An Invariance Principle for Semimartingale Reflecting Brownian Motions in an Orthant[1]

R. J. Williams
Department of Mathematics
University of California, San Diego
La Jolla CA 92093-0112

ABSTRACT

Semimartingale reflecting Brownian motions in an orthant (SRBMs) are of interest in applied probability because of their role as heavy traffic approximations for open queueing networks. It is shown in this paper that a process which satisfies the definition of an SRBM, except that small random perturbations in the defining conditions are allowed, is close in distribution to an SRBM. This perturbation result is called an invariance principle by analogy with the invariance principle of Stroock and Varadhan for diffusions with boundary conditions. A crucial ingredient in the proof of this result is an oscillation inequality for solutions of a perturbed Skorokhod problem. In a subsequent paper, the invariance principle is used to give general conditions under which a heavy traffic limit theorem holds for open multiclass queueing networks.

Keywords: Semimartingale reflecting Brownian motion, diffusions, invariance principle, Skorokhod problem, oscillation inequality, open multiclass queueing networks.
AMS 1991 Subject Classification: 60F17, 60J60, 60J65, 60K25, 34C35.

NOTE TO THE COPYEDITOR AND TYPESETTER: *The numbering of Sections, Definitions, Assumptions, Lemmas, Propositions and Theorems should not be changed since these are referred to by number in a subsequent paper by the same author submitted to this volume of QUESTA.*

1 Introduction

Semimartingale reflecting Brownian motions in an orthant (SRBMs) have been shown to approximate many single class and some multiclass open queueing networks under conditions of heavy traffic. While it is known that not all open multiclass networks can be approximated in heavy traffic by SRBMs (cf. [6, 24, 5]), one of the outstanding challenges in contemporary research on queueing networks is to identify a broad collection of multiclass networks that can be approximated by SRBMs and to prove heavy traffic limit theorems justifying such approximations.

[1]Research supported in part by NSF grants GER 9023335 and DMS 9703891.

Heavy traffic limit theorems for open queueing networks such as those of Reiman [18] for single class FIFO networks, of Peterson [17] for feedforward multiclass networks with preemptive resume static priorities, and of Chen and Zhang [3] for re-entrant lines with a first-buffer-first-served priority service discipline, have used a continuous mapping argument, in conjunction with functional central limit theorems for the arrival, service time and routing processes, to prove convergence in distribution of the normalized queue length or workload processes to SRBMs. These limit theorems rely on the existence and uniqueness of a continuous path-to-path mapping which can be applied to obtain a semimartingale reflecting Brownian motion (SRBM) path as the continuous image of a Brownian motion path (cf. [12, 9]). A major difficulty in generalizing these results to multiclass networks with feedback is that uniqueness does not always hold for such a path-to-path mapping [1, 16], even when existence and uniqueness *in law* of the proposed SRBM approximant is known. Thus, it is natural to seek a limit theorem or perturbation result for SRBMs to take the place of the continuous mapping argument. Such a result might also be useful for proving that other processes approximate SRBMs, such as those arising in the control of some Brownian network models (cf. [11, 13, 14]) or in numerical approximation of SRBMs.

The aim of this paper is to formulate and prove such a perturbation result for SRBMs. We call our perturbation result an *invariance principle* for SRBMs. The reason for this choice of terminology is that the perturbation result given here is similar in spirit to other results for diffusions which have been called invariance principles. In particular, in their work on diffusion processes with boundary conditions, Stroock and Varadhan [22] showed that diffusions with smooth boundary conditions can be approximated by Markov chains whose generators are "close" to those of the diffusions in an appropriate sense. They called this result an invariance principle for the diffusions. Furthermore, one may use the results of this paper, in conjunction with Donsker's invariance principle for Brownian motion, to obtain an invariance principle for SRBMs in the sense that one can show how to approximate an SRBM by a sequence of reflected random walks.

An outline of the paper is as follows. Some notation is given in Section 2. The definition of an SRBM and some of its properties are delineated in Section 3. Here an SRBM is allowed to have an arbitrary initial distribution, which is a slight generalization of the definition given in Taylor and Williams [23], where a fixed initial value was used. The main result of this paper, namely the invariance principle for SRBMs, is stated in Section 4. Intuitively, the idea of the result is as follows. Suppose that $\{W^n\}_{n=1}^{\infty}$ is a sequence of d-dimensional stochastic processes with r.c.l.l. (right continuous with finite left limits) paths such that each W^n satisfies the conditions for an SRBM, except that random perturbations in the requirements are allowed, where these perturbations converge to zero in distribution as $n \to \infty$. Then $\{W^n\}_{n=1}^{\infty}$ converges in distribution to an SRBM. While the general form of the result is not difficult to guess, subtleties arise in the exact formulation so that any weak limit point of the sequence $\{W^n\}_{n=1}^{\infty}$ can be identified as an SRBM. The fact that only weak existence and uniqueness of an SRBM is known in general is the source of this subtlety. Several sufficient conditions are given in Section 4 for identifying any weak limit point as an SRBM. In particular, the third sufficient condition (III) given in Proposition 4.2 anticipates use in a subsequent paper [26] by this author where general conditions are given under which a heavy traffic limit theorem holds for open multiclass queueing networks. A key element for the proof of the invariance principle, namely, an oscillation inequality for solutions of a perturbed Skorokhod problem, is established in Section 5. Using this, the results of Section 4 are proved in Sections 6 and 7.

Motivated by applications for closed and capacitated multiclass queueing networks, it is natural to ask whether the results of this paper can be extended to semimartingale reflecting Brownian motions living in convex polyhedrons, as defined in Dai and Williams [7]. Indeed, contemporaneously and independently of the work of the current author, J. G. Dai and W. Dai [4, 8] established an oscillation inequality and tightness result for semimartingale reflecting Brownian motions living in convex polyhedrons. Their results and proofs are similar in nature to those associated with Theorems 4.1 and 5.1 of this paper. However, their assumptions are somewhat different from those used here, and as a consequence neither result can be derived directly from the other.

2 Notation

The m-dimensional ($m \geq 1$) Euclidean space will be denoted by $\mathbb{R}^m$ and $\mathbb{R}_+$ will denote $[0, \infty)$. The superscript $'$ will be used to denote the operation of taking the transpose of a vector or matrix. Inequalities between vectors in $\mathbb{R}^m$ should be interpreted componentwise. The set of bounded continuous functions defined from $\mathbb{R}^m$ into $\mathbb{R}$ will be denoted by $C_b(\mathbb{R}^m)$.

A triple $(\Omega, \mathcal{F}, \{\mathcal{F}_t, t \geq 0\})$ will be called a filtered space if Ω is a set, $\mathcal{F}$ is a σ-field of subsets of Ω, and $\{\mathcal{F}_t, t \geq 0\}$ is an increasing family of sub-σ-fields of $\mathcal{F}$, i.e., a filtration. If in addition, P is a probability measure on $(\Omega, \mathcal{F})$, then $(\Omega, \mathcal{F}, \{\mathcal{F}_t, t \geq 0\}, P)$ is called a filtered probability space. A filtration $\{\mathcal{F}_t, t \geq 0\}$ will often be simply written as $\{\mathcal{F}_t\}$. Given a fixed filtration, an m-dimensional process will be called a martingale (relative to the filtration) if and only if each component is a martingale (relative to the filtration).

For each positive integer m, let D^m be the space of "Skorokhod paths" in $\mathbb{R}^m$ having time domain $[0, \infty)$. That is, D^m is the set of all functions $w : [0, \infty) \to \mathbb{R}^m$ that are right continuous on $[0, \infty)$ and have finite left limits on $(0, \infty)$. The subspace of D^m consisting only of continuous functions is denoted by C^m. When $m = 1$, D, C will be used instead of D^m, C^m, respectively. Consider D^m to be endowed with the Skorokhod $\mathbf{J}_1$-topology (see Skorokhod [20] and Ethier and Kurtz [10]). When the $\mathbf{J}_1$-topology on D^m is relativized to C^m, it is the topology of uniform convergence on compact time intervals. The abbreviation *u.o.c.* will stand for *uniformly on compacts,* to indicate that a sequence of functions in D^m (or C^m) is converging uniformly on compact time intervals to a limit in D^m (or C^m). Let $\mathcal{M}^m$ denote the Borel σ-algebra on D^m (or C^m) associated with the Skorokhod topology. This is the same as the σ-algebra generated by the coordinate maps, i.e., $\mathcal{M}^m = \sigma\{w(s) : 0 \leq s < \infty\}$.

Similarly, the filtration $\{\mathcal{M}_t^m, t \geq 0\}$ is defined by $\mathcal{M}_t^m = \sigma\{w(s) : 0 \leq s \leq t\}$, $t \geq 0$. The canonical process on D^m (or C^m) is the process W defined by $W(t, w) = w(t)$, $t \geq 0$, $w \in D^m$ (or C^m).

All (stochastic) processes in this paper will be assumed to have r.c.l.l. (right continuous with finite left limits) paths. That is, a process is a measurable mapping from some probability space $(\Omega, \mathcal{F}, P)$ into the space $(D^m, \mathcal{M}^m)$ for a suitable value of m. To indicate the dependence on m, we shall call such a process an m-dimensional process. Consider $W^1, W^2, \ldots, W$, each of which is an m-dimensional process. The filtration generated by W is taken to be $\{\mathcal{F}_t, t \geq 0\}$ where $\mathcal{F}_t = \sigma\{W(s) : 0 \leq s \leq t\}$ for all $t \geq 0$. The sequence $\{W^n\}_{n=1}^{\infty}$ is said to be tight if and only if the probability measures induced by the W^n on $(D^m, \mathcal{M}^m)$ form a tight sequence, i.e., they form a weakly relatively compact sequence in the space of probability measures on $(D^m, \mathcal{M}^m)$. The notation "$W^n \Rightarrow W$" will mean that the probability measures induced by the W^n on $(D^m, \mathcal{M}^m)$ converge weakly to the probability measure induced on $(D^m, \mathcal{M}^m)$ by W; this same state of affairs may be expressed by the statement "W^n converges in distribution to W as $n \to \infty$". For more on tightness and convergence in distribution of processes taking values in D^m, see Chapter 3 of Ethier and Kurtz [10].

3 Definition and properties of an SRBM

Throughout this paper, d is a positive integer, $S \equiv \{x \in \mathbb{R}^d : x_i \geq 0 \text{ for } i = 1, \ldots, d\}$, $\mathcal{B}$ denotes the σ-algebra of Borel subsets of S, θ is a constant vector in $\mathbb{R}^d$, Γ is a $d \times d$ non-degenerate covariance matrix (symmetric and strictly positive definite), and R is a $d \times d$ matrix.

The following definition of a semimartingale reflecting Brownian motion generalizes that in Taylor and Williams [23] to allow an arbitrary initial distribution ν.

Definition 3.1 (Semimartingale Reflecting Brownian Motion) *Given a probability measure ν on $(S, \mathcal{B})$, a semimartingale reflecting Brownian motion (abbreviated as SRBM) associated with the data $(S, \theta, \Gamma, R, \nu)$ is an $\{\mathcal{F}_t\}$-adapted, d-dimensional process W defined on some filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}, P)$ such that*

(i) $W = X + RY$, *P-a.s.,*

(ii) *P-a.s., W has continuous paths and $W(t) \in S$ for all $t \geq 0$,*

(iii) *under P,*

- **(a)** *X is a d-dimensional Brownian motion with drift vector θ, covariance matrix Γ and $X(0)$ has distribution ν,*
- **(b)** $\{X(t) - X(0) - \theta t, \mathcal{F}_t, t \geq 0\}$ *is a martingale,*

(iv) *Y is an $\{\mathcal{F}_t\}$-adapted, d-dimensional process such that P-a.s. for $i = 1, \ldots, d$,*

(a) $Y_i(0) = 0$,

(b) *Y_i is continuous and non-decreasing,*

(c) *Y_i can increase only when W is on the face $F_i \equiv \{x \in S : x_i = 0\}$, i.e., $\int_0^\infty 1_{(0,\infty)}(W_i(s))\, dY_i(s) = 0$.*

We shall often refer to Y as the "pushing process" associated with the SRBM W. When $\nu \equiv \delta_x$, the unit mass at $x \in S$, we may alternatively say that W is an SRBM associated with (S, θ, Γ, R) that starts from x.

Remark. Note that the filtration $\{\mathcal{F}_t, t \geq 0\}$ really only comes into play in the martingale condition (iii)(b). Indeed, a triple of processes (W, X, Y) defined on some probability space $(\Omega, \mathcal{F}, P)$ satisfies the definition of an SRBM if and only if it satisfies the definition where the filtration $\{\mathcal{F}_t, t \geq 0\}$ is taken to be the one generated by (W, Y) and the P-null sets. For if (W, X, Y) satisfy the above conditions for some filtration $\{\mathcal{F}_t\}$, then they will still satisfy them when the filtration is augmented by the P-null sets. Furthermore, by (i) and the adaptedness assumptions on (W, Y), we have that X is adapted to the filtration generated by (W, Y) and the P-null sets and this filtration is contained in the aforementioned augmented filtration. Since a martingale is still a martingale with respect to any smaller filtration to which it is adapted, it follows that (iii)(b) still holds when $\{\mathcal{F}_t, t \geq 0\}$ is replaced by the filtration generated by (W, Y) and the P-null sets.

Loosely speaking, an SRBM behaves like a Brownian motion in the interior of the orthant S and it is confined to the orthant by instantaneous "reflection" (or "pushing") at the boundary, where the direction of "reflection" on the i^{th} face F_i is given by the i^{th} column of the *reflection matrix* R. The results of Reiman and Williams [19] and Taylor and Williams [23] show that a necessary and sufficient condition for the existence of an SRBM associated with $(S, \theta, \Gamma, R, \delta_x)$ for each $x \in S$ is that the reflection matrix R be completely-$\mathcal{S}$, i.e., for each principal submatrix $\tilde{R}$ of R there is a vector $\tilde{x} \geq 0$ such that $\tilde{R}\tilde{x} > 0$. (Here a principal submatrix of R is a matrix obtained by deleting all rows and columns of R with indices in some set $\mathcal{I} \subset \{1, \ldots, d\}$, where $\mathcal{I}$ has no more than $d-1$ elements but it may be empty.) The completely-$\mathcal{S}$ condition on R can be interpreted geometrically as requiring that at each point on the boundary of S there is a positive linear combination of the directions of reflection that can be used there which points into the interior of S. The following existence and uniqueness result is a slight extension of Theorem 1.3 of Taylor and Williams [23]. Only initial distributions $\nu = \delta_x$ for $x \in S$ were treated in that paper. The way in which the more general result can be deduced from that in [23] is indicated after the statement of the theorem.

Theorem 3.1 *Suppose that R is completely-$\mathcal{S}$ and let ν be a probability measure on $(S, \mathcal{B})$. There exists an SRBM associated with $(S, \theta, \Gamma, R, \nu)$. Let W on $(\Omega,\ \mathcal{F},\ \{\mathcal{F}_t\},\ P_\nu)$ be such* *an SRBM with associated "pushing process" Y. Then the probability measure Q_ν induced on $(C^{2d}, \mathcal{M}^{2d})$ by the pair (W, Y) under P_ν is unique, and hence the law of any SRBM, together with its associated pushing process, for the data $(S, \theta, \Gamma, R, \nu)$ is unique. When $\nu = \delta_x$ for $x \in S$, we simply write Q_x in place of Q_{δ_x}. Then the family $\{Q_x, x \in S\}$ together with the canonical process $W(w, y) = w$ on $C^{2d} = \{(w, y) : w \in C^d, y \in C^d\}$ defines a Feller continuous strong Markov process.*

Proof of Theorem 3.1 for general ν. From Taylor and Williams [23], it is known that Theorem 3.1 holds when ν is allowed to run over the unit masses δ_x, $x \in S$. In particular, one has the Feller continuity of the measures $\{Q_x, x \in S\}$, and so for any probability measure ν on $(S, \mathcal{B})$ one can define a probability measure Q_ν on $(C^{2d}, \mathcal{M}^{2d})$ by

$$Q_\nu(A) = \int_S Q_x(A)\nu(dx) \quad \text{for any } A \in \mathcal{M}^{2d}. \tag{1}$$

Now, by the result in Taylor and Williams [23], for each $x \in S$, properties (i)–(iv) of Definition 3.1 hold on $(C^{2d}, \mathcal{M}^{2d}, \{\mathcal{M}_t^{2d}\}, Q_x)$ with δ_x in place of ν, $P = Q_x$, $W(w, y) = w$, $Y(w, y) = y$, $X(w, y) = w - Ry$ for each $(w, y) \in C^{2d} = C^d \times C^d$. It can then be readily verified that these properties also hold with ν in place of δ_x and Q_ν in place of Q_x. This establishes the existence of an SRBM associated with $(S, \theta, \Gamma, R, \nu)$.

To establish uniqueness in law for any SRBM associated with the data $(S, \theta, \Gamma, R, \nu)$, let P_ν be the probability measure induced on $(C^{2d}, \mathcal{M}^{2d})$ by such an SRBM and its associated pushing process. For $\mathcal{F}_0 = \sigma\{w(0) : (w, y) \in C^{2d}\}$, let $\{P^0_{(w,y)} : (w, y) \in C^{2d}\}$ denote a regular conditional probability distribution for P_ν given $\mathcal{F}_0$ (cf. Stroock and Varadhan [21], pages 16, 34). In particular, there is a P_ν-null set $N \in \mathcal{F}_0$ such that for each $(w, y) \in C^{2d} \setminus N$, under $P^0_{(w,y)}$, $W(\tilde{w}, \tilde{y}) = \tilde{w}$, $Y(\tilde{w}, \tilde{y}) = \tilde{y}$ and $X(\tilde{w}, \tilde{y}) = \tilde{w} - R\tilde{y}$ for $(\tilde{w}, \tilde{y}) \in C^{2d}$, define an SRBM on $(C^{2d}, \mathcal{M}^{2d}, \{\mathcal{M}_t^{2d}\})$ associated with the data $(S, \theta, \Gamma, R, \delta_{w(0)})$, i.e., the conditions of Definition 3.1 are satisfied with $\delta_{w(0)}$ and $P^0_{(w,y)}$ in place of ν and P. Then by the uniqueness in law established by Taylor and Williams [23], it follows that for each $(w, y) \in C^{2d} \setminus N$, $P^0_{(w,y)} = Q_{w(0)}$. Since the distribution of $w(0)$ under P_ν is ν, it then follows that for all $A \in \mathcal{M}^{2d}$,

$$P_\nu(A) \equiv \int_{C^{2d}} P^0_{(w,y)}(A) P_\nu(d(w, y)) \tag{2}$$

$$= \int_{C^{2d}} Q_{w(0)}(A) P_\nu(d(w, y)) \tag{3}$$

$$= \int_S Q_x(A)\nu(dx). \tag{4}$$

This establishes the desired uniqueness. □

4 Invariance principle for SRBMs

In this section, all of the processes

$$W, X, Y, W^n, X^n, Y^n, \tilde{W}^n, \tilde{Y}^n, \check{W}^n, \check{X}^n, \check{Y}^n, \alpha^n, \gamma^n, \epsilon_1^n, \epsilon_2^n, \epsilon_3^n$$

are assumed to have paths in D^d. The notation $\mathbf{0}$ will denote the constant deterministic process whose constant value is the zero vector in $\mathbb{R}^d$.

Theorem 4.1 (Invariance principle for SRBMs) *Suppose R is completely-$\mathcal{S}$ and let ν be a probability measure on $(S, \mathcal{B})$. For each positive integer n, let R^n be a $d \times d$ matrix and W^n, X^n, Y^n be d-dimensional r.c.l.l. processes defined on some probability space $(\Omega^n, \mathcal{F}^n, P^n)$ such that*

(i) $W^n = X^n + R^n Y^n$,

(ii) $W^n = \tilde{W}^n + \alpha^n$ *where $\tilde{W}^n(t) \in S$ for all $t \geq 0$, P^n-a.s., and α^n converges to $\mathbf{0}$ in probability as $n \to \infty$,*

(iii) *X^n converges in distribution as $n \to \infty$ to a d-dimensional Brownian motion with drift θ, covariance matrix Γ and initial distribution ν,*

(iv) $Y^n = \tilde{Y}^n + \gamma^n$ *where γ^n converges to $\mathbf{0}$ in probability as $n \to \infty$ and there are constants $\delta^n \geq 0$ such that $\delta^n \to 0$ as $n \to \infty$ and P^n-a.s. for $i = 1, \ldots, d$,*

  (a) $\tilde{Y}_i^n(0) = 0$,

  (b) *$\tilde{Y}_i^n$ is non-decreasing,*

  (c) $\int_0^\infty 1_{(\delta^n, \infty)}(\tilde{W}_i^n(s))\, d\tilde{Y}_i^n(s) = 0$.

(v) *R^n converges to R (componentwise) as $n \to \infty$.*

Define $\mathcal{Z}^n = (W^n, X^n, Y^n)$ for each n. The sequence of processes $\{\mathcal{Z}^n\}_{n=1}^\infty$ is tight, i.e., the sequence of probability measures induced by $\{\mathcal{Z}^n\}_{n=1}^\infty$ on $(D^{3d}, \mathcal{M}^{3d})$ is tight. Any (weak) limit point of this sequence, i.e., any process $\mathcal{Z}$ such that $\mathcal{Z}^{n_k} \Rightarrow \mathcal{Z}$ as $k \to \infty$ for some subsequence $\{n_k\}$ of $\{n\}$, is of the form $\mathcal{Z} = (W, X, Y)$ where W, X, Y are P-a.s. continuous d-dimensional processes defined on some probability space $(\Omega, \mathcal{F}, P)$ such that conditions (i), (ii), (iii)(a) and (iv) of Definition 3.1 hold with $\mathcal{F}_t = \sigma\{\mathcal{Z}(s) : 0 \leq s \leq t\}$, $t \geq 0$. Furthermore, if the following condition (vi) holds, then $\mathcal{Z}^n \Rightarrow \mathcal{Z}$ as $n \to \infty$ where $\mathcal{Z} = (W, X, Y)$ satisfies all of the conditions of Definition 3.1 and in particular, W is an SRBM associated with $(S, \theta, \Gamma, R, \nu)$.

(vi) *For each (weak) limit point $\mathcal{Z} = (W, X, Y)$ of $\{\mathcal{Z}^n\}_{n=1}^\infty$, $\{X(t) - X(0) - \theta t, \mathcal{F}_t, t \geq 0\}$ is a martingale where $\mathcal{F}_t = \sigma\{\mathcal{Z}(s) : 0 \leq s \leq t\}$ for all $t \geq 0$.*

Remark. The Lebesgue-Stieltjes integral condition in (iv)(c) above may be written more explicitly as follows. Let $\tilde{Y}_i^n = {}^c\tilde{Y}_i^n + {}^j\tilde{Y}_i^n$ be the unique decomposition of $\tilde{Y}_i^n$ into its continuous part ${}^c\tilde{Y}_i^n$ and jump part ${}^j\tilde{Y}_i^n$, and let $\tilde{J}_i^n$ denote the set of times t at which $\Delta\tilde{Y}_i^n(t) \equiv \tilde{Y}_i^n(t) - \tilde{Y}_i^n(t-) > 0$, i.e., at which $\tilde{Y}_i^n$ jumps. Then the integral condition in (iv)(c) above is equivalent to the two conditions that $\int_0^\infty 1_{(\delta^n, \infty)}(\tilde{W}_i^n(s))\, d\,{}^c\tilde{Y}_i^n(s) = 0$ and for every $t \in \tilde{J}_i^n$, $\tilde{W}_i^n(t) \le \delta^n$. Note that care is needed in paraphrasing condition (iv)(c) when the paths of $\tilde{W}^n, X^n, \tilde{Y}^n$ may have jumps. The verbal description that "$\tilde{Y}_i^n$ can have a point of increase at t only if $\tilde{W}_i^n(t) \le \delta^n$", although pleasing to the ear, is not correct, because for example it precludes the scenario where $\alpha^n = \gamma^n = \mathbf{0}$ and W_i^n leaves the threshold region $\{w \in S : w_i \le \delta^n\}$ by a positive jump of X_i^n at time t and Y_i^n does not have a jump at t but increases continuously from the *left* at t. This scenario, which is relevant to the approximation of workload processes in queueing networks, is however allowed by the integral form of the condition given above. Note, in particular, that the (at most countably many) times at which $\tilde{W}_i^n$ is discontinuous are not charged by the measure associated with the continuous non-decreasing function ${}^c\tilde{Y}_i^n$.

Some sufficient conditions for (vi) to hold are given in the following proposition and then for the convenience of the reader, the combination of Theorem 4.1 with this proposition is formally stated as Corollary 4.3 below.

Proposition 4.2 *Assume the hypotheses (excluding (vi)) of Theorem 4.1 hold. If, in addition, any of the following conditions (I)–(III) holds, then condition (vi) of Theorem 4.1 is satisfied.*

(I) *For any triple of d-dimensional $\{\mathcal{F}_t\}$-adapted processes (W, X, Y) defined on some filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}, P)$ and satisfying conditions (i), (ii), (iii)(a) and (iv) of Definition 3.1, the pair (W, Y) is adapted to the filtration generated by X and the P-null sets.*

(II) *$R = I + Q$, where I is the $d \times d$ identity matrix and Q is a $d \times d$ matrix such that $|Q|$, the matrix obtained by replacing all of the entries in Q by their absolute values, has spectral radius strictly less than one.*

(III) *$X^n = \check{X}^n + \epsilon_1^n$, $Y^n = \check{Y}^n + \epsilon_2^n$, $W^n = \check{W}^n + \epsilon_3^n$, where $\epsilon_1^n, \epsilon_2^n, \epsilon_3^n$ converge to $\mathbf{0}$ in probability as $n \to \infty$, $\{\check{X}^n(t) - \check{X}^n(0)\}_{n=1}^{\infty}$ is uniformly integrable for each $t \ge 0$, and there is a sequence of constants $\{\theta^n\}_{n=1}^{\infty}$ in $\mathbb{R}^d$ such that $\lim_{n\to\infty} \theta^n = \theta$ and for each n, $\{\check{X}^n(t) - \check{X}^n(0) - \theta^n t, t \ge 0\}$ is a P^n-martingale with respect to the filtration generated by $(\check{W}^n, \check{X}^n, \check{Y}^n)$.*

Remark. Condition (I) is most likely to be verified in situations where a strong form of pathwise uniqueness for SRBMs holds. Condition (II) is a known sufficient condition for such strong pathwise uniqueness. Condition (III) is oriented towards situations where only weak uniqueness is known and in particular it is used in a subsequent paper [26] where general conditions are given under which a heavy traffic limit theorem holds for open multiclass queueing networks.

Corollary 4.3 *Assume that R is completely-$\mathcal{S}$, ν is a probability measure on $(S, \mathcal{B})$, conditions (i)–(v) of Theorem 4.1 hold and at least one of the conditions (I)–(III) of Proposition 4.2 holds. Then $\mathcal{Z}^n \equiv (W^n, X^n, Y^n) \Rightarrow \mathcal{Z}$ as $n \to \infty$ where $\mathcal{Z} \equiv (W, X, Y)$ satisfies the conditions of Definition 3.1 and in particular, W is an SRBM associated with $(S, \theta, \Gamma, R, \nu)$.*

5 Oscillation inequality for solutions of a perturbed Skorokhod problem

The following oscillation inequality for solutions of a perturbed Skorokhod problem is the key to the proof of the tightness result claimed in Theorem 4.1. With $\delta = 0$ (and often $x(\cdot)$ restricted to be continuous), such problems are called Skorokhod problems and they can sometimes be used to give "strong" constructions of SRBMs (cf. [12, 9]). The oscillation inequality given below is a generalization of one given by Bernard and El Kharroubi [1] for the case where $x(\cdot)$ is continuous and $\delta = 0$.

In this section, for any $0 \le t_1 < t_2 < \infty$, $D([t_1,t_2],\mathbb{R}^d)$ denotes the set of functions $w : [t_1,t_2] \to \mathbb{R}^d$ that are right continuous on $[t_1,t_2)$ and have finite left limits on $(t_1,t_2]$.

Theorem 5.1 *Assume that R is completely-$\mathcal{S}$. Suppose that $\delta \ge 0$, $0 \le t_1 < t_2 < \infty$ and $w, x, y \in D([t_1,t_2],\mathbb{R}^d)$ are such that*

(i) $w(t) = x(t) + Ry(t)$ *for all* $t \in [t_1,t_2]$,

(ii) $w(t) \in S$ *for all* $t \in [t_1,t_2]$,

(iii) *for* $i = 1,\dots,d$,

- **(a)** $y_i(t_1) \ge 0$,
- **(b)** y_i *is non-decreasing,*
- **(c)** $\int_{[t_1,t_2]} 1_{(\delta,\infty)}(w_i(s))\, dy_i(s) = 0$.

Then there is a constant $C > 0$ depending only on R such that

$$Osc(y,[t_1,t_2]) \le C(Osc(x,[t_1,t_2]) + \delta), \tag{5}$$

$$Osc(w,[t_1,t_2]) \le C(Osc(x,[t_1,t_2]) + \delta), \tag{6}$$

where

$$Osc(f,[t_1,t_2]) \equiv \sup\{|f(t) - f(s)| : t_1 \le s < t \le t_2\} \tag{7}$$

for any $f \in D([t_1,t_2],\mathbb{R}^d)$ and $|a| = \max_{i=1}^{d} |a_i|$ for any $a \in \mathbb{R}^d$.

Remark. As in the Remark following Theorem 4.1, the Lebesgue-Stieltjes integral condition in (iii)(c) above can be written more explicitly in terms of the continuous and jump parts of y_i.

Proof. It suffices to prove the result for all $\delta > 0$. For if the conditions of the theorem hold with $\delta = 0$, then they hold for any $\delta > 0$, and assuming the result holds for strictly positive δ, by letting $\delta \to 0$ in the inequalities (5)–(6), one obtains the result for $\delta = 0$. (The fact that C only depends upon R is used here.) Thus in the following proof, δ will always be strictly positive.

The proof is a modification of that of Bernard and El Kharroubi [1] which applies when w, x, y are continuous and $\delta = 0$. It proceeds by induction on the dimension d.

For $d = 1$, $R = r > 0$ by the completely-$\mathcal{S}$ assumption. For $0 \le t \le t_2 - t_1 \equiv \bar{t}$ and w, x, y in $D([t_1, t_2], \mathbb{R})$ satisfying (i)–(iii), define

$$\begin{aligned}\bar{w}(t) &= w(t + t_1),\\ \bar{x}(t) &= w(t_1) + x(t + t_1) - x(t_1),\\ \bar{y}(t) &= y(t + t_1) - y(t_1),\end{aligned}$$

so that $\bar{w}, \bar{x}, \bar{y} \in D([0, \bar{t}\,], \mathbb{R})$ and

(i)′ $\bar{w}(t) = \bar{x}(t) + r\bar{y}(t)$ for all $t \in [0, \bar{t}\,]$,

(ii)′ $\bar{w}(t) \in \mathbb{R}_+$ for all $t \in [0, \bar{t}\,]$,

(iii)′ **(a)** $\bar{y}(0) = 0$,

(b) $\bar{y}$ is non-decreasing,

(c) $\int_{[0,\bar{t}]} 1_{(\delta,\infty)}(\bar{w}(s))\, d\bar{y}(s) = 0$.

For $z \in \mathbb{R}$, let $z^- = \max(-z, 0)$. Then we define $\hat{y}(t) = \sup\{(\bar{x}(s))^- : 0 \le s \le t\}/r$ and $\tilde{y}(t) = \hat{y}(t) + \frac{\delta}{r}$ for all $t \in [0, \bar{t}\,]$. It is well known and easy to verify that given $\bar{x} \in D([0, \bar{t}\,]), \mathbb{R})$ satisfying $\bar{x}(0) \in \mathbb{R}_+$, with $\bar{y} = \hat{y}$ and $\bar{w}$ given by (i)′, the above conditions (i)′–(iii)′ hold with $\delta = 0$ in (iii)′(c). It is claimed that with $\bar{x}$ fixed, any solution $\bar{y}, \bar{w}$ of (i)′–(iii)′ satisfies

$$\hat{y}(t) \le \bar{y}(t) \le \tilde{y}(t) \quad \text{for all } t \in [0, \bar{t}\,]. \tag{8}$$

The right hand inequality will be verified here, since that is all that we need and in any case the proof of the left hand inequality is similar. Let $\epsilon > 0$ and $\tau_\epsilon = \inf\{t \in [0, \bar{t}\,] : \bar{y}(t) > \tilde{y}(t) + \epsilon\}$ with $\inf \emptyset = \infty$. If $\tau_\epsilon < \infty$, then $\bar{y}(\tau_\epsilon) \ge \tilde{y}(\tau_\epsilon) + \epsilon$ by the right continuity of paths and

$$\begin{aligned}\bar{w}(\tau_\epsilon) &= \bar{x}(\tau_\epsilon) + r\bar{y}(\tau_\epsilon) \ge \bar{x}(\tau_\epsilon) + r\tilde{y}(\tau_\epsilon) + r\epsilon = \bar{x}(\tau_\epsilon) + r\hat{y}(\tau_\epsilon) + \delta + r\epsilon\\ &\ge 0 + \delta + r\epsilon.\end{aligned}$$

However, by the definition of τ_ϵ and (iii)′(a)–(b), $\bar{y}$ must either have a positive jump at time τ_ϵ or $\bar{y}$ must have a point of increase to the right there. In either case, since $\bar{w}(\tau_\epsilon) > \delta$, this contradicts (iii)′(c). Thus $\tau_\epsilon = \infty$ for each $\epsilon > 0$ and hence $\bar{y}(t) \le \tilde{y}(t)$ for all $t \in [0, \bar{t}\,]$, as desired. From this we conclude that for any w, x, y satisfying (i)–(iii),

$$\begin{aligned}Osc(y, [t_1, t_2]) &= y(t_2) - y(t_1) = \bar{y}(\bar{t}\,) \le \tilde{y}(\bar{t}\,)\\ &\le \frac{1}{r}(Osc(\bar{x}, [0, \bar{t}\,]) + \delta)\\ &= \frac{1}{r}(Osc(x, [t_1, t_2]) + \delta),\end{aligned}$$

where we used the fact that $\bar{x}(0) \geq 0$ to obtain the second line above. Then, we have

$$\begin{aligned} Osc(w,[t_1,t_2]) &\leq Osc(x,[t_1,t_2]) + r\,Osc(y,[t_1,t_2]) \\ &\leq 2(Osc(x,[t_1,t_2])+\delta). \end{aligned}$$

Thus the desired result holds for $d=1$ with $C = \max(2, \frac{1}{r})$.

Now suppose the result has been proved for dimension $d-1 \geq 1$. The aim is to prove that it holds for dimension d.

First it will be shown that there is a constant C_{d-1} depending only on R such that if w, x, y with paths in $D([t_1,t_2], \mathbb{R}^d)$, for some $0 \leq t_1 < t_2 < \infty$, satisfy (i)–(iii) and $i \in \{1,\dots,d\}$ such that y_i does not increase on $[t_1,t_2]$, then

$$Osc(y,[t_1,t_2]) \leq C_{d-1}\left(Osc(x,[t_1,t_2])+\delta\right), \tag{9}$$

$$Osc(w,[t_1,t_2]) \leq C_{d-1}\left(Osc(x,[t_1,t_2])+\delta\right). \tag{10}$$

Assume w, x, y satisfy (i)–(iii) and that there is an index i such that y_i does not increase on $[t_1,t_2]$. Let $\mathcal{I} = \{1,\dots,d\}\backslash\{i\}$, let $w^{\mathcal{I}}$, $x^{\mathcal{I}}$, $y^{\mathcal{I}}$ denote the $(d-1)$-dimensional paths obtained by deleting the i^{th} component from w, x, y, respectively. Let $R^{\mathcal{I}}$ denote the $(d-1)\times(d-1)$ matrix obtained by deleting the i^{th} row and i^{th} column from R, and let r^i denote the i^{th} row of R with the i^{th} (column) element removed. Then for $t_1 \leq t \leq t_2$, since y_i does not increase on $[t_1,t_2]$,

$$w^{\mathcal{I}}(t) = w^{\mathcal{I}}(t_1) + x^{\mathcal{I}}(t) - x^{\mathcal{I}}(t_1) + R^{\mathcal{I}}(y^{\mathcal{I}}(t) - y^{\mathcal{I}}(t_1)), \tag{11}$$

$$w_i(t) = w_i(t_1) + x_i(t) - x_i(t_1) + r^i \cdot (y^{\mathcal{I}}(t) - y^{\mathcal{I}}(t_1)). \tag{12}$$

Now $R^{\mathcal{I}}$ is completely-$\mathcal{S}$ being a principal submatrix of R and so it follows from applying the induction hypothesis to the $(d-1)$-dimensional equation (11) that there is a constant $C^{\mathcal{I}}$ depending only on $R^{\mathcal{I}}$ such that

$$\begin{aligned} Osc(y^{\mathcal{I}},[t_1,t_2]) &= Osc(y^{\mathcal{I}}(\cdot) - y^{\mathcal{I}}(t_1),[t_1,t_2]) \\ &\leq C^{\mathcal{I}}(Osc(w^{\mathcal{I}}(t_1) + x^{\mathcal{I}}(\cdot) - x^{\mathcal{I}}(t_1),[t_1,t_2]) + \delta) \\ &= C^{\mathcal{I}}\Big(Osc(x^{\mathcal{I}},[t_1,t_2]) + \delta\Big), \end{aligned} \tag{13}$$

and similarly,

$$Osc(w^{\mathcal{I}},[t_1,t_2]) \leq C^{\mathcal{I}}(Osc(x^{\mathcal{I}},[t_1,t_2]) + \delta). \tag{14}$$

Combining (12) and (13) yields

$$\begin{aligned} Osc(w_i,[t_1,t_2]) &\leq Osc(x_i,[t_1,t_2]) + \sum_{j=1}^{d-1} |r^i_j| \cdot Osc(y^{\mathcal{I}},[t_1,t_2]) \\ &\leq \left(1 + C^{\mathcal{I}} \sum_{j=1}^{d-1} |r^i_j|\right)(Osc(x,[t_1,t_2]) + \delta). \end{aligned} \tag{15}$$

Since y_i does not increase on $[t_1, t_2]$, we have $Osc(y, [t_1, t_2]) = Osc(y^{\mathcal{I}}, [t_1, t_2])$. Using this and combining the inequalities (13)–(15) above, we obtain (9)–(10) with $C_{d-1} = 1 + C^{\mathcal{I}}\left(1 + \sum_{j=1}^{d-1} |r_j^i|\right)$. By maximizing over $i \in \{1, \ldots, d\}$, one can find a constant C_{d-1} that is good for all i and that depends only on R.

The main result can now be proved for dimension d. For this, let w, x, y be in $D([t_1, t_2], \mathbb{R}^d)$ and satisfy (i)–(iii). Define

$$\eta = Osc(x, [t_1, t_2]) + \delta. \tag{16}$$

By the assumption that R is completely-$\mathcal{S}$ and the invariance of this property under transpose [19], there is $\lambda \in \mathbb{R}^d_+$ with $\lambda_i > 0$ for $i = 1, \ldots, d$, such that $(\lambda' R)_i \geq 1$ for $i = 1, \ldots, d$. For $t_1 \leq s < t \leq t_2$, since y is non-decreasing in each component we have

$$Osc(y, [s, t]) = \max_{i=1}^{d} |y_i(t) - y_i(s)|, \tag{17}$$

and then

$$\begin{aligned} \lambda'(w(t) - w(s)) &= \lambda'(x(t) - x(s)) + \lambda' R(y(t) - y(s)) \\ &\geq \lambda'(x(t) - x(s)) + Osc(y, [s, t]) \end{aligned}$$

and hence

$$Osc(y, [s, t]) \leq \lambda'(w(t) - w(s)) + \left(\sum_{i=1}^{d} \lambda_i\right) Osc(x, [s, t]). \tag{18}$$

Let $\bar{\lambda} = \sum_{i=1}^{d} \lambda_i$ and $\bar{R} = \max_{i=1}^{d} \sum_{j=1}^{d} |R_{ij}|$. From (i), (18) and the facts that $\lambda_i > 0$ and $w_i(s), w_i(t) \geq 0$ for all i, it follows that

$$\begin{aligned} Osc(w, [s, t]) &\leq Osc(x, [s, t]) + \bar{R}\, Osc(y, [s, t]) && (19) \\ &\leq Osc(x, [s, t]) + \bar{R}\left(\lambda' w(t) + \bar{\lambda}\, Osc(x, [s, t])\right) \\ &\leq Osc(x, [s, t]) + \bar{R}\,\bar{\lambda}\left(\max_{i=1}^{d} w_i(t) + Osc(x, [s, t])\right) \\ &\leq (1 + \bar{R}\,\bar{\lambda})\Big(Osc(x, [s, t]) + \max_{i=1}^{d} w_i(t)\Big). && (20) \end{aligned}$$

Let $K \geq 1$ be a positive constant whose exact value will be determined in the course of the discussion of Case 1 below.

Case 1: There is $j \in \{1, \ldots, d\}$ such that $w_j(t_1) \geq K\eta$.

Let $\tau = \inf\{t \in [t_1, t_2] : w_j(t) \leq \delta\}$, where $\inf \emptyset = \infty$. Then for any $t_2' \in [t_1, \tau) \cap [t_1, t_2]$, $w_j(t) > \delta$ for all $t \in [t_1, t_2']$ and so by (iii)(c), y_j cannot increase on $[t_1, t_2']$. It then follows from (10) with j, t_2' in place of i, t_2, respectively, that

$$Osc(y, [t_1, t_2']) \leq C_{d-1}\eta, \tag{21}$$

$$Osc(w, [t_1, t_2']) \leq C_{d-1}\eta. \tag{22}$$

If $\tau = \infty$, the desired inequalities (5)–(6) hold with $C = C_{d-1}$. On the other hand, suppose $\tau < \infty$. Then, either

(a) there is some i such that y_i does not jump at τ, or

(b) y_i jumps at τ for $i = 1, \ldots, d$.

In case (a), from (21) and the continuity from the left of y_i at τ, one has

$$Osc(y_i, [t_1, \tau]) = \lim_{t_2' \uparrow \tau} Osc(y_i, [t_1, t_2']) \le C_{d-1}\eta. \tag{23}$$

Furthermore, using the same notation as that following equations (9)–(10) and letting s^i denote the i^{th} column of R with the i^{th} (row) element removed, one has for $t \in [t_1, \tau]$,

$$w^{\mathcal{I}}(t) = x^{\mathcal{I}}(t) + s^i y_i(t) + R^{\mathcal{I}} y^{\mathcal{I}}(t). \tag{24}$$

This equation can be expressed in the same form as (11), but with $x^{\mathcal{I}} + s^i y_i, \tau$, in place of $x^{\mathcal{I}}, t_2$, respectively. Then as for (13), using the fact that the constant $C^{\mathcal{I}}$ depends only on $R^{\mathcal{I}}$, it follows that

$$\begin{aligned} Osc(y^{\mathcal{I}}, [t_1, \tau]) &\le C^{\mathcal{I}}(Osc(x^{\mathcal{I}} + s^i y_i, [t_1, \tau]) + \delta) \\ &\le C^{\mathcal{I}}(Osc(x^{\mathcal{I}}, [t_1, \tau]) + |s^i| Osc(y_i, [t_1, \tau]) + \delta). \end{aligned}$$

Combining this with (23) yields

$$Osc(y^{\mathcal{I}}, [t_1, \tau]) \le C^{\mathcal{I}}(1 + C_{d-1}|s^i|)\eta. \tag{25}$$

Note from the definition of C_{d-1} that $C^{\mathcal{I}} \le C_{d-1}$ and so combining this with (23) and (25), on setting $\bar{C} = C_{d-1}(1 + C_{d-1}\max_{i=1}^{d}|s^i|)$, one has

$$Osc(y, [t_1, \tau]) \le \bar{C}\eta, \tag{26}$$

and so (cf. (19)),

$$Osc(w, [t_1, \tau]) \le (1 + \bar{R}\bar{C})\eta. \tag{27}$$

Since $w_j(t_1) \ge K\eta$, if one chooses $K \ge 3 + \bar{R}\bar{C}$, then by (27),

$$w_j(\tau) \ge w_j(t_1) - Osc(w, [t_1, \tau]) \ge (K - 1 - \bar{R}\bar{C})\eta \ge 2\eta > \delta, \tag{28}$$

but this contradicts the assumption that $\tau < \infty$.

In case (b), y_i has a jump at τ for all i and so by (iii)(c), $w_i(\tau) \le \delta$ for all i. Then by (20),

$$\begin{aligned} Osc(w, [t_1, \tau]) &\le (1 + \bar{R}\,\bar{\lambda})(Osc(x, [t_1, \tau]) + \delta) \\ &\le (1 + \bar{R}\,\bar{\lambda})\eta, \end{aligned}$$

and if $K \ge 3 + \bar{R}\bar{\lambda}$,

$$w_j(\tau) \ge (K - 1 - \bar{R}\,\bar{\lambda})\eta \ge 2\eta > \delta, \tag{29}$$

which again contradicts the assumption that $\tau < \infty$.

Thus, by choosing $K = 3 + \bar{R}(\bar{C} + \bar{\lambda})$, one ensures that the only valid option is that $\tau = \infty$ and in this case (5)–(6) hold with $C = C_{d-1}$.

Case 2: $w_i(t_1) < K\eta$ **for** $i = 1, \ldots, d$.
Then there are two possibilities:

(a) $\max_{i=1}^{d} w_i(t) \leq K\eta$ for all $t \in [t_1, t_2]$, or

(b) there is $t \in [t_1, t_2]$ and $j \in \{1, \ldots, d\}$ such that $w_j(t) > K\eta$.

In case (a), since $w_i(t) \geq 0$ for all i and $t \in [t_1, t_2]$,

$$Osc(w, [t_1, t_2]) \leq K\eta \tag{30}$$

and it follows from (18) that

$$\begin{aligned} Osc(y, [t_1, t_2]) &\leq \bar{\lambda}(Osc(w, [t_1, t_2]) + Osc(x, [t_1, t_2])) \\ &\leq \bar{\lambda}(K+1)\eta. \end{aligned} \tag{31}$$

Then (5)–(6) hold with $C = \max(\bar{\lambda}(K+1), K)$.

Let $\tau = \inf\{t \in [t_1, t_2] : w_i(t) > K\eta \text{ for some } i\}$. Assuming we are in case (b), we have $\tau < \infty$. Then for $t_2' \in [t_1, \tau) \cap [t_1, t_2]$, by the same reasoning as for case (a), the oscillation estimates (30)–(31) hold with t_2' in place of t_2, i.e.,

$$\begin{aligned} Osc(w, [t_1, t_2']) &\leq K\eta \\ Osc(y, [t_1, t_2']) &\leq \bar{\lambda}(K+1)\eta. \end{aligned} \tag{32}$$

Since $\tau < \infty$, then $w_j(\tau) \geq K\eta$ for some j and the analysis of Case 1 implies that

$$\begin{aligned} Osc(y, [\tau, t_2]) \leq C_{d-1}\eta, \\ Osc(w, [\tau, t_2]) \leq C_{d-1}\eta. \end{aligned}$$

It remains to consider the size of any jump of w (or y) at the time $\tau < \infty$. Now for j as indicated above, $w_j(\tau) \geq K\eta > \delta$ and so by (iii)(c), y_j cannot have a jump at τ. Then by letting $t_2' \uparrow \tau$ in a similar manner to that for Case 1(a), one obtains from (32) that

$$Osc(y_j, [t_1, \tau]) \leq \bar{\lambda}(K+1)\eta,$$

and using this in place of (23) one sees that (26)–(27) hold with

$$\hat{C} = \bar{\lambda}(K+1) + C_{d-1}(1 + \bar{\lambda}(K+1)\max_{i=1}^{d}|s^i|) \tag{33}$$

in place of $\bar{C}$. Consolidating the above yields

$$
\begin{aligned}
Osc(w,[t_1,t_2]) &\le Osc(w,[t_1,\tau]) + Osc(w,[\tau,t_2]) \\
&\le (1+\bar{R}\hat{C}+C_{d-1})\eta,
\end{aligned}
$$

and

$$
Osc(y,[t_1,t_2]) \le (\hat{C}+C_{d-1})\eta.
$$

Combining all of the cases, one sees that (5)–(6) hold where the constant C can be chosen to depend only on R. ▫

6 Proof of Theorem 4.1

Assume the hypotheses (i)–(v) of Theorem 4.1 hold. First the tightness of $\{\mathcal{Z}^n\}_{n=1}^{\infty}$ will be proved. Simple algebraic manipulations yield

$$
\tilde{W}^n = X^n + \epsilon^n + S^n\tilde{Y}^n + R\tilde{Y}^n \tag{34}
$$

where

$$
\begin{aligned}
\epsilon^n &\equiv -\alpha^n + R^n\gamma^n, \\
S^n &\equiv R^n - R.
\end{aligned}
$$

The hypotheses on α^n, γ^n and R^n imply that $\epsilon^n \to \mathbf{0}$ in probability as $n\to\infty$ and S^n converges componentwise to the $d\times d$ matrix of all zeroes as $n\to\infty$.

Now for each n, P^n-a.s., for any $0\le t_1<t_2<\infty$, the hypotheses of Theorem 5.1 are satisfied with $\tilde{W}^n$, $X^n+\epsilon^n+S^n\tilde{Y}^n$, $\tilde{Y}^n$, δ^n in place of w, x, y, δ, respectively. Hence,

$$
Osc(\tilde{Y}^n,[t_1,t_2]) \le C(Osc(X^n+\epsilon^n+S^n\tilde{Y}^n,[t_1,t_2])+\delta^n) \tag{35}
$$

$$
Osc(\tilde{W}^n,[t_1,t_2]) \le C(Osc(X^n+\epsilon^n+S^n\tilde{Y}^n,[t_1,t_2])+\delta^n) \tag{36}
$$

where $C>0$ is a constant depending only on R. Now (35) implies that

$$
Osc(\tilde{Y}^n,[t_1,t_2]) \le C\left(Osc(X^n+\epsilon^n,[t_1,t_2]) + |S^n|Osc(\tilde{Y}^n,[t_1,t_2])+\delta^n\right)
$$

where $|S^n| \equiv \max_{i=1}^{d}\sum_{j=1}^{d}|S^n_{ij}| \to 0$ as $n\to\infty$. Then for all n sufficiently large that $C|S^n|<\frac{1}{2}$,

$$
Osc(\tilde{Y}^n,[t_1,t_2]) \le 2C(Osc(X^n+\epsilon^n,[t_1,t_2])+\delta^n). \tag{37}
$$

Substituting (37) in (36) one sees that for the same sufficiently large n,

$$
Osc(\tilde{W}^n,[t_1,t_2]) \le 2C(Osc(X^n+\epsilon^n,[t_1,t_2])+\delta^n). \tag{38}
$$

The assumed convergence of $\{X^n\}_{n=1}^\infty$ in distribution, together with the convergence of $\{\epsilon^n\}_{n=1}^\infty$ in probability to $\mathbf{0}$, implies that $\{X^n + \epsilon^n\}_{n=1}^\infty$ converges in distribution to a Brownian motion, and hence this last sequence is tight. Combining this with the facts that $\delta^n \to 0$ as $n \to \infty$, $\tilde{Y}^n(0) = 0$ P^n-a.s., $\tilde{W}^n(0) = X^n(0) + \epsilon^n(0)$ P^n-a.s., and the oscillation estimates (37)–(38), one sees that the sequence $\{(\tilde{W}^n, X^n, \tilde{Y}^n)\}_{n=1}^\infty$ inherits tightness from $\{X^n + \epsilon^n\}_{n=1}^\infty$. More precisely, the necessary and sufficient conditions for tightness given in Corollary 3.7.4 of Ethier and Kurtz [10] can be verified. For completeness, some details of this verification are given below.

The compact containment condition (a) of Corollary 3.7.4 in [10] follows from the tightness for $\{X^n + \epsilon^n\}_{n=1}^\infty$ (cf., (3.7.9) of Ethier and Kurtz [10]) combined with (37)–(38) and the properties of $\tilde{Y}^n$, $\tilde{W}^n$ which imply that P^n-a.s. for any $t \geq 0$,

$$
\begin{aligned}
|\tilde{Y}^n(t)| &\leq |\tilde{Y}^n(t) - \tilde{Y}^n(0)| \leq 2C(Osc(X^n + \epsilon^n, [0,t]) + \delta^n), \\
|\tilde{W}^n(t)| &\leq |\tilde{W}^n(t) - \tilde{W}^n(0)| + |\tilde{W}^n(0)| \\
&\leq 2C(Osc(X^n + \epsilon^n, [0,t]) + \delta^n) + |X^n(0) + \epsilon^n(0)|
\end{aligned}
$$

where $Osc(X^n + \epsilon^n, [0,t]) \leq 2 \sup_{0\leq s\leq t} |(X^n + \epsilon^n)(s)|$. Control on the modulus of continuity required for (b) of Corollary 3.7.4 in [10] follows directly from that for $\{X^n + \epsilon^n\}_{n=1}^\infty$ together with the oscillation estimates (37)–(38). It then follows that $\{(\tilde{W}^n, X^n, \tilde{Y}^n)\}_{n=1}^\infty$ is tight. Since $\mathcal{Z}^n = (\tilde{W}^n, X^n, \tilde{Y}^n) + (\alpha^n, \mathbf{0}, \gamma^n)$ where $\alpha^n, \gamma^n \to \mathbf{0}$ in probability as $n \to \infty$, $\{\mathcal{Z}^n\}_{n=1}^\infty$ is also tight.

Now $\{X^n + \epsilon^n\}_{n=1}^\infty$ converges in distribution to a continuous process and so for

$$
J(x(\cdot)) \equiv \int_0^\infty e^{-t} \left(\sup_{0<s\leq t} |x(s) - x(s-)| \wedge 1 \right) dt, \quad x(\cdot) \in D^d,
$$

one has $J(X^n + \epsilon^n) \Rightarrow 0$ as $n \to \infty$ (cf. [10], Theorem 3.10.2). By letting $t_1 \uparrow t_2 = s$ in (37)–(38), one has for each n and $s \geq 0$,

$$
|\tilde{Y}^n(s) - \tilde{Y}^n(s-)| + |\tilde{W}^n(s) - \tilde{W}^n(s-)| \leq 4C(|(X^n + \epsilon^n)(s) - (X^n + \epsilon^n)(s-)| + \delta^n).
$$

It then follows from the facts that $J(X^n + \epsilon^n) \Rightarrow 0$ as $n \to \infty$ and $\delta^n \to 0$ as $n \to \infty$, that $\{J(\tilde{W}^n)\}_{n=1}^\infty$, $\{J(\tilde{Y}^n)\}_{n=1}^\infty$ converge to zero in distribution and hence in probability as $n \to \infty$. Furthermore, α^n, $\gamma^n \to \mathbf{0}$ in probability as $n \to \infty$, and so it readily follows that $\{J(W^n)\}_{n=1}^\infty$, $\{J(Y^n)\}_{n=1}^\infty$ converge to zero in probability as $n \to \infty$. It can then be concluded from the above that any (weak) limit point of $\{\mathcal{Z}^n\}_{n=1}^\infty$ has continuous paths a.s. (cf. [10], Theorem 3.10.2).

Now suppose that $\mathcal{Z} = (W, X, Y)$ is a (weak) limit point of $\{\mathcal{Z}^n\}_{n=1}^\infty$, i.e., there is a subsequence $\{\mathcal{Z}^{n_k}\}_{k=1}^\infty$ such that $\mathcal{Z}^{n_k} \Rightarrow \mathcal{Z}$ as $k \to \infty$. Let $\mathcal{F}_t = \sigma\{\mathcal{Z}(s) : 0 \leq s \leq t\}$ for each $t \geq 0$. From the last paragraph above, one knows that a.s. the paths of $\mathcal{Z}$ are continuous. Combining this with the assumed convergence in probability to $\mathbf{0}$ of α^n, γ^n as $n \to \infty$, one has that the sequence $\{(\mathcal{Z}^{n_k}, \alpha^{n_k}, \gamma^{n_k})\}_{k=1}^\infty$ of D^{5d}-valued processes converges in distribution to $(\mathcal{Z}, \mathbf{0}, \mathbf{0})$. For the purpose of verifying that $\mathcal{Z}$ satisfies properties (i), (ii), (iii)(a) and (iv) of Definition 3.1, one may use the Skorokhod representation theorem ([10], Theorem 3.1.8) and the fact that the limit is a.s. continuous to replace the above sequence of processes by one that is term-by-term equivalent in distribution to the original one and which a.s. converges u.o.c. (uniformly on each compact time interval). With this simplification, it is easily verified that the limit process $\mathcal{Z} = (W, X, Y)$ inherits properties (i), (ii), (iii)(a), (iv)(a)–(b) of Definition 3.1 from properties (i), (ii), (iii), (iv)(a)–(b), (v) of Theorem 4.1. For property (iv)(c) of Definition 3.1, note that it suffices to show that

$$\int_0^t f_m(W_i(s))dY_i(s) = 0 \quad \text{a.s. for each } t \geq 0, \quad i = 1, \ldots, d, \quad \text{and} \quad m = 1, 2, \ldots, \tag{39}$$

where $\{f_m\}_{m=1}^\infty$ is a sequence of continuous functions such that for each m, $f_m : \mathbb{R} \to [0,1]$, $f_m(x) = 0$ for $x \leq \frac{1}{m}$, and $f_m(x) = 1$ for $x \geq \frac{2}{m}$. For this, fix $i \in \{1, \ldots, d\}$, $m \in \{1, 2, \ldots\}$, and $t \geq 0$. By property (iv)(c) of Theorem 4.1, one has for all n_k such that $\delta^{n_k} \leq \frac{1}{m}$,

$$\int_{[0,t]} f_m(\tilde{W}_i^{n_k}(s))d\tilde{Y}_i^{n_k}(s) = 0 \quad \text{a.s.} \tag{40}$$

Now, the almost sure convergence assumed above implies that a.s. as $k \to \infty$,

$$\tilde{W}_i^{n_k} \equiv W_i^{n_k} - \alpha_i^{n_k} \to W_i \qquad \text{and} \qquad \tilde{Y}_i^{n_k} \equiv Y_i^{n_k} - \gamma_i^{n_k} \to Y_i, \tag{41}$$

uniformly on compact time intervals, and since f_m is uniformly continuous, it follows that a.s.,

$$f_m(\tilde{W}_i^{n_k}) \to f_m(W_i) \quad \text{u.o.c. as } k \to \infty. \tag{42}$$

We will show that a.s.,

$$\int_{[0,t]} f_m(\tilde{W}_i^{n_k}(s))d\tilde{Y}_i^{n_k}(s) \to \int_{[0,t]} f_m(W_i(s))dY_i(s) \quad \text{as } k \to \infty. \tag{43}$$

Our argument is a modification of the proof of Lemma 2.4 of Dai-Williams [7]. Note that

$$\begin{aligned} &\int_{[0,t]} f_m(\tilde{W}_i^{n_k}(s))\, d\tilde{Y}_i^{n_k}(s) - \int_{[0,t]} f_m(W_i(s))dY_i(s) \\ &= \int_{[0,t]} \left(f_m(\tilde{W}_i^{n_k}(s)) - f_m(W_i(s))\right) d\tilde{Y}_i^{n_k}(s) + \int_{[0,t]} f_m(W_i(s))d(\tilde{Y}_i^{n_k} - Y_i)(s). \end{aligned} \tag{44}$$

The first term in the last line above converges to zero a.s. as $k \to \infty$, because its absolute value is dominated by

$$\sup_{0 \leq s \leq t} |f_m(\tilde{W}_i^{n_k}(s)) - f_m(W_i(s))|\, \tilde{Y}_i^{n_k}(t), \tag{45}$$

which tends to zero a.s. as $k \to \infty$, by (41) and (42). For the last term in (44), note that since $f_m(W_i(\cdot))$ is a.s. continuous, if we define

$$z^l(s) = \sum_{j=0}^{2^l - 1} f_m\left(W_i\left(\frac{(j+1)t}{2^l}\right)\right) 1_{\left(\frac{jt}{2^l}, \frac{(j+1)t}{2^l}\right]}(s) + f_m(W_i(0))1_{\{0\}}(s), \tag{46}$$

for all $s \geq 0$ and $l = 1, 2, \ldots$, then a.s., $z^l(\cdot) \to f_m(W_i(\cdot))$ uniformly on $[0, t]$ as $l \to \infty$. We then have

$$\begin{aligned}
&\left| \int_{[0,t]} f_m(W_i(s)) d(\tilde{Y}_i^{n_k} - Y_i)(s) \right| \\
&\leq \left| \int_{[0,t]} (f_m(W_i(s)) - z^l(s)) d(\tilde{Y}_i^{n_k} - Y_i)(s) \right| + \left| \int_{[0,t]} z^l(s) d(\tilde{Y}_i^{n_k} - Y_i)(s) \right| \\
&\leq \sup_{0 \leq s \leq t} \left| f_m(W_i(s)) - z^l(s) \right| (\tilde{Y}_i^{n_k}(t) + Y_i(t)) \\
&\quad + \left| \sum_{j=0}^{2^l - 1} f_m \left(W_i \left(\frac{(j+1)t}{2^l} \right) \right) \left(\tilde{Y}_i^{n_k} \left(\frac{(j+1)t}{2^l} \right) - Y_i \left(\frac{(j+1)t}{2^l} \right) - \left(\tilde{Y}_i^{n_k} \left(\frac{jt}{2^l} \right) - Y_i \left(\frac{jt}{2^l} \right) \right) \right) \right| .
\end{aligned}$$

By first fixing l and letting $k \to \infty$, and then letting $l \to \infty$, we see that a.s. the lim sup as $k \to \infty$ of the first line above is zero. This proves (43). With this, (39) follows from (40). This completes the proof that any (weak) limit point of $\{\mathcal{Z}^n\}_{n=1}^\infty$ satisfies properties (i), (ii), (iii)(a) and (iv) of Definition 3.1 with $\{\mathcal{F}_t\}$ equal to the filtration generated by $\mathcal{Z} = (W, X, Y)$.

Now consider such a (weak) limit point $\mathcal{Z}$ defined on some probability space $(\Omega, \mathcal{F}, P)$ and let $\mathcal{F}_t = \sigma\{\mathcal{Z}(s) : 0 \leq s \leq t\}$, $t \geq 0$. If condition (vi) of Theorem 4.1 holds then all of the conditions of Definition 3.1 will hold and W is an SRBM associated with $(S, \theta, \Gamma, R, \nu)$. By Theorem 3.1, the law of the pair (W, Y) is unique and since $X = W - RY$ a.s., it follows that all (weak) limit points of $\{\mathcal{Z}^n\}_{n=1}^\infty$ have the same law. Combining this uniqueness with the tightness of $\{\mathcal{Z}^n\}_{n=1}^\infty$, it follows by a standard argument that the whole sequence $\{\mathcal{Z}^n\}_{n=1}^\infty$ converges in distribution to a $(3d)$-dimensional process $\mathcal{Z} = (W, X, Y)$ that satisfies the conditions of Definition 3.1 with $\mathcal{F}_t = \sigma\{\mathcal{Z}(s) : 0 \leq s \leq t\}$, $t \geq 0$, and in particular, W is an SRBM associated with $(S, \theta, \Gamma, R, \nu)$. □

7 Proof of Proposition 4.2

Let $\mathcal{Z} = (W, X, Y)$, defined on a probability space $(\Omega, \mathcal{F}, P)$, be a weak limit of the sequence $\{\mathcal{Z}^n\}_{n=1}^\infty$ defined in Theorem 4.1. For each $t \geq 0$, let $\mathcal{F}_t = \sigma\{\mathcal{Z}(s) : 0 \leq s \leq t\}$ and $\mathcal{G}_t = \sigma\{X(s) : 0 \leq s \leq t\} \vee \mathcal{N}$, where $\mathcal{N}$ denotes the collection of P-null sets in $\mathcal{F}$. Then $\{X(t) - X(0) - \theta t, t \geq 0\}$ is a driftless d-dimensional Brownian motion that starts from the origin and it is well known to be a martingale with respect to $\{\mathcal{G}_t, t \geq 0\}$.

Now suppose condition (I) of Proposition 4.2 holds. Then $\mathcal{F}_t \subset \mathcal{G}_t$ and since $\{X(t) - X(0) - \theta t, t \geq 0\}$ is adapted to $\{\mathcal{F}_t, t \geq 0\}$, it remains a martingale with respect to this (possibly) smaller filtration. Thus, condition (vi) of Theorem 4.1 has been verified.

Now suppose that instead of condition (I), it is assumed that condition (II) holds. Then there is a continuous (path-to-path) mapping π from the space of continuous d-dimensional paths $x(\cdot)$ that start in S into the space $C^d \times C^d$ of continuous paths $(w, y)(\cdot)$ living in $\mathbb{R}^d \times \mathbb{R}^d$ such that for each $x(\cdot)$, $(w, y)(\cdot) = \pi(x(\cdot))$ satisfies conditions (i), (ii), (iv)(a)–(c) of

Definition 3.1 with $(w,x,y)(\cdot)$ in place of $(W,X,Y)(\cdot)$ and the P-a.s. omitted. Furthermore, for each $x(\cdot)$, $(w,y)(\cdot)=\pi(x(\cdot))$ is the unique pair in $C^d\times C^d$ with these properties, and the mapping π is non-anticipating in the sense that the values of $(w,y)(\cdot)=\pi(x(\cdot))$ on $[0,t]$ depend only on the values of $x(\cdot)$ on $[0,t]$, for each $t\geq 0$. (The latter is equivalent to the property that for each $t\geq 0$, the mapping $x(\cdot)\to\pi(x(\cdot))(t)$ is measurable with respect to the σ-field $\mathcal{M}_t$ on the space of $x(\cdot)$-paths.) The result cited above is an easy generalization of one due to Harrison and Reiman [12]. They considered the case where the entries of Q are all non-positive and Q has zeros on its diagonal, but their contraction argument is easily verified to extend to the case cited above (cf. Dupuis and Ishii [9]).

Now, to verify that the conclusion of (vi) of Theorem 4.1 holds, in view of the discussion under condition (I), it suffices to show that (W,Y) is adapted to $\{\mathcal{G}_t,\, t\geq 0\}$. By the path-to-path mapping result cited above,

$$(W,Y)=\pi(X)\qquad P\text{-a.s.,}\tag{47}$$

and by the non-anticipating property of π, (W,Y) is adapted to the filtration $\{\mathcal{G}_t,\, t\geq 0\}$. (The addition of the P-null sets is needed because (47) only holds P-a.s.) This completes the verification that condition (II) implies condition (vi) of Theorem 4.1.

Finally, suppose that condition (III) holds. Then assuming $\mathcal{Z}^{n_k}\Rightarrow\mathcal{Z}$, one obtains $\check{\mathcal{Z}}^{n_k}\equiv(\check{W}^{n_k},\check{X}^{n_k},\check{Y}^{n_k})\Rightarrow\mathcal{Z}$ as $k\to\infty$. For any positive integer m, $0\leq s_1<s_2<\ldots<s_m=s<t<\infty$, $f_1,\ldots,f_m\in C_b(\mathbb{R}^{3d})$ and $u=s$ or t, by the uniform integrability of $\{\check{X}^{n_k}(u)-\check{X}^{n_k}(0)\}_{k=1}^{\infty}$ and the assumed convergence in distribution,

$$\begin{aligned}&E[(\check{X}^{n_k}(u)-\check{X}^{n_k}(0)-\theta^{n_k}u)f_1(\check{\mathcal{Z}}^{n_k}(s_1))\cdots f_m(\check{\mathcal{Z}}^{n_k}(s_m))]\\&\to E[(X(u)-X(0)-\theta u)f_1(\mathcal{Z}(s_1))\cdots f_m(\mathcal{Z}(s_m))]\end{aligned}$$

as $k\to\infty$. By the martingale property of $\check{X}^{n_k}$ stated in condition (III), in the above, the expectation involving $\check{X}^{n_k}$ has the same value for $u=s$ and $u=t$, and so the same is true for the expectation involving X. The latter property implies that $\{X(t)-X(0)-\theta t,\mathcal{F}_t,t\geq 0\}$ is a martingale. □

Acknowledgements. The author thanks J. Michael Harrison for pointing out the need for an invariance principle for SRBMs, and Jim Dai for bringing his work with Wanyang Dai [4, 8] to her attention.

References

[1] Bernard, A., and El Kharroubi, A. (1991). Régulation de processus dans le premier orthant de $\mathbb{R}^n$. *Stochastics and Stochastics Reports,* **34,** 149–167.

[2] Bramson, M. (1998). State space collapse with application to heavy traffic limits for multiclass queueing networks. To appear in *Queueing Systems: Theory and Applications.*

[3] Chen, H., and Zhang, H. (1996). Diffusion approximations for re-entrant lines with a first-buffer-first-served priority discipline. *Queueing Systems: Theory and Applications,* **23,** 177–195.

[4] Dai, J. G., and Dai, W. (1997). A heavy traffic limit theorem for a class of open queueing networks with finite buffers. Submitted to *Queueing Systems: Theory and Applications.*

[5] Dai, J. G., and Nguyen, V. (1994). On the convergence of multiclass queueing networks in heavy traffic. *Annals of Applied Probability*, **4**, 26–42.

[6] Dai, J. G., and Wang, Y. (1993). Nonexistence of Brownian models of certain multiclass queueing networks. *Queueing Systems: Theory and Applications*, **13**, 41–46.

[7] Dai, J. G. and Williams, R. J. (1995). Existence and uniqueness of semimartingale reflecting Brownian motions in convex polyhedrons. *Theory of Probability and Its Applications,* **40,** 1–40.

[8] Dai, W. (1996). *Brownian Approximations for Queueing Networks with Finite Buffers: Modeling, Heavy Traffic Analysis and Numerical Implementations.* Ph.D. Dissertation, School of Mathematics, Georgia Institute of Technology, Atlanta, GA.

[9] Dupuis, P., and Ishii, H. (1991). On the Lipschitz continuity of the solution mapping to the Skorokhod problem. *Stochastics and Stochastics Reports,* **35,** 31–62.

[10] Ethier, S. N., and Kurtz, T. G. (1986). *Markov Processes: Characterization and Convergence.* Wiley, New York.

[11] Harrison, J. M. (1996). The BIGSTEP approach to flow management in stochastic processing networks. In *Stochastic Networks: Theory and Applications,* F. P. Kelly, S. Zachary, and I. Ziedins (eds.), Oxford University Press, Oxford, 1996, pp. 57–90.

[12] Harrison, J. M., and Reiman, M. I. (1981). Reflected Brownian motion on an orthant. *Annals of Probability,* **9,** 302–308.

[13] Harrison, J. M., and Wein, L. M. (1989). Scheduling networks of queues: heavy traffic analysis of a simple open network. *Queueing Systems: Theory and Applications,* **5,** 265–280.

[14] Kelly, F. P., and Laws, C. N. (1993). Dynamic routing in open queueing networks: Brownian models, cut constraints and resource pooling. *Queueing Systems: Theory and Applications*, **13**, 47–86.

[15] Johnson, D. P. (1983). *Diffusion Approximations for Optimal Filtering of Jump Processes and for Queueing Networks*. Ph.D. dissertation, Department of Mathematics, University of Wisconsin, Madison, WI.

[16] Mandelbaum, A., and Van der Heyden, L. (1987, unpublished work). Complementarity and reflection.

[17] Peterson, W. P. (1991). Diffusion approximations for networks of queues with multiple customer types. *Mathematics of Operations Research*, **9**, 90–118.

[18] Reiman, M. I. (1984). Open queueing networks in heavy traffic. *Mathematics of Operations Research* **9**, 441–458.

[19] Reiman, M. I., and Williams, R. J. (1988–89). A boundary property of semimartingale reflecting Brownian motions. *Probability Theory and Related Fields,* **77,** 87–97, and **80,** 633.

[20] Skorokhod, A. V. (1956). Limit Theorems for Stochastic Processes. *Theory of Probability and Its Applications*, **1**, 261–290.

[21] Stroock, D. W., and Varadhan, S. R. S. (1979). *Multidimensional Diffusion Processes.* Springer-Verlag, New York, N.Y.

[22] Stroock, D. W., and Varadhan, S. R. S. (1971). Diffusion processes with boundary conditions. *Comm. Pure Appl. Math.* **24,** 147–225.

[23] Taylor, L. M., and Williams, R. J. (1993). Existence and uniqueness of semimartingale reflecting Brownian motions in an orthant. *Probability Theory and Related Fields*, **96**, 283–317.

[24] Whitt, W. (1993). Large fluctuations in a deterministic multiclass network of queues. *Management Science,* **39,** 1020–1028.

[25] Williams, R. J. (1996). On the approximation of queueing networks in heavy traffic. In *Stochastic Networks: Theory and Applications,* F. P. Kelly, S. Zachary, and I. Ziedins (eds.), Oxford University Press, Oxford, pp. 35–56.

[26] Williams, R. J. (1998). Diffusion approximations for open multiclass queueing networks: sufficient conditions involving state space collapse. To appear in *Queueing Systems: Theory and Applications.*