \equiv \lim_{\alpha \uparrow 0} f(\alpha \boldsymbol{c}_A).$$

The following proposition is one of the main tools used to prove Theorem 2.1. It will allow us to show that whenever the sequence of the MGFs of $Z^{(n)}(\infty)$ has a limit, it must be left continuous at 0.

PROPOSITION 5.1. *Assume $\psi \in \mathcal{C}$ and $\{\psi_j, j \in \mathcal{J}\} \subset \mathcal{C}$ satisfy*

$$\gamma(\theta)\psi(\theta) + \sum_{j \in \mathcal{J}} b_j \gamma_j(\theta)\psi_j(\theta) = 0, \quad \text{for } \theta \le 0, \tag{5.1}$$

and that $\psi_j(\theta)$ is independent of θ_j. For any $A \subset \mathcal{J}$,

$$\psi(0_A-) \ge \psi_j(0_A-), \quad j \in \mathcal{J}. \tag{5.2}$$

Furthermore, when $A = \mathcal{J}$ we have

$$\psi(0_{\mathcal{J}}-) = \psi_j(0_{\mathcal{J}}-), \quad j \in \mathcal{J}. \tag{5.3}$$

PROOF. Recall the definitions of $\gamma(\theta)$ and $\gamma_j(\theta)$ from (2.27). For any $A \subset \mathcal{J}$, $\boldsymbol{c}_A > 0$, and $\alpha < 0$, we set $\theta = \alpha \boldsymbol{c}_A$ in (5.1) to get

$$\Big(\frac{1}{2}\alpha^2 \langle \boldsymbol{c}_A, \Sigma \boldsymbol{c}_A \rangle - \alpha \sum_{j \in \mathcal{J}} b_j \langle \boldsymbol{c}_A, R^{(j)} \rangle\Big)\psi(\alpha \boldsymbol{c}_A) + \alpha \sum_{j \in \mathcal{J}} b_j \langle \boldsymbol{c}_A, R^{(j)} \rangle \psi_j(\alpha \boldsymbol{c}_A) = 0.$$

We divide both sides by α and let $\alpha \uparrow 0$, which yields

$$-\sum_{j \in \mathcal{J}} \langle \boldsymbol{c}_A, R^{(j)} \rangle b_j \psi(0_A-) + \sum_{j \in \mathcal{J}} \langle \boldsymbol{c}_A, R^{(j)} \rangle b_j \psi_j(0_A-) = 0.$$

By Lemma 5.1, $\boldsymbol{c}_A = (c_1, ..., c_d)^{\mathrm{T}}$ can be arbitrary as long as $\boldsymbol{c}_A > 0$. For each fixed $i \in A$, we set $c_i = 1$ and let $c_k \downarrow 0$ for $k \in A \setminus \{i\}$. We arrive at

$$\sum_{j \in \mathcal{J}} r_{ij} b_j \big(\psi(0_A-) - \psi_j(0_A-)\big) = 0, \quad i \in A,$$

where r_{ij} is the (i, j)th entry of the reflection matrix R. Next, we split the summation in this formula into two parts:

$$\sum_{j \in A} r_{ij} b_j \big(\psi(0_A-) - \psi_j(0_A-)\big) \tag{5.4}$$

$$+\sum_{j\in\mathcal{J}\setminus A} r_{ij}b_j\big(\psi(0_A-)-\psi_j(0_A-)\big)=0,\quad i\in A.$$

We consider these equations as an $|A|$-dimensional vector equation. Let R^A be the principal sub-matrix of R whose entry indices are taken from A. Since R is an $\mathcal{M}$-matrix, R^A is also an $\mathcal{M}$-matrix, so it has an inverse whose entries are all nonnegative. Denote the (i,j)th entry of the inverse of R^A by $\widetilde{r}^A_{ij}$. Multiplying the vector version of (5.4) from the left by $(R^A)^{-1}$, we have

$$\begin{aligned}(5.5)\qquad & b_k\big(\psi(0_A-)-\psi_k(0_A-)\big)\\ &\quad+\sum_{i\in A}\sum_{j\in\mathcal{J}\setminus A}\widetilde{r}^A_{ki}r_{ij}b_j\big(\psi(0_A-)-\psi_j(0_A-)\big)=0,\qquad k\in A.\end{aligned}$$

Set $A=\mathcal{J}\setminus B$, where $B\subset\mathcal{J}$ and $B\neq\mathcal{J}$. Using induction on the size of the set B, we will show that

$$\psi(0_{\mathcal{J}\setminus B}-)-\psi_k(0_{\mathcal{J}\setminus B}-)\geq 0,\quad k\in\mathcal{J}.$$

We first take $B=\emptyset$, meaning $A=\mathcal{J}$. Since the summation in (5.5) vanishes, we have

$$\psi(0-)=\psi_k(0-),\quad k\in\mathcal{J},\tag{5.6}$$

which proves (5.3). Hence, the base case is true and we now justify the inductive step. For a general set $B\subset\mathcal{J}$, (5.5) becomes

$$\begin{aligned}(5.7)\qquad & b_k\big(\psi(0_{\mathcal{J}\setminus B}-)-\psi_k(0_{\mathcal{J}\setminus B}-)\big)\\ &\quad+\sum_{i\in\mathcal{J}\setminus B}\sum_{j\in B}\widetilde{r}^{\mathcal{J}\setminus B}_{ki}r_{ij}b_j\big(\psi(0_{\mathcal{J}\setminus B}-)-\psi_j(0_{\mathcal{J}\setminus B}-)\big)=0\end{aligned}$$

for each $k\in\mathcal{J}\setminus B$. For $j\in B$ we have

$$\begin{aligned}\psi(0_{\mathcal{J}\setminus B}-)-\psi_j(0_{\mathcal{J}\setminus B}-)&=\psi(0_{\mathcal{J}\setminus B}-)-\psi_j(0_{\mathcal{J}\setminus(B\setminus\{j\})}-)\\ &\geq\psi(0_{\mathcal{J}\setminus(B\setminus\{j\})}-)-\psi_j(0_{\mathcal{J}\setminus(B\setminus\{j\})}-),\end{aligned}$$

where

$$\psi(0_{\mathcal{J}\setminus B}-)\geq\psi(0_{\mathcal{J}\setminus(B\setminus\{j\})}-)$$

follows because $\psi\in\mathcal{C}$. We now assume that

$$\psi(0_{\mathcal{J}\setminus(B\setminus\{j\})}-)-\psi_k(0_{\mathcal{J}\setminus(B\setminus\{j\})}-)\geq 0,\quad k\in\mathcal{J}$$

Downloaded from informs.org by [216.73.217.10] on 05 November 2026, at 13:43 . For personal use only, all rights reserved.

Downloaded from informs.org by [216.73.217.10] on 05 November 2026, at 13:43 . For personal use only, all rights reserved.

to immediately obtain

$$\psi(0_{\mathcal{J}\setminus B}-) - \psi_j(0_{\mathcal{J}\setminus B}-) \geq 0, \quad j \in B. \tag{5.8}$$

Note that $r_{ij} \leq 0$ for $j \neq i$ and $\widetilde{r}_{ki}^{\mathcal{J}\setminus B} \geq 0$ because R and R^A are $\mathcal{M}$-matrices. Hence,

$$\widetilde{r}_{ki}^{\mathcal{J}\setminus B} r_{ij} b_j \big(\psi(0_{\mathcal{J}\setminus B}-) - \psi_j(0_{\mathcal{J}\setminus B}-)\big) \leq 0, \quad j \in B,$$

which together with (5.7) immediately gives us

$$\psi(0_{\mathcal{J}\setminus B}-) - \psi_k(0_{\mathcal{J}\setminus B}-) \geq 0, \quad k \in \mathcal{J}\setminus B. \tag{5.9}$$

We combine (5.8) and (5.9) to complete the proof. □

6. Proof of Theorem 2.1. For $n \geq 1$, recall that $\varphi^{(n)}(\theta)$ and $\varphi_j^{(n)}(\theta)$ are MGFs defined in (4.1). To prove Theorem 2.1, it suffices to prove

$$\begin{aligned} &\lim_{n\to\infty} \varphi^{(n)}(\theta) = \varphi(\theta) \\ &\lim_{n\to\infty} \varphi_j^{(n)}(\theta) = \varphi_j(\theta), \quad j \in \mathcal{J}, \qquad \text{for } \theta \leq 0, \end{aligned} \tag{6.1}$$

where φ and φ_j are MGFs defined in (2.26).

We now state a lemma that will be used frequently in this section. Its proof is given in Section B.

LEMMA 6.1. *Let $\{\nu^{(n)}, n \geq 1\}$ be a sequence of probability measures on $\mathbb{R}_+^d$ with corresponding MGFs*

$$f^{(n)}(\theta) = \int_{\mathbb{R}_+^d} e^{\langle\theta,x\rangle} \nu^{(n)}(dx), \qquad \text{for } \theta \in \mathbb{R}^d \text{ with } \theta \leq 0.$$

Suppose that there exists a function $f(\theta)$ such that

$$f^{(n)}(\theta) \to f(\theta) \qquad \text{pointwise for all } \theta \in \mathbb{R}^d \text{ with } \theta \leq 0.$$

(a) $f(0_{\mathcal{J}}-) = 1$ if and only if $\{\nu^{(n)}, n \geq 1\}$ is tight.
(b) If $f(0_{\mathcal{J}}-) = 1$, then $f(\theta)$ is the MGF of some probability measure ν on $\mathbb{R}_+^d$, which immediately implies that

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

Downloaded from informs.org by [216.73.217.10] on 05 November 2026, at 13:43 . For personal use only, all rights reserved.

PROOF OF THEOREM 2.1. It follows from the weak compactness of a set of probability distributions (e.g. see [34, Lemma 1]) that for every subsequence $\{n'\} \subset \{n\}_{n=1}^{\infty}$, there exists a further subsequence $\{n''\}$ and some non-decreasing functions $\psi(\theta)$ and $\psi_j(\theta)$, such that for every $\theta \leq 0$,

$$\varphi^{(n'')}(\theta) \to \psi(\theta) \quad \text{and} \quad \varphi_j^{(n'')}(\theta) \to \psi_j(\theta), \quad j \in \mathcal{J}. \tag{6.2}$$

Both $\psi(\theta)$ and $\psi_j(\theta)$ are functions in $\mathcal{C}$, which was introduced in Section 5. In particular we note that $\psi(0) = \psi_j(0) = 1$. Furthermore, Proposition 4.1 implies that $\psi(\theta)$ and $\psi_j(\theta)$ satisfy BAR (5.1). Suppose we know that the limiting functions ψ and ψ_j satisfy

$$\psi(0_{\mathcal{J}}-) = 1, \tag{6.3}$$

$$\psi_j(0_{\mathcal{J}}-) = 1, \quad j \in \mathcal{J}. \tag{6.4}$$

Then Lemma 6.1, BAR (5.1) and the MGF version of Lemma 2.1 imply that $\psi(\theta) = \varphi(\theta)$ and $\psi_j(\theta) = \varphi_j(\theta)$. Since $\{n'\}$ was an arbitrary subsequence, we have proved (6.1), and consequently Theorem 2.1.

In the remainder of the proof, we prove (6.3), which implies (6.4) by (5.3). Let $A = \{j\}$, then (5.2) implies

$$\psi(0_{\{j\}}-) - \psi_j(0_{\{j\}}-) \geq 0, \quad j \in \mathcal{J}. \tag{6.5}$$

Since $\psi_j(0_{\{j\}}-) = \psi_j(0) = 1$, we have

$$\psi(0_{\{j\}}-) = 1, \quad j \in \mathcal{J}. \tag{6.6}$$

Recall that

$$Z^{(n'')}(\infty) = (Z_1^{(n'')}(\infty), \ldots, Z_d^{(n'')}(\infty))$$

is the scaled steady state queue length vector of the n''th GJN. Applying Lemma 6.1 to (6.6), we see that $\{Z_j^{(n'')}(\infty)\}$ is tight for each $j \in \mathcal{J}$. Since all the marginals are tight, $\{Z^{(n'')}(\infty)\}$ is tight as well. We invoke Lemma 6.1 again to conclude that (6.3) holds. □

7. Concluding remarks. To summarize, this paper uses a novel approach to justify the steady-state diffusion approximation of GJNs. Our method does not rely on first using the process limit $Z^{(n)}(\cdot) \Rightarrow Z(\cdot)$, followed by the tightness of $\{Z^{(n)}(\infty),\ n \geq 1\}$ as in [17] or [7]. Rather, we work directly with the basic adjoint relationships of the GJN and the SRBM, which characterize their respective stationary distributions. By applying the

Downloaded from informs.org by [216.73.217.10] on 05 November 2026, at 13:43 . For personal use only, all rights reserved.

BAR (3.15) of the GJN to a carefully designed exponential test function, we show in Proposition 4.1 that the sequence of moment generating functions of the GJN (together with a corresponding sequence of boundary measures) asymptotically satisfies the BAR (2.30) of the SRBM.

In addition to Proposition 4.1, we also require tightness of the sequence $\{Z^{(n)}(\infty),\ n \geq 1\}$ like in [17] and [7]. However, unlike the previous two papers, we do not rely on constructing Lyapunov functions to prove this tightness. Our tightness argument from Section 6, relies on algebraic manipulations of the BAR (2.30) of the SRBM, the bulk of which are presented as Proposition 5.1. A critical condition for this proposition is that the reflection matrix of the SRBM is an $\mathcal{M}$-matrix, and in particular, that its diagonal entries are non-negative and off diagonal entries are non-positive.

An important direction for future research is generalizing the methodology presented in this paper to the multiclass queueing network setting, where new sources of difficulty emerge. One source of difficulty is the need to handle state space collapse. In a multiclass network with d stations, the vector of queue lengths is of a higher dimension than d, while the approximating diffusion process remains d-dimensional. This will manifest itself as an extra error term on the right hand side of (4.2) in Proposition 4.1, and we would need to have some way to deal with it. Another source of difficulty is to relax the $\mathcal{M}$-matrix currently needed to prove tightness, as the $\mathcal{M}$-matrix structure is a feature unique to GJNs. Preliminary experiments with several multiclass networks suggest that a variant of Proposition 5.1 can be established even when the reflection matrix is not an $\mathcal{M}$-matrix. However, it remains an open problem to find the minimal conditions on the reflection matrix under which our approach can be carried out.

APPENDIX A: PROOFS OF LEMMAS IN SECTION 4

In this section, we prove the lemmas that were stated but not proved in Section 4.

A.1. Proof of Lemma 4.1. The fact that $f(x)$ is well defined and finite is argued in the same way as in Lemma 3.2. We want to perform Taylor expansion on $f(x)$ and to do so we first show that it is infinitely differentiable. Observe that the function

$$G(y) = \mathbb{E}(e^{yH}), \qquad y \in \mathbb{R},$$

is well defined and belongs to $C^\infty(\mathbb{R})$. Hence, $F(x, y) \equiv G(y) - e^{-x}$ belongs to $C^\infty(\mathbb{R}^2)$. Since

$$\frac{\partial}{\partial y} F(x, y) = \mathbb{E}(He^{yH}) \neq 0, \qquad (x, y) \in \mathbb{R}^2,$$

Downloaded from informs.org by [216.73.217.10] on 05 November 2026, at 13:43 . For personal use only, all rights reserved.

the implicit function theorem [27, Theorem 3.3.1] implies that there is a single valued and infinitely differentiable function $y(x)$ of x which is the solution of $F(x, y(x)) = 0$ for each $x \in \big(-\infty, -\log(\mathbb{P}(H = 0))\big)$. Therefore $f(x) = y(x)$ for $x \in \big(-\infty, -\log(\mathbb{P}(H = 0))\big)$, proving (a). The first and second derivatives of $f(x)$ are

$$f'(x) = -\frac{\mathbb{E}(e^{f(x)H})}{\mathbb{E}(He^{f(x)H})}, \tag{A.1}$$

$$f''(x) = -f'(x)\left(1 - \frac{\mathbb{E}(e^{f(x)H})\mathbb{E}(H^2e^{f(x)H})}{\big(\mathbb{E}(He^{f(x)H})\big)^2}\right).$$

Observe that $f'(x) < 0$, which implies that $f(x)$ is decreasing. To prove concavity, we wish to show that $f''(x) \le 0$. This follows from

$$\big(E(He^{f(x)H})\big)^2 \le \mathbb{E}(e^{f(x)H})\mathbb{E}(H^2e^{f(x)H}),$$

which is just the Cauchy-Schwarz inequality. A second order Taylor expansion of $f(x)$ combined with the facts that $f(0) = 0$, $f'(0) = -\lambda_H$ and $f''(0) = -\lambda_H^3\sigma_H^2$ immediately implies (4.3). Lastly, since $f(x)$ is concave,

$$f(x) \le f'(0)x \le \lambda_H|x|, \quad x \le 0.$$

Furthermore, the function $|f(x)| = -f(x)$ is convex when restricted to $x \ge 0$. Therefore,

$$|f(x)| \le \frac{|f(K)| - |f(0)|}{K - 0}x = \frac{|f(K)|}{K}x, \quad 0 \le x \le K.$$

A.2. Proof of Lemma 4.2. We begin by proving (4.15), or

$$\lim_{n\to\infty} \sup_{0<|\theta_i|<K} \frac{c_{e,i}^{(n)}(r_n\theta_i)}{r_n^2\theta_i^2} \tag{A.2}$$

$$= \lim_{n\to\infty} \sup_{|y|\le K} \frac{1}{2}\big|\big(\eta_i^{(n)}\big)''(r_n y) - \big(\eta_i^{(n)}\big)''(0)\big| = 0, \qquad i \in \mathcal{E}.$$

Using (A.1), we see that $\big(\eta_i^{(n)}\big)''(\theta_i)$ can be written symbolically as

$$\big(\eta_i^{(n)}\big)''(\theta_i) = \frac{a_1(\theta_i)}{a_2(\theta_i)}\left(1 - \frac{a_1(\theta_i)a_3(\theta_i)}{(a_2(\theta_i))^2}\right),$$

where

$$a_1(\theta_i) = \mathbb{E}\big(e^{\eta_i^{(n)}(\theta_i)g^{(n)}(T_{e,i}^{(n)})}\big) = e^{-\theta_i},$$

$$a_2^{(n)}(\theta_i) = \mathbb{E}\big(g^{(n)}(T_{e,i}^{(n)})e^{\eta_i^{(n)}(\theta_i)g^{(n)}(T_{e,i}^{(n)})}\big),$$
$$a_3^{(n)}(\theta_i) = \mathbb{E}\big((g^{(n)}(T_{e,i}^{(n)}))^2 e^{\eta_i^{(n)}(\theta_i)g^{(n)}(T_{e,i}^{(n)})}\big).$$

To prove (A.2), we want to replace each $a_i(r_n y)$ by $a_i(0)$ one at a time. For this, we need to show that

$$\sup_{\substack{n\geq 1\\ |y|\leq K}} a_1(r_n y) < \infty, \quad \sup_{\substack{n\geq 1\\ |y|\leq K}} a_2^{(n)}(r_n y) < \infty,$$
$$\sup_{\substack{n\geq 1\\ |y|\leq K}} \frac{1}{a_2^{(n)}(r_n y)} < \infty, \quad \sup_{\substack{n\geq 1\\ |y|\leq K}} a_3^{(n)}(r_n y) < \infty,$$

and

$$\lim_{n\to\infty} \sup_{|y|\leq K} \big|a_1(r_n y) - a_1(0)\big| = 0,$$
$$\lim_{n\to\infty} \sup_{|y|\leq K} \big|a_2^{(n)}(r_n y) - a_2^{(n)}(0)\big| = 0, \tag{A.3}$$
$$\lim_{n\to\infty} \sup_{|y|\leq K} \big|a_3^{(n)}(r_n y) - a_3^{(n)}(0)\big| = 0. \tag{A.4}$$

We begin with $a_1(\theta_i) = e^{-\theta_i}$:

$$\sup_{\substack{n\geq 1\\ |y|\leq K}} e^{r_n y} < \infty \quad \text{and} \quad \lim_{n\to\infty} \sup_{|y|\leq K} \big|e^{-r_n y} - 1\big| = 0.$$

Moving on to $a_2^{(n)}(\theta_i) = \mathbb{E}\big(g^{(n)}(T_{e,i}^{(n)})e^{\eta_i^{(n)}(\theta_i)g^{(n)}(T_{e,i}^{(n)})}\big)$; we first prove (A.3). By (4.8), we have

$$|\eta_i^{(n)}(r_n y)g^{(n)}(T_{e,i}^{(n)})| \leq \hat{c}_{e,i}^{(n)}(r_n K)|y|, \quad |y| \leq K. \tag{A.5}$$

Now observe that

$$\begin{aligned}
&\lim_{n\to\infty} \sup_{|y|\leq K} \left|\mathbb{E}g^{(n)}(T_{e,i}^{(n)}) - \mathbb{E}g^{(n)}(T_{e,i}^{(n)})e^{\eta_i^{(n)}(r_n y)g^{(n)}(T_{e,i}^{(n)})}\right| \\
&\leq \lim_{n\to\infty} \sup_{|y|\leq K} \left|\mathbb{E}\big(g^{(n)}(T_{e,i}^{(n)})\big)^2 \eta_i^{(n)}(r_n y)e^{\hat{c}_{e,i}^{(n)}(r_n K)|y|}\right| \\
&\leq \lim_{n\to\infty} r_n K \hat{c}_{e,i}^{(n)}(r_n K)e^{\hat{c}_{e,i}^{(n)}(r_n K)K}\mathbb{E}\big(T_{e,i}^{(n)}\big)^2 = 0.
\end{aligned}$$

Downloaded from informs.org by [216.73.217.10] on 05 November 2026, at 13:43 . For personal use only, all rights reserved.

In the first inequality, we use (A.5) together with the bound

$$|e^x - 1| \le |x| e^{|x|}, \quad x \in \mathbb{R}. \tag{A.6}$$

The second inequality is by (4.8) and the fact that the limit equals zero follows from (2.11) and Lemma 4.3, proving (A.3). By (2.9), we know that

$$\lim_{n\to\infty} a_2^{(n)}(0) = \frac{1}{\lambda_{e,i}} > 0.$$

Combined with (A.3), this implies

$$\sup_{\substack{n\ge 1\\ |y|\le K}} a_2^{(n)}(r_n y) < \infty \quad \text{and} \quad \sup_{\substack{n\ge 1\\ |y|\le K}} \frac{1}{a_2^{(n)}(r_n y)} < \infty.$$

Lastly, we tackle $a_3^{(n)}(\theta_i) = \mathbb{E}\big((g^{(n)}(T_{e,i}^{(n)}))^2 e^{\eta_i^{(n)}(\theta_i) g^{(n)}(T_{e,i}^{(n)})}\big)$, first proving (A.4). We have

$$\begin{aligned}
&\sup_{|y|\le K} \Big| \mathbb{E}\big(g^{(n)}(T_{e,i}^{(n)})\big)^2 e^{\eta_i^{(n)}(r_n y) g^{(n)}(T_{e,i}^{(n)})} - \mathbb{E}\big(g^{(n)}(T_{e,i}^{(n)})\big)^2 \Big| \\
&\quad \le \sup_{|y|\le K} \mathbb{E}\Big| \big(g^{(n)}(T_{e,i}^{(n)})\big)^3 \eta_i^{(n)}(r_n y) e^{\hat{c}_{e,i}^{(n)}(r_n K) K} \Big| \\
&\quad \le r_n K \hat{c}_{e,i}^{(n)}(r_n K) e^{\hat{c}_{e,i}^{(n)}(r_n K) K} \mathbb{E}\big(T_{e,i}^{(n)} \wedge 1/r_n\big)^3 \\
&\quad \le K \hat{c}_{e,i}^{(n)}(r_n K) e^{\hat{c}_{e,i}^{(n)}(r_n K) K} \\
&\qquad\qquad \times \mathbb{E}\big[\big(T_{e,i}^{(n)}\big)^2 \big(r_n T_{e,i}^{(n)} \wedge 1\big)\big(1(T_{e,i}^{(n)} > r_n^{-1/2}) + 1(T_{e,i}^{(n)} \le r_n^{-1/2})\big)\big] \\
&\quad \le K \hat{c}_{e,i}^{(n)}(r_n K) e^{\hat{c}_{e,i}^{(n)}(r_n K) K} \mathbb{E}\big[\big(T_{e,i}^{(n)}\big)^2 1(T_{e,i}^{(n)} > r_n^{-1/2}) + \big(T_{e,i}^{(n)}\big)^2 r_n^{1/2}\big] \to 0
\end{aligned}$$

as $n \to \infty$. The first inequality is obtained by (A.5) and (A.6) and the second one by (4.8). Convergence to zero is justified by (2.11) and Lemma 4.3 (whose proof in Section A.3 does not use Lemma 4.2), which proves (A.4). Finally, (2.9) implies

$$\lim_{n\to\infty} a_3^{(n)}(0) < \infty.$$

Combined with (A.4), this gives us

$$\sup_{\substack{n\ge 1\\ |y|\le K}} a_3^{(n)}(r_n y) < \infty,$$

which concludes the proof of (A.2).

Downloaded from informs.org by [216.73.217.10] on 05 November 2026, at 13:43 . For personal use only, all rights reserved.

Downloaded from informs.org by [216.73.217.10] on 05 November 2026, at 13:43 . For personal use only, all rights reserved.

The second claim of this lemma, (4.16), is now simple to verify. Indeed, recalling the form of $c_{s,j}^{(n)}(\theta)$ from (4.13) and the definition of $c_{Lip,j}(K)$ from (4.14), we have

$$\begin{aligned}
&\lim_{n\to\infty} \sup_{0<\|\theta\|<K} \frac{c_{s,j}^{(n)}(r_n\theta)}{r_n^2\|\theta\|^2} \\
&\quad = \lim_{n\to\infty} \sup_{|y|\le c_{Lip,j}(r_nK)r_nK} \frac{\log^2 t_j(r_n\theta)}{r_n^2\|\theta\|^2} \left|\left(\chi_j^{(n)}\right)''(y) - \left(\chi_j^{(n)}\right)''(0)\right| \\
&\qquad + \lim_{n\to\infty} \sup_{0<\|\theta\|<K} \widetilde{\lambda}_{s,j}^{(n)} \frac{c_{1,j}(r_n\theta)}{r_n^2\|\theta\|^2} + \lim_{n\to\infty} \sup_{0<\|\theta\|<K} \frac{1}{2}\left(\widetilde{\lambda}_{s,j}^{(n)}\right)^3 \left(\widetilde{\sigma}_{s,j}^{(n)}\right)^2 \frac{c_{2,j}(r_n\theta)}{r_n^2\|\theta\|^2}.
\end{aligned}$$

From (4.6), (4.10) and the fact that $\log t_j(\theta) \in C^\infty(\mathbb{R}^d)$ its easy to see that

$$\lim_{n\to\infty} \sup_{0<\|\theta\|\le K} \frac{c_{2,j}(r_n\theta)}{r_n^2\|\theta\|^2} = 0 \quad \text{and} \quad \lim_{n\to\infty} \sup_{0<\|\theta\|\le K} \frac{c_{2,j}(r_n\theta)}{r_n^2\|\theta\|^2} = 0.$$

Furthermore,

$$\begin{aligned}
&\lim_{n\to\infty} \sup_{|y|\le c_{Lip,j}(r_nK)r_nK} \frac{\log^2 t_j(r_n\theta)}{r_n^2\|\theta\|^2} \left|\left(\chi_j^{(n)}\right)''(y) - \left(\chi_j^{(n)}\right)''(0)\right| \\
&\quad \le \lim_{n\to\infty} \sup_{|y|\le c_{Lip,j}(r_nK)r_nK} c_{Lip,j}^2(r_nK) \left|\left(\chi_j^{(n)}\right)''(y) - \left(\chi_j^{(n)}\right)''(0)\right| = 0,
\end{aligned}$$

because by the definition of $\chi_j^{(n)}$ in (3.24), it is not hard to repeat the arguments used in this proof to see that a version (A.2) holds for $\chi_j^{(n)}$ as well.

A.3. Proof of Lemma 4.3. We want to show that for all $i \in \mathcal{E}$ and all $K \in (0, r_nM]$,

$$\sup_{n\ge1} \hat{c}_{e,i}^{(n)}(r_nK) = \sup_{n\ge1} \left\{ \max\left\{\widetilde{\lambda}_{e,i}^{(n)}, |\eta_i^{(n)}(r_nK)|/r_nK\right\}\right\} < \infty. \tag{A.7}$$

By (4.6), $\widetilde{\lambda}_{e,i}^{(n)} \to \lambda_{e,i} < \infty$. Now since $\eta_i^{(n)}(\theta_i)$ is finite for each fixed n, it is enough to show that

$$\limsup_{n\to\infty} \frac{|\eta_i^{(n)}(r_nK)|}{r_nK} < \infty.$$

Recalling the definition of $\eta_i^{(n)}(\theta_i)$ in (3.22) and the truncation function $g^{(n)}(y) = \min(y, 1/r_n)$ defined for $y \in \mathbb{R}$, we observe that

$$\begin{aligned}\frac{|\eta_i^{(n)}(r_nK)|}{r_nK} &\le \sup_{0\le y\le r_nK} \left|\left(\eta_i^{(n)}\right)'(y)\right| \\ &= \sup_{0\le y\le r_nK} \frac{e^{-y}}{\mathbb{E}\left(g^{(n)}(T_{e,i}^{(n)})e^{g^{(n)}(T_{e,i}^{(n)})\eta_i^{(n)}(y)}\right)} \\ &\le \frac{1}{\mathbb{E}\left(g^{(n)}(T_{e,i}^{(n)})e^{g^{(n)}(T_{e,i}^{(n)})\eta_i^{(n)}(r_nK)}\right)},\end{aligned} \tag{A.8}$$

where in the last inequality, we used part (b) of Lemma 4.1, which states that $\eta_i^{(n)}(\theta_i)$ is a decreasing function of θ_i. To bound this term, we first argue that there exists an $\epsilon \in (0,1)$ and $y_1 > 0$ such that

$$\liminf_{n\to\infty} \mathbb{P}(T_{e,i}^{(n)} \ge y_1) \ge \epsilon. \tag{A.9}$$

Suppose (A.9) is not true. Then $\liminf_{n\to\infty} \mathbb{P}(T_{e,i}^{(n)} \ge y) < \epsilon$ for any $\epsilon \in (0,1)$ and $y > 0$. On the other hand, it follows from the uniform integrability assumption (2.11) that for any ϵ', there exists a_0 such that for any $a \ge a_0$,

$$\mathbb{E}\left(T_{e,i}^{(n)} 1(T_{e,i}^{(n)} > a)\right) < \epsilon', \quad n \ge 1.$$

Hence,

$$\begin{aligned}\mathbb{E}(T_{e,i}^{(n)}) &= \mathbb{E}\left(T_{e,i}^{(n)} 1(T_{e,i}^{(n)} \le y)\right) + \mathbb{E}\left(T_{e,i}^{(n)} 1(T_{e,i}^{(n)} \in (y,a])\right) \\ &\quad + \mathbb{E}\left(T_{e,i}^{(n)} 1(T_{e,i}^{(n)} > a)\right) \\ &\le y + a\mathbb{P}(T_{e,i}^{(n)} > y) + \epsilon',\end{aligned}$$

and

$$\lim_{n\to\infty} \mathbb{E}(T_{e,i}^{(n)}) = \liminf_{n\to\infty} \mathbb{E}(T_{e,i}^{(n)}) \le y + \liminf_{n\to\infty} a\mathbb{P}(T_{e,i}^{(n)} > y) + \epsilon' \le y + a\epsilon + \epsilon'.$$

Taking $\epsilon, y \to 0$ and then $\epsilon' \to 0$, we conclude that $\lim_{n\to\infty} \mathbb{E}(T_{e,i}^{(n)}) = 0$, which contradicts $\lambda_{e,i} < \infty$ in (2.9) and therefore proves (A.9). In addition to (A.9), the uniform integrability assumption (2.11) implies that there exists a $y_2 > y_1$ such that

$$\sup_{n\ge 1} \mathbb{P}(T_{e,i}^{(n)} \ge y_2) < \epsilon/2. \tag{A.10}$$

Downloaded from informs.org by [216.73.217.10] on 05 November 2026, at 13:43 . For personal use only, all rights reserved.

Downloaded from informs.org by [216.73.217.10] on 05 November 2026, at 13:43 . For personal use only, all rights reserved.

Fixing such y_1 and y_2, and recalling that $\eta_i^{(n)}(r_nK) < 0$ for all $n \geq 1$, we see that

$$\begin{aligned}
&\frac{1}{\mathbb{E}\big(g^{(n)}(T_{e,i}^{(n)})e^{g^{(n)}(T_{e,i}^{(n)})\eta_i^{(n)}(r_nK)}\big)} \\
&\quad\leq \frac{1}{\mathbb{E}\big(g^{(n)}(T_{e,i}^{(n)})e^{g^{(n)}(T_{e,i}^{(n)})\eta_i^{(n)}(r_nK)}1(y_1 \leq T_{e,i}^{(n)} < y_2)\big)} \\
&\quad\leq \frac{1}{\mathbb{E}\big(g^{(n)}(y_1)e^{g^{(n)}(y_2)\eta_i^{(n)}(r_nK)}1(y_1 \leq T_{e,i}^{(n)} < y_2)\big)} \\
&\quad= \frac{e^{g^{(n)}(y_2)|\eta_i^{(n)}(r_nK)|}}{g^{(n)}(y_1)\mathbb{P}(y_1 \leq T_{e,i}^{(n)} < y_2)}.
\end{aligned}$$

We recall our choices of y_1 and y_2 from (A.9) and (A.10) to arrive at

$$\begin{aligned}
\limsup_{n\to\infty} \frac{|\eta_i^{(n)}(r_nK)|}{r_nK} &\leq \limsup_{n\to\infty} \frac{e^{g^{(n)}(y_2)|\eta_i^{(n)}(r_nK)|}}{g^{(n)}(y_1)\mathbb{P}(y_1 \leq T_{e,i}^{(n)} < y_2)} \\
&\leq \limsup_{n\to\infty} \frac{2}{y_1\epsilon} e^{y_2|\eta_i^{(n)}(r_nK)|}.
\end{aligned}$$

Since $\eta_i^{(n)}(r_nK) < 0$, it remains to show that

$$\liminf_{n\to\infty} \eta_i^{(n)}(r_nK) > -\infty.$$

Now for any $y > 0$,

$$\begin{aligned}
e^{-r_nK} = \mathbb{E}\big(e^{g^{(n)}(T_{e,i}^{(n)})\eta_i^{(n)}(r_nK)}\big) &\leq e^{g^{(n)}(y)\eta_i^{(n)}(r_nK)}\mathbb{P}(T_{e,i}^{(n)} > y) + \mathbb{P}(T_{e,i}^{(n)} < y) \\
&\leq e^{g^{(n)}(y)\eta_i^{(n)}(r_nK)} + \mathbb{P}(T_{e,i}^{(n)} < y).
\end{aligned}$$

We choose $y = y_1$ from (A.9), and use the fact that $e^{-r_nK} \to 1$, to see that

$$\liminf_{n\to\infty} \Big(e^{-r_nK} - \mathbb{P}(T_{e,i}^{(n)} < y_1)\Big) = 1 - \limsup_{n\to\infty} \mathbb{P}(T_{e,i}^{(n)} < y_1) \geq 1 - (1-\epsilon) = \epsilon.$$

Therefore,

$$\begin{aligned}
\liminf_{n\to\infty} \eta_i^{(n)}(r_nK) &\geq \liminf_{n\to\infty} \frac{1}{g^{(n)}(y_1)} \log\big(e^{-r_nK} - \mathbb{P}(T_{e,i}^{(n)} < y_1)\big) \\
&= \frac{1}{y_1} \liminf_{n\to\infty} \log\big(e^{-r_nK} - \mathbb{P}(T_{e,i}^{(n)} < y_1)\big),
\end{aligned}$$

$$\geq \frac{1}{y_1}\log(\epsilon),$$

where the last inequality is obtained by taking the limit infimum inside the logarithm. This proves (A.7). The argument to show that

$$\sup_{n\geq 1}\hat{c}^{(n)}_{s,j}(r_nK) = \sup_{n\geq 1} c_{Lip,j}(r_nK)\max\left\{\widetilde{\lambda}^{(n)}_{s,j}, |\chi^{(n)}_j(\widehat{r_nK})|/\widehat{r_nK}\right\} < \infty$$

is nearly identical to the one for $\hat{c}^{(n)}_{e,i}(r_nK)$, and requires two slight modifications. First, observe that

$$\sup_{n\geq 1} c_{Lip,j}(r_nK) < \infty,$$

which is a trivial consequence of its definition in (4.14). Second, the equality in (A.8) would be modified to

$$\begin{aligned}\sup_{\substack{\|y\|\leq r_nK\\ y\in\mathbb{R}^d}} \left|\left(\chi^{(n)}_j\right)'(\log t_j(y))\right| &= \sup_{\substack{\|y\|\leq r_nK\\ y\in\mathbb{R}^d}} \frac{e^{-\log t_j(y)}}{\mathbb{E}\left(g^{(n)}(T^{(n)}_{s,j})e^{g^{(n)}(T^{(n)}_{s,j})\chi^{(n)}_j(\log t_j(y))}\right)}\\ &\leq \sup_{\substack{\|y\|\leq r_nK\\ y\in\mathbb{R}^d}} \frac{e^{C}}{\mathbb{E}\left(g^{(n)}(T^{(n)}_{s,j})e^{g^{(n)}(T^{(n)}_{s,j})\chi^{(n)}_j(\log t_j(y))}\right)},\end{aligned}$$

where $C>0$ is an upper bound for $\sup_{\|y\|\leq r_nK}|\log t_j(y)|$ that is independent of n, and $t_j(\theta)$ is defined in (3.23).

A.4. Proofs of Lemmas 4.4 and 4.5 . Let $\nu^{(n)}$ be the probability measure corresponding to $X^{(n)}(\infty)$. For $i\in\mathcal{E}$ and $j\in\mathcal{J}$, recall the external arrival process $\{E^{(n)}_i(t), t\geq 0\}$ and the departure process $\{D^{(n)}_j(t), t\geq 0\}$ defined in (3.7) and (3.8), respectively. Set

$$E^{(n)}_i(0,t] = E^{(n)}_i(t) - E^{(n)}_i(0) \quad \text{and} \quad D^{(n)}_j(0,t] = D^{(n)}_j(t) - D^{(n)}_j(0).$$

In this section we prove the following five statements, which imply Lemmas 4.4 and 4.5. We show that for all $i\in\mathcal{E}$ and $j\in\mathcal{J}$, any real number $K>0$ and any integer $n\geq 1$,

$$\text{(A.11)}\quad \mathbb{P}(L^{(n)}_j(\infty)>0) = \frac{\mathbb{E}_{\nu^{(n)}}\left[D^{(n)}_j(0,1]\right]}{\lambda^{(n)}_{s,j}},$$

$$\text{(A.12)}\quad \mathbb{E}_{\nu^{(n)}}\left[E^{(n)}_i(0,t]\right] = \lambda^{(n)}_{e,i}t,$$

Downloaded from informs.org by [216.73.217.10] on 05 November 2026, at 13:43 . For personal use only, all rights reserved.

Downloaded from informs.org by [216.73.217.10] on 05 November 2026, at 13:43 . For personal use only, all rights reserved.

$$\text{(A.13)}\quad \mathbb{E}_{\nu^{(n)}}\big[D_j^{(n)}(0,t]\big] = \lambda_{a,j}^{(n)} t,$$

$$\text{(A.14)}\quad \mathbb{E}\big[R_{s,j}^{(n)}(\infty)1(R_{s,j}^{(n)}(\infty) > K)1(L_j^{(n)}(\infty) > 0)\big] = \frac{\mathbb{E}_{\nu^{(n)}}\big[D_j^{(n)}(0,1]\big]}{2}\mathbb{E}\big[(T_{s,j}^{(n)})^2 1(T_{s,j}^{(n)} > K) - K^2 1(T_{s,j}^{(n)} > K)\big],$$

$$\text{(A.15)}\quad \mathbb{E}\big[R_{e,i}^{(n)}(\infty)1(R_{e,i}^{(n)}(\infty) > K)\big] = \frac{\mathbb{E}_{\nu^{(n)}}\big[E_i^{(n)}(0,1]\big]}{2}\mathbb{E}\big[(T_{e,i}^{(n)})^2 1(T_{e,i}^{(n)} > K) - K^2 1(T_{e,i}^{(n)} > K)\big].$$

Assuming these five equations are correct, we see that Lemma 4.4 is immediately implied by (A.11) and (A.13). To see why Lemma 4.5 is true, we write

$$\begin{aligned}
&\mathbb{E}\big[R_{s,j}^{(n)}(\infty)1(R_{s,j}^{(n)}(\infty) > K)\big] \\
&= \mathbb{E}\big[R_{s,j}^{(n)}(\infty)1(R_{s,j}^{(n)}(\infty) > K)1(L_j^{(n)}(\infty) > 0)\big] \\
&\quad + \mathbb{E}\big[R_{s,j}^{(n)}(\infty)1(R_{s,j}^{(n)}(\infty) > K)1(L_j^{(n)}(\infty) = 0)\big] \\
&= \mathbb{E}\big[R_{s,j}^{(n)}(\infty)1(R_{s,j}^{(n)}(\infty) > K)1(L_j^{(n)}(\infty) > 0)\big] \\
&\quad + \mathbb{E}\big[T_{s,j}^{(n)}1(T_{s,j}^{(n)} > K)\big]\mathbb{P}(L_j^{(n)}(\infty) = 0),
\end{aligned}$$

where in the last equality we use $\big[R_{s,j}^{(n)}(\infty)|L_j^{(n)}(\infty) = 0\big] \stackrel{d}{=} T_{s,j}^{(n)}$, which is true by construction. Uniform integrability of $\{R_{e,i}^{(n)}(\infty), n \geq 1\}$ and $\{R_{s,j}^{(n)}(\infty), n \geq 1\}$ then follows from (2.9), (2.11)–(2.13), and (A.12)–(A.15), thereby proving Lemma 4.5.

In the rest of this section, we prove (A.11)–(A.15) using the BAR (3.15), which is repeated here as

$$\text{(A.16)}\quad \mathbb{E}_\nu\big[\mathcal{A}f(X(0))\big] + \frac{1}{t}\sum_{m=1}^{\infty}\mathbb{E}_\nu\Big[\big(f(X_{\delta_m}) - f(X_{\delta_m-})\big)1(0 < T(\delta_m) \leq t)\Big] = 0, \quad f \in \mathcal{D},$$

where $\delta_m^{(n)}$, $T(\delta_m^{(n)})$, $X_{\delta_m^{(n)}}^{(n)}$ and $X_{\delta_m^{(n)}-}^{(n)}$ are introduced below (3.13) in Section 3.1, the operator $\mathcal{A}$ is defined in (3.11), and the class of functions $\mathcal{D}$ is defined above (3.11).

Fix an integer $\kappa \geq 1$ and $j \in \mathcal{J}$. We now prove (A.11) by applying the BAR (A.16) to $f_{\kappa,1}(x) = v_j \wedge \kappa$ (which belongs to $\mathcal{D}$) to obtain

$$\mathbb{E}\big[1(L_j^{(n)}(\infty) > 0)1(R_{s,j}^{(n)}(\infty) < \kappa)\big]$$
$$= \frac{1}{t}\sum_{m=1}^{\infty} \mathbb{E}_{\nu^{(n)}}\Big[\big(f_{\kappa,1}(X^{(n)}_{\delta_m^{(n)}}) - f_{\kappa,1}(X^{(n)}_{\delta_m^{(n)}-})\big)1(0 < T(\delta_m^{(n)}) \le t)\Big].$$

For station j, we recall the sequence of service times $\{T_{s,j}^{(n)}(q),\ q \ge 1\}$ and the sequence of departure times $\{S_j^{(n)}(q),\ q \ge 1\}$, defined in (2.2) and (3.6), respectively. From the form of the test function, we see that $f_{\kappa,1}(X^{(n)}_{\delta_m^{(n)}}) - f_{\kappa,1}(X^{(n)}_{\delta_m^{(n)}-}) = 0$ for all events $\delta_m^{(n)}$ that do not correspond to departures from station j. Therefore,

$$\text{(A.17)}\qquad \sum_{m=1}^{\infty} \mathbb{E}_{\nu^{(n)}}\Big[\big(f_{\kappa,1}(X^{(n)}_{\delta_m^{(n)}}) - f_{\kappa,1}(X^{(n)}_{\delta_m^{(n)}-})\big)1(0 < T(\delta_m^{(n)}) \le t)\Big]$$
$$= \sum_{q=1}^{\infty} \mathbb{E}_{\nu^{(n)}}\Big[f_{\kappa,1}\big(T_{s,j}^{(n)}(q)\big)1(S_j^{(n)}(q) \le t)\Big]$$
$$= \sum_{q=1}^{\infty} \mathbb{E}\big[f_{\kappa,1}(T_{s,j}^{(n)})\big]\mathbb{E}_{\nu^{(n)}}\Big[1(S_j^{(n)}(q) \le t)\Big]$$
$$= \mathbb{E}_{\nu^{(n)}}\big[D_j^{(n)}(0,t]\big]\mathbb{E}\big[f_{\kappa,1}(T_{s,j}^{(n)})\big].$$

In the second equality, we used the independence of $T_{s,j}^{(n)}(q)$ and $S_j^{(n)}(q)$, and in the last equality we use the Fubini-Tonelli theorem, justified by the fact that the summands are non-negative. We therefore have

$$\mathbb{E}\big[1(L_j^{(n)}(\infty) > 0)1(R_{s,j}^{(n)}(\infty) < \kappa)\big]$$
$$= \frac{1}{t}\mathbb{E}_{\nu^{(n)}}\big[D_j^{(n)}(0,t]\big]\mathbb{E}\big[f_{\kappa,1}(T_{s,j}^{(n)})\big], \qquad j \in \mathcal{J}.$$

We set $t = 1$, take the limit as $\kappa \to \infty$ on each side, and apply the monotone convergence theorem to conclude the proof of (A.11).

To prove (A.12), we use the test function $f_{\kappa,2}(x) = u_i \wedge \kappa$ (with $\kappa \ge 1$, and $i \in \mathcal{E}$) and repeat the argument for (A.11), but instead of $S_j^{(n)}(q)$, we use $U_i^{(n)}(q)$, as defined in (3.2) to be the qth external arrival time to station $i \in \mathcal{E}$. An alternative way to prove (A.12) would be to verify that the process $E_i^{(n)}$ is stationary when $E_i^{(n)}(0)$ is initialized accoridng to $\nu^{(n)}$. Then (A.12) would follow from [33, Theorem 76].

We now move on to prove (A.13). We recall from (3.9) that on every sample path,

$$\text{(A.18)}\qquad L_j^{(n)}(t) - L_j^{(n)}(0) = E_j^{(n)}(0,t] - D_j^{(n)}(0,t] + \sum_{k\in\mathcal{J}} \Phi_j^{(k)}(D_k^{(n)}(0,t]),$$

Downloaded from informs.org by [216.73.217.10] on 05 November 2026, at 13:43 . For personal use only, all rights reserved.

Downloaded from informs.org by [216.73.217.10] on 05 November 2026, at 13:43 . For personal use only, all rights reserved.

where $E_j(t) \equiv 0$ if $j \in \mathcal{J} \setminus \mathcal{E}$ and $\{\Phi^{(k)}(m), m \geq 1\}$ is the customer routing process from (3.4). We first show that

$$\mathbb{E}_{\nu^{(n)}}\left[L_j^{(n)}(t) - L_j^{(n)}(0)\right] = 0,$$

which would have been immediate provided we knew that $\mathbb{E}\big[L_j^{(n)}(\infty)\big] < \infty$. Instead, we use the stationarity of $\nu^{(n)}$ to see that for every $\kappa \geq 1$,

$$\mathbb{E}_{\nu^{(n)}}\big[(L_j^{(n)}(t) \wedge \kappa) - (L_j^{(n)}(0) \wedge \kappa)\big] = 0.$$

We wish to take the limit as $\kappa \to \infty$ and apply the dominated convergence theorem. To justify doing so, observe that

$$\big|(L_j^{(n)}(t) \wedge \kappa) - (L_j^{(n)}(0) \wedge \kappa)\big| \leq E_j^{(n)}(0,t] + \sum_{k \in \mathcal{J}} D_k^{(n)}(0,t],$$

and by (3.12), the expectation of the right hand side (with respect to $\nu^{(n)}$) is finite. We apply the dominated convergence theorem to get

$$0 = \lim_{\kappa \to \infty} \mathbb{E}_{\nu^{(n)}}\big[(L_j^{(n)}(t) \wedge \kappa) - (L_j^{(n)}(0) \wedge \kappa)\big] = \mathbb{E}_{\nu^{(n)}}\left[L_j^{(n)}(t) - L_j^{(n)}(0)\right],$$

implying

$$\mathbb{E}_{\nu^{(n)}}\big[E_j^{(n)}(0,t]\big] - \mathbb{E}_{\nu^{(n)}}\big[D_j^{(n)}(0,t]\big] + \sum_{k \in \mathcal{J}} \mathbb{E}_{\nu^{(n)}}\left[\Phi_j^{(k)}(D_k^{(n)}(0,t])\right] = 0.$$

Suppose we knew that

$$\mathbb{E}_{\nu^{(n)}}\left[\Phi_j^{(k)}(D_k^{(n)}(0,t])\right] = p_{kj}\mathbb{E}_{\nu^{(n)}}\big[D_k^{(n)}(0,t]\big]. \tag{A.19}$$

Then

$$\mathbb{E}_{\nu^{(n)}}\big[E_j^{(n)}(0,t]\big] - \sum_{k \in \mathcal{J}} \mathbb{E}_{\nu^{(n)}}\big[D_k^{(n)}(0,t]\big](\delta_{kj} - p_{kj}) = 0, \quad j \in \mathcal{J}.$$

Letting $D^{(n)}(0,t] \in \mathbb{R}^d$ be the vector whose elements are $D_k^{(n)}(0,t]$, we would use (2.4) and (A.12) to conclude that

$$\mathbb{E}\big[D^{(n)}(0,t]\big] = (I - P^{\mathrm{T}})^{-1}\lambda_e^{(n)}t = \lambda_a^{(n)}t,$$

thereby completing the proof of (A.13). It remains to verify (A.19). Observe that

$$\mathbb{E}_{\nu^{(n)}}\left[\Phi_j^{(k)}(D_k^{(n)}(0,t])\right] = \mathbb{E}_{\nu^{(n)}}\left[\sum_{q=1}^{\infty} \phi_j^{(k)}(q)1(S_j^{(n)}(q) \leq t)\right]$$

$$= \sum_{q=1}^{\infty} \mathbb{E}_{\nu^{(n)}}\left[\phi_j^{(k)}(q) 1(S_j^{(n)}(q) \le t)\right],$$

where $\phi_j^{(k)}$ is the indicator that the qth customer departing station k routes to station j, as defined in (3.1), and in the second equality we apply the Fubini-Tonnelli theorem, justified by the non-negativity of the summands. By definition, $\phi_j^{(k)}(q)$ is independent of $S_j^{(n)}(q)$ for all $q \ge 1$, and

$$\mathbb{E}_{\nu^{(n)}}\left[\phi_j^{(k)}(q)\right] = p_{kj}, \quad q \ge 1.$$

Repeating the arguments used to obtain (A.17), we see that

$$\mathbb{E}_{\nu^{(n)}}\left[\Phi_j^{(k)}(D_k^{(n)}(0,t])\right] = p_{kj}\mathbb{E}_{\nu^{(n)}}\left[D_k^{(n)}(0,t]\right].$$

This proves (A.19) and concludes the proof of (A.13).

We move on to prove (A.14). For $j \in \mathcal{J}$ and $K > 0$, we introduce the test function

$$f_{\kappa,3}(x) = [v_j^2 1(v_j > K) + K^2 1(v_j \le K)] \wedge \kappa^2, \quad \kappa \ge \lceil K \rceil.$$

Observe that

$$f'_{\kappa,3}(v_j) = 2v_j 1(K < v_j < \kappa).$$

Plugging $f_{\kappa,3}(x)$ into (A.16) and repeating the arguments used to get (A.17), we get

$$\begin{aligned}&\mathbb{E}\left[1(L_j^{(n)}(\infty) > 0) R_{s,j}^{(n)}(\infty) 1(K < R_{s,j}^{(n)}(\infty) < \kappa)\right] \\ &\quad = \frac{1}{2}\mathbb{E}_{\nu^{(n)}}\left[D_j^{(n)}(0,1]\right]\mathbb{E}\left[f_{\kappa,3}(T_{s,j}^{(n)}) - K^2\right].\end{aligned}$$

Taking $\kappa \to \infty$ and applying the monotone convergence theorem proves (A.14). Finally, for $i \in \mathcal{E}$, the argument for (A.15) is identical once we use the test function

$$f_{\kappa,4}(x) = [u_i^2 1(u_i > K) + K^2 1(u_i \le K)] \wedge \kappa^2, \quad \kappa \ge \lceil K \rceil.$$

This concludes the proofs of Lemmas 4.4 and 4.5.

A.5. Proof of Lemma 4.6. We begin with (4.18). Recall the definition of $\gamma(\theta)$ from (2.27) and for $i \in \mathcal{E}$ and $j \in \mathcal{J}$, the quadratic approximations

Downloaded from informs.org by [216.73.217.10] on 05 November 2026, at 13:43 . For personal use only, all rights reserved.

Downloaded from informs.org by [216.73.217.10] on 05 November 2026, at 13:43 . For personal use only, all rights reserved.

of $\eta_i^{(n)}(\theta_i)$ and $\zeta_j^{(n)}(\theta)$ from (4.7) and (4.11). Lemma 4.2 tells us that the error from these quadratic approximations vanishes. To conclude (4.18), it remains to show that for all $i \in \mathcal{E}$ and $j \in \mathcal{J}$,

$$\lim_{n\to\infty} \frac{1}{r_n}|\widetilde{\lambda}_{e,i}^{(n)} - \lambda_{e,i}^{(n)}| = 0, \quad \lim_{n\to\infty} \frac{1}{r_n}|\widetilde{\lambda}_{s,j}^{(n)} - \lambda_{s,j}^{(n)}| = 0,$$
$$\lim_{n\to\infty} |\widetilde{\sigma}_{e,i}^{(n)} - \sigma_{e,i}^{(n)}| = 0, \quad \lim_{n\to\infty} |\widetilde{\sigma}_{s,j}^{(n)} - \sigma_{s,j}^{(n)}| = 0.$$

The latter two statements are a consequence of (2.9), (2.10), and (4.6). We observe that

$$\begin{aligned}\frac{1}{r_n}|\widetilde{\lambda}_{e,i}^{(n)} - \lambda_{e,i}^{(n)}| &= \widetilde{\lambda}_{e,i}^{(n)}\lambda_{e,i}^{(n)}\frac{1}{r_n}\left|\mathbb{E}T_{e,i}^{(n)} - \mathbb{E}(T_{e,i}^{(n)} \wedge 1/r_n)\right| \\ &= \widetilde{\lambda}_{e,i}^{(n)}\lambda_{e,i}^{(n)}\frac{1}{r_n}\mathbb{E}(T_{e,i}^{(n)}1(T_{e,i}^{(n)} > 1/r_n)) \\ &\le \widetilde{\lambda}_{e,i}^{(n)}\lambda_{e,i}^{(n)}\left|\mathbb{E}((T_{e,i}^{(n)})^2 1(T_{e,i}^{(n)} > 1/r_n))\right| \to 0 \quad \text{as } n \to \infty,\end{aligned}$$

where the convergence to zero is by (2.9), (4.6) and the uniform integrability of $\{(T_{e,i}^{(n)})^2, n \ge 1\}$. The same argument holds for $\widetilde{\lambda}_{s,j}^{(n)}$, proving (4.18).

We move on to verify (4.19), or

$$\lim_{n\to\infty} \sup_{\substack{\theta<0\\ \|\theta\|\le K}} \left|\mathbb{E}\big[f_{r_n\theta}^{(n)}(X^{(n)}(\infty))\big] - \varphi^{(n)}(\theta)\right| = 0. \tag{A.20}$$

Recalling the definitions of $f_{r_n\theta}^{(n)}$ and $\varphi^{(n)}(\theta)$ from (3.21) and (4.1), we have

(A.21)

$$\begin{aligned}&\sup_{\substack{\theta<0\\ \|\theta\|\le K}} \left|\mathbb{E}\big[f_{r_n\theta}^{(n)}(X^{(n)}(\infty)) - e^{\langle\theta, Z^{(n)}(\infty)\rangle}\big]\right| \\ &\quad\le \sup_{\substack{\theta<0\\ \|\theta\|\le K}} \mathbb{E}\left|e^{\sum_{i\in\mathcal{E}}\eta_i^{(n)}(r_n\theta_i)g^{(n)}(R_{e,i}^{(n)}(\infty))+\sum_{j\in\mathcal{J}}\zeta_j^{(n)}(r_n\theta)g^{(n)}(R_{s,j}^{(n)}(\infty))} - 1\right| \\ &\quad\le c_f(K)\sup_{\substack{\theta<0\\ \|\theta\|\le K}} \mathbb{E}\Big|\sum_{i\in\mathcal{E}}(R_{e,i}^{(n)}(\infty)\wedge 1/r_n)\eta_i^{(n)}(r_n\theta_i) \\ &\qquad\qquad + \sum_{j\in\mathcal{J}}(R_{s,j}^{(n)}(\infty)\wedge 1/r_n)\zeta_j^{(n)}(r_n\theta_i)\Big|.\end{aligned}$$

To obtain the last inequality, we used (A.6) and

$$\sup_{\|\theta\|\le K} e^{|\sum_{i\in\mathcal{E}}\eta_i^{(n)}(r_n\theta_i)g^{(n)}(R_{e,i}^{(n)}(\infty))+\sum_{j\in\mathcal{J}}\zeta_j^{(n)}(r_n\theta)g^{(n)}(R_{s,j}^{(n)}(\infty))|} \le c_f(K),$$

where $c_f(K)$ is defined in (4.17). For $i \in \mathcal{E}$ we have

$$\begin{aligned}\lim_{n\to\infty} \sup_{\substack{\theta<0\\ \|\theta\|\le K}} \mathbb{E}\Big|(R_{e,i}^{(n)}(\infty)\wedge 1/r_n)\eta_i^{(n)}(r_n\theta_i)\Big| \\ \le \lim_{n\to\infty} \hat{c}_{e,i}^{(n)}(r_nK) r_n K \mathbb{E}\big[R_{e,i}^{(n)}(\infty)\big] = 0,\end{aligned}$$

where the inequality is by (4.8) and the fact that the limit equals zero is by Lemma 4.3 and Lemma 4.5. A similar argument holds for $R_{s,j}^{(n)}(\infty)$, implying (A.20).

We move on to verify (4.20), or

$$\begin{aligned}\text{(A.22)}\quad \lim_{n\to\infty} \sup_{\substack{\theta<0\\ \|\theta\|\le K}} \frac{1}{r_n}\Big|\mathbb{E}\big[1(L_k^{(n)}(\infty)=0)\big(f_{r_n\theta}^{(n)}(X^{(n)}(\infty)) - e^{\langle\theta, Z^{(n)}(\infty)\rangle}\big)\big]\Big| \\ = 0, \quad k\in\mathcal{J}.\end{aligned}$$

Repeating the steps used to obtain (A.21), we have

$$\begin{aligned}\text{(A.23)}\quad &\sup_{\substack{\theta<0\\ \|\theta\|\le K}} \frac{1}{r_n}\Big|\mathbb{E}\big[1(L_k^{(n)}(\infty)=0)\big(f_{r_n\theta}^{(n)}(X^{(n)}(\infty)) - e^{\langle\theta, Z^{(n)}(\infty)\rangle}\big)\big]\Big| \\ &\le c_f(K)\sum_{i\in\mathcal{E}} \mathbb{E}\Big|K\hat{c}_{e,i}^{(n)}(K)1(L_k^{(n)}(\infty)=0)(R_{e,i}^{(n)}(\infty)\wedge 1/r_n)\Big| \\ &\quad + c_f(K)\sum_{j\in\mathcal{J}} \mathbb{E}\Big|K\hat{c}_{s,j}^{(n)}(r_nK)1(L_k^{(n)}(\infty)=0)(R_{s,j}^{(n)}(\infty)\wedge 1/r_n)\Big|.\end{aligned}$$

To show that (A.23) vanishes as $n\to\infty$, observe that by (2.11) and Lemma 4.4,

$$\begin{aligned}&\mathbb{E}\big[1(L_k^{(n)}(\infty)=0)(R_{e,i}^{(n)}(\infty)\wedge 1/r_n)\big] \\ &\quad = \mathbb{E}\big[1(L_k^{(n)}(\infty)=0)(R_{e,i}^{(n)}(\infty)\wedge 1/r_n)\big(1(R_{e,i}^{(n)}(\infty) > r_n^{-1/2}) \\ &\qquad + 1(R_{e,i}^{(n)}(\infty)\le r_n^{-1/2})\big)\big] \\ &\quad \le \mathbb{E}\big[R_{e,i}^{(n)}(\infty)1(R_{e,i}^{(n)}(\infty) > r_n^{-1/2})\big] + \mathbb{E}\big[1(L_k^{(n)}(\infty)=0)r_n^{-1/2}\big] \\ &\quad \to 0 \quad \text{as } n\to\infty.\end{aligned}$$

Repeating similar arguments for $R_{s,j}^{(n)}(\infty)$ proves (A.22).

Downloaded from informs.org by [216.73.217.10] on 05 November 2026, at 13:43 . For personal use only, all rights reserved.

Downloaded from informs.org by [216.73.217.10] on 05 November 2026, at 13:43 . For personal use only, all rights reserved.

APPENDIX B: PROOF OF LEMMA 6.1

We begin by proving (a). Fix $M > 0$ and $\delta > 0$. Setting

$$\theta = -\delta(1, 1, ..., 1)^{\mathrm{T}} = -\delta\mathbf{1},$$

we see that

$$\begin{aligned}(\mathrm{B.1}) \qquad \phi^{(n)}(\theta) &= \phi^{(n)}(\delta) \\ &= \int_{\mathbb{R}^d_+} e^{-\delta\langle\mathbf{1},x\rangle}\nu^{(n)}(dx) \\ &= \int_{\{||x||_\infty \le M\}} e^{-\delta\langle\mathbf{1},x\rangle}\nu^{(n)}(dx) + \int_{\{||x||_\infty \le M\}^c} e^{-\delta\langle\mathbf{1},x\rangle}\nu^{(n)}(dx).\end{aligned}$$

Hence, we have

$$\phi^{(n)}(\theta) \ge e^{-\delta M d}\nu^{(n)}\big\{||x||_\infty \le M\big\}.$$

Assuming that $\{\nu^{(n)}\}$ is tight, we can fix $\epsilon > 0$ and take M large enough and then δ small enough such that

$$\phi(\delta) = \lim_{n\to\infty} \phi^{(n)}(\delta) \ge 1 - \epsilon.$$

Since each $\phi^{(n)}$ is monotone, ϕ is monotone too, which immediately implies

$$\lim_{\theta\uparrow 0} \phi(\theta) = 1.$$

Furthermore, (B.1) also implies that

$$\phi^{(n)}(\delta) \le \nu^{(n)}\big\{||x||_\infty \le M\big\} + e^{-\delta M},$$

or

$$\phi(\delta) - e^{-\delta M} \le \liminf_{n\to\infty} \nu^{(n)}\big\{||x||_\infty \le M\big\}.$$

If $\lim_{\theta\uparrow 0} \phi(\theta) = 1$, we can fix $\epsilon > 0$ and choose a δ small enough so that $\phi(\delta) > 1 - \epsilon/2$. Then take M large enough so that $e^{-\delta M} < \epsilon/2$ to conclude that

$$\liminf_{n\to\infty} \nu^{(n)}\big\{||x||_\infty \le M\big\} \ge 1 - \epsilon,$$

which proves (a). Statement (b) is an immediate consequence of the continuity theorem for moment generating functions [34, Lemma 4].

Acknowledgements. This research is supported in part by NSF Grants CNS-1248117, CMMI-1335724, and CMMI-1537795, and by JSPS KAKENHI Grant Number 16H027860001.

REFERENCES

[1] Baccelli, F. and Brémaud, P. (2003). *Elements of queueing theory: Palm martingale calculus and stochastic recurrences*, vol. 26 of *Applications of Mathematics*. 2nd ed. Springer, Berlin. MR1957884

[2] Berman, A. and Plemmons, R. J. (1979). *Nonnegative matrices in the mathematical sciences*. Academic Press, New York. MR0544666

[3] Boucherie, R. J. and van Dijk, N. M. (2011). *Queueing networks: a fundamental approach*. International Series in Operations Research & Management Science, Springer. http://dx.doi.org/10.1007/978-1-4419-6472-4. MR2829413

[4] Bramson, M. and Dai, J. G. (2001). Heavy traffic limits for some queueing networks. *Ann. Appl. Probab.*, **11** 49–90. http://dx.doi.org/10.1214/aoap/998926987. MR1825460

[5] Braverman, A. and Dai, J. G. (2017). Stein's method for steady-state diffusion approximations of $M/Ph/n+M$ systems. *Ann. Appl. Probab.*, **27** 550–581. http://dx.doi.org/10.1214/16-AAP1211. MR3619795

[6] Braverman, A., Dai, J. G. and Feng, J. (2016). Stein's method for steady-state diffusion approximations: an introduction through the Erlang-A and Erlang-C models. *Stochastic Systems*, **6** 301–366. http://www.i-journals.org/ssy/viewarticle.php?id=212&layout=abstract.

[7] Budhiraja, A. and Lee, C. (2009). Stationary distribution convergence for generalized Jackson networks in heavy traffic. *Math. Oper. Res.*, **34** 45–56. http://dx.doi.org/10.1287/moor.1080.0353. MR2542988

[8] Chen, H. and Mandelbaum, A. (1991). Discrete flow networks: bottleneck analysis and fluid approximations. *Math. Oper. Res.*, **16** 408–446. http://dx.doi.org/10.1287/moor.16.2.408. MR1106809

[9] Chen, L. H. Y., Goldstein, L. and Shao, Q.-M. (2011). *Normal approximation by Stein's method*. Probability and its Applications (New York), Springer, Heidelberg. http://dx.doi.org/10.1007/978-3-642-15007-4. MR2732624

[10] Chung, K. L. (2001). *A Course in Probability Theory*. 3rd ed. Academic Press (An Imprint of Elsevier). MR1796326

[11] Dai, J. G. (1995). On positive Harris recurrence of multiclass queueing networks: a unified approach via fluid limit models. *Annals of Applied Probability*, **5** 49–77. http://links.jstor.org/sici?sici=1050-5164(199502)5:1<49:OPHROM>2.0.CO;2-4&origin=MSN. MR1325041

[12] Dai, J. G., Dieker, A. and Gao, X. (2014). Validity of heavy-traffic steady-state approximations in many-server queues with abandonment. *Queueing Systems*, **78** 1–29. http://dx.doi.org/10.1007/s11134-014-9394-x. MR3238006

[13] Dai, J. G. and Kurtz, T. G. (1994). Characterization of the stationary distribution for a semimartingale reflecting Brownian motion in a convex polyhedron. Preprint. MR1346729

[14] Dai, J. G., Miyazawa, M. and Wu, J. (2014). A multi-dimensional SRBM: geometric views of its product form stationary distribution. *Queueing Systems*, **78** 313–335. (arXiv version can be downloaded at the website), http://arxiv.org/abs/1312.1758. MR3269261

Downloaded from informs.org by [216.73.217.10] on 05 November 2026, at 13:43 . For personal use only, all rights reserved.

[15] Down, D. and Meyn, S. (1994). Piecewise linear test functions for stability of queueing networks. *Proceedings of the 33rd Conference on Decision and Control* 2069–2074. http://dx.doi.org/10.1109/CDC.1994.411432.

[16] Gamarnik, D. and Stolyar, A. L. (2012). Multiclass multiserver queueing system in the Halfin-Whitt heavy traffic regime: asymptotics of the stationary distribution. *Queueing Systems*, **71** 25–51. http://dl.acm.org/citation.cfm?id=2339029. MR2925789

[17] Gamarnik, D. and Zeevi, A. (2006). Validity of heavy traffic steady-state approximation in generalized Jackson networks. *Ann. Appl. Probab.*, **16** 56–90. http://dx.doi.org/10.1214/105051605000000638. MR2209336

[18] Gurvich, I. (2014). Diffusion models and steady-state approximations for exponentially ergodic Markovian queues. *The Annals of Applied Probability*, **24** 2527–2559. http://dx.doi.org/10.1214/13-AAP984. MR3262510

[19] Gurvich, I. (2014). Validity of heavy-traffic steady-state approximations in multiclass queueing networks: the case of queue-ratio disciplines. *Mathematics of Operations Research*, **39** 121–162. http://dx.doi.org/10.1287/moor.2013.0593. MR3173006

[20] Gurvich, I., Huang, J. and Mandelbaum, A. (2014). Excursion-based universal approximations for the Erlang-A queue in steady-state. *Mathematics of Operations Research*, **39** 325–373. http://dx.doi.org/10.1287/moor.2013.0606. MR3205550

[21] Harrison, J. M. and Reiman, M. I. (1981). Reflected Brownian motion on an orthant. *Ann. Probab.*, **9** 302–308. http://links.jstor.org/sici?sici=0091-1798(198104)9:2<302:RBMOAO>2.0.CO;2-P&origin=MSN. MR0606992

[22] Harrison, J. M. and Williams, R. J. (1987). Brownian models of open queueing networks with homogeneous customer populations. *Stochastics*, **22** 77–115. http://dx.doi.org/10.1080/17442508708833469. MR0912049

[23] Huang, J. and Gurvich, I. (2016). Beyond heavy-traffic regimes: universal bounds and controls for the single-server queue. Submitted for publication, http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2784752.

[24] Johnson, D. P. (1983). *Diffusion approximations for optimal filtering of jump processes and for queueing networks*. Ph.D. thesis, University of Wisconsin. http://gateway.proquest.com/openurl?url_ver=Z39.88-2004&rft_val_fmt=info:ofi/fmt:kev:mtx:dissertation&res_dat=xri:pqdiss&rft_dat=xri:pqdiss:8316215.

[25] Katsuda, T. (2010). State-space collapse in stationarity and its application to a multiclass single-server queue in heavy traffic. *Queueing Syst.*, **65** 237–273. http://dx.doi.org/10.1007/s11134-010-9178-x. MR2652044

[26] Kingman, J. F. C. (1962). On queues in heavy traffic. *J. Roy. Statist. Soc. Ser. B*, **24** 383–392. http://links.jstor.org/sici?sici=0035-9246(1962)24:2<383:OQIHT>2.0.CO;2-A&origin=MSN. MR0148146

[27] Krantz, S. and Parks, H. (2013). *The implicit function theorem: history, theory, and applications*. Birkhäuser. http://link.springer.com/book/10.1007%2F978-1-4614-5981-1. MR2977424

[28] Miyazawa, M. (2015). Diffusion approximation for stationary analysis of queues and their networks: a review. *J. Oper. Res. Soc. Japan*, **58** 104–148. http://dx.doi.org/10.15807/jorsj.58.104. MR3363214

[29] Miyazawa, M. (2017). A unified approach for large queue asymptotics in a heterogeneous multiserver queue. *Adv. in Appl. Probab.*, **49** 182–220. http://dx.doi.org/10.1017/apr.2016.84. MR3631221

[30] Reiman, M. I. (1984). Open queueing networks in heavy traffic. *Mathematics*

Downloaded from informs.org by [216.73.217.10] on 05 November 2026, at 13:43 . For personal use only, all rights reserved.

of Operations Research, **9** 441–458. http://dx.doi.org/10.1287/moor.9.3.441. MR0757317

[31] Resnick, S. I. (1992). *Adventures in stochastic processes*. Birkhäuser, Boston. MR1181423

[32] Ross, N. (2011). Fundamentals of Stein's method. *Probab. Surv.*, **8** 210–293. http://dx.doi.org/10.1214/11-PS182. MR2861132

[33] Serfozo, R. (2009). *Basics of applied stochastic processes*. Probability and its Applications (New York), Springer-Verlag, Berlin. MR2484222

[34] Stadtmüller, U. and Trautner, R. (1981). Tauberian theorems for Laplace transforms in dimension D > 1. *Journal für die reine und angewandte Mathematik*, **323** 127–138. http://eudml.org/doc/152337. MR0611447

[35] Stein, C. (1986). Approximate computation of expectations. *Lecture Notes-Monograph Series*, **7**. http://www.jstor.org/stable/4355512. MR0882007

[36] Tezcan, T. (2008). Optimal control of distributed parallel server systems under the Halfin and Whitt regime. *Mathematics of Operations Research*, **33** 51–90. http://search.proquest.com/docview/212618995?accountid=10267. MR2393541

[37] Williams, D. (1991). *Probability with Martingales*. Cambridge mathematical textbooks, Cambridge University Press. MR1155402

[38] Ye, H.-Q. and Yao, D. D. (2016). Diffusion limit of fair resource control—stationarity and interchange of limits. *Mathematics of Operations Research*, **41** 1161–1207. http://dx.doi.org/10.1287/moor.2015.0773, http://dx.doi.org/10.1287/moor.2015.0773. MR3544792

[39] Zhang, J. and Zwart, B. (2008). Steady state approximations of limited processor sharing queues in heavy traffic. *Queueing Systems: Theory and Applications*, **60** 227–246. http://dx.doi.org/10.1007/s11134-008-9095-4. MR2461617

Anton Braverman
Department of Operations Research
and Information Engineering
Cornell University
206 Rhodes Hall
Ithaca, NY
14850
E-mail: ab2329@cornell.edu

J. G. Dai
Department of Operations Research
and Information Engineering
Cornell University
226 Rhodes Hall
Ithaca, NY
14850
E-mail: jd694@cornell.edu

Masakiyo Miyazawa
Department of Information Sciences
Tokyo University of Science
2641 Yamazaki
Noda, Chiba
278-8510
Japan
E-mail: miyazawa@rs.tus.ac.jp

Downloaded from informs.org by [216.73.217.10] on 05 November 2026, at 13:43 . For personal use only, all rights reserved.