STEADY-STATE ANALYSIS OF REFLECTED BROWNIAN MOTIONS: CHARACTERIZATION, NUMERICAL METHODS AND QUEUEING APPLICATIONS

A DISSERTATION
SUBMITTED TO THE DEPARTMENT OF MATHEMATICS
AND THE COMMITTEE ON GRADUATE STUDIES
OF STANFORD UNIVERSITY
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

By
Jiangang Dai
July 1990

I certify that I have read this dissertation and that in my opinion it is fully adequate, in scope and in quality, as a dissertation for the degree of Doctor of Philosophy.

J. Michael Harrison
(Graduate School of Business)
(Principal Adviser)

I certify that I have read this dissertation and that in my opinion it is fully adequate, in scope and in quality, as a dissertation for the degree of Doctor of Philosophy.

Joseph Keller

I certify that I have read this dissertation and that in my opinion it is fully adequate, in scope and in quality, as a dissertation for the degree of Doctor of Philosophy.

David Siegmund
(Statistics)

Approved for the University Committee on Graduate Studies:

Dean of Graduate Studies

Abstract

This dissertation is concerned with multidimensional diffusion processes that arise as approximate models of queueing networks. To be specific, we consider two classes of *semimartingale reflected Brownian motions* (SRBM's), each with polyhedral state space. For one class the state space is a two-dimensional rectangle, and for the other class it is the general d-dimensional non-negative orthant $\mathbb{R}^d_+$.

SRBM in a rectangle has been identified as an approximate model of a two-station queueing network with finite storage space at each station. Until now, however, the foundational characteristics of the process have not been rigorously established. Building on previous work by Varadhan and Williams [53] and by Taylor and Williams [50], we show how SRBM in a rectangle can be constructed by means of localization. The process is shown to be unique in law, and therefore to be a Feller continuous strong Markov process. Taylor and Williams [50] have proved the analogous foundational result for SRBM in the non-negative orthant $\mathbb{R}^d_+$, which arises as an approximate model of a d-station open queueing network with infinite storage space at every station.

Motivated by the applications in queueing theory, our focus is on steady-state analysis of SRBM, which involves three tasks: (a) determining when a stationary distribution exists; (b) developing an analytical characterization of the stationary distribution; and (c) computing the stationary distribution from that characterization. With regard to (a), we give a sufficient condition for the existence of a stationary distribution in terms of Liapunov functions. With regard to (b), for a special class of SRBM's in an orthant, Harrison and Williams [26] showed that the stationary distribution must satisfy a weak form of an *adjoint* linear elliptic partial differential equation with *oblique* derivative boundary conditions, which they called the *basic adjoint relationship* (BAR). They further conjectured that (BAR) characterizes the stationary distribution. We give two proofs of their conjecture. For an SRBM in a rectangle, using Echeverria's Theorem [10], we give a direct proof of their

conjecture. For an SRBM in a general dimensional space, we first characterize an SRBM as a solution to a *constrained* martingale problem. Then, using Kurtz's recent theorem [32], we prove their conjecture for general SRBM's.

The most novel contribution of this dissertation relates to the computational task (c) above. To make practical use of SRBM's as approximate models of queueing networks, one needs practical methods for determining stationary distributions, and it is very unlikely that general analytical solutions will ever be found. We describe an approach to computation of stationary distributions that seems to be widely applicable. That approach gives rise to a family of algorithms, and we investigate one version of the algorithm. Under one mild assumption, we are able to provide a full proof of the algorithm's convergence. We compare the numerical results from our algorithm with known analytical results for SRBM, and also use the algorithm to estimate the performance measures of several illustrative open queueing networks. All the numerical comparisons show that our method gives reasonably accurate estimates and the convergence is relatively fast.

The algorithms that we have thus far implemented in computer code are quite limited as tools for analysis of queueing systems, but the underlying computational approach is widely applicable. Our ultimate goal is to implement this approach in a general routine for computing the stationary distribution of SRBM in an arbitrary polyhedral state space. The plan is to combine that routine with appropriate "front end" and "back end" program modules to form a software package, tentatively called QNET, for analysis of complex queueing networks. This package would be much more widely applicable than the commercial packages currently available for performance analysis of queueing networks, such as PANACEA [38] and QNA [54], but only the first tentative steps in its development have been taken thus far.

Acknowledgments

I am grateful to my advisor, Professor Michael Harrison, for introducing me to reflected Brownian motions and their connections with queueing theory, and for his advice and many contributions to this dissertation. I am also grateful for his role in making the presentation of this dissertation clearer.

I would like to thank Professor Ruth Williams in the Department of Mathematics, University of California at San Diego, for her help throughout my graduate studies at Stanford. She has been generous with her time and ideas, and has given me access to her important current work with Lisa Taylor. In addition, Professor Williams has read portions of this dissertation in preliminary draft form, always returning the manuscript promptly, often with comments and corrections as extensive as the text itself. Her influence is particularly apparent in Section 2.2.

I am grateful to Professor Thomas Kurtz in the Department of Mathematics and Statistics, University of Wisconsin at Madison, for several valuable conversations while he was visiting Stanford in the Spring of 1989. Before leaving Stanford, he wrote me a note that basically provided a proof for Theorem 3.5. I am also grateful to Gary Lieberman in the Department of Mathematics, Iowa State University, for many communications on elliptic equations with oblique boundary conditions. Although many of the results he produced for me are not quoted here, his generous help has been important.

I thank Stanford Professors Joseph Keller and David Siegmund for taking time to read my dissertation, and Professor T. L. Lai for his guidance and encouragement in the early stages of my career at Stanford.

My special thanks go to Professors Xuan-Da Hu and Hu-Mian Zhong at Nanjing University, and Professor Jia-An Yan at the Institute of Applied Mathematics, Academica Sinica, for introducing me to the stochastic calculus, which plays an essential role in this dissertation.

I would like to thank my fellow students for proofreading parts of this dissertation, and their friendship. Among them are Ken Dutch, David Ho, Christoph Loch, Vien Nguyen, Mike Pich, Richard Tobias and Heping Zhang.

The final group of people whom I wish to thank for their support is my family: my father, my sister, my late mother and my late brother. They have been a constant source of emotional support throughout the years of my education. This dissertation is dedicated to my beloved mother and brother.

Although many people have helped me along the way, none is more responsible for my completing this dissertation than my wife, Li-Qin. Giving up a graduate study opportunity in biochemistry, she came to Stanford and provided love, support, and understanding through many disappointments which accompany accomplishments. Her support was particularly important during the difficult period when I lost my beloved brother at the end of my first year at Stanford.

Thanks go to Stanford Mathematics Department and Graduate School of Business for their financial support over the past four years.

Contents

List of Tables

List of Figures

Frequently Used Notation

I. Topological Notation. Let X be a subset of the d-dimensional Euclidean Space $\mathbb{R}^d$.

1. $\overline{X}$ is the closure of $X \subset \mathbb{R}^d$.
2. ∂X is the boundary of $X \subset \mathbb{R}^d$.
3. $\mathcal{B}_X$ is the Borel field of subsets of X.
4. $C_b(X)$ is the set of bounded continuous functions $f : X \to \mathbb{R}$.
5. $B(X)$ is the set of bounded $\mathcal{B}_X$-measurable $f : X \to \mathbb{R}$.
6. $\mathcal{B}(X)$ is the set of $\mathcal{B}_X$-measurable $f : X \to \mathbb{R}$.
7. $\mathcal{P}(X)$ is the set of probability measures on $(X, \mathcal{B}_X)$.
8. $||f|| \equiv \sup_{x \in X} |f(x)|$ for $f \in B(X)$.
9. $C_K(X)$ is the set of $f \in C_b(X)$ having compact support.
10. $C_b^m(X)$ is the set of $f : A \to \mathbb{R}$ possessing bounded derivatives of order up to and including m for some *open* set A containing X.
11. $C_K^m(X) \equiv C_b^m(X) \cap C_K(X)$.
12. $C_0(X)$ is the set of $f \in C_b(X)$ with $\lim_{x \to \infty} f(x) = 0$, see page 63.

II. Special Notation for Euclidean Space

1. $x \in \mathbb{R}^d$ is understood as a column vector.
2. $|x| \equiv (\sum_{i=1}^d x_i^2)^{1/2}$ for $x \in \mathbb{R}^d$.
3. $B(x, r) \equiv \{y \in \mathbb{R}^d : |x - y| < r\}$.
4. $\mathbb{R}_+^d \equiv \{x = (x_1, \ldots, x_d)' \in \mathbb{R}^d, x_i \geq 0, i = 1, \ldots, d\}$.
5. $\langle x, y \rangle \equiv \sum_{i=1}^d x_i y_i$ for $x, y \in \mathbb{R}^d$; $x \cdot y \equiv \langle x, y \rangle$ for $x, y \in \mathbb{R}^d$.

III. State Space and Path Space Notation for Reflected Brownian Motion

1. S is the state space, either a two-dimensional rectangle or the orthant $\mathbb{R}^d_+$.
2. $\mathcal{O}$ is the interior of S.
3. F_i is the ith face of boundary ∂S.
4. v_i is a d-dimensional vector on face F_i, direction of reflection on F_i.
5. Γ is a $d \times d$ positive definite matrix, the covariance matrix.
6. μ is a d-dimensional vector, the drift vector.
7. R is a matrix whose columns are the vectors v_i, the reflection matrix.
8. dx is Lebesgue measure on S.
9. $d\sigma_i$ is Lebesgue measure on boundary face F_i.
10. ∇ is the gradient operator.
11. $D_i f \equiv v_i \cdot \nabla f$.
12. $Gf = \frac{1}{2}\sum_{i,j=1}^d \Gamma_{ij} \frac{\partial^2 f}{\partial x_i \partial x_j} + \sum_{i=1}^d \mu_i \frac{\partial f}{\partial x_i}$.
13. $\mathcal{A}f$, see page 34 and page 73.
14. $d\lambda$, see page 34 and page 73.
15. $d\eta$, see page 74.
16. $C_S \equiv \{\omega : [0,\infty) \to S,\ \omega \text{ is continuous }\}$.
17. ω is a generic element of C_S.
18. $Z = \{Z(t,\cdot), t \geq 0\}$ is the canonical process on C_S, where $Z(t,\omega) \equiv \omega(t)$, $\omega \in C_S$.
19. $\mathcal{M}_t \equiv \sigma\{Z(s), 0 \leq s \leq t\}$.
20. $\mathcal{M} \equiv \sigma\{Z(s), 0 \leq s < \infty\}$.

IV. Miscellaneous Notation and Terminology

1. $\Rightarrow$ weak convergence, see page 29.
2. $\approx$ equivalence of measures, see page 30.
3. $\doteq$ approximate equality.
4. A' is the transpose of a matrix A.
5. completely-$\mathcal{S}$ matrix, see Definition 3.4.

6. Minkowski matrix, see Definition 3.3.
7. $\inf \emptyset = +\infty$ and $\sup \emptyset = 0$, where $\emptyset$ is the empty set.
8. $\sigma(\mathcal{C})$, where $\mathcal{C}$ is a set of functions on X into a measurable space, is the smallest σ-algebra over X with respect to which every element of $\mathcal{C}$ is measurable.
9. $r \cdot q$ not an inner product, see page 73.
10. E_i: Erlang distribution.
11. M: exponential distribution.
12. D: deterministic distribution.
13. H_k: hyperexponential distribution.

V. Abbreviations

1. RBM—reflected Brownian motion, see page 11.
2. SRBM— semimartingale reflected Brownian motion, see page 9 and page 53.
3. (BAR)—basic adjoint relationship, see (2.35) and (3.30).
4. SCV— squared coefficient of variation (variance over squared mean).
5. QNA—queueing network analyzer, see page v.
6. QNET, see page v.
7. SIM—simulation estimate.

Chapter 1

Introduction

1.1 Motivation

This dissertation is concerned with a class of multidimensional diffusion processes, variously known as reflected Brownian motions, regulated Brownian motions, or just RBM's, that arise as approximate models of queueing networks. More specifically, we consider two special classes of *semimartingale* RBM's (SRBM's), which have by far the widest applications in queueing network theory. The state space of the first class is a two-dimensional rectangle, and the state space of the second class is an arbitrary d-dimensional orthant $\mathbb{R}^d_+$. SRBM in a rectangle arises as an approximate model of a queueing network of two stations with finite buffer size at each station [9]. The analysis of this two-dimensional SRBM can be generalized to deal with SRBM in a higher dimensional "box". SRBM in an orthant can be used as an approximate model for open queueing networks [39, 36, 26, 24]. To make the presentation clearer, we consider the two classes separately, and throughout the remainder of this introduction, only SRBM's in an orthant will be discussed. Readers are referred to Section 2.5.3 for the motivation of studying SRBM in a rectangle.

Given a $d \times d$ positive definite matrix Γ, a d-dimensional vector μ and a $d \times d$ matrix R (whose i-th column is denoted by v_i) with ones on the diagonal, a d-dimensional continuous stochastic process Z is said to be an SRBM in the orthant $\mathbb{R}^d_+$ associated with data (Γ, μ, R) if

(1.1) $Z(t) = X(t) + RL(t) = X(t) + v_1 L_1(t) + \ldots + v_d L_d(t)$, $t \geq 0$;

(1.2) $Z_i(t) \geq 0$, $t \geq 0$, $i = 1, \ldots, d$;

(1.3) $X = \{X(t)\}$ is a d-dimensional Brownian motion with covariance matrix Γ and drift vector μ;

(1.4) For $i = 1, \ldots, d$, $L_i(0) = 0$, L_i is non-decreasing and $L_i(\cdot)$ increases only at times t such that $Z_i(t) = 0$.

This definition suggests that the SRBM Z behaves like an ordinary Brownian motion with covariance matrix Γ and drift vector μ in the interior of the orthant. When Z hits the boundary $\{x_i = 0\}$, the process (local time) $L_i(\cdot)$ increases, causing an overall pushing in the direction v_i. The magnitude of the pushing is the minimal amount required to keep Z inside the orthant.

The motivation for our study of SRBM in an orthant comes from the theory of open queueing networks, that is, networks of interacting processors or service stations where customers arrive from outside, visit one or more stations, perhaps repeatedly, in an order that may vary from one customer to the next, and then depart. (In contrast, a closed queueing network is one where a fixed customer population circulates perpetually through the stations of the network, with no new arrivals and no departures.) It was shown by Reiman [39] that the d-dimensional queue length process associated with a certain type of open d-station network, if properly normalized, converges under "heavy traffic" conditions to a corresponding SRBM with state space $\mathbb{R}^d_+$. Peterson [36] proved a similar "heavy traffic limit theorem" for open queueing networks with multiple customer types and deterministic, feedforward customer routing; Peterson's assumptions concerning the statistical distribution of customer routes are in some ways more general and in some ways more restrictive than Reiman's. The upshot of this work on limit theorems is to show that SRBM's with state space $\mathbb{R}^d_+$ may serve as good approximations, at least under heavy traffic conditions, for the queue length processes, the workload processes, and the waiting time processes associated with various types of open d-station networks. Recently Harrison and Nguyen [24] have defined a very general class of open queueing networks and articulated a systematic procedure for approximating the associated stochastic processes by SRBM's. This general approximation scheme subsumes those suggested by the limit theorems of both Reiman and Peterson, but it has not yet been buttressed by a rigorous and equally general heavy traffic limit theory.

1.2 A Tandem Queue

To illustrate the role of SRBM in queueing network theory, let us consider a network of two stations in tandem as pictured in Figure 1.1. After describing this queueing model in mathematical terms, we will explain how one can use a two-dimensional SRBM to approximate the workload process of the tandem queue under *heavy traffic* conditions. This is basically a recapitulation of Reiman's [39] heavy traffic limit theorem, which has also been discussed at some length in the survey papers of Lemoine [33], Flores [12], Coffman-Reiman [7], and Glynn [18]. We follow the treatment in [24]. It is hoped that this description of the heavy traffic approximation will motivate the study of SRBM's for readers who are not familiar with diffusion approximations.

The network pictured in Figure 1.1 consists of two single–server stations arranged in series, each with a first–in–first–out discipline. Arriving customers go to station 1 first, after completing service there they go to station 2, and after completing service at station 2 they exit the system. The inter-arrival times of the customers to station 1 are assumed to be independent, identically distributed (i.i.d.) positive random variables with mean one and *squared coefficient of variation* (defined to be variance over squared mean) C_a^2. Similarly, the service times at station i are assumed to be i.i.d. random variables with mean ρ_i and squared coefficient of variation $C_{s_i}^2$, $i = 1, 2$. This network is a *generalized Jackson network*; in a classical Jackson network, both the inter-arrival times and service times are assumed to be exponentially distributed, implying $C_a^2 = C_{s_1}^2 = C_{s_2}^2 = 1$. The steady-state performance measures we focus on are

(1.5) $w_i \equiv$ the long-run average waiting time (excluding service time) that customers experience in queue i, $i = 1, 2$.

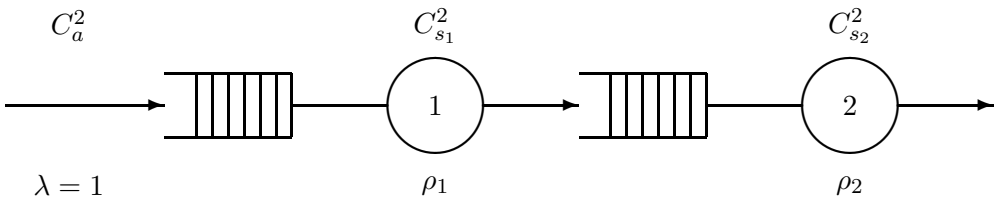

Figure 1.1: Two Queues in Tandem

When $\rho_i < 1$ $(i = 1, 2)$, it is known that the network is stable (or ergodic), that is, $w_i < \infty$.

Despite its apparent simplicity, the tandem queue described above is not amenable to exact mathematical analysis, except for the Jackson network case. But as an alternative to simulation one may proceed with the following approximate analysis. Let $W_i(t)$ be the *current workload* at time t for server i, that is, the sum of the impending service times for customers waiting at station i at time t, plus the remaining service time for the customer currently in service there (if any). One can also think of $W_i(t)$ as a *virtual waiting time:* if a customer arrived at station i at time t, this customer would have to wait $W_i(t)$ units of time before gaining access to server i. The tandem queue is said to be *in heavy traffic* if ρ_1 and ρ_2 are both close to one, and the heavy traffic limit theory referred to earlier suggests that under such conditions the workload process $W(t)$ can be well approximated by an SRBM with state space $\mathbb{R}^2_+$ and certain data (Γ, μ, R). To be more specific, Harrison and Nguyen [24] propose that $W(t)$ be approximated by an SRBM $Z(t)$ with data

$$\Gamma = \begin{pmatrix} \rho_1^2(C_a^2 + C_{s_1}^2) & -\rho_1\rho_2 C_{s_1}^2 \\ -\rho_1\rho_2 C_{s_1}^2 & \rho_2^2(C_{s_1}^2 + C_{s_2}^2) \end{pmatrix}, \quad \mu = \begin{pmatrix} \rho_1 - 1 \\ \rho_2/\rho_1 - 1 \end{pmatrix}, \quad R = \begin{pmatrix} 1 & 0 \\ -\rho_2/\rho_1 & 1, \end{pmatrix}.$$

The directions of reflection for the SRBM Z are pictured in Figure 1.2 below; recall that in general v_i denotes the ith column of R, which is the direction of reflection associated with the boundary surface $\{x_i = 0\}$.

If the steady-state mean $m = (m_1, m_2)'$ of the SRBM Z can be calculated, then m_i can be used to estimate the long run average *virtual waiting time*, i.e.,

$$m_i \doteq \lim_{t\to\infty} \frac{\int_0^t E[W_i(s)]\, ds}{t}, \quad i = 1, 2.$$

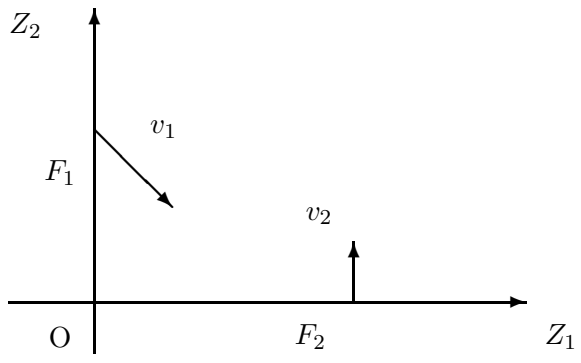

Figure 1.2: State Space and Directions of Reflection for the Approximating SRBM

It is suggested in [24] that this long run average virtual waiting time be used to estimate the long run average waiting time w_i, i.e., $w_i \doteq m_i$ $(i = 1, 2)$. Notice that the SRBM uses only the first two moments of the primitive queueing network data. This is typical of "Brownian approximations". In this dissertation we will focus on the analysis of an SRBM instead of the original queueing network model.

1.3 Overview

There is now a substantial literature on Brownian models of queueing networks, and virtually all papers in that literature are devoted to one or more of the following tasks.

(a) Identify the Brownian analogs for various types of conventional queueing models, explaining how the data of the approximating SRBM are determined from the structure and the parameters of the conventional model; prove limit theorems that justify the approximation of conventional models by their Brownian analogs under "heavy traffic" conditions.

(b) Show that the SRBM exists and is uniquely determined by an appropriate set of axiomatic properties.

(c) Determine the analytical problems that must be solved in order to answer probabilistic questions associated with the SRBM. These are invariably partial differential equation problems (PDE problems) with oblique derivative boundary conditions. A question of central importance, given the queueing applications that motivate the theory, is which PDE problem one must solve in order to determine the stationary distribution of an SRBM.

(d) Solve the PDE problems of interest, either analytically or numerically.

Most research to date has been aimed at questions (a) through (c) above. Topic (a) has been discussed in Section 1.2 above. With regard to (b), Harrison and Reiman [25] proved, using a unique path-to-path mapping, the existence and uniqueness of SRBM in an orthant when the reflection matrix R is a Minkowski matrix (see Definition 3.3). This class of SRBM's corresponds to open queueing networks with *homogeneous customer populations*, which means that customers occupying any given node or station of the network are essentially indistinguishable from one another. Recently Taylor and Williams [50] proved the existence

and uniqueness of SRBM in an orthant when the reflection matrix R is completely-$\mathcal{S}$ (see Definition 3.4). In an earlier paper, Reiman and Williams [41] showed that the reflection matrix R being completely-S is necessary for the existence of an SRBM in an orthant. Thus, when the state space is an orthant, category (b) is completely resolved. For the two-dimensional case, Varadhan and Williams [53] considered driftless RBM in a general wedge, and existence and uniqueness were resolved there. In that setting, they actually considered the more general class of RBM's which may not have a semimartingale representation. Taylor and Williams [51] showed that under a condition on the directions of reflection, which corresponds to the completely-S condition in the orthant case, the RBM constructed in [53] actually is a *semimartingale* RBM. For a network of two stations with finite buffer size, the corresponding SRBM lives in a rectangle, see [9]. There has been no literature on the explicit construction of such an SRBM. In Section 2.2 we show, using a detailed localization argument, that there exists an unique SRBM in the rectangle when the completely-S condition is satisfied at each corner locally (see condition (2.4)).

With regard to (c), for a driftless RBM in two dimensions the work of Harrison, Landau and Shepp [23] gives an analytical expression for the stationary distribution. For the two-dimensional case with drift, Foddy [13] found analytical expressions for the stationary distributions for certain special domains, drifts, and directions of reflection, using Riemann-Hilbert techniques. For a special class of SRBM in an orthant, Harrison and Williams [26] gave a criterion for the existence of the stationary distribution. Furthermore, they showed that the stationary distribution together with the corresponding boundary measures must satisfy a *basic adjoint relationship* (BAR). In that paper, the authors conjectured that (BAR) characterizes the stationary distribution as well. In Chapter 3 we prove the conjecture to be true for the most general class of SRBM's in an orthant (Theorem 3.6). We also establish a sufficient condition for the existence of stationary distribution in terms of Liapunov functions for a general SRBM.

With regard to research category (d), the availability of a package for evaluation of Schwartz-Christoffel transformations based on the result in [23] makes the evaluation of associated performance measures for a driftless RBM in two dimensions numerically feasible, cf. [52]. In dimensions three and more, RBM's having stationary distributions of exponential form were identified in [27, 58] and these results were applied in [26, 28] to SRBM's arising as approximations to open and closed queueing networks with homogeneous customer populations. However, until now there has been no general method for solving the

PDE problems alluded to in (d).

If Brownian system models are to have an impact in the world of practical performance analysis, task (d) above is obviously crucial. In particular, practical methods are needed for determining stationary distributions, and it is very unlikely that general analytical solutions will ever be found. Thus we are led to the problem of computing the stationary distribution of RBM in an orthant, or at least computing summary statistics of the stationary distribution. As we will explain later, the stationary distribution is the solution of a certain highly structured partial differential equation problem (an adjoint PDE problem expressed in weak form). In this dissertation, we describe an approach to computation of stationary distributions that seems to be widely applicable. The method will be developed and tested for two–dimensional SRBM's with a rectangular state space in Section 2.4.2 through Section 2.5.3, and for higher dimensional SRBM's in an orthant in Chapter 4. We should point out that the proof of convergence would be complete if we could prove that any solution to (BAR) does not change sign (see Conjecture 2.1 and Conjecture 4.1). As readers will see, the method we use actually gives rise to a family of converging algorithms. One particular implementation of our algorithm is tested against known analytical results for SRBM's as well as simulation results for queueing network models. The testing results show that both the accuracy and speed of convergence are impressive for small networks. We must admit that, currently, we do not have a general method to choose one "best" algorithm from this family. In Appendix A, we will describe in detail how to implement the version of the algorithm we use in the orthant case. As a tool for analysis of queueing systems, the computer program described in this dissertation is obviously limited in scope, but our ultimate goal is to implement the same basic computational approach in a general routine that can compete with software packages like PANACEA [38] and QNA [54] in the analysis of large, complicated networks.

1.4 Notation and Terminology

Here and later the symbol "$\equiv$" means "equals by definition". We assume some basic notation and terminology in probability as in Billingsley [4]. We denote the characteristic function of a set A by 1_A, i.e., $1_A(x) = 1$ if $x \in A$ and $1_A(x) = 0$ otherwise. For a random variable X and an event set A, we use $E[X; A]$ to denote $E[X1_A]$. Given a filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}, P)$, a real-valued process $X = \{X(t), t \geq 0\}$ defined on this space is said to be

adapted if for each $t \geq 0$, $X(t)$ is $\mathcal{F}_t$-measurable. The process X is said to be an $\{\mathcal{F}_t\}$-(sub)martingale under P if X is adapted, $E^P[|X(t)|] < \infty$ for each $t \geq 0$ and for each pair $0 \leq s < t$ and each $A \in \mathcal{F}_s$, $E^P[X(s); A](\leq) = E^P[X(t); A]$, where E^P is the expectation with respect to the probability measure P.

The following are used extensively in the sequel. $d \geq 1$ is an integer. S always denotes a state space, either a two-dimensional rectangle or a d-dimensional orthant $\mathbb{R}^d_+ \equiv \{x = (x_1, x_2, \ldots, x_d)' \in \mathbb{R}^d : x_i \geq 0, i = 1, 2, \ldots, d\}$, where the prime is the transpose operator. If no dimensionality is explicitly specified, a vector and a process are considered to be d-dimensional. Vectors are treated as column vectors. Inequalities involving matrices or vectors are interpreted componentwise. The set of continuous functions $\omega : [0, \infty) \to S$ is denoted by C_S. The canonical process on C_S is $Z = \{Z(t, \cdot), t \geq 0\}$ defined by

$$Z(t, \omega) = \omega(t), \quad \text{for } \omega \in C_S.$$

The symbol ω is often suppressed in Z. The natural filtration associated with C_S is $\{\mathcal{M}_t\}$, where $\mathcal{M}_t \equiv \sigma\{Z(s, \cdot) : 0 \leq s \leq t\}$, $t \geq 0$. For $t \geq 0$, $\mathcal{M}_t$ can also be characterized as the smallest σ-algebra of subsets of C_S which makes $Z(s)$ measurable for each $0 \leq s \leq t$. The natural σ-algebra associated with C_S is $\mathcal{M} \equiv \sigma\{Z(s, \cdot) : 0 \leq s < \infty\} = \bigvee_{t=0}^{\infty} \mathcal{M}_t$ ($\bigvee_{t=0}^{\infty} \mathcal{M}_t$ is defined to be the smallest σ-algebra containing $\mathcal{M}_t$ for each $t \geq 0$). For commonly used notation, readers are referred to "Frequently Used Notation" on page xiv. Other notation and terminology will be introduced as we proceed.

Chapter 2

SRBM in a Rectangle

2.1 Definition

Let S be a closed two–dimensional rectangle and $\mathcal{O}$ be the interior of the rectangle. For $i = 1, 2, 3, 4$ let F_i be the i^{th} boundary face of S and let v_i be an inward–pointing vector on F_i with unit normal component (see Figure 2.1). Also, let us define the 2×4 matrix $R \equiv (v_1, v_2, v_3, v_4)$. Remember $Z = \{Z(t, \omega), t \geq 0\}$ is the canonical process on C_S. Throughout this chapter, we use Γ to denote a 2×2 positive definite matrix and μ to denote a two-dimensional vector.

Definition 2.1 *Z together with a family of probability measures $\{P_x, x \in S\}$ on the filtered space $(C_S, \mathcal{M}, \{\mathcal{M}_t\})$ is said to be a semimartingale reflected Brownian motion (abbreviated as SRBM) associated with data (S, Γ, μ, R) if for each $x \in S$ we have*

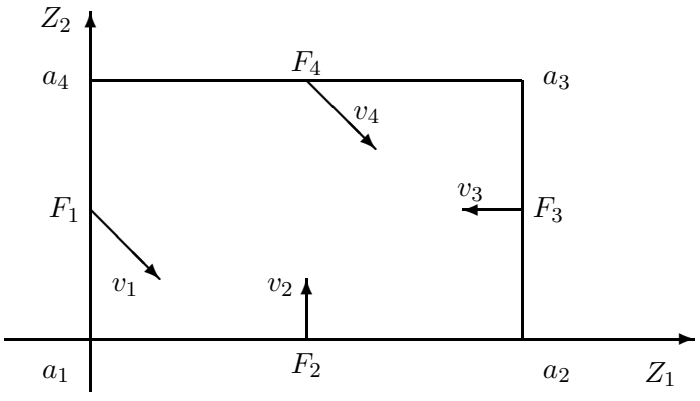

Figure 2.1: State Space S and Directions of Reflection of an SRBM in a Rectangle

(2.1) $Z(t) = X(t) + RL(t) = X(t) + \sum_{i=1}^{4} v_i L_i(t)$ $\forall t \geq 0$, P_x-*a.s., where*

(2.2) $X(0) = x$ P_x-*a.s. and* X *is a 2-dimensional Brownian motion with covariance matrix* Γ *and drift vector* μ *such that* $\{X(t) - \mu t, \mathcal{M}_t, t \geq 0\}$ *is a martingale under* P_x*, and*

(2.3) L *is a continuous* $\{\mathcal{M}_t\}$*-adapted four-dimensional process such that* $L(0) = 0$*,* L *is non-decreasing, and* P_x*-almost surely* L_i *can increase only at times* t *such that* $Z(t) \in F_i, i = 1, 2, 3, 4$.

An SRBM Z as defined above behaves like a two-dimensional Brownian motion with drift vector μ and covariance matrix Γ in the interior $\mathcal{O}$ of its state space. When the boundary face F_i is hit, the process L_i (sometimes called the local time of Z on F_i) increases, causing an instantaneous displacement of Z in the direction given by v_i; the magnitude of the displacement is the minimal amount required to keep Z always inside S. Therefore, we call Γ, μ and R the covariance matrix, the drift vector and the reflection matrix of Z, respectively.

SRBM in a rectangle can be used as an approximate model of a two station queueing network with finite buffer sizes at each station. Readers are referred to Section 2.5.3 and [9] for more details.

Throughout this dissertation, when the state space S is a rectangle, we always assume the given directions of reflection satisfy the following condition:

(2.4) there are positive constants a_i and b_i such that $a_i v_i + b_i v_{i+1}$ points into the interior of S from the vertex where F_i and F_{i+1} meet ($i = 1, 2, 3, 4$), where $v_5 \equiv v_1$ and $F_5 \equiv F_1$.

Because a Brownian motion can reach every region in the plane, it can be proved as in Reiman and Williams [41] that (2.4) is a necessary condition for the existence of an SRBM. In the following section, we will prove that there is a unique family $\{P_x, x \in S\}$ on $(C_S, \mathcal{M})$ such that Z together with $\{P_x, x \in S\}$ is an SRBM when the columns of the reflection matrix R satisfy (2.4).

Definition 2.2 *An SRBM* Z *is said to be unique (in law) if the corresponding family of probability measures* $\{P_x, x \in S\}$ *is unique.*

2.2 Existence and Uniqueness of an SRBM

Theorem 2.1 *Let there be given a rectangle S, a non-degenerate covariance matrix Γ, a drift vector μ and a reflection matrix R whose columns satisfy (2.4). Then there is a unique family of probability measures $\{P_x, x \in S\}$ on $(C_S, \mathcal{M}, \{\mathcal{M}_t\})$ such that the canonical process Z together with $\{P_x, x \in S\}$ is an SRBM associated with the data (S, Γ, μ, R). Furthermore, the family $\{P_x, x \in S\}$ is Feller continuous, i.e., $x \to E_x[f(Z(t))]$ is a continuous for all $f \in C_b(S)$ and $t \geq 0$, and Z together with $\{P_x, x \in S\}$ is a strong Markov process. Moreover, $\sup_{x \in S} E_x[L_i(t)] < \infty$ for each $t \geq 0$ and $i = 1, 2, 3, 4$.*

Remark. In this chapter, E_x is the expectation operator associated with the unique probability measure P_x. We leave the proof of this theorem to the end of Section 2.2.3. To this end, we first consider a class of reflected Brownian motions (RBM's) as solutions to certain *submartingale* problems as considered in Varadhan and Williams [53]. The main difference between an RBM and an SRBM is that an RBM may not have a semimartingale representation as in (2.1). When $\mu = 0$, the authors in [53] showed the existence and uniqueness of an RBM in a wedge for $\Gamma = I$, which immediately implies the existence and uniqueness of an RBM in a quadrant with general non-degenerate covariance matrix Γ. In Section 2.2.1, taking four RBM's in four appropriate quadrants, we will carry out a detailed patching argument to construct an RBM in the rectangle S. In Section 2.2.2, we show when the reflection matrix R satisfies (2.4) that such an RBM actually has the semimartingale representation (2.1). For $\mu \neq 0$, the existence of an SRBM follows from that for $\mu = 0$ and Girsanov's Theorem. Finally, in Section 2.2.3, we prove the uniqueness and Feller continuity of an SRBM, and hence prove that Z together with $\{P_x, x \in S\}$ is a strong Markov process.

2.2.1 Construction of an RBM

Throughout this section we assume $\mu = 0$.

Theorem 2.2 *Let there be given a covariance matrix Γ, drift vector $\mu = 0$ and reflection matrix R whose columns satisfy (2.4). Then there is a family of probability measures $\{P_x, x \in S\}$ on $(C_S, \mathcal{M}, \{\mathcal{M}_t\})$ such that for each $x \in S$,*

(2.5) $P_x\{Z(0) = x\} = 1$,

(2.6) $P_x\left\{\int_0^\infty 1_{\{Z(s) \in \partial S\}}\, ds = 0\right\} = 1$.

(2.7) *For each* $f \in C^2(S)$ *with* $D_i f \geq 0$ *on* F_i $(i = 1, 2, 3, 4)$,

$$\left\{ f(Z(t)) - \int_0^t Gf(Z(s))\, ds, \mathcal{M}_t, t \geq 0 \right\}$$

is a P_x*-submartingale, where*

$$Gf = \frac{1}{2} \sum_{i,j=1}^{2} \Gamma_{ij} \frac{\partial^2 f}{\partial x_i \partial x_j} + \sum_{i=1}^{2} \mu_i \frac{\partial f}{\partial x_i} \tag{2.8}$$

$$D_i f = v_i \cdot \nabla f, \quad i = 1, 2, 3, 4. \tag{2.9}$$

In order to carry out the construction of $\{P_x\}$ we need more notation and some preliminary results. Let

$$\Omega_S \equiv \{\omega : [0, \infty) \to \mathbb{R}^2, \omega(0) \in S,\ \omega \text{ is continuous}\}.$$

The *canonical process* on Ω_S is $w = \{w(t, \omega), t \geq 0\}$ defined by

$$w(t, \omega) \equiv \omega(t), \quad \text{for } \omega \in \Omega_S.$$

The natural filtration on Ω_S is $\{\mathcal{F}_t \equiv \sigma\{w(s) : 0 \leq s \leq t\}, t \geq 0\}$ and the natural σ-field is $\mathcal{F} \equiv \sigma\{w(s) : 0 \leq s < \infty\}$. We intentionally use w to denote our canonical process instead of Z used before, because they are the canonical processes on two different spaces. Obviously, we have

$$w|_{C_S} = Z,$$

and $\mathcal{M}_t = \mathcal{F}_t \cap C_S$ and $\mathcal{M} = \mathcal{F} \cap C_S$.

Without loss of generality, by rescaling of coordinates if necessary, we assume S to be the *unit* square, with sides parallel to the coordinates axes and lower left corner at the origin of the coordinate system. Let a_i denote the i-th corner of the square, counting counterclockwise starting from the origin $(i = 1, 2, 3, 4)$, see Figure 2.1. For $i = 1, 2, 3, 4$, define

$$A_i \equiv S \cap B(a_i, 0.9) \quad \text{and} \quad B_i \equiv S \cap B(a_i, 0.8),$$

where $B(x, r) \equiv \{y \in \mathbb{R}^2 : |x - y| < r\}$. Note that the B_i's, hence A_i's cover S. Let $S_i \supset S$ be the quadrant with vertex a_i defined in an obvious way. Assume the drift vector $\mu = 0$. It follows from Varadhan and Williams [53] and Williams [57] or Taylor and Williams [51] that there exists a family of probability measures $\{P_x^i, x \in S_i\}$ on $(\Omega_S, \mathcal{F}, \{\mathcal{F}_t\})$ which, together with the canonical process $\{w(t, \cdot), t \geq 0\}$ is a $(S_i, \Gamma, \mu, (v_i, v_{i+1}))$-SRBM on S_i $(i = 1, 2, 3, 4)$. That is, for each $i \in \{1, 2, 3, 4\}$ and $x \in S_i$, one has

(a) $P_x^i(C_{S_i}) = 1$, $P_x^i(w(0) = x) = 1$,

(b) P_x-almost surely, $w(t) = X(t) + v_i y_1^i(t) + v_{i+1} y_2^i(t)$, $\forall t \geq 0$, where

(c) X is a Brownian motion, and $\{X(t)\}$ is an $\{\mathcal{F}_t\}$-martingale under P_x^i,

(d) For $j = 1, 2$, $y_j^i(0) = 0$, and y_j^i is nondecreasing and P_x^i-almost surely, $y_j^i(\cdot)$ can increase only at times t such that $w_j(t) \in F'_{i-1+j}$, where F'_{i-1+j} is the half line passing through F_{i-1+j} with a_i as one end point.

In particular, it follows from Lemma 7.2 of [41], for each $i = 1, 2, 3, 4$, and $x \in S_i$,

(e) $P_x^i \{\int_0^\infty 1_{\partial S_i}(w(s))\, ds = 0\} = 1$,

(f) For each $f \in C^2(S)$ satisfying $D_k f \geq 0$ on F'_k $k = i - 1 + j$, $j = 1, 2$

$$\left\{ f(w(t)) - \int_0^t Gf(w(s))\, ds, \mathcal{F}_t, t \geq 0 \right\},$$

is a P_x^i-submartingale.

These four families of $\{P_x^i, x \in S\}$ $(i = 1, 2, 3, 4)$ are the building blocks in our construction of $\{P_x, x \in S\}$.

For $\omega \in \Omega_S$, define an increasing sequence of times $\tau_n(\omega)$, $n = 0, 1, 2, \ldots$ and an associated sequence of neighborhood indices $k_n(\omega)$, $n = 1, 2, \ldots$ by induction as follows. (From now on, the symbol ω will be suppressed) Set $\tau_0 = 0$, let k_0 be the smallest index such that $w(0) \in B_{k_0}$, and define $\tau_1 = \inf\{t \geq 0 : w(t) \in S \backslash A_{k_0}\}$. Then, assuming τ_0, k_0, $\tau_1, \ldots, k_{n-1}, \tau_n$ have been defined for some $n \geq 1$, if $\tau_n < \infty$, let k_n be the smallest index such that $w(\tau_n) \in B_{k_n}$ and define

$$\tau_{n+1} = \inf\{t \geq \tau_n : w(t) \in S \backslash A_{k_n}\},$$

or if $\tau_n = \infty$, let $k_n = k_{n-1}$ and $\tau_{n+1} = \tau_n$. Then $\tau_n \to \infty$ as $n \to \infty$. We have the following lemma.

Lemma 2.1 *For $x \in B_i \backslash \cup_{j<i} B_j$, $P_x^i(\tau_1 < \infty) = 1$, $(i = 1, 2, 3, 4)$.*

Proof. It is enough to prove the case when $i = 1$. The other cases can be proved similarly. Since the matrix (v_1, v_2) is completely-$\mathcal{S}$ (cf. [41]), there is $u \in \mathbb{R}_+^2$ such that $u > 0$, $u \cdot v_1 > 0$ and $u \cdot v_2 > 0$. Hence, for $x \in B_1$, using representation (b) above,

$$|u| \sup_{0 \leq s < \infty} |w(s)| \geq \sup_{0 \leq s < \infty} u \cdot w(s) \geq \sup_{0 \leq s < \infty} u \cdot X(s), \quad P_x\text{-a.s.}$$

Since $u \cdot X(t)$ is a Brownian motion with variance $u'\Gamma u$ and drift 0, we have

$$\sup_{0 \le s < \infty} u \cdot X(s) = +\infty, \quad P_x^1\text{-a.s.}$$

This lemma follows since P_x^1-a.s.,

$$\tau_1 = \inf\{t \ge 0 : w(t) \in S \backslash A_1\}.$$

□

The following relies heavily on a compactness result of Bernard and El Kharroubi [2, Lemma 1].

Lemma 2.2 *For each $t > 0$ there exists $\epsilon(t) > 0$ such that*

$$\max_{1 \le i \le 4} \sup_{x \in B_i} P_x^i(\tau_1 \le t) \le 1 - \epsilon(t).$$

Proof. Fix $t > 0$. It is enough to prove that there is $\epsilon(t) > 0$ such that $\sup_{x \in B_1} P_x^1(\tau_1 \le t) \le 1 - \epsilon(t)$. Since the matrix (v_1, v_2) is completely-$\mathcal{S}$, there is a constant K (cf. [2, Lemma 1]) such that

$$\sup_{0 \le s \le t} |w(s) - w(0)| \le K \sup_{0 \le s \le t} |X(s) - X(0)|, \quad P_x^1\text{-a.s.}$$

Hence, for $x \in B_1$,

$$\begin{aligned} P_x^1(\tau_1 \le t) &\le P_x^1\left(\sup_{0 \le s \le t} |w(s)| \ge 0.9\right) \\ &\le P_x^1\left(\sup_{0 \le s \le t} |w(s) - x| \ge 0.1\right) \\ &\le P_x^1\left(\sup_{0 \le s \le t} |X(s) - x| \ge 0.1/K\right) \\ &= 1 - \epsilon(t), \quad \text{for some } \epsilon(t) > 0, \end{aligned}$$

since $\{X(t) - x, t \ge 0\}$ is a Brownian motion staring from the origin. □

It is time for us to introduce *shift operators*. They are used only in this section. For each $t \ge 0$, the shift operator $\theta_t : \Omega_S \to \Omega_S$ is defined as

$$\theta_t(\omega) \equiv \omega(t + \cdot), \quad \omega \in \Omega_S.$$

Obviously, for the canonical process w on Ω_S,

$$w(s, \theta_t(\omega)) = w(s + t, \omega).$$

We sometimes use $[w(s) \circ \theta_t](\omega)$ to denote $w(s, \theta_t(\omega))$ and often with ω being suppressed. For a $[0, \infty]$-valued random variable T, $w(s) \circ \theta_T$ can be defined as

$$[w(s) \circ \theta_T](\omega) \equiv \begin{cases} [w(s) \circ \theta_{T(\omega)}](\omega), & \omega \in \{T(\omega) < \infty\}, \\ \Delta, & \omega \in \{T(\omega) = \infty\}, \end{cases}$$

where Δ is a cemetery state, disjoint from $\mathbb{R}^2$. Readers should be warned, however, when the notation $w(s) \circ \theta_{T(\omega)}$ is used, it is understood that ω is fixed, i.e.,

$$[w(s) \circ \theta_T(\omega)](\omega') \equiv [w(s) \circ \theta_{T(\omega)}](\omega') \quad \text{for } \omega' \in \Omega_S.$$

Define

$$\mathcal{F}^{[\tau_n, \tau_{n+1}]} \equiv \sigma \left\{ w((\tau_n + t) \wedge \tau_{n+1}) 1_{\{\tau_n < \infty\}} + 1_{\{\tau_n = +\infty\}} \Delta, t \geq 0 \right\}.$$

Lemma 2.3 *For each $n \geq 1$,*

$$\mathcal{F}_{\tau_{n+1}} = \mathcal{F}_{\tau_n} \vee \mathcal{F}^{[\tau_n, \tau_{n+1}]}, \tag{2.10}$$

and $\theta_{\tau_n}^{-1} \left(\mathcal{F}^{[\tau_n, \tau_{n+1}]} \right) = \mathcal{F}_{\tau_1}$.

Proof. First, it follows from Lemma 1.3.3 of [48] that, for any stopping time τ,

$$\mathcal{F}_\tau = \sigma\{w(t \wedge \tau) : t \geq 0\}. \tag{2.11}$$

Equality (2.10) follows from (2.11). The rest of the proof uses the definition of the shift operator. □

Lemma 2.4 *Let $\{\tau_n : n \geq 1\}$ be a nondecreasing sequence of stopping times and for each n suppose P_n is a probability measure on $(\Omega_S, \mathcal{F}_{\tau_n})$. Assume that P_{n+1} equals P_n on $\mathcal{F}_{\tau_n}$ for each $n \geq 1$. If $\lim_{n \to \infty} P_n(\tau_n \leq t) = 0$ for all $t \geq 0$, then there is a unique probability measure P on $(\Omega_S, \mathcal{F})$ such that P equals P_n on $\mathcal{F}_{\tau_n}$ for all $n \geq 1$.*

Proof. See the proof of Theorem 1.3.5 of [48]. □

Lemma 2.5 *Let $s \geq 0$ be given and suppose that P is a probability measure on $(\Omega, \mathcal{F}^s)$, where $\mathcal{F}^s \equiv \sigma\{w(t) : t \geq s\}$. If $\eta \in C([0, s], \mathbb{R}^d)$ and $P(w(s) = \eta(s)) = 1$, then there is a unique probability measure $\delta_\eta \otimes_s P$ on $(\Omega, \mathcal{F})$ such that $\delta_\eta \otimes_s P(w(t) = \eta(t), 0 \leq t \leq s) = 1$ and $\delta_\eta \otimes_s P(A) = P(A)$ for all $A \in \mathcal{F}^s$.*

Proof. See the proof of Lemma 6.1.1 of [48]. □

Theorem 2.3 *For each $x \in S$, there is a unique probability measure Q_x on $(\Omega_S, \mathcal{F})$ such that $Q_x(C_S) = 1$, $Q_x(\int_0^\infty 1_{\partial S}(w(s))\,ds = 0) = 1$, $Q_x(\tau_n < \infty) = 1$ for each $n \geq 1$, $Q_x = P_x^{k_0}$ on $\mathcal{F}_{\tau_1}$, and moreover, for each n and Q_x-a.s. $\omega \in \{\tau_n < \infty\}$, $\left(P^{k_n}_{w(\tau_n(\omega),\omega)} \circ \theta^{-1}_{\tau_n}\right)(\cdot)$ is equal to $Q^n_\omega(\cdot)$ on $\mathcal{F}^{[\tau_n,\tau_{n+1}]}$, where Q^n_ω is a regular conditional probability distribution (r.c.p.d.) of*

$$Q_x\left(\cdot \,|\, \mathcal{F}_{\tau_n}\right)(\omega).$$

Proof. For $x \in S$, define $Q_x^1 \equiv P_x^{k_0}$ on $\mathcal{F}_{\tau_1}$. Then from Lemma 2.1, $Q_x^1(\tau_1 < \infty) = 1$ and from the definition of τ_1, $Q_x^1(w(\cdot \wedge \tau_1) \in C_S) = 1$, $Q_x^1(\int_0^{\tau_1} 1_{\partial S}(w(s))\,ds = 0) = 1$. Suppose Q_x^n on $\mathcal{F}_{\tau_n}$ has been defined, and $Q_x^n(\tau_n < \infty) = 1$, $Q_x^n(w(\cdot \wedge \tau_n) \in C_S) = 1$, $Q_x^n(\int_0^{\tau_n} 1_{\partial S}(w(s))\,ds = 0) = 1$ and for Q_x^n-a.e. $\omega \in \{\tau_{n-1} < \infty\}$, $\left(P^{k_{n-1}}_{w(\tau_{n-1}(\omega),\omega)} \circ \theta^{-1}_{\tau_{n-1}}\right)(\cdot)$ is equal to $Q^{n-1}_\omega(\cdot)$ on $\mathcal{F}^{[\tau_{n-1},\tau_n]}$ where Q^{n-1}_ω is an r.c.p.d. of $Q_x^{n-1}\left(\,\cdot \,|\, \mathcal{F}_{\tau_{n-1}}\right)(\omega)$. We want to define Q_x^{n+1} on $\mathcal{F}_{\tau_{n+1}}$ such that

$$Q_x^{n+1} = Q_x^n, \quad \text{on } \mathcal{F}_{\tau_n},$$

$Q_x^{n+1}(\tau_{n+1} < \infty) = 1$, $Q_x^{n+1}(w(\cdot \wedge \tau_{n+1}) \in C_S) = 1$, $Q_x^{n+1}(\int_0^{\tau_{n+1}} 1_{\partial S}(w(s))\,ds = 0) = 1$ and for Q_x^n-a.e. $\omega \in \{\tau_n < \infty\}$, $\left(P^{k_n}_{w(\tau_n(\omega),\omega)} \circ \theta^{-1}_{\tau_n}\right)(\cdot)$ is equal to $Q^n_\omega(\cdot)$ on $\mathcal{F}^{[\tau_n,\tau_{n+1}]}$, where Q^n_ω is a r.c.p.d. of

$$Q_x^n\left(\cdot \,|\, \mathcal{F}_{\tau_n}\right)(\omega).$$

Fix a $\omega \in \Omega_S$ such that $\tau_n(\omega) < \infty$. Now, $P^{k_n(\omega)}_{w(\tau_n(\omega),\omega)}$ is a probability measure on $(\Omega_S, \mathcal{F})$, therefore $\left(P^{k_n(\omega)}_{w(\tau_n(\omega),\omega)} \circ \theta^{-1}_{\tau_n(\omega)}\right)(\cdot)$ is a probability measure on $\mathcal{F}^{\tau_n(\omega)}$. Therefore by Lemma 2.5, for each ω, we can define a probability measure on $(\Omega, \mathcal{F})$ via

$$\delta_\omega \otimes_{\tau_n(\omega)} \left(P^{k_n(\omega)}_{w(\tau_n(\omega),\omega)} \circ \theta^{-1}_{\tau_n(\omega)}\right)(\cdot).$$

For any $A \in \mathcal{F}_{\tau_n}$ and $B \in \mathcal{F}_{\tau_1}$, since by Lemma 2.5,

$$\delta_\omega \otimes_{\tau_n(\omega)} \left(P^{k_n(\omega)}_{w(\tau_n(\omega),\omega)} \circ \theta^{-1}_{\tau_n(\omega)}\right)(\tau_n(\cdot) = \tau_n(\omega)) = 1,$$

we have $\theta_{\tau_n}(B) = \theta_{\tau_n(\omega)}(B)$ almost surely under $\delta_\omega \otimes_{\tau_n(\omega)} \left(P^{k_n(\omega)}_{w(\tau_n(\omega),\omega)} \circ \theta^{-1}_{\tau_n(\omega)}\right)$. Hence

$$\delta_\omega \otimes_{\tau_n(\omega)} \left(P^{k_n(\omega)}_{w(\tau_n(\omega),\omega)} \circ \theta^{-1}_{\tau_n(\omega)}\right)(A \cap \theta_{\tau_n}(B)) = 1_A(\omega) P^{k_n(\omega)}_{w(\tau_n(\omega),\omega)}(B), \tag{2.12}$$

which is of course $\mathcal{F}_{\tau_n}$-measurable. It follows from Lemma 2.3, that for any $A \in \mathcal{F}_{\tau_{n+1}}$

$$\delta_\omega \otimes_{\tau_n(\omega)} \left(P^{k_n(\omega)}_{w(\tau_n(\omega),\omega)} \circ \theta^{-1}_{\tau_n(\omega)}\right)(A)$$

is $\mathcal{F}_{\tau_n}$-measurable. For each $A \in \mathcal{F}_{\tau_{n+1}}$, define

$$Q_x^{n+1}(A) \equiv E^{Q_x^n}\left[\delta_\omega \otimes_{\tau_n(\omega)} \left(P^{k_n(\omega)}_{w(\tau_n(\omega),\omega)} \circ \theta^{-1}_{\tau_n(\omega)}\right)(A)\right]. \tag{2.13}$$

On the Q_x^n null set where $\tau_n = \infty$, the integrand in the right member above is defined to be δ_ω, the Dirac measure at point ω. Then Q_x^{n+1} is a probability measure on $\mathcal{F}_{\tau_{n+1}}$. It then follows from (2.12) and (2.13) that $Q_x^{n+1} = Q_x^n$ on $\mathcal{F}_{\tau_n}$, and

$$\begin{aligned} Q_x^{n+1}(\tau_{n+1} < \infty) &= Q_x^{n+1}(\tau_n(\cdot) < \infty, \tau_1(\theta_{\tau_n}(\cdot)) < \infty) \\ &= E^{Q_x^n}\left[1_{\{\tau_n<\infty\}} P^{k_n(\cdot)}_{w(\tau_n(\cdot),\cdot)}(\tau_1 < \infty)\right] \\ &= 1. \end{aligned}$$

Also,

$$\begin{aligned} &Q_x^{n+1}(w(\cdot \wedge \tau_{n+1}) \in C_S) \\ &= Q_x^{n+1}(\tau_n < \infty, w(\cdot \wedge \tau_n) \in C_S, w((\tau_n + \cdot) \wedge \tau_{n+1}) 1_{\{\tau_n<\infty\}} \in C_S) \\ &\quad + Q_x^{n+1}(\tau_n = \infty, \omega(\cdot) \in C_S) \\ &= E^{Q_x^n}\left[\tau_n < \infty, w(\cdot \wedge \tau_n) \in C_S; P^{k_n}_{w(\tau_n)}\left(w(\cdot \wedge \tau_1) \in C_S\right)\right] \\ &= 1, \end{aligned}$$

and similarly, we have

$$Q_x^{n+1}\left\{\int_0^{\tau_{n+1}} 1_{\partial S}(w(s))\, ds = 0\right\} = 1.$$

If we can show that for each $t \geq 0$,

$$\lim_{n\to\infty} Q_x^n(\tau_n \leq t) = 0, \tag{2.14}$$

then it follows from Lemma 2.4 that there is a Q_x on $(\Omega_S, \mathcal{F})$ with the desired properties. We leave the proof of (2.14) to the following lemma. □

For the rest of this section, we use E_x^n to denote $E^{Q_x^n}$.

Lemma 2.6 *For each $t \geq 0$,*

$$\lim_{n\to\infty} \sup_{x\in S} Q_x^n(\tau_n \leq t) = 0. \tag{2.15}$$

Proof. Fix $t > 0$ and let $\epsilon(t)$ be as in Lemma 2.2. We will prove by induction on n that

$$\sup_{x\in S} Q_x^n(\tau_n \leq t) \leq (1 - \epsilon(t))^n.$$

This is clearly true for $n = 0$. Suppose it holds for some $n \geq 0$. Then, for any $x \in S$,

$$\begin{aligned} Q_x^{n+1}\{\tau_{n+1} \leq t\} &= Q_x^{n+1}\{\tau_n \leq t, \tau_n + \tau_1 \circ \theta_{\tau_n} \leq t\} \\ &= \int_0^t Q_x^{n+1}\{\tau_n \in ds, \tau_1 \circ \theta_{\tau_n} \leq t - s\} \\ &= \int_0^t E^{Q_x^n}\left\{1_{\{\tau_n \in ds\}} P^{k_n}_{w(\tau_n)}\{\tau_1 \leq t - s\}\right\} \\ &\leq (1 - \epsilon(t)) \int_0^t E^{Q_x^n}\left\{1_{\{\tau_n \leq t\}}\right\} = (1 - \epsilon(t)) Q_x^n\{\tau_n \leq t\} \\ &\vdots \\ &\leq (1 - \epsilon(t))^{n+1}, \end{aligned}$$

where from the third equality to the following inequality, we have used Lemma 2.2. Hence

$$\lim_{n\to\infty} \sup_{x \in S} Q_x^n(\tau_n \leq t) = 0. \tag{2.16}$$

□

Theorem 2.4 *The family of probability measures $\{Q_x, x \in S\}$ defined in Theorem 2.3 has the following properties:*

(i) $Q_x(w(\cdot) \in C_S) = 1$,

(ii) $Q_x\left(\int_0^\infty 1_{\{w(s) \in \partial S\}}\, ds = 0\right) = 1$,

(iii) *for each $x \in S$, any $f \in C_b^2(S)$ with $D_i f \geq 0$ on F_i $(i = 1, 2, 3, 4)$*

$$m_f(t) \equiv f(w(t)) - \int_0^t Gf(w(s))\, ds \tag{2.17}$$

is an $\{\mathcal{F}_t\}$-submartingale under Q_x.

Proof. Properties (i) and (ii) have already been established for Q_x. For (iii), since $m_f(\cdot)$ is bounded on each finite interval and $\tau_n \to \infty$, Q_x-a.s., it is enough to show that, for each $n \geq 1$, $\{m_f(\cdot \wedge \tau_n), \mathcal{F}_{t\wedge\tau_n}, t \geq 0\}$ is a submartingale under Q_x. We prove this by induction. When $n = 1$, it follows from the definition of $Q_x = Q_x^1$ on $\mathcal{F}_{\tau_1}$ and since $m_f(t \wedge \tau_1) \in \mathcal{F}_{\tau_1}$ for each $t \geq 0$, $\{m_f(t \wedge \tau_1), \mathcal{F}_{t\wedge\tau_1}, t \geq 0\}$ is a submartingale. Assume $\{m_f(t \wedge \tau_n), \mathcal{F}_{t\wedge\tau_n}, t \geq 0\}$ is a submartingale under Q_x, and hence under Q_x^n. We first show that $\{m_f(t \wedge \tau_{n+1}), \mathcal{F}_{t\wedge\tau_{n+1}}, t \geq 0\}$ is a submartingale under Q_x^{n+1}. Because $Q_x|_{\mathcal{F}_{\tau_{n+1}}} = Q_x^{n+1}$, it follows that $\{m_f(t \wedge \tau_{n+1}), \mathcal{F}_{t\wedge\tau_{n+1}}, t \geq 0\}$ is a submartingale under Q_x, and this would finish our proof.

For any $0 \le t_1 < t_2$, and any $A \in \mathcal{F}_{t_1 \wedge \tau_{n+1}}$, we want to show

$$E_x^{n+1}\left[m_f(t_2 \wedge \tau_{n+1}); A\right] \ge E_x^{n+1}\left[m_f(t_1 \wedge \tau_{n+1}); A\right].$$

From the definition of Q_x^{n+1}, we have

$$E_x^{n+1}\left[m_f(t_2 \wedge \tau_{n+1}); A\right] = E_x^n\left[E^{\delta_\omega \otimes_{\tau_n(\omega)} \left(P^{k_n(\omega)}_{w(\tau_n(\omega,\omega))} \circ \theta^{-1}_{\tau_n(\omega)}\right)}\left[m_f(t_2 \wedge \tau_{n+1}); A\right]\right]. \tag{2.18}$$

For notational convenience, in this part of the proof only, we denote for each ω

$$P_\omega^n \equiv P^{k_n(\omega)}_{w(\tau_n(\omega,\omega))}.$$

For each fixed ω, we consider three cases. For $\tau_n(\omega) \ge t_2$ we have

$$E^{\delta_\omega \otimes_{\tau_n(\omega)} P^n_\omega \circ \theta^{-1}_{\tau_n(\omega)}}[m_f(t_2 \wedge \tau_{n+1}), A] = m_f(t_2 \wedge \tau_{n+1}) 1_A = m_f(t_2 \wedge \tau_n) 1_A.$$

The remaining cases are (a) $t_1 \le \tau_n(\omega) < t_2$ and (b) $\tau_n(\omega) < t_1$. We first consider case (a). By the definition of $\delta_\omega \otimes_{\tau_n(\omega)} P_\omega^n \circ \theta^{-1}_{\tau_n(\omega)}$, we have,

$$\begin{aligned}
&E^{\delta_\omega \otimes_{\tau_n(\omega)} P^n_\omega \circ \theta^{-1}_{\tau_n(\omega)}}\left[m_f(t_2 \wedge \tau_{n+1}), A\right] \\
&= 1_A(\omega) E^{\delta_\omega \otimes_{\tau_n(\omega)} P^n_\omega \circ \theta^{-1}_{\tau_n(\omega)}}\left[m_f(t_2 \wedge \tau_{n+1})\right] \\
&= 1_A(\omega) E^{P^n_\omega}\left[(m_f((t_2 - \tau_n(\omega)) \wedge \tau_1) - f(w(0)))\right] + 1_A m_f(\tau_n(\omega)) \\
&\ge 1_A(\omega) m_f(\tau_n(\omega)) = 1_A(\omega) m_f(t_2 \wedge \tau_n(\omega)),
\end{aligned}$$

where we have used the fact that

$$\begin{aligned}
m_f(t_2 \wedge \tau_{n+1}) &= f(w(t_2 \wedge \tau_{n+1})) - \int_0^{t_2 \wedge \tau_{n+1}} Gf(w(s))\, ds \\
&= f(w(t_2 \wedge \tau_{n+1})) - f(w(\tau_n(\omega))) - \int_{\tau_n(\omega)}^{t_2 \wedge \tau_{n+1}} Gf(w(s))\, ds \\
&\quad + f(w(\tau_n(\omega))) - \int_0^{\tau_n(\omega)} Gf(w(s))\, ds \\
&= (m_f((t_2 - \tau_n(\omega)) \wedge \tau_1 - f(w(0))) \circ \theta_{\tau_n(\omega)} + m_f(\tau_n(\omega)),
\end{aligned}$$

and $E^{P^n_\omega}\left[m_f(t \wedge \tau_1) - f(w(0))\right] \ge E^{P^n_\omega}\left[m_f(0) - f(w(0))\right] = 0$ since $m_f(t \wedge \tau_1) - f(w(0))$ is a submartingale starting from zero under P^n_ω. This follows in a similar manner to that in Lemma 7.2 of [26] using Itô's formula and the decomposition (b) of w under P_x^i, stopped

at τ_1. For case (b), suppose $A = C_1 \cap C_2$, where $C_1 \in \mathcal{F}_{\tau_n}$ and $C_2 = \theta_{\tau_n}(B)$ for $B \in \mathcal{F}_{(t_1-\tau_n(\omega))\wedge\tau_1}$. Then,

$$\begin{aligned}
&E^{\delta_\omega \otimes_{\tau_n(\omega)} P^n_\omega \circ \theta^{-1}_{\tau_n(\omega)}} [m_f(t_2 \wedge \tau_{n+1}); A] \\
&= 1_{C_1}(\omega) E^{\delta_\omega \otimes_{\tau_n(\omega)} P^n_\omega} \circ \theta^{-1}_{\tau_n(\omega)} [m_f(t_2 \wedge \tau_{n+1}); C_2] \\
&= 1_{C_1}(\omega) \left\{ E^{P^n_\omega} \left[(m_f((t_2 - \tau_n(\omega)) \wedge \tau_1) - f(w(0))) ; B \right] + m_f(\tau_n(\omega)) P^n_\omega(B) \right\} \\
&\geq 1_{C_1}(\omega) \left\{ E^{P^n_\omega} \left[(m_f((t_1 - \tau_n(\omega)) \wedge \tau_1) - f(w(0))) ; B \right] + m_f(\tau_n(\omega)) P^n_\omega(B) \right\} \\
&= 1_{C_1}(\omega) E^{\delta_\omega \otimes_{\tau_n(\omega)} P^n_\omega} \left[(m_f((t_1 - \tau_n(\omega)) \wedge \tau_1) - f(w(0))) \circ \theta_{\tau_n(\omega)}, \theta_{\tau_n(\omega)}(B) \right] \\
&\quad + E^{\delta_\omega \otimes_{\tau_n(\omega)} P^n_\omega} [m_f(\tau_n(\omega)); A] \\
&= E^{\delta_\omega \otimes_{\tau_n(\omega)} P^n_\omega} \circ \theta^{-1}_{\tau_n(\omega)} [m_f(t_1 \wedge \tau_{n+1}), A] .
\end{aligned}$$

Since sets of the form $A = C_1 \cap C_2$ generate $\mathcal{F}_{t_1 \wedge \tau_{n+1}}$, it follows that the left member above is greater than or equal to the last member above for all $A \in \mathcal{F}_{t_1 \wedge \tau_{n+1}}$. Putting these cases together yields

$$\begin{aligned}
&E^{n+1}_x [m_f(t_2 \wedge \tau_{n+1}); A] \\
&= E^n_x \left[1_{\{t_1 \leq \tau_n\}} E^{\delta_\omega \otimes_{\tau(\omega)} P^{k_n(\omega)}_{w(\tau_n(\omega,\omega))} \theta^{-1}_{\tau_n(\omega)}} [m_f(t_2 \wedge \tau_{n+1}); A] \right] \\
&\quad + E^n_x \left[1_{\{t_1 > \tau_n\}} E^{\delta_\omega \otimes_{\tau(\omega)} P^{k_n(\omega)}_{w(\tau_n(\omega,\omega))} \theta^{-1}_{\tau_n(\omega)}} [m_f(t_2 \wedge \tau_{n+1}); A] \right] \\
&\geq E^n_x \left[1_{\{t_1 \leq \tau_n\}} m_f(t_2 \wedge \tau_n); A \right] + E^{n+1}_x \left[1_{\{t_1 > \tau_n\}} m_f(t_1 \wedge \tau_{n+1}); A \right] \\
&\geq E^{n+1}_x \left[m_f(t_1 \wedge \tau_n) 1_{\{t_1 \leq \tau_n\} \cap A} \right] + E^{n+1}_x \left[1_{\{t_1 > \tau_n\}} m_f(t_1 \wedge \tau_{n+1}); A \right] \\
&= E^{n+1}_x [m_f(t_1 \wedge \tau_{n+1}); A] ,
\end{aligned}$$

where for the last inequality we have used the submartingale property of $\{m_f(t \wedge \tau_n), \mathcal{F}_{t\wedge\tau_n}, t \geq 0\}$ under Q_x and the fact that $A \cap \{t_1 \leq \tau_n\} \in \mathcal{F}_{t_1 \wedge \tau_n}$. Thus, $\{m_f(t \wedge \tau_{n+1}), \mathcal{F}_{t\wedge\tau_{n+1}}, t \geq 0\}$ is a Q_x-submartingale. □

Proof of Theorem 2.2. For each $x \in S$, if we define $P_x \equiv Q_x|_{C_S}$, noticing that $Z = w|_{C_S}$, $\mathcal{M}_t = \mathcal{F}_t \cap C_S$, $\mathcal{M} = \mathcal{F} \cap C_S$ and $Q_x(C_S) = 1$, it is easy to check that $\{P_x, x \in S\}$ has the desired properties as a family of probability measures on $(C_S, \mathcal{M})$. □

2.2.2 Semimartingale Representation

Assume $\mu = 0$. We first prove Z together with the family $\{P_x, x \in S\}$ in Theorem 2.2 is an (S, Γ, μ, R)-SRBM. Our approach follows the general line of Stroock and Varadhan [47], in

which only smooth domains were considered. We begin with a few lemmas. In this section, for any subset $U \subset S$, $Df \geq g$ on $U \cap \partial S$ means $D_i f \geq g$ on $U \cap F_i$ $(i = 1, 2, 3, 4)$, and all the (sub)martingales are with respect to the filtration $\{\mathcal{M}_t\}$.

Lemma 2.7 *There exists an $f_0 \in C_b^2(S)$ such that $Df_0 \geq 1$ on ∂S.*

Proof. Fix $x \in S$. If $x \in F_i^o$, (the part of F_i without corner points) for some i, let (r, θ) denote polar coordinates with origin at x and polar axis along the side F_i in the direction from x towards a_i. Let θ_x denote the angle between v_i and the inward unit normal n_x to F_i^o, where θ_x is taken as positive if v_i points towards a_{i-1} and is negative otherwise. Define

$$\psi_x(r, \theta) = re^{\theta \tan \theta_x}.$$

Then ψ_x is a continuous function on S that is infinitely differentiable in $S \backslash \{x\}$. Also, $v_i \cdot \nabla \psi_x = 0$ on F_i^o. Let $d_x = \text{dist}(x, \partial S \backslash F_i^o)$ and $c_x = 1/2\, d_x \exp(-\pi/2\, |\tan \theta_x|)$. Let h be a C^2, non-increasing function on $\mathbb{R}$ such that

$$h_x(y) = \begin{cases} 1 & \text{for} \quad y \leq 1/2\, c_x \\ 0 & \text{for} \quad y \geq c_x. \end{cases} \tag{2.19}$$

Define

$$\phi_x(z) = (n_x \cdot (z - x))\, h_x(\psi_x(z)) \quad \text{for all } z \in S.$$

Note that $\phi_x \in C_b^2(S)$ and $\phi_x(\cdot) = 0$ in a neighborhood of $S \backslash F_i^o$, by the choice of c_x. Now,

$$U_x \equiv \left\{ z \in S : \psi_x(z) < \frac{1}{2} c_x \right\}$$

is an open neighborhood of x in S where $\psi_x(z) = n_x \cdot (z - x)$ and hence $v_i \cdot \nabla \phi_x = 1$ on U_x and on F_i^o,

$$\begin{aligned} v_i \cdot \nabla \phi_x &= (v_i \cdot n_x)\, h_x(\psi_x(z)) + (n_x \cdot (z - x))\, h_x'(\psi_x(z)) v_i \cdot \nabla \psi_x(z) \\ &= 1\, h_x(\psi_x(z)) + 0 \\ &\geq 0. \end{aligned}$$

It follows that $D\phi_x \geq 0$ on ∂S.

On the other hand, if $x = a_i$ for some i, let (r, θ) be polar coordinates centered at x with polar axis in the direction of F_{i+1}. Let θ_1 be the angle that v_i makes with the inward normal to F_i and θ_2 be the angle that v_{i+1} makes with inward normal to F_{i+1}. Either of

these angles is positive if it points towards the corner a_i. Let $\alpha = 2(\theta_1 + \theta_2)/\pi$. Then (2.4) implies $\alpha < 1$. Define, for $r > 0$,

$$\psi_x(r,\theta) \equiv \begin{cases} r^\alpha \cos(\alpha\theta - \theta_2), & \alpha > 0, \\ r \exp(\theta \tan\theta_2), & \alpha = 0, \\ 1/(r^\alpha \cos(\alpha\theta - \theta_2)), & \alpha < 0. \end{cases}$$

Define $\psi_x(o) = 0$ where o denotes the origin of the polar coordinates (r,θ). Observe that $c \equiv \min_{0\le\theta\le\pi/2} \cos(\alpha\theta - \theta_2) \ge \cos(|\theta_1| \vee |\theta_2|) > 0$ and so ψ_x is continuous on S, infinitely differentiable on $S\backslash\{x\}$, $\psi_x \ge 0$ on S and on each ray emanating from x, ψ_x is an increasing function of r. Moreover (cf. Varadhan and Williams [53]),

$$v_j \cdot \nabla\psi_x = 0 \quad \text{on } F_j^o, j = i, i+1.$$

By condition (2.4), there is $u_x \in S_i$ (the quadrant with vertex at $x = a_i$ that contains S) such that $u_x \cdot v_i \ge 1$ and $u_x \cdot v_{i+1} \ge 1$. Let $d_x = \text{dist}(x, \partial S)\backslash\left(F_i^o \cup F_{i+1}^o \cup \{a_i\}\right)$ and

$$c_x = \begin{cases} 1/2\, d_x^\alpha c & \text{if} \quad \alpha > 0 \\ 1/2\, d_x \exp(-\pi/2|\tan\theta_2|) & \text{if} \quad \alpha = 0 \\ 1/2\, d_x^{-\alpha} & \text{if} \quad \alpha < 0. \end{cases}$$

Let h_x be defined as in (2.19) for this c_x and define

$$\phi_x(z) = (u_x \cdot (z - x))\, h_x(\psi_x(z)) \quad \text{for all } z \in S.$$

Then, in a similar manner to that for the case $x \in F_i^o$, we have $\phi_x \in C_b^2(S)$, $\phi_x \equiv 0$ in some neighborhood of $\partial S\backslash\left(F_i^o \cup F_{i+1}^o \cup \{a_i\}\right)$ and $D\phi_x \ge 0$ on ∂S and $v_j \nabla\phi_x \ge 1$ on $F_j \cap U_x$, $j = i, i+1$ where

$$U_x \equiv \left\{ z \in S : \psi_x(z) < \frac{1}{2} c_x \right\}.$$

Now, $\{U_x : x \in \partial S\}$ is an open cover of ∂S and so it has a finite subcover $\{U_{x_1}, \ldots, U_{x_n}\}$. Define

$$f_0(z) = \sum_{i=1}^n \phi_{x_i}(z) \quad \text{for all } z \in S.$$

Then f_0 has the desired properties. □

Suppose that $f \in C^2(S)$ (Since S is bounded, $C^2(S) = C_b^2(S)$), and $Df \ge 0$ on ∂S. Recall the definition of $\{m_f(t)\}$ in (2.17). Since we are restricting ourself on the space C_S, the canonical process is Z instead of w. Therefore

$$m_f(t) = f(Z(t)) - \int_0^t Gf(Z(s))\, ds,$$

and for each $x \in S$, $m_f(t)$ is a bounded P_x-submartingale. Hence by the Doob–Meyer decomposition theorem (cf. Theorem 6.12 and 6.13 of Ikeda and Watanabe [29, Chapter 1]), there exists an integrable, non-decreasing, adapted continuous function $\xi_f : [0,\infty) \times C_S \to [0,\infty)$ such that $\xi_f(0) = 0$ and $m_f(t) - \xi_f(t)$ is a P_x-martingale. In general, for $f \in C^2(S)$, we can find a constant c such that $D\bar{f} \geq 0$ on ∂S, where $\bar{f} = f + cf_0$, hence we can choose a $\xi_{\bar{f}}$ for $\bar{f}$. If we set $\xi_f \equiv \xi_{\bar{f}} - c\xi_{f_0}$, then we see that

(2.20) $\xi_f(t)$ is an adapted continuous function of bounded variation such that

1. $\xi_f(0) = 0$ and $E_x\left[|\xi_f|(t)\right] < \infty$ for $t \geq 0$, and
2. $m_f(t) - \xi_f(t)$ is a P_x-martingale.

Lemma 2.8 *For $f \in C^2(S)$, there is at most one ξ_f satisfying (2.20). Moreover, for each $t \geq 0$*

$$\int_0^t 1_{\mathcal{O}}(Z(s))\, d|\xi_f|(s) = 0, \quad P_x\text{-a.s.}$$

Proof. See Lemma 2.4 of [47]. □

Lemma 2.9 *If $f \in C^2(S)$ and U is an open neighborhood of a point $x \in \partial S$ such that $f \equiv c$ on U, then*

$$\int_0^t 1_U(Z(s))\, d|\xi_f|(s) = 0.$$

Proof. See the proof of Lemma 2.4 of [47]. □

Lemma 2.10 *Let $f \in C^2(S)$ and let U be a neighborhood of a point $x \in \partial S$ such that $Df \geq 0$ on $U \cap \partial S$. Then*

$$\int_0^t 1_U(Z(s))\, d\xi_f(s) \geq 0.$$

Proof. See the proof of Lemma 2.5 of [47]. □

Theorem 2.4 *Define*

$$\xi_0(t) = \int_0^t \frac{1}{Df_0(Z(s))} d\xi_{f_0}(s),$$

Then $\xi_0(0) = 0$, $E_x[\xi_0(t)] < \infty$,

$$\xi_0(t) = \int_0^t 1_{\partial S}(Z(s))\, d\xi_0(s)$$

and

$$m_f(t) - \int_0^t Df(Z(s))\, d\xi_0(s)$$

is a P_x-martingale for all $f \in C^2(S)$ which is constant in a neighborhood of each corner point.

Proof. It is obvious from the properties of ξ_{f_0} and Lemma 2.8 and 2.10 that ξ_0 defined satisfies all the conditions in the theorem except the last expression being a martingale. Therefore it is enough to show that

$$\xi_f(t) = \int_0^t Df(Z(s))\, d\xi_0(s) = \int_0^t \frac{Df(Z(s))}{Df_0(Z(s))}\, d\xi_{f_0}(s). \tag{2.21}$$

Notice that since $Df_0 \geq 1$ on ∂S and $f \in C^2(S)$ is constant near corners, the expression $Df(Z(s))/Df_0(Z(s))$ is continuous in s and so by Lemma 2.8, the integral in the right member of (2.21) is well defined and (2.21) itself is equivalent to

(a) $d\xi_f(t)$ being absolutely continuous with respect to $d\xi_{f_0}$, and

(b) $\dfrac{d\xi_f(t)}{d\xi_{f_0}(t)} = \dfrac{Df(Z(t))}{Df_0(Z(t))}$.

For (a), let $\bar{f} = f + cf_0$ and $\hat{f} = -f + cf_0$. Choose a large c such that $D\bar{f} \geq 0$ and $D\hat{f} \geq 0$ on ∂S, and so $-c\, d\xi_{f_0}(t) \leq d\xi_f(t) \leq c\, d\xi_{f_0}(t)$. Therefore (a) is true. To prove (b), let $\alpha(t) \equiv d\xi_f(t)/d\xi_{f_0}(t)$. For any $x \in \partial S$, let

$$\beta = \frac{Df(x)}{Df_0(x)}.$$

Since f is flat near corners, $Df(x)/Df_0(x)$ is a continuous function on S. Hence, for any $\epsilon > 0$, there is an open set $U \subset S$ containing x such that

$$(\beta - \epsilon)Df_0(y) \leq Df(y) \leq (\beta + \epsilon)Df_0(y), \quad y \in U, \quad d\xi_{f_0}\text{-a.e.}$$

Then it follows from Lemma 2.10 that

$$(\beta - \epsilon) \int_u^t 1_U(Z(s)) d\xi_{f_0}(s) \leq \int_u^t 1_U(Z(s))\alpha(s)\, d\xi_{f_0}(s) \leq (\beta + \epsilon) \int_u^t 1_U(Z(s)) d\xi_{f_0}(s)$$

for any $0 \leq u < t$. Hence

$$(\beta - \epsilon)1_U(Z(s)) \leq 1_U(Z(s))\alpha(s) \leq (\beta + \epsilon)1_U(Z(s)).$$

It follows that

$$\alpha(t) = \frac{Df(Z(t))}{Df_0(Z(t))},$$

which proves (b) and hence the theorem. □

Theorem 2.5 *For $i = 1, 2, 3, 4$, let*

$$L_i(t) \equiv \int_0^t 1_{F_i \setminus \{a_{i+1}\}}(Z(s))\, d\xi_0(s).$$

Then P_x-a.s. $L_i(0) = 0$, L_i is a non-decreasing, adapted continuous. L_i increases only at times when $Z(\cdot) \in F_i$, i.e.,

$$\int_0^t 1_{\{Z(s) \notin F_i\}}\, dL_i(s) = 0, \quad (i = 1, 2, 3, 4),$$

and for any $f \in C^2(S)$

$$f(Z(t)) - \int_0^t Gf(Z(s))\, ds - \sum_{i=1}^4 \int_0^t D_i f(Z(s))\, dL_i(s) \tag{2.22}$$

is a P_x-martingale.

Proof. It is clear, except (2.22), the defined L_i has all the desired properties. When f is flat near corners, (2.22) is proved in Theorem 2.4. Suppose $f \in C^2(S)$, and f is flat near corners except near corner a_1. We need to prove (2.22) is a martingale for such an f. This can be proved basically in the same was as in Theorem 5.5 and Theorem 6.2 of [56]. □

Theorem 2.6 *Define*

$$X(t) \equiv Z(t) - \sum_{i=1}^4 v_i L_i(t),$$

then $P_x(X(0) = x) = 1$, and under P_x, X is an (Γ, μ)-Brownian motion, and $X(t) - \mu t$ is an $\{\mathcal{F}_t\}$-martingale. Therefore

$$Z(t) = X(t) + RL(t)$$

is an (S, Γ, μ, R)-SRBM.

Proof. To prove X is a Brownian motion and a $\{\mathcal{F}_t\}$-martingale, it can be accomplished in an exact same way as the proof of Theorem 3.3. That Z is an SRBM follows from Theorem 2.5 and X being a Brownian motion. □

We have proved, when $\mu = 0$, there is a family $\{P_x, x \in S\}$ such that Z with $\{P_x, x \in S\}$ is an SRBM, that is, Z has the following semimartingale representation

$$Z(t) = X(t) + \sum_{i=1}^4 v_i L_i(t), \tag{2.23}$$

where X and L_i's satisfy (2.2) and (2.3). For arbitrary μ we have the following theorem.

Theorem 2.7 *Let there be given a rectangle S, covariance matrix Γ, a drift vector μ and a reflection matrix R whose columns satisfy (2.4). Then there is a family of probability measures $\{P_x, x \in S\}$ on $(C_S, \mathcal{M}, \{\mathcal{M}_t\})$ such that the canonical process Z together with $\{P_x, x \in S\}$ is an SRBM associated with the data (S, Γ, μ, R).*

Proof. For this proof only, to avoid confusion among different families of probability measures, we use $\{P_x^\mu, x \in S\}$ to denote the family corresponding to data (S, Γ, μ, R). Let $\mu_0 = 0$, it follows from Theorem 2.6 that there is a family $\{P_x^{\mu_0}, x \in S\}$ such that Z together with this family is an SRBM. In particular, Z has the representation (2.23). Fixing an $x \in S$, for each $t \geq 0$, let

$$\alpha(t) \equiv \exp\left(\mu \cdot (X(t) - x) - \frac{1}{2}|\mu|^2 t\right).$$

Then $\{\alpha_t, t \geq 0\}$ is a martingale on $(C_S, \mathcal{M}, \{\mathcal{M}_t\}, P_x^{\mu_0})$ and it follows from Girsanov's Theorem (cf. [6, Chapter 9]) that there exists a unique probability measure P_x^μ on $(C_S, \mathcal{M})$ such that

$$\frac{dP_x^\mu}{dP_x^{\mu_0}} = \alpha(t) \quad \text{on } \mathcal{M}_t \text{ for all } t \geq 0.$$

Since X is a (Γ, μ_0)-Brownian motion and $\{\mathcal{M}_t\}$-martingale starting from x under $P_x^{\mu_0}$, it also follows from Girsanov's Theorem that X is a (Γ, μ)-Brownian motion starting with x under P_x^μ, and $\{X(t) - \mu t, \mathcal{M}_t, t \geq 0\}$ is a martingale on $(C_S, \mathcal{M}, P_x^\mu)$. It remains to show (2.3) is true under P_x^μ, i.e., for each $t \geq 0$,

$$\int_0^t 1_{\{Z_i(s) \notin F_i\}}\, dL_i(s) = 0, \quad P_x^\mu\text{-a.s.}$$

This is true because

$$\int_0^t 1_{\{Z_i(s) \notin F_i\}}\, dL_i(s) = 0, \quad P_x^{\mu_0}\text{-a.s.}$$

and P_x^μ is equivalent to $P_x^{\mu_0}$ on $\mathcal{M}_t$. Thus, for each $x \in S$ we have constructed P_x^μ such that (2.1), (2.2) and (2.3) are satisfied under P_x^μ. □

2.2.3 Uniqueness

In this section, we prove that the family $\{P_x, x \in S\}$ is unique. Let F^o denote the smooth part of the boundary ∂S, i.e., F^o is obtained by taking out four corner points from ∂S. Let

$$\mathcal{D}_0 = \Big\{ f : f \in C^1(S) \cap C^2(\mathcal{O} \cup F^o), \quad D_i f(x) = 0 \quad \text{on} \quad F_i, \\ i = 1, 2, 3, 4, \text{ and } Gf \text{ has a continuous extension onto } S \Big\}. \tag{2.24}$$

Definition 2.3 *Let π be a probability measure on S. By a solution of the martingale problem for (G,π) we mean a probability measure P on $(C_S,\mathcal{M})$ such that $PZ(0)^{-1}=\pi$ and for each $f\in\mathcal{D}_0$,*

$$f(Z(t))-\int_0^t Gf(Z(s))\,ds \tag{2.25}$$

is a P-martingale with respect to the filtration $\{\mathcal{M}_t\}$.

Remark. From now on, if no filtration is explicitly given, every martingale considered will be a martingale with respect to the filtration $\{\mathcal{M}_t\}$.

Proposition 2.1 *For any probability measure π on S, the measure $P_\pi\equiv\int_S P_x\,\pi(dx)$ is a solution of the martingale problem for (G,π).*

Proof. It is enough to show that for each $f\in\mathcal{D}_0$ and each $x\in S$,

$$f(Z(t))-\int_0^t Gf(Z(s))\,ds \tag{2.26}$$

is a P_x-martingale. By a standard convolution argument [56, p.30], there is a sequence $\{f_n\}$ of functions in $C^2(S)$ such that f_n and ∇f_n converge uniformly on S to f and ∇f, respectively, and $\{Gf_n\}$ is bounded on S and converges pointwise to Gf on $\mathcal{O}\cup F^o$. Since Z has the semimartingale representation (2.1), applying Itô's formula with f_n on the completion $(C_S,\tilde{\mathcal{M}},P_x)$ of $(C_S,\mathcal{M},P_x)$, we obtain P_x-a.s. for all $t\ge 0$:

$$\begin{aligned} f_n(Z(t)) &= f_n(Z(0))+\int_0^t \nabla f_n(Z(s))\,d\xi(s)+\sum_{i=1}^2\int_0^t D_i f_n(Z(s))\,dL_i(s) \\ &\quad +\int_0^t Gf_n(Z(s))\,ds, \end{aligned} \tag{2.27}$$

where $\xi(t)\equiv X(t)-\mu t$. By the uniform convergence of $\{\nabla f_n\}$ on S, the stochastic integral (with respect to $d\xi$) in (2.27) converges in $L^2(C_S,\tilde{\mathcal{M}},P_x)$ to that with f in place of f_n. Moreover, since $\{Gf_n(Z(s))\}$ converges boundedly to $Gf(Z(s))$ on $\{s\in[0,t]: Z(s)\in\mathcal{O}\cup F^o\}$, and by (2.6),

$$\zeta\{s\in[0,t]: Z(s)\notin\mathcal{O}\cup F^o\}=0 \quad P_x\text{-a.s.},$$

where ζ is Lebesgue measure on $\mathbb{R}$, then it follows by bounded convergence that the last integral in (2.27) converges P_x-a.s. to that with f in place of f_n. The remaining terms in (2.27) converge in a similar manner. Hence, (2.27) holds with f in place of f_n. Then it follows by $D_if=0$ on F_i $(i=1,2,3,4)$ and (2.3) that

$$f(Z(t))-\int_0^t Gf(Z(s))\,ds \tag{2.28}$$

is a martingale on $(C_S, \tilde{\mathcal{M}}, \{\tilde{\mathcal{M}}_t\}, P_x)$ where $\tilde{\mathcal{M}}_t$ denotes the augmentation of $\mathcal{M}_t$ by the P_x-null sets in $\tilde{\mathcal{M}}$. But since (2.28) is adapted to $\{\mathcal{M}_t\}$, it is in fact a martingale on $(C_S, \mathcal{M}, \{\mathcal{M}_t\}, P_x)$. This proves the proposition.

Lemma 2.11 *The operator $(G, \mathcal{D}_0)$ is* dissipative, *i.e., for every $f \in \mathcal{D}_0$ and every $\lambda > 0$:*

$$||\lambda f - Gf|| \geq \lambda ||f|| \tag{2.29}$$

where the norm $||\cdot||$ is the supremum norm on $C(S)$.

Proof. For $x \in S$, let δ_x be the Dirac measure at x. By Proposition 2.1, P_x is a solution of the martingale problem for (G, δ_x). Hence for $f \in \mathcal{D}_0$,

$$f(Z(t)) - f(Z(0)) - \int_0^t Gf(Z(s))\, ds \tag{2.30}$$

is a P_x-martingale. It follows that for $\lambda > 0$

$$e^{-\lambda t} f(Z(t)) - f(Z(0)) - \int_0^t e^{-\lambda s}(\lambda - G) f(Z(s))\, ds \tag{2.31}$$

is also a P_x-martingale. Therefore, by taking expectation E_x with respect to P_x in (2.31), we obtain

$$E_x\left[e^{-\lambda t} f(Z(t))\right] - f(x) = E_x\left[\int_0^t e^{-\lambda s}(\lambda - G) f(Z(s))\, ds\right]. \tag{2.32}$$

Letting $t \to \infty$, we yields

$$f(x) = E_x\left[\int_0^\infty e^{-\lambda s}(\lambda - G) f(Z(s))\, ds\right], \tag{2.33}$$

and from (2.33) one immediately gets $\lambda||f|| \leq ||(\lambda - G)f||$.

Theorem 2.8 *For every probability measure π on S, the martingale problem (G, π) has the unique solution P_π.*

Proof. It has been proved in Proposition 2.1 that P_π is a solution of the martingale problem for (G, π). Now we will show that the solution is unique. For every Hölder continuous function g on S and every $\lambda > 0$, by using Lieberman's theorem [34, Theorem 1], there is $u \in C^1(S) \cap C^2(\mathcal{O})$ such that $(\lambda - G)u = g$ on $\mathcal{O}$ and $D_i u(x) = 0$ on F_i $(i = 1, 2, 3, 4)$. By the classical regularity properties of elliptic partial differential equations (cf. Gilbarg and Trudinger [17, Lemma 6.18]), u is twice differentiable on the smooth part of the boundary F^o. Since $Gu(x) = \lambda u(x) - g(x)$ for $x \in \mathcal{O}$, Gu has continuous extension to S, and therefore

we have $u \in \mathcal{D}_0$ and $(\lambda - G)u = g$. Because the set of Hölder continuous functions is dense in $C(S)$ (with the sup norm topology), the range of $\lambda - G$ is dense in $C(S)$ for every $\lambda > 0$. Also, by Lemma 2.11, the operator $(G, \mathcal{D}_0)$ is dissipative. Therefore we can apply the uniqueness theorem of Ethier and Kurtz [11, Theorem 4.1 of Chapter 4] to assert that the solution of the martingale problem for (G, π) is unique.

Now we are ready to prove Theorem 2.1.

Proof of Theorem 2.1. Existence of a family $\{P_x, \in S\}$ is given in Theorem 2.7 and the uniqueness is given in Theorem 2.8. To show Feller continuity, it is enough to show for any $x \in S$ and sequence $\{x_n\}$ in S such that $x_n \to x \in S$, that one has $P_{x_n} \Rightarrow P_x$, where the symbol "$\Rightarrow$" means that the left member converges *weakly* to the right member. To see this, notice that since the state space S is compact, by *tightness*, the family $\{P_{x_n}\}$ is tight, and hence it is precompact in the topology of weak convergence, see Billingsley [3]. Assume $P_{x_{n_k}} \Rightarrow P_*$ for some subsequence $\{n_k\}$. Using an argument similar to that in the proof of Theorem 3.1 later in this dissertation and using the uniqueness of a solution to the martingale problem for (G, δ_x) (Theorem 2.8), we can show $P_* = P_x$. Therefore $P_{x_{n_k}} \Rightarrow P_x$ for any convergent subsequence $P_{x_{n_k}}$, hence $P_{x_n} \Rightarrow P_x$ and this proves the Feller continuity. It follows from uniqueness for the martingale problem (G, δ_x), Feller continuity and Theorem 4.2 in Chapter 4 of [11] that Z with $\{P_x, s \in S\}$ is a strong Markov process, i.e.,

$$E_x\left[f(Z(\tau + t))|\mathcal{M}_\tau\right] = E_{Z(\tau)} f(Z(t)), \quad P_x\text{-a.s.}$$

for any $f \in B(S)$, $t \geq 0$, and P_x-a.s. finite $\{\mathcal{M}_t\}$-stopping time τ.

It remains to prove that

$$\sup_{x \in S} E_x\left[L_i(t)\right] < \infty$$

for each $t \geq 0$ $(i = 1, 2, 3, 4)$. To see this, for the function f_0 defined in Lemma 2.7,

$$f_0(Z(t)) - f_0(Z(0)) - \int_0^t G f_0(Z(s))\, ds - \sum_{i=1}^4 \int_0^t D_i f_0(Z(s))\, dL_i(s)$$

is a martingale. Taking expectation with respect P_x, we have

$$E_x\left[f_0(Z(t))\right] - f_0(x) - E_x\left[\int_0^t G f_0(Z(s))\, ds\right] = E_x\left[\sum_{i=1}^4 \int_0^t D_i f_0(Z(s))\, dL_i(s)\right]$$

Because $D_i f_0 \geq 1$ on F_i, we have

$$\sup_{x \in S} E_x\left[\sum_{i=1}^4 L_i(t)\right] \leq 2||f_0|| + ||G f_0|| t.$$

□

2.3 Stationary Distribution

2.3.1 The Basic Adjoint Relationship (BAR)

For a probability measure π on S, recall that P_π has been defined as $P_\pi(A) \equiv \int_S P_x(A)\,\pi(dx)$. Let E_π denote the expectation with respect to P_π. A probability measure π on S is called a *stationary distribution* of the SRBM Z if for every bounded Borel function f on S and every $t > 0$

$$\int_S E_x[f(Z_t)]\,\pi(dx) = \int_S f(x)\,\pi(dx).$$

Because the state space S is compact, there is a stationary distribution for Z (see Dai [8]). Also, noticing from Theorem 2.1 that $\sup_{x\in S} E_x\,[L_i(t)] < \infty$ $(i = 1, 2, 3, 4)$ and using arguments virtually identical to those in [26], one can show that

Proposition 2.2 *Any stationary distribution for an SRBM Z is unique. If π is the stationary distribution,*

(a) *π is equivalent to Lebesgue measure dx on S, denoted as $\pi \approx dx$, and for each $x \in S$ and $f \in C(S)$*

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} E_x\,[f(Z(i))] = \int_S f(z)\,d\pi(z).$$

(b) *there is a finite Borel measure ν_i on F_i such that $\nu_i \approx \sigma_i$, where σ_i is Lebesgue measure on F_i, and for each bounded Borel function f on F_i and $t \geq 0$,*

$$E_\pi \left[\int_0^t f(Z(s))\,dL_i(s) \right] = t \int_{F_i} f\,d\nu_i, \quad (i = 1, 2, 3, 4).$$

□

For an $f \in C^2(S)$, applying Itô's formula to the process Z exactly as in [26], one has that

$$\begin{aligned} f(Z(t)) &= f(Z(0)) + \sum_{i=1}^{2} \int_0^t \frac{\partial}{\partial x_i} f(Z(s))\,d\xi_i(s) + \int_0^t Gf(Z(s))\,ds \\ &\quad + \sum_{i=1}^{4} \int_0^t \mathcal{D}_i f(Z(s))\,dL_i(s) \end{aligned} \tag{2.34}$$

where $\xi_i(t) = X_i(t) - \mu_i t$. Again proceeding exactly as in [26], we can then take E_π of both sides of (2.34) to prove the following theorem.

Theorem 2.9 *The stationary density* p_0 $(\equiv d\pi/dx)$ *and the boundary densities* p_i $(\equiv d\nu_i/d\sigma)$ $(i = 1, 2, 3, 4)$ *jointly satisfy the following basic adjoint relationship (BAR):*

$$\int_S (Gf \cdot p_0)\, dx + \sum_{i=1}^{4} \int_{F_i} (D_i f \cdot p_i)\, d\sigma_i = 0 \quad \text{for all } f \in C^2(S). \tag{2.35}$$

□

2.3.2 Sufficiency of (BAR)—A First Proof

The argument given in the previous section shows that (2.35) is *necessary* for p_0 to be the stationary density of Z. The following theorem says that the converse is true. It is an essential part of an algorithm that we are going to develop to compute the stationary density numerically. Note that π is not initially assumed to have a density, nor are $\nu_1, \ldots, \nu_4$ initially assumed to have densities.

Theorem 2.10 *Suppose that* π *is a probability measure on* S *and* $\nu_1, \ldots, \nu_4$ *are positive finite Borel measures on* $F_1, \ldots, F_4$ *respectively. If they jointly satisfy*

$$\int_S Gf\, d\pi + \sum_{i=1}^{4} \int_{F_i} \mathcal{D}_i f\, d\nu_i = 0 \quad \text{for all } f \in C^2(S), \tag{2.36}$$

then π *is the stationary distribution* $p_0\, dx$ *of* Z *and the* ν_i *are the corresponding boundary measures defined in Proposition 2.2.*

Remark. We are going to give a more or less *direct* proof of the main part of this theorem. This proof establishes that π is the stationary distribution but does not show that $\nu_1, \ldots, \nu_4$ are the corresponding boundary measures defined in Proposition 2.2. Nevertheless, the theorem is true. By considering a corresponding *constrained martingale problem*, we are able to provide a complete proof of the theorem. That general proof is left to Chapter 4 where we deal with SRBM in an orthant. Before we present the proof of Theorem 2.10, two lemmas are needed.

Lemma 2.12 *The operator* $(G, \mathcal{D}_0)$ *satisfies the positive maximum principle, i.e., whenever* $f \in \mathcal{D}_0$, $x_0 \in S$, *and* $\sup_{x \in S} f(x) = f(x_0) \geq 0$, *we have* $Gf(x_0) \leq 0$.

Proof. Suppose that $f \in \mathcal{D}_0$, $x_0 \in S$, and $\sup_{x \in S} f(x) = f(x_0)$, Because (2.26) is a P_{x_0}-martingale, by taking expectations under P_{x_0}, we have

$$E_{x_0}\left[f(Z(t))\right] - f(x_0) = E_{x_0}\left[\int_0^t Gf(Z(s))\, ds\right]. \tag{2.37}$$

The left hand side of (2.37) is non-positive because of x_0 being a maximum point of f. Dividing the right hand side of (2.37) by t and taking $t \to 0$, by the continuity of Gf and the continuity of the process Z, we get $Gf(x_0) \leq 0$. □

The proof of the following lemma is adapted from Williams [56, Lemma 4.4]. Note that the symbol θ is used in the following proof to denote the angle in polar coordinates.

Lemma 2.13 *$\mathcal{D}_0$ is dense in $C(S)$ (with the sup norm topology).*

Proof. It is easy to check that $\mathcal{D}_0$ is an algebra, i.e., for any pair $f, g \in \mathcal{D}_0$, $\alpha f + \beta g \in \mathcal{D}_0$ and $f \cdot g \in \mathcal{D}_0$ for any real constants α and β. Since S is compact and all the constant functions are in $\mathcal{D}_0$, by the Stone–Weierstrass theorem [42, p.174], it is enough to show that $\mathcal{D}_0$ separates points in S, i.e., for any distinct pair z_0 and z^* in S, there is an $f \in \mathcal{D}_0$ such that $f(z_0) = 0$ and $f(z^*) = 1$.

If one of the z's is in the interior $\mathcal{O}$, then such a function f separating z_0 and z^* can be trivially constructed. Now assume both the z_0 and z^* are in ∂S, but at least one of them, say z_0, is in the interior part of the boundary F^o. We can further assume z_0 is in the interior of F_1, the proof for $z_0 \in F_i$ $(i = 2, 3, 4)$ is similar. Let $v_1^{\perp} = (v_{12}, -1)'$ be a vector perpendicular to v_1, where v_{12} is the second component of v_1. For any $\epsilon > 0$, choose $g : \mathbb{R} \to [0, 1]$ to be twice continuously differentiable function satisfying

$$g(y) = \begin{cases} 1 \text{ for } |y| \leq \frac{\epsilon}{2}, \\ 0 \text{ for } |y| \geq \epsilon. \end{cases} \tag{2.38}$$

Define

$$f(z) = 1 - g(z_1) g\left((z - z_0) \cdot v_1^{\perp}\right), \quad z = (z_1, z_2). \tag{2.39}$$

Then $f \in C^2(S)$, $f(z_0) = 0$ and for $z \in F_1$,

$$\begin{aligned} D_1 f(z) &= -g'(0) g\left((z - z_0) \cdot v_1^{\perp}\right) - g(0) g'\left((z - z_0) \cdot v_1^{\perp}\right) v_1^{\perp} \cdot v_1 \\ &= 0 \end{aligned}$$

since $g'(0) = 0$ and $v_1^{\perp} \cdot v_1 = 0$. Also, on $\{z : |z - z_0| > (|v_{12}| + 2)\epsilon\} \cap S$, $f(z) \equiv 1$. Therefore, by choosing a small enough ϵ, we have $f(z^*) = 1$ and $D_j f(z) = 0$ on F_j $(j = 2, 3, 4)$. Thus, $f \in \mathcal{D}_0$ and f separating z_0 and z^*.

The remaining cases are when both z_0 and z^* are at corners. Without loss of generality, we can assume z_0 to be the origin, and $|z^*| > 1$. Let θ_i denote the angle that the direction of reflection on F_i makes with the inward normal to the side F_i $(i = 1, 2)$, positive angles

being toward the origin ($-\frac{\pi}{2} < \theta_1, \theta_2 < \frac{\pi}{2}$). Then $v_1 = (1, -\tan\theta_1)'$ and $v_2 = (-\tan\theta_2, 1)'$. Also, let $\alpha \equiv 2(\theta_1 + \theta_2)/\pi$. Then condition (2.4) implies that $\alpha < 1$.

Let us first assume that $\alpha > 0$. Define

$$\Phi(r,\theta) = r^\alpha \cos(\alpha\theta - \theta_2) \quad \text{for} \quad (r,\theta) \in S. \tag{2.40}$$

Proceeding exactly as in [56], we have for $r > 0$

$$D_1\Phi\left(r, \frac{\pi}{2}\right) = 0, \quad D_2\Phi(r,0) = 0. \tag{2.41}$$

If we define

$$c \equiv \min_{\theta\in[0,\frac{\pi}{2}]} \cos(\alpha\theta - \theta_2), \tag{2.42}$$

then c is strictly positive. Let $g : [0,\infty) \to [0,1]$ be a twice continuously differentiable function satisfying

$$g(y) = \begin{cases} 0 & \text{for} \quad 0 \le y \le \frac{c}{2} \\ 1 & \text{for} \quad y \ge c, \end{cases} \tag{2.43}$$

and let $f(z) = g(\Phi(z))$. It is easy to check that f is identically zero when $|z| < (\frac{c}{2})^{1/\alpha}$, therefore $f \in C^2(S)$ and $f(z_0) = 0$. Also one can check that

$$f(z) \equiv 1, \quad \text{for} \quad |z| \ge 1. \tag{2.44}$$

Therefore $f(z^*) = 1$, and by (2.41), (2.41) and (2.44), we get $D_j f = 0$ $(i = 1,2,3,4)$. Hence $f \in \mathcal{D}_0$ separates z_0 from z^*.

For $\alpha < 0$, we let

$$\Phi(r,\theta) = r^{-\alpha} \cos(\alpha\theta - \theta_2) \quad \text{for} \quad (r,\theta) \in S, \tag{2.45}$$

and construct f as in the previous case. Proceeding almost exactly as in the previous case, we can show that the function $f \in \mathcal{D}_0$ separates z_0 from z^*.

The last case we are now considering is when $\alpha = 0$. In this case, we let

$$\Phi(r,\theta) = re^{\theta\tan\theta_2}, \quad \text{for} \quad (r,\theta) \in S, \tag{2.46}$$

and use the same g as in (2.43) with $c \equiv \inf_{\theta\in[0,\pi/2]} e^{\theta\tan\theta_2}$. It can be checked that $f \equiv g(\Phi) \in \mathcal{D}_0$ separating z_0 from z^*. This finishes the proof of the lemma. □

Proof of Theorem 2.10. For $f \in \mathcal{D}_0$, the basic adjoint relationship (2.35) reduces to

$$\int_S Gf(x)\, d\pi(x) = 0. \tag{2.47}$$

As shown in the proof of Lemma 2.13, $\mathcal{D}_0$ is an algebra and $\mathcal{D}_0$ is dense in $C(S)$. Moreover, by Lemma 2.12, the operator $(G, \mathcal{D}_0)$ satisfies the positive maximum principle, and therefore Echeverria's theorem (see [10] or [11, Theorem 9.17 of Chapter 4]) can be applied to assert that π is a stationary distribution for a solution of the martingale problem for (G, π). By Theorem 2.8, the law of $Z = \{Z(t)\}$ under P_π is the unique solution to the martingale problem for (G, π). Therefore, π is a stationary distribution for Z. Because the stationary distribution of Z is unique, $d\pi(x) = p_0\, dx$ on S. □

2.4 Numerical Method for Steady-State Analysis

In this section we develop an algorithm for computing the stationary density p_0 and the boundary densities p_i $(i = 1, 2, 3, 4)$. The higher dimensional analog will be discussed in Chapter 4. The following conjecture is vital assumption in our proof of the convergence of the algorithm that we develop.

Conjecture 2.1 *Suppose that p_0 is an integrable Borel function in S such that $\int_S p_0\, dx = 1$ and $p_1, \ldots, p_4$ are integrable on $F_1, \ldots, F_4$ respectively. If they jointly satisfy the basic adjoint relationship (2.36), then p_i is non-negative* $(i = 0, 1, 2, 3, 4)$.

2.4.1 Inner Product Version of (BAR) and a Least Squares Problem

Readers might naturally assume that it is best to convert (2.35) into a direct PDE for p_0, but that gets very complicated because of auxiliary conditions associated with the singular parts of the boundary; we are just going to work with (2.35) directly. We start this section by converting (2.35) into a compact form that will be used later. Let

$$\mathcal{A}f = (Gf; D_1 f, D_2 f, D_3 f, D_4 f), \tag{2.48}$$

$$d\lambda = (dx; d\sigma_1, d\sigma_2, d\sigma_3, d\sigma_4). \tag{2.49}$$

We also incorporate the stationary density p_0 with the boundary densities p_i into a new function p, i.e.

$$p = (p_0; p_1, p_2, p_3, p_4). \tag{2.50}$$

Hereafter, we simply call p the stationary density of the corresponding SRBM. For a subset E of $\mathbb{R}^d$, let $\mathcal{B}(E)$ denote the set of functions which are $\mathcal{B}_E$ measurable. For $i = 1, 2$, let

$$L^i(S, d\lambda) \equiv \left\{ g = (g_0; g_1, \ldots, g_4) \in \mathcal{B}(S) \times \mathcal{B}(F_1) \times \cdots \times \mathcal{B}(F_4) : \int_S |g_0|^i \, dx + \sum_{j=1}^{4} \int_{F_j} |g_j|^i \, d\sigma_j < \infty \right\}, \tag{2.51}$$

and for $g \in L^1(S, d\lambda)$, let

$$\int_S g d\lambda \equiv \int_S g_0 \, dx + \sum_{i=1}^{4} \int_{F_i} g_i \, d\sigma_i.$$

For $g, h \in \mathcal{B}(S) \times \mathcal{B}(F_1) \times \cdots \times \mathcal{B}(F_4)$, we put $g \cdot h \equiv (g_0 h_0; g_1 h_1, \ldots, g_4 h_4)$, and for $h > 0$ ($h_i > 0, i = 0, 1, \ldots, 4$), we put $g/h \equiv (g_0/h_0; g_1/h_1, \ldots, g_4/h_4)$. With these notation the basic adjoint relationship (2.35) can be rewritten as

$$\int_S (\mathcal{A}f \cdot p) \, d\lambda = 0, \qquad \text{for all } f \in C^2(S). \tag{2.52}$$

Next we convert the problem of solving (2.52) into a *least squares problem*, and then propose an algorithm to solve the least squares problem. Our approach is similar in spirit to that of Bramble and Schatz [5], who considered a Rayleigh-Ritz-Galerkin method for solutions of the Dirichlet problem using a subspace without boundary conditions. The purpose of their method was to avoid finding boundary elements when the boundary of the domain is complicated. In our problem, the domain is not complicated at all except that it is non-smooth, but the boundary condition is implicit in (2.52) and is not known to us.

We start with the compact form (2.52) of the basic adjoint relationship (2.35). Let $L^2 = L^2(S, d\lambda)$, and denote by $||\cdot||$ the usual L^2 norm and by $(\cdot\,, \cdot)$ the usual inner product. It is evident that $\mathcal{A}f \in L^2$ for any $f \in C^2(S)$. Hence we can define

$$H = \overline{\{\mathcal{A}f : \quad f \in C^2(S)\}},$$

where the closure is taken in L^2. If one assumes that the unknown density p is in L^2, then (2.52) says simply that $\mathcal{A}f \perp p$ for all $f \in C^2(S)$, or equivalently $p \in H^\perp$. Conversely if $w \in H^\perp$, then w satisfies (2.52).

Let us assume for the moment that the unknown density function p defined by (2.50) is in L^2. That is, assume p_0 is square integrable with respect to Lebesgue measure in S, and p_i

is square integrable with respect to one-dimensional Lebesgue measure on F_i $(i = 1, 2, 3, 4)$. In order to construct a function $w \in H^\perp$, let

$$\phi_0(x) = (1; 0, 0, 0, 0). \tag{2.53}$$

Because p_0 is a probability density, we have $(p, \phi_0) = \int_S (p \cdot \phi_0)\, d\lambda = \int_S p_0\, dx = 1$, so p is *not* orthogonal to ϕ_0. On the other hand, we know from (2.52) that $p \perp h$ for all $h \in H$, and therefore ϕ_0 is *not* in H. Let $\bar{\phi_0}$ be the projection of ϕ_0 onto H. That is,

$$\bar{\phi_0} \equiv \text{argmin}_{\phi \in H} ||\phi_0 - \phi||^2. \tag{2.54}$$

Because ϕ_0 is not in H, we know that

$$\tilde{\phi_0} \equiv \phi_0 - \bar{\phi_0} \neq 0. \tag{2.55}$$

Obviously, $\tilde{\phi_0} \in H^\perp$. Simple algebra gives

$$\alpha \equiv \int_S \left(\tilde{\phi_0}(x) \cdot \phi_0(x)\right) d\lambda(dx) = (\tilde{\phi_0}, \tilde{\phi_0}) > 0.$$

Proposition 2.3 *Suppose that $p \in L^2$. Then $w = (w_0, w_1, \ldots, w_4) \equiv \frac{1}{\alpha}\tilde{\phi_0}$ satisfies the basic adjoint relationship (2.52) and $\int_S w_0\, dx = 1$. Therefore, assuming that Conjecture 2.1 is true, we have $w = p$ almost everywhere with respect to $d\lambda$.*

Proof. Let $w \equiv \frac{1}{\alpha}\tilde{\phi_0}$. Then, by construction, w satisfies the basic adjoint relationship (2.52). Because $\int_S w_0\, dx = 1$, Conjecture 2.1 asserts that w_i are non-negative. Thus Theorem 2.10 can be applied to assert $w = p$ almost everywhere with respect to $d\lambda$ □

As we will see later, the assumption that p is in L^2 is *not* satisfied in all cases of practical interest. However, when that assumption is satisfied, Proposition 2.3 says that in order to find the unknown stationary density p, it suffices to solve the *least squares problem* (2.54).

We now define some quantities that are of interest in the queueing theoretic applications of SRBM. Let $m_i = \int_S (x_i \cdot w_0(x))\, dx$ $(i = 1, 2)$ and $\delta_i = \int_{F_i} w_i(x)\, d\sigma_i$ $(i = 1, 2, 3, 4)$. Assuming Conjecture 2.1 is true, $m_i = \int_S (x_i \cdot p_0(x))\, dx$, which represents the long-run average value of Z_i, and $\delta_i = \int_{F_i} p_i(x)\, d\sigma_i$, which represents the long-run average amount of pushing per unit of time needed on boundary F_i in order to keep Z inside the rectangle S. That is, $E_x[L_i(t)] \sim \delta_i t$ as $t \to \infty$ for each $x \in S$ $(k = 1, 2, 3, 4)$, see Proposition 2.2 for details.

2.4.2 An Algorithm

Given Proposition 2.3, we will now propose an algorithm for approximate computation of p based on L^2 projection. In the examples presented later, it will be seen that the algorithm works well even in cases where p is known *not* to be in L^2.

Proposition 2.4 *Suppose that we can construct a sequence of finite dimensional subspaces* $\{H_n\}$ *of* H *such that* $H_n \uparrow H$ *as* $n \uparrow \infty$ ($H_n \uparrow H$ *means that* $H_1, H_2, \ldots$ *are increasing and every* $h \in H$ *can be approximated by a sequence* $\{h_n\}$ *with* $h_n \in H_n$ *for each* n). *Let*

$$\psi_n \equiv \text{argmin}_{\phi \in H_n} ||\phi_0 - \phi||^2.$$

Then $||\bar{\phi_0} - \psi_n||^2 \to 0$ *as* $n \to \infty$. *Furthermore, if*

$$w_n \equiv \phi_0 - \psi_n, \tag{2.56}$$

then $w_n \to \tilde{\phi_0}$ *in* $L^2(S, d\lambda)$ *as* $n \to \infty$.

Proof. We can find an orthonormal basis $\{\phi_i\}_{i \geq 1}$ in H, such that $\{\phi_1, \ldots, \phi_n\}$ is an orthonormal basis for H_n. Then

$$\bar{\phi_0} = \sum_{i=1}^{\infty} (\phi_0, \phi_i)\phi_i, \quad \text{and} \quad \psi_n = \sum_{i=1}^{n} (\phi_0, \phi_i)\phi_i.$$

Hence

$$||\bar{\phi_0} - \psi_n||^2 = \sum_{i=n+1}^{\infty} (\phi_0, \phi_i)^2 \to 0.$$

Let $w_n \equiv \phi_0 - \psi_n$. Then

$$||w_n - \tilde{\phi_0}||^2 = ||(\phi_0 - \psi_n) - (\phi_0 - \bar{\phi_0})||^2 = ||\psi_n - \bar{\phi_0}||^2 \to 0.$$

□

Now the problem is to find the projections ψ_n. The way to find the projection is standard. Suppose that $\{\phi_1, \ldots, \phi_n\}$ is a basis for H_n (it need not be an orthonormal basis). Let

$$\psi_n = \sum_{i=1}^{n} a_i \phi_i$$

where $(a_1, a_2, \ldots, a_n)'$ is the unique solution x of the following *normal* equations:

$$Ax = b, \quad \text{where} \tag{2.57}$$

$$A = \begin{pmatrix} (\phi_1,\phi_1) & \dots & (\phi_1,\phi_n) \\ \vdots & \ddots & \vdots \\ (\phi_n,\phi_1) & \dots & (\phi_n,\phi_n) \end{pmatrix}, \quad \text{and} \quad \mathrm{b} = \begin{pmatrix} (\phi_0,\phi_1) \\ \vdots \\ (\phi_0,\phi_n) \end{pmatrix}. \tag{2.58}$$

Because A is positive definite, the normal equations do have a unique solution. Finally, we have

$$w_n = \phi_0 - \sum_{i=1}^{n} a_i \phi_i.$$

As pointed out in Serbin [44, 43], the normal matrix A in the normal equations (2.57) is generally ill conditioned. There are many alternatives for solving the normal equations. However, we have chosen to use Gram-Schmidt orthogonalization to find the projections ψ_n directly.

There are many ways to choose the approximating subspaces H_n, each of which yields a different version of the algorithm. We choose H_n as

$$H_n = \text{span of } \{\mathcal{A} f_{k,i} : \quad k = 1, 2, \dots, n; \quad i = 0, 1, \dots, k\}$$

where $f_{k,i} = x_1^i x_2^{k-i}$. The dimension of H_n is $\dfrac{(n+1)(n+2)}{2} - 1$.

Proposition 2.5 *If H_n is defined as above, then $H_n \uparrow H$. Let w_n be defined as in (2.56); then $\alpha_n \equiv (w_n, w_n) \neq 0$. Therefore we can define $p^n \equiv \frac{1}{\alpha_n} w_n$. Furthermore, if $p \in L^2$, then $p^n \to \frac{1}{\alpha}\tilde{\phi}_0$ in L^2 as $n \to \infty$.*

Proof. The proof of $H_n \uparrow H$ is an immediate consequence of Proposition 7.1 and Remark 6.2 in the appendices of Ethier and Kurtz [11]. Because $\phi_0 \notin H_n$ for each n, we know that $w_n \neq 0$ and $\alpha_n \equiv (w_n, w_n) \neq 0$. Hence we can define $p^n \equiv \frac{1}{\alpha_n} w_n$. If we assume $p \in L^2$, then $\alpha \neq 0$. Because $w_n \to \tilde{\phi}_0$ and $\alpha_n \to \alpha$, it is immediate that $p^n \to \frac{1}{\alpha}\tilde{\phi}_0$. □

Several considerations lie behind our choice of approximating subspaces. First, more complicated subspaces (e.g., those used in the finite element method) may lead to problems in higher dimensions. Second, when low order polynomials are substituted into (2.35), one obtains *exact* relations among some quantities associated with the stationary density p. These relations resemble *energy preserving* relations in the finite element method. We believe that this property will enhance the accuracy of our computational method. Finally as the following section will show, our choice seems to give reasonably good results.

Proposition 2.6 *Suppose that $p \in L^2$. Let p^n be defined as in Proposition 2.5. Let $m_1^{(n)} = \int_{\mathcal{O}} (x_1 \cdot p^n(x))\, dx$ and $m_2^{(n)} = \int_{\mathcal{O}} (x_2 \cdot p^n(x))\, dx$. Then $m_1^{(n)} \to m_1$ and $m_2^{(n)} \to m_2$ as $n \to \infty$.*

Proof.

$$\begin{aligned}\left|m_1^{(n)} - m_1\right| &\leq \int_{\mathcal{O}} x_1 \left|p^n(x) - w_0(x)\right| dx = \int_{\mathcal{O}} x_1 \left|p^n(x) - \frac{1}{\alpha}\tilde{\phi}_0(x)\right| dx \\ &\leq \left(\int_S x_1^2\, dx\right)^{\frac{1}{2}} \left(\int_S \left|p^n(x) - \frac{1}{\alpha}\tilde{\phi}_0(x)\right|^2 dx\right)^{\frac{1}{2}} \\ &\leq \left(\frac{a_2^3 a_4}{3}\right)^{\frac{1}{2}} \left\|p^n - \frac{1}{\alpha}\tilde{\phi}_0\right\|^2 \to 0,\end{aligned}$$

as $n \to \infty$, where the constants a_2 and a_4 are indicated in Figure 2.3. Similarly we have $m_2^{(n)} \to m_2$ as $n \to \infty$. #

Remark. If $p \notin L^2$ and Conjecture 2.1 holds, then $\tilde{\phi}_0$ in (2.55) must be zero in L^2. However, as stated in Proposition 2.5, each p^n is well defined. I conjecture that p^n still converges to p in a weak sense, and the weak convergence would imply $m_i^{(n)} \to m_i$ as $n \to \infty$ $(i = 1, 2)$.

2.5 Numerical Comparisons

2.5.1 Comparison with SC Solutions

In this section we consider a special case of the SRBM described in Section 2.1, comparing results obtained with our algorithm against a known analytic solution. The special case considered has $\mu = 0$ and $\Gamma = 2I$ (I is the 2×2 identity matrix), so our differential operator G is the ordinary Laplacian. Before going further, we introduce some additional notation. Let n_i be the unit normal vector on F_i, and θ_i be the angle between the vector v_i and the normal n_i, with θ_i being positive when v_i lies to the right of n_i as one traces the boundary counter clockwise, and non-positive otherwise $(i = 1, 2, 3, 4)$. Let $\beta_i = 2(\theta_{i+1} - \theta_i)/\pi$ $(i = 1, 2, 3, 4)$ with $\theta_5 \equiv \theta_1$. It can be shown that (2.4) is equivalent to $\beta_i > -1$ for all i. From the results in [23] it follows that p_0 is always square integrable in S with respect to Lebesgue measure, whereas p_i is square integrable on F_i with respect to one-dimensional Lebesgue measure if and only if

$$\beta_i > -\frac{1}{2}, \quad (i = 1, 2, 3, 4). \tag{2.59}$$

Hence we conclude that $p \in L^2(S, d\lambda)$ if and only if (2.59) is true.

In addition to the restrictions mentioned earlier, we assume that $\theta_1 = \pi/4$, $\theta_2 = 0$,

$\theta_3 = 0$ and $\theta_4 = -\pi/4$. The corresponding reflection matrix is

$$R = \begin{pmatrix} 1 & 0 & -1 & 1 \\ -1 & 1 & 0 & -1 \end{pmatrix}. \tag{2.60}$$

We fix the height of the rectangle at $b = 1$ and let the length a of the rectangle change freely. The SRBM in this case corresponds to the heavy traffic limit of two balanced finite queues in tandem. It is easy to calculate that $\beta_1 = -1/2$, $\beta_2 = 0$, $\beta_3 = -1/2$ and $\beta_4 = 1$. Therefore (2.59) is not satisfied, implying that $p \notin L^2(S, d\lambda)$. Readers will see that our algorithm gives very accurate approximations even in this case. This is consistent with our conjecture explained at the end of Section 2.4.

For various values of the length parameter a, Table 2.1 compares two different estimates of $m_1, m_2, \delta_1, \delta_2, \delta_3, \delta_4$. The QNET estimate is that obtained with our algorithm (see Section 2.4), using $n = 6$. The SC estimate is that obtained by Trefethen and Williams [52] using a software package called SCPACK; for the special case under discussion (the restriction to two dimensions and the assumption of zero drift are both essential), Harrison, Landau and Shepp [23] used complex variable methods to compute the stationary density function p_0 in terms of a certain Schwartz–Christoffel transformation, and then SCPACK allows numerical evaluation of these formulas. The SC estimates on our Table 2.1 are taken from Table 2 on page 244 of [52], and the rows labelled DIFF give the differences between those SC estimates and our QNET estimates. It should be mentioned that our algorithm also applies to problems with non–zero drift, and its basic logic extends readily to higher dimensions; neither of those statements is true of the methods used in [23] and [52]. Incidentally, the QNET estimates in Table 2.1 were obtained using $n = 6$ and double precision on a VAX machine; about 32 seconds of CPU time were required to generate all the numbers in the table.

2.5.2 Comparison with Exponential Solutions

In this section, we first derive a criterion for the stationary density p to be of exponential form. Under the criterion, the stationary density is of exponential form and all of the performance measures have explicit formulas. Therefore we can compare our QNET estimates with the ones calculated from those exponential densities. Recall from Section 2.5.1 the definition of the angle θ_i between v_i and the normal n_i on F_i $(i = 1, 2, 3, 4)$. Let $t_i = \tan(\theta_i)$

	a	m_1	m_2	δ_1	δ_2	δ_3	δ_4
QNET	0.5	0.258229	0.380822	1.848991	2.413695	2.413695	0.564704
SC		0.258585	0.380018	1.871418	2.412890	2.412890	0.541472
DIFF		−0.000356	0.000804	−0.022427	0.000805	0.000805	0.023232
QNET	1.0	0.551325	0.448675	0.805813	1.611625	1.611625	0.805813
SC		0.551506	0.448494	0.805295	1.610589	1.610589	0.805295
DIFF		−0.000181	0.000181	0.000518	0.000036	0.000036	0.000518
QNET	1.5	0.878800	0.471640	0.466710	1.340876	1.340876	0.874166
SC		0.879534	0.471624	0.446669	1.340225	1.340225	0.893557
DIFF		−0.000734	0.000016	0.020041	0.000651	0.000651	−0.019391
QNET	2.0	1.238442	0.483103	0.292077	1.206981	1.206981	0.914904
SC		1.239964	0.482830	0.270736	1.206445	1.206445	0.935709
DIFF		−0.001522	0.000273	0.021341	0.000536	0.000536	0.020805
QNET	2.5	1.625775	0.489845	0.188642	1.131142	1.131142	0.942499
SC		1.628342	0.489146	0.171214	1.130587	1.130587	0.959373
DIFF		−0.002567	0.000699	0.017428	0.000555	0.000555	−0.016874
QNET	3.0	2.036371	0.494084	0.122836	1.085136	1.085136	0.962300
SC		2.040075	0.492970	0.110891	1.084582	1.084582	0.973691
DIFF		−0.003704	0.001114	0.011945	0.000554	0.000554	0.003308
QNET	3.5	2.466108	0.496881	0.079113	1.056112	1.056112	0.976999
SC		2.471022	0.495381	0.072873	1.055585	1.055585	0.982712
DIFF		−0.004914	0.001500	0.006240	0.000527	0.000527	−0.005713
QNET	4.0	2.911243	0.498826	0.048974	1.037364	1.037364	0.988391
SC		2.917572	0.496936	0.048334	1.036868	1.036868	0.988534
DIFF		−0.006329	0.001890	0.000640	0.000496	0.000496	−0.000143

Table 2.1: Comparisons with SCPACK when $n = 6$

($i = 1, 2, 3, 4$). Then the reflection matrix is

$$R = \begin{pmatrix} 1 & t_2 & -1 & t_4 \\ t_1 & 1 & t_3 & -1 \end{pmatrix}. \tag{2.61}$$

Using a result in [27, Theorem 6.1] and following the derivation in [26, Section 9], we can get the following proposition.

Proposition 2.7 *The stationary density p_0 is of exponential form if and only if*

$$\begin{cases} t_1\Gamma_{11} + t_2\Gamma_{22} = 2\Gamma_{21}, \\ t_3 = -t_1,\ t_4 = -t_2. \end{cases} \tag{2.62}$$

In this case, the stationary density is an exponential function

$$x \to c \cdot \exp(\lambda \cdot x), \tag{2.63}$$

where

$$\lambda = \begin{pmatrix} \lambda_1 \\ \lambda_2 \end{pmatrix} \quad \text{with} \quad \lambda_1 = \frac{2(\mu_1 - t_2\mu_2)}{(1-t_1t_2)\Gamma_{11}} \quad \text{and} \quad \lambda_2 = \frac{2(\mu_2 - t_1\mu_1)}{(1-t_1t_2)\Gamma_{22}} \tag{2.64}$$

and c is a normalizing constant such that $\int_S p_0(x)\,dx = 1$.

Remark. The denominators in the expressions for λ_1 and λ_2 are not zero because $1 - t_1t_2 = (t_1^2\,\Gamma_{11} - 2t_1\,\Gamma_{12} + \Gamma_{22})/\Gamma_{22} > 0$ by the positive definiteness of Γ.

Proof. Define two matrices

$$N = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ \text{-1} & 0 \\ 0 & \text{-1} \end{pmatrix} \quad \text{and} \quad Q = \begin{pmatrix} 0 & t_1 \\ t_2 & 0 \\ 0 & t_3 \\ t_4 & 0 \end{pmatrix}.$$

The i-th column of N' and the i-th column of Q' are the unit normal and the tangential component of v_i, respectively, on F_i $(i = 1, 2, 3, 4)$, that is,

$$R = N' + Q'.$$

Let U be the rotation matrix whose rows are the orthonormal eigenvectors of the covariance matrix Γ and let A be the corresponding diagonal matrix of eigenvalues such that $\Gamma = U'AU$, where $U' = U^{-1}$. Let $V = A^{-1/2}U$ and define $\tilde{Z} = VZ$, $\tilde{X} = VX$ and $\tilde{R} = VR$. Then $\tilde{X}$ is a $(I, V\mu)$-Brownian motion. The i-th row of the matrix

$$\tilde{N} \equiv \Lambda^{-1/2} N U' A^{1/2} \tag{2.65}$$

is the inward unit normal to the face of the state space of $\tilde{Z}$ on which L_i increases, where

$$\Lambda \equiv \begin{pmatrix} \Gamma_{11} & 0 & 0 & 0 \\ 0 & \Gamma_{22} & 0 & 0 \\ 0 & 0 & \Gamma_{11} & 0 \\ 0 & 0 & 0 & \Gamma_{22} \end{pmatrix} \equiv \begin{pmatrix} \Lambda_1 & 0 \\ 0 & \Lambda_1 \end{pmatrix}.$$

Similar to the derivation in [26, Theorem 9.23], we have $\tilde{R} = (\tilde{N}' + \tilde{Q}')\Lambda^{-1/2}$, where

$$\tilde{Q}' = A^{-1/2}U(N' + Q')\Lambda^{1/2} - A^{1/2}UN'\Lambda^{-1/2}. \tag{2.66}$$

In a similar manner to that in the proof of Theorem 8.2 of [28], it follows from [27, 58] that $\tilde{Z}$ has an exponential form stationary distribution if and only if the following skew symmetric condition holds

$$\tilde{N}\tilde{Q}' + \tilde{Q}\tilde{N}' = 0. \tag{2.67}$$

In this case $\tilde{Z}$ has a stationary distribution $\tilde{\pi}$ such that

$$d\tilde{\pi} = c\exp(\tilde{\lambda}\cdot\tilde{z})d\tilde{z},$$

where $\tilde{\lambda} = 2(I - \bar{N}^{-1}\bar{Q})^{-1}V\mu$ and $\bar{N}$ is any nonsingular 2×2 submatrix of $\tilde{N}$ obtained by deleting 2 rows from $\tilde{N}$ and $\bar{Q}$ is the submatrix obtained by deleting the corresponding two rows from $\tilde{Q}$. A simple algebraic manipulation gives that (2.67) is equivalent to

$$2N\Gamma N' = N(N' + Q')\Lambda + \Lambda(N + Q)N'. \tag{2.68}$$

Writing $N' = (I, -I)$ and $R = (R_1, R_2)$, where R_1 is a 2×2 matrix consisting of the first two columns of R and R_2 is defined similarly, we have that the left hand side of (2.68) is equal to

$$\begin{pmatrix} 2\Gamma & -2\Gamma \\ -2\Gamma & 2\Gamma \end{pmatrix},$$

and the right hand side of (2.68) is equal to

$$\begin{pmatrix} R_1\Lambda_1 + \Lambda_1 R_1' & R_2\Lambda_1 - \Lambda_1 R_1' \\ -R_1\Lambda_1 + \Lambda_1 R_2' & -R_2\Lambda_1 - \Lambda_1 R_2' \end{pmatrix}.$$

Therefore we have that (2.68) is equivalent to

$$2\Gamma = R_1\Lambda_1 + \Lambda_1 R_1' \tag{2.69}$$

$$-2\Gamma = R_2\Lambda_1 - \Lambda_1 R_1' \tag{2.70}$$

$$-2\Gamma = -R_1\Lambda_1 + \Lambda_1 R_2' \tag{2.71}$$

$$2\Gamma = -R_2\Lambda_1 - \Lambda_1 R_2'. \tag{2.72}$$

It is easy to check that (2.69) is equivalent to

$$t_1\Gamma_{11} + t_2\Gamma_{22} = 2\Gamma_{21}. \tag{2.73}$$

Equations (2.70) and (2.71) are redundant and they are equivalent to

$$t_4\Gamma_{22} - t_1\Gamma_{11} = -2\Gamma_{21} \tag{2.74}$$

$$t_3\Gamma_{11} - t_2\Gamma_{22} = -2\Gamma_{21}. \tag{2.75}$$

From equations (2.73)–(2.75), we see that they are equivalent to (2.62). Notice that (2.72), being equivalent to $t_3\Gamma_{11} + t_4\Gamma_{22} = -2\Gamma_{21}$, is satisfied when (2.62) holds.

When (2.62) holds, by taking

$$N_1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \text{and} \quad Q_1 = \begin{pmatrix} 0 & t_1 \\ t_2 & 0 \end{pmatrix},$$

we can get a particular pair of $\bar{N}, \bar{Q}$ by replacing N and Q in (2.65) (2.66) by N_1 and Q_1 respectively, i.e.,

$$\bar{N} = \Lambda_1^{1/2} U' A^{1/2}, \quad \bar{Q} = \Lambda_1^{1/2} R_1' U' A^{-1/2} - \Lambda_1^{-1/2} U' A^{1/2}.$$

Hence

$$\begin{aligned} \tilde{\lambda} &= 2\left(I - \bar{N}^{-1}\bar{Q}\right)^{-1} V\mu \\ &= 2\left(2I - A^{-1/2} U \Lambda_1 R_1' U' A^{-1/2}\right)^{-1} V\mu \\ &= 2\left(2I - V\Lambda_1 R_1' V'\right)^{-1} V\mu \\ &= 2\left(2I - V(2\Gamma - R_1\Lambda_1)V'\right)^{-1} V\mu \\ &= 2\left(2I - 2V\Gamma V' + VR_1\Lambda_1 V'\right)^{-1} V\mu \\ &= 2V'^{-1}\Lambda_1^{-1} R_1^{-1}\mu, \end{aligned}$$

where the fourth equality is obtained from (2.69). It then follows that Z has a product form stationary distribution with density given by

$$z \to c\exp(\lambda \cdot z),$$

where

$$\lambda = V'\tilde{\lambda} = 2\Lambda_1^{-1} R_1^{-1}\mu = \left(\frac{2(\mu_1 - t_2\mu_2)}{(1 - t_1t_2)\Gamma_{11}}, \frac{2(\mu_2 - t_1\mu_1)}{(1 - t_1t_2)\Gamma_{22}}\right)'. \tag{2.76}$$

This proves the proposition. □

Let c_1 and c_2 satisfy

$$c_1 \int_0^a e^{\lambda_1 x_1} dx_1 = 1 \quad \text{and} \quad c_2 \int_0^b e^{\lambda_2 x_2} dx_2 = 1,$$

where a is the length of the rectangle and b is the height of the rectangle. Then c_1c_2 is the normalizing constant for the density p_0 and

$$m_1 = c_1 \int_0^a x_1 e^{\lambda_1 x_1} dx_1, \quad \delta_1 = c_1, \quad \delta_3 = c_1 e^{\lambda_1 a}, \tag{2.77}$$

$$m_2 = c_2 \int_0^b x_2 e^{\lambda_2 x_2} dx_2, \quad \delta_2 = c_2, \quad \delta_4 = c_2 e^{\lambda_2 b}. \tag{2.78}$$

n	μ_1	μ_2	m_1–error	m_2–error	m_1	m_2
6	0.0	0.0	$6.938894e-18$	$6.938894e-18$	0.500000	0.500000
6	0.5	0.5	$-5.495242e-09$	$-5.495242e-09$	0.581977	0.581977
6	-0.5	-0.5	$5.495242e-09$	$5.495242e-09$	0.418023	0.418023
6	-0.5	0.0	$-1.988788e-09$	$5.551115e-17$	0.418023	0.500000
8	1.0	1.0	$-3.648693e-07$	$-3.648693e-07$	0.656518	0.656518
6	-1.0	1.0	$3.648693e-07$	$-3.648693e-07$	0.343482	0.656518
6	2.0	2.0	$1.345184e-04$	$1.345184e-04$	0.768523	0.768657
6	2.0	-2.0	$1.345184e-04$	$-1.345184e-04$	0.768523	0.231477
6	0.0	-2.0	$-5.551115e-16$	$-7.098788e-05$	0.500000	0.231414
7	3.0	-3.0	$1.821861e-04$	$-1.821861e-04$	0.835636	0.164364
8	4.0	-4.0	$2.562974e-04$	$-2.562999e-04$	0.875079	0.124921
9	5.0	-5.0	$2.986824e-04$	$-2.983432e-04$	0.899747	0.100253

Table 2.2: Comparisons with Exponential Solutions when $\Gamma = I$ and $t_3 = 0.0$

In this section, we choose S to be the unit rectangle, i.e., both the length a and height b of the rectangle are equal to one. Tables 2.2 through 2.5 give estimates of m_1 and m_2 computed with our algorithm for various test problems having exponential stationary distributions. All four of these test problems have $\Gamma = I$, and each corresponds to a different choice of t_2; in order to assure an exponential stationary distribution, (2.62) shows that one must choose $t_1 = -t_3$, $t_2 = t_3$ and $t_4 = -t_3$. The rows of Tables 2.2 through 2.5 correspond to different choices of the drift vector μ, and the columns labelled m_1–error and m_2–error give differences between estimates computed with our algorithm and the exact values derived from (2.77) and (2.78). Note that our algorithm converges more slowly (that is, larger n values are needed to get comparable accuracy) as the absolute magnitude of the drift vector increases.

Table 2.6 gives similar computational results for test problems whose covariance matrix is not diagonal. The first four columns show the t_3 values chosen and the covariance matrix chosen; the values of t_1, t_2 and t_4 are then uniquely determined by formula (2.62). The drift vector μ is $(1, -1)'$ in all cases and our algorithm was run with $n = 6$ for all cases.

2.5.3 A Tandem Queue with Finite Buffers

Consider the simple queueing network pictured in Figure 2.2. The network consists of two single–server stations arranged in series, each with a first–in–first–out discipline; arriving customers go to station 1 first, after completing service there they go to station 2, and after

n	μ_1	μ_2	m_1–error	m_2–error	m_1	m_2
6	0.0	0.0	$1.387779e-16$	$4.163336e-17$	0.500000	0.500000
6	1.0	1.0	$-5.846718e-04$	$1.031703e-03$	0.566551	0.682070
6	1.0	−1.0	$1.031703e-03$	$5.846718e-04$	0.682070	0.433449
6	−1.0	−1.0	$5.846718e-04$	$-1.031703e-03$	0.433449	0.317930
6	−1.0	1.0	$-1.031703e-03$	$-5.846718e-04$	0.317930	0.566551
6	2.0	−2.0	$3.490021e-03$	$5.032398e-04$	0.796475	0.371526
6	−2.0	−2.0	$5.032398e-04$	$-3.490021e-03$	0.371526	0.203525
6	2.0	2.0	$-5.032398e-04$	$3.490021e-03$	0.628474	0.796475
6	3.0	3.0	$9.441234e-03$	$7.367338e-03$	0.673661	0.854491
7	−3.0	3.0	$-3.485767e-03$	$3.424234e-03$	0.141628	0.679678
8	3.0	−3.0	$2.171832e-03$	$-2.469559e-03$	0.859686	0.319367

Table 2.3: Comparisons with Exponential Solutions when $\Gamma = I$ and $t_3 = 0.5$

n	μ_1	μ_2	m_1–error	m_2–error	m_1	m_2
6	1.0	−1.0	$2.612041e-03$	$1.767830e-03$	0.653906	0.498232
6	2.0	−2.0	$8.701640e-03$	$4.264274e-03$	0.759956	0.495736
7	3.0	−3.0	$7.411813e-03$	$1.042879e-03$	0.828406	0.498957
7	−3.0	3.0	$-7.411813e-03$	$-1.042879e-03$	0.171594	0.501043
8	4.0	−4.0	$6.360089e-03$	$8.693270e-05$	0.868975	0.499913

Table 2.4: Comparisons with Exponential Solutions when $\Gamma = I$ and $t_3 = 1.0$

n	μ_1	μ_2	m_1–error	m_2–error	m_1	m_2
7	1.0	−1.0	$1.972118e-03$	$7.033465e-04$	0.621445	0.524897
7	2.0	−2.0	$6.229480e-03$	$2.065347e-03$	0.717099	0.548896
8	3.0	−3.0	$6.109330e-03$	$2.797207e-03$	0.787221	0.573055
8	4.0	4.0	$-4.322155e-03$	$1.115873e-02$	0.404257	0.828471
9	4.0	4.0	$-3.768640e-03$	$6.286212e-03$	0.403704	0.833344

Table 2.5: Comparisons with Exponential Solutions when $\Gamma = I$ and $t_3 = 1.5$

t_3	Γ_{11}	Γ_{12}	Γ_{22}	m_1–error	m_2–error	m_1	m_2
0.0	1.0	-0.5	1.0	$4.829933e-07$	$2.622528e-03$	0.656517	0.497377
1.0	1.0	-0.5	1.0	$1.370037e-04$	$2.637569e-03$	0.607815	0.552511
-1.0	1.0	-0.5	1.0	$-2.622528e-03$	$-4.829933e-07$	0.502623	0.343483
-1.0	3.0	-1.0	2.0	$-2.083261e-03$	$-6.640263e-03$	0.502083	0.424664
-1.0	0.4	-0.1	0.75	$-6.254477e-03$	$7.554294e-03$	0.506254	0.292774
-1.0	0.85	-0.1	2.0	$-5.007373e-03$	$6.429745e-03$	0.505007	0.411594

Table 2.6: Comparisons with Other Exponential Solutions when $n = 6$

completing service at station 2 they exit the system. The input process to station 1 is a Poisson process with average arrival rate λ. Service times at station 1 are deterministic of duration $\tau_1 = 1$, and service times at station 2 are exponentially distributed with mean $\tau_2 = 1$. There is a storage buffer in front of station k that can hold $b_k = 24$ waiting customers ($k = 1, 2$), in addition to the customer occupying the service station. When the buffer in front of station 1 is full, the Poisson input process is simply turned off, and in similar fashion, server 1 stops working when the buffer in front of station 2 is full, although a customer may still occupy station 1 when the server is idle because of such blocking. (In the literature of queueing theory, this is called "communications blocking".) The steady–state performance measures on which we focus are

$$\begin{aligned} \gamma &= \text{the long-run average throughput rate, and} \\ q_k &= \text{the long-run average queue length at station } k \ (k = 1, 2). \end{aligned}$$

In these definitions "queue length" means the number of customers at the station, either waiting or being served, and the "average throughput rate" may be equivalently viewed as (a) the average rate at which new arrivals are accepted into the system, or as (b) the average rate at which services are completed at the first station, or as (c) the average rate

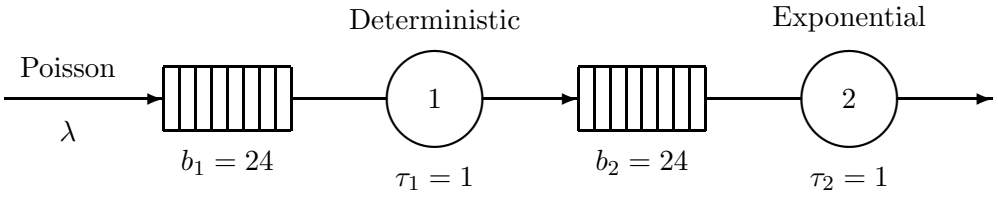

Figure 2.2: Finite Queues in Tandem

at which customers depart from the system.

Despite its apparent simplicity, the tandem queue described above is not amenable to exact mathematical analysis, but as an alternative to simulation one may construct and analyze what we call an "approximate Brownian system model". This is a diffusion approximation of the general type suggested by "heavy traffic limit theorems" for queueing networks (see below). However, no limit theorem to justify our particular approximation has been proved thus far, and we will not try to provide such a formal justification here. As described in Section 7 of [9], the two-dimensional queue length process associated with the tandem queue is represented in our approximate model by the (S, Γ, μ, R) SRBM Z, where $\Gamma = \gamma I$, $\mu = (\lambda - 1, 0)'$ and the reflection matrix is given by

$$R = \begin{pmatrix} 1 & 0 & -1 & 1 \\ -1 & 1 & 0 & -1 \end{pmatrix},$$

whose columns (directions of reflection) are portrayed in Figure 2.3 below. Readers who are interested in knowing the motivation of the approximate model in direct, intuitive terms, and how the data (S, Γ, μ, R) are derived can consult [9] for details.

Again it can be checked as in Section 2.5.1 that the stationary density p is not square integrable on ∂S. Nevertheless, the stationary density of Z can be computed using the algorithm developed in Section 2.4. Then one can derive from it approximate values for γ, m_1 and m_2. Additional complication arises from the appearance of the unknown γ in the covariance matrix Γ. An iterative procedure to compute γ was proposed in [9]. The results of that analysis are summarized in Table 2.5.3 below, where we give performance

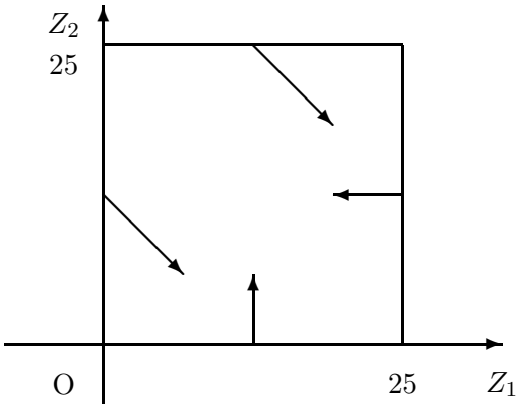

Figure 2.3: State Space and Directions of Reflection for the Approximating SRBM

	$\lambda = 0.9$			$\lambda = 1.0$		
	γ	m_1	m_2	γ	m_1	m_2
SIM	0.8991	5.1291	6.2691	0.9690	13.87	11.07
QNET	0.8995	4.8490	6.3184	0.9688	13.75	11.25
	$\lambda = 1.1$			$\lambda = 1.2$		
	γ	m_1	m_2	γ	m_1	m_2
SIM	0.9801	20.4801	12.3801	0.9804	22.4804	12.4804
QNET	0.9801	20.5239	12.4445	0.9807	22.2688	12.4676

Table 2.7: Performance Estimates for the Queueing Model Pictured in Figure 2.2

estimates derived from the approximate Brownian model, identified in the table as QNET estimates, as well as estimates obtained via simulation. None of the QNET estimates of average queue length differs from the corresponding simulation estimate by more than five percent, and the accuracy of our throughput rate estimates is equally impressive: when $\lambda = 0.9$ both simulation and QNET predict a throughput loss rate below one–tenth of one percent; when $\lambda = 1.0$, simulation and QNET predict throughput loss rates of 3.10% and 3.14%, respectively; when $\lambda = 1.1$, the limiting factor on system throughput is the average service rate of 1.0, and both simulation and QNET predict a throughput rate 1.99% below this maximum; when $\lambda = 1.2$, the maximum possible throughput rate is again 1.0, and the simulation and QNET estimates of γ are 1.96% and 1.93% below this maximum, respectively.

Readers should be warned that there are two stages of error incurred in estimating the performance measures of the queueing networks. The first stage of error occur when we replace a queueing network model by a corresponding Brownian model. This replacement has been justified only for certain classes of queueing networks under conditions of heavy traffic. The second stage of error comes from our numerical computation of the stationary density of the corresponding SRBM. Presumably, error at the second stage is smaller than that at the first stage. Our analysis of tandem queues in series is just an illustration of the usefulness of a Brownian system approximation. It can be extended to allow an arbitrary renewal input process and arbitrary service time distributions (the first two moments of the interarrival and service time distributions determine the drift vector and covariance matrix of the corresponding Brownian system model), and to general open and closed queueing networks with single customer type as well as multiple customer types. Table 2.7 shows in a concrete way how useful Brownian system models can be, *if* one can compute their

stationary distributions.

2.6 Concluding Remarks

Let us return to the setting of Section 2.4, where the problem of computing the stationary density p was cast as a least squares problem. The treatment given there can be generalized in the following way, which may be important for both practical and theoretical purposes. (In this section the letter q will be re–used with a new meaning, but that should cause no confusion.) Let q_0 be a strictly positive function on the interior S of the rectangle S, and let $q_1, \ldots, q_4$ be strictly positive functions on the boundary surfaces $F_1, \ldots, F_4$ respectively. Defining

$$q(x) = (q_0; q_1, q_2, q_3, q_4). \tag{2.79}$$

we call q a "reference density" and we define a corresponding "reference measure" η via

$$\eta(dx) \equiv q d\lambda = (q_0\, dx; q_1\, d\sigma_1, \ldots, q_4\, d\sigma_4). \tag{2.80}$$

If we work in the Hilbert space $L^2(S, \eta)$ rather than the space $L^2(S, d\lambda)$ used in Section 2.4, then the focus is on the unknown function r defined by

$$r(x) \equiv p/q = (p_0/q_0; p_1/q_1, \cdots, p_4/q_4). \tag{2.81}$$

That is, with the inner product defined by $(f, g) = \int_S (f \cdot g)\, d\eta$, our basic adjoint relationship (2.35) says that $\mathcal{A}f \perp r$ for all $f \in C^2(S)$, and hence one may proceed exactly as in Sections 2.4 to devise an algorithm for approximate computation of r by projection in $L^2(S, \eta)$. Of course, the final estimate of r is converted to an estimate of p via $p = rq$, where q is the reference density chosen.

A different computational procedure is obtained depending on how one chooses the reference density q and the functions $f_1, f_2, \ldots$ that are used to build up the approximating subspaces $H_1, H_2, \ldots$ via $H_n = \text{span}\{\mathcal{A}f_1, \ldots, \mathcal{A}f_n\}$. Recall that in Section 2.4 we took $f_1, f_2, \ldots$ to be polynomial functions, but other choices are obviously possible. One wants to choose q and $f_1, f_2, \ldots$ in such a way that the inner products $(\mathcal{A}f_m, \mathcal{A}f_n)$ can be determined analytically, and in such a way as to accelerate convergence of the algorithm. From a theoretical standpoint, the freedom to choose q is important because one may have $r \in L^2(S, \eta)$ even though $p \notin L^2(S, d\lambda)$ (e.g. by choosing $q = p$, we have $r = 1 \in L^2(S, \eta)$), and thus a judicious choice of reference density enables a rigorous proof of convergence in

$L^2(S, \eta)$. From a practical standpoint, one may be able to choose q in such a way that convergence is accelerated, taking q to be a "best guess" of the unknown density p based on either theory or prior computations. In Chapter 4, we will discuss computation of stationary distributions on *unbounded* regions, where a proper choice of reference density is essential to efficient computation.

Chapter 3

SRBM in an Orthant

3.1 Introduction and Definitions

Notation. Let $d \geq 1$ be an integer, and $S \equiv \mathbb{R}^d_+$ be the orthant in a d-dimensional Euclidean space $\mathbb{R}^d$. For $i = 1, 2, \ldots, d$, let $F_i \equiv \{x \in S : x_i = 0\}$ be the i-th face of ∂S, and v_i be a vector on F_i with unit normal component, pointing into S (see Figure 3.1 when $d = 2$). Let $R \equiv (v_1, v_2, \ldots, v_d)$ be a $d \times d$ matrix, Γ be a $d \times d$ positive definite matrix and μ a d-dimensional vector. As before, the continuous sample path space C_S is defined as

$$C_S = \{\omega : [0, \infty) \to S, \quad \omega \text{ is continuous}\},$$

with natural filtration $\{\mathcal{M}_t\}$ and natural σ-algebra $\mathcal{M}$. The canonical process on C_S is

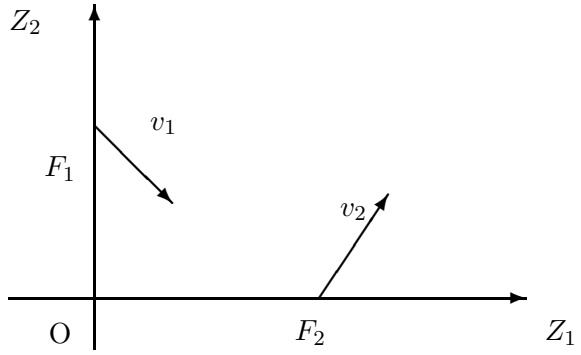

Figure 3.1: State Space and Directions of Reflection of an SRBM when $d = 2$

denoted by $Z = \{Z(t,\omega), t \geq 0\}$ defined by

$$Z(t,\omega) = \omega(t).$$

The symbol ω is often suppressed in Z. The SRBM in an orthant S is defined as follows.

Definition 3.1 *Z is said to be a semimartingale reflected Brownian motion (abbreviated as SRBM) associated with data (S, Γ, μ, R) if there is a family of probability measures $\{P_x, x \in S\}$ defined on the filtered probability space $(C_S, \mathcal{M}, \{\mathcal{M}_t\})$ such that for each $x \in S$ we have*

(3.1) *$Z(t) = X(t) + RL(t) = X(t) + \sum_{i=1}^{d} L_i(t) \cdot v_i \ \forall t \geq 0$, P_x-a.s.,*

(3.2) *$X(0) = x$, P_x-a.s. and X is a d-dimensional Brownian motion with covariance matrix Γ and drift vector μ such that $\{X(t) - \mu t, \mathcal{M}_t, t \geq 0\}$ is a martingale under P_x,*

(3.3) *L is a continuous $\{\mathcal{M}_t\}$-adapted d-dimensional process such that $L(0) = 0$ P_x-a.s, L is non-decreasing, and L_i increases only at times t when $Z_i(t) = 0$, $i = 1, \ldots, d$.*

Remark. This is the definition of SRBM used by Reiman and Williams [41]. It was pointed out by those authors that $\{X(t) - \mu t\}$ being an $\{\mathcal{M}_t\}$-martingale is necessary for an SRBM to have certain desired properties.

The SRBM Z defined above behaves like a d-dimensional Brownian motion with drift vector μ and covariance matrix Γ in the interior of its state space. When the boundary face F_i is hit, the process L_i (sometimes called the local time of Z on F_i) increases, causing an instantaneous displacement of Z in the direction given by v_i; the magnitude of the displacement is the minimal amount required to keep Z always inside S. Therefore, we call Γ, μ and R the covariance matrix, the drift vector and the reflection matrix of Z, respectively.

Definition 3.2 *The SRBM is said to be unique if the family $\{P_x, x \in S\}$ is unique.*

Definition 3.3 *The matrix R is said to be Minkowski if $I - R \geq 0$ and $I - R$ is transient, that is, all the eigenvalues of $I - R$ are less than one, where I is the $d \times d$ identity matrix.*

Definition 3.4 *A $d \times d$ matrix A is said to be an $\mathcal{S}$ matrix if there exists a d-dimensional vector $u \geq 0$ such that $Au > 0$, and to be a completely-$\mathcal{S}$ matrix if each of its principal submatrices is an $\mathcal{S}$ matrix.*

Harrison and Reiman [25], by using a unique path-to-path mapping, defined an SRBM in S when the *reflection* matrix R is Minkowski. This class of SRBM's arise naturally from queueing networks with homogeneous customers. It is known [24] that the reflection matrix R of an SRBM which arises from a queueing network with heterogeneous customers is, in general, not Minkowski. Reiman and Williams [41] proved that R being *completely-$\mathcal{S}$* is a necessary condition for the existence of an SRBM in S.

Recently Taylor and Williams [50], by considering solutions of *local* submartingale problems, proved the existence and uniqueness of an SRBM when R is a completely-$\mathcal{S}$ matrix. We state their result in the following proposition.

Proposition 3.1 *Assume R is a completely-$\mathcal{S}$ matrix. For any positive definite matrix Γ and vector μ, there exists a unique family of probability measures $\{P_x, x \in S\}$ on the filtered space $(C_S, \mathcal{M}, \{\mathcal{M}_t\})$ such that Z together with $\{P_x, x \in S\}$ is an (S, Γ, μ, R)-SRBM.*

In this chapter, E_x will denote the expectation operator with respect to the probability measure P_x, and for a probability measure π on S, define

$$P_\pi(\cdot) \equiv \int_S P_x(\cdot)\, \pi(dx);$$

then E_π denotes the corresponding expectation.

Readers should keep in mind that there are RBM's which are not SRBM's, see Harrison, Landau and Shepp [23] and Varadhan and Williams [53]. However, in this dissertation we *only* consider the class of RBM's which have the semimartingale representation (3.1) as defined in Definition 3.1 above.

In this chapter we first prove the Feller continuity of an SRBM, and therefore prove that an SRBM is a strong Markov process. It is reported that Taylor and Williams [50] give a different proof of the Feller continuity of an SRBM. Then we give an alternative characterization of an SRBM via a solution to a *constrained martingale problem*. This alternative characterization is critical in proving sufficiency of a basic adjoint relationship governing the stationary distribution of an SRBM, which was conjectured by Harrison and Williams [26] when R is Minkowski.

3.2 Feller Continuity and Strong Markov Property

Definition 3.5 *The Z with $\{P_x, x \in S\}$ is said to be Feller continuous if for any $f \in C_b(S)$, $T_t f \in C_b(S)$ for each $t \geq 0$. Here $T_t f(x) \equiv E_x f(Z(t))$.*

Theorem 3.1 *Let Z with $\{P_x, x \in S\}$ be an SRBM. Suppose $\{x_n\}$ is a sequence in S which converges to $x \in S$. Then $\{P_{x_n}\}$ converges weakly to P_x.*

Proof. In this proof only, we need a bigger probability space. Let $C^+_{\mathbb{R}^d}$ denote the space of continuous functions $x(\cdot) : [0,\infty) \to \mathbb{R}^d$ with $x(0) \geq 0$ and Λ_d denote the space of continuous functions $l(\cdot) : [0,\infty) \to \mathbb{R}^d_+$ such that $l(0) = 0$ and each component of $l(\cdot)$ is a non-decreasing function. Both $C_{\mathbb{R}^d_+}$ and Λ_d are endowed with the Skorohod topology. For $z(\cdot) \in C_S$, $x(\cdot) \in C^+_{\mathbb{R}^d}$ and $l(\cdot) \in \Lambda_d$, define $\omega(t) = (z(t), x(t), l(t))$ for each $t \geq 0$. Then ω is a generic element of $\Omega_S \equiv C_S \times C^+_{\mathbb{R}^d} \times \Lambda_d$. Define three *canonical processes* Z, X and L via $Z(t,\omega) = z(t)$, $X(t,\omega) = x(t)$ and $L(t,\omega) = l(t)$, and filtration $\{\mathcal{M}^0_t\}$ via $\mathcal{M}^0_t \equiv \sigma\{Z(s), X(s), L(s) : 0 \leq s \leq t\}$, $t \geq 0$. It is obvious that the family $\{P_x, x \in S\}$ on C_S induces a family of probability measures $\{Q_x, x \in S\}$ on the *sample space* $(\Omega_S, \{\mathcal{M}^0_t\})$, such that, for each $x \in S$, the following holds.

(3.4) For each $t \geq 0$, $Z(t) = X(t) + RL(t)$, Q_x-a.s.

(3.5) Under Q_x, X is a (Γ, μ)-Brownian motion starting from x and $\{X(t) - \mu t\}$ is a $\{\mathcal{M}^0_t\}$-martingale.

(3.6) L_i increases only when at times $Z(\cdot) \in F_i$ almost surely in Q_x, i.e.,

$$\int_0^\infty Z_i(t)\, dL_i(t) = 0, \quad Q_x\text{-a.s.} \quad (i = 1, 2, \ldots, d).$$

Let x_n be a sequence in S such that $x_n \to x$, obviously $Q_{x_n}X^{-1} \Rightarrow Q_xX^{-1}$, therefore $\{Q_{x_n}X^{-1}\}$ is tight. Hence for any $\epsilon > 0$, there exists a compact set $A \subset C^+_{\mathbb{R}^d}$ such that

$$Q_{x_n}X^{-1}(A) > 1 - \epsilon, \quad \text{for all } n.$$

Let $\tilde{A} \subset \Omega_S$ be defined as $\omega = (z, x, l) \in \tilde{A}$ if and only if $x \in A$ and

$$z = x + Rl, \qquad \int_0^\infty z(t)\, dl(t) = 0, \tag{3.7}$$

where

$$\int_0^t z(s)\, dl(s) \equiv \left(\int_0^t z_1(s)\, dl_1(s), \ldots, \int_0^t z_d(s)\, dl_d(s) \right)', \quad t \geq 0.$$

It follows from Proposition 1 of [2] that $\tilde{A}$ is precompact. Because

$$Q_{x_n}(\tilde{A}) = Q_{x_n}X^{-1}(A) > 1 - \epsilon,$$

we have proved that the family $\{Q_{x_n}\}$ is tight, and Prohorov's Theorem, see [3, Theorem 6.1 of Chapter 1], asserts that $\{Q_{x_n}\}$ is weakly relatively compact. Let Q_* be any accumulation point of $\{Q_{x_n}\}$. There is a subsequence of Q_{x_n} that converges to Q_* weakly. For notational convenience, we assume the sequence itself converges, that is, $Q_{x_n} \Rightarrow Q_*$. We are going to prove that Z is an SRBM starting from x under Q_*, i.e.,

$$Q_* Z^{-1} = P_x, \tag{3.8}$$

Then it follows from (3.8) that

$$P_{x_n} = Q_{x_n} Z^{-1} \Rightarrow Q_* Z^{-1} = P_x.$$

To prove (3.8), we prove that (3.4) through (3.6) hold under Q_*. It is clear that under Q_*, X is a (Γ, μ)-Brownian motion starting from x. If $X(t)$ were bounded for each t, because $\{X(t) - \mu t\}$ is a $\{\mathcal{M}_t^0\}$-martingale under each Q_{x_n} and $Q_{x_n} \Rightarrow Q_*$, $\{X(t) - \mu t\}$ is a $\{\mathcal{M}_t^0\}$-martingale under each Q_*. A general argument can be obtained through standard localizing arguments. Therefore (3.5) holds under Q_*. To show that (3.4) and (3.6) hold under Q_*, define two functions $\Omega_S \to C_{\mathbb{R}^d}$ as

$$\begin{aligned} f_1(\omega)(t) &\equiv z(t) - x(t) - l(t), \quad t \geq 0, \\ f_2(\omega)(t) &\equiv \int_0^t z(s)\, dl(s), \quad t \geq 0. \end{aligned} \tag{3.9}$$

It is obvious that f_1 is continuous and it follows from the following Lemma 3.1 that f_2 is continuous. Hence $A_i \equiv \{\omega : f_i(\omega) = 0\}$ is a closed set in Ω_S $(i = 1, 2)$. Because (3.4) and (3.6) hold under Q_{x_n}, $Q_{x_n}(A_i) = 1$ for each n $(i = 1, 2)$. Therefore, see [3, Theorem 2.1 of Chapter 1],

$$1 = \limsup_{n \to \infty} Q_{x_n}(A_i) \leq Q_*(A_i), \quad (i = 1, 2).$$

Thus $Q_*(A_i) = 1$ $(i = 1, 2)$, which implies that (3.4) and (3.6) hold under Q_*. Therefore we have proved Z under Q_* is an SRBM starting from x, and by the uniqueness of the SRBM, $P_x = Q_* Z^{-1}$. This proves (3.8) and thus proves the theorem. □

Lemma 3.1 *The function $f_2 : \Omega_S \to C_{\mathbb{R}^d}$ defined in (3.9) is continuous.*

Proof. Let $z_n \in C_S$ be a sequence converging to $z \in C_S$ and $l_n \in \Lambda_d$ be a sequence converging to $l \in \Lambda_d$. Fix a $T > 0$, we like to show

$$\sup_{0 \leq t \leq T} \left| \int_0^t z_n(s)\, dl_n(s) - \int_0^t z(s)\, dl(s) \right| \to 0,$$

as $n \to \infty$. To see this, for each positive integer k, define a step function $z_{(k)}$ as

$$z_{(k)}(t) \equiv z\left(iT/k\right), \quad \text{if } \frac{i}{k}T \le t < \frac{i+1}{k}T.$$

It is clear that $\sup_{0\le t<T} |z_{(k)}(t) - z(t)| \to 0$ as $k \to \infty$, and

$$\begin{aligned}
&\sup_{0\le t\le T}\left|\int_0^t z_n(s)\,dl_n(s) - \int_0^t z(s)\,dl(s)\right| \\
&\quad\le \sup_{0\le t\le T}\int_0^t |z_n(s)-z(s)|\,dl_n(s) + \sup_{0\le t\le T}\left|\int_0^t z(s)\,d\left(l_n(s)-l(s)\right)\right| \\
&\quad\le \sup_{0\le s\le T}|z_n(s)-z(s)|l_n(T) + \int_0^T |z(s)-z_{(k)}(s)|d\left(l_n(s)+l(s)\right) \\
&\qquad + \sup_{0\le t\le T}\left|\int_0^t z_{(k)}(s)\,d(l_n(s)-l(s))\right| \\
&\quad\le \sup_{0\le s\le T}|z_n(s)-z(s)|l_n(T) + \sup_{0\le s<T}|z(s)-z_{(k)}(s)|(l_n(T)+l(T)) \\
&\qquad + \sum_{i=1}^{k} z\left(iT/k\right)\left|\left(l_n\left((i+1)T/k\right) - l\left((i+1)T/k\right)\right) - \left(l_n\left(iT/k\right) - l\left(iT/k\right)\right)\right|.
\end{aligned}$$

Since $l_n(T)$ is bounded in n, for any $\epsilon > 0$, choose k big enough so that the middle term in the previous expression is less than $\epsilon/2$, then for n large enough we have the first term adding the third term is less than $\epsilon/2$. This proves our lemma. □

Corollary 3.1 *Z with $\{P_x, x \in S\}$ is Feller continuous, i.e., for any $f \in C_b(S)$ and any $t \ge 0$, $x \to E_x\left[f(Z(t))\right]$ is continuous.*

Proof. Obviously, for any $f \in C_b(S)$, $g : C_S \to \mathbb{R}$ defined by $g(z(\cdot)) \equiv f(z(t))$ is a bounded continuous function on C_S, hence, from Theorem 3.1 and the definition of weak convergence, $T_t f(x) \equiv E_x\left[f(Z(t))\right] = E_x\left[g(Z(\cdot))\right]$ is a continuous function of x. □

It is now easy to show that the probability family associated with an SRBM Z is Borel measurable, i.e., for any $A \in \mathcal{M}$, $x \to P_x(A)$ is a Borel measurable function on S. In fact, we are going to show

Corollary 3.2 *Suppose $h : C_S \to \mathbb{R}$ is bounded and $\mathcal{M}$-measurable. Then the function*

$$x \to E_x\left[h(Z(\cdot))\right], \quad \forall x \in S$$

is a Borel measurable function on S.

Proof. The corollary can be obtained from Theorem 3.1 directly. □

Now we have the following theorem, which ensures that an SRBM is a strong Markov process.

Theorem 3.2 *An SRBM Z together with a measurable family $\{P_x, \in S\}$ on a filtered space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\})$ is a strong Markov process, i.e., for each $x \in S$,*

$$E_x\left[f(Z(\tau+t))|\mathcal{F}_\tau\right] = T(t)f(Z(\tau)), \quad P_x\text{-a.s.}$$

for all $f \in B(S)$, $t \geq 0$, and P_x-a.s. finite $\{\mathcal{F}_t\}$-stopping time τ.

Proof. Since the SRBM Z with the family $\{P_x, x \in S\}$ is unique and Feller continuous, it follows from the proof of Theorem 4.2 in Chapter 4 of [11] that Z together with $\{P_x, x \in S\}$ is strong Markov. □

3.3 Constrained Martingale Problem

Using the ideas of Kurtz [32, 31], we now characterize an SRBM as a solution of a corresponding *constrained* martingale problem. This alternative characterization of an SRBM plays a key role in proving the sufficiency of a basic adjoint relationship for the stationary distribution in Section 3.5. For $f \in C^2(S)$, define

$$Gf \equiv \frac{1}{2}\sum_{i=1}^{d}\sum_{j=1}^{d}\Gamma_{ij}\frac{\partial^2 f}{\partial x_i \partial x_j} + \sum_{i=1}^{d}\mu_i\frac{\partial f}{\partial x_i}, \tag{3.10}$$

$$D_i f(x) \equiv v_i \cdot \nabla f(x) \text{ for } x \in F_i \quad (i = 1, 2, \ldots, d). \tag{3.11}$$

The operators G and D_i defined in (3.10) and (3.11), respectively, can be viewed as mappings from $C_K^2(\mathbb{R}^d) \to C_K(\mathbb{R}^d)$. Denote $D \equiv (D_1, \ldots, D_d)$. From now on, $C_K^2(\mathbb{R}^d)$ will be taken implicitly as the domain of (G, D). In the following, $\mathcal{P}(S)$ denotes the set of probability measures on S.

Definition 3.6 For any $\pi \in \mathcal{P}(S)$, by a *local time solution* of the constrained martingale problem for $(S, G, D; \pi)$ we mean a pair of d-dimensional continuous processes (Z, L) on some filtered probability space $(\Omega, \{\mathcal{F}_t\}, \mathcal{F}, P)$ such that

(3.12) $Z(t) \in S$ for all $t \geq 0$ and $PZ(0)^{-1} = \pi$.

(3.13) $f(Z(t)) - \int_0^t Gf(Z(s))\,ds - \sum_{i=1}^d \int_0^t D_i f(Z(s))\,dL_i(s)$
is a $\{\mathcal{F}_t\}$-martingale under P.

(3.14) P-almost surely, $L_i(0) = 0$, $L_i(\cdot)$ non-decrease, and $L_i(\cdot)$ increases only at times t when $Z(t) \in F_i$ $(i = 1, \ldots, d)$.

The following theorem gives the equivalence of an SRBM to a local time solution of the corresponding constrained martingale problem.

Theorem 3.3 *For $\pi \in \mathcal{P}(S)$, suppose Z together with $\{P_x, x \in S\}$ on $(C_S, \{\mathcal{M}_t\}, \mathcal{M})$ is an SRBM and L is the associated local time defined in (3.3). Then, under P_π, (Z, L) on $(C_S, \{\mathcal{M}_t\}, \mathcal{M})$ is a solution of the constrained martingale problem for $(S, G, D; \pi)$. Conversely, suppose (Z, L) on a filtered probability space $(\Omega, \{\mathcal{F}_t\}, \mathcal{F}, P)$ is a local time solution of the constrained martingale problem for $(S, G, D; \pi)$. Then Z is an SRBM starting with π, i.e.,*

(a) $Z(t) = X(t) + RL(t) = X(t) + \sum_{i=1}^d L_i(t) \cdot v_i \in S\ \forall t \geq 0$, P-*a.s., where*

(b) $PX(0)^{-1} = \pi$ *and X is a d-dimensional Brownian motion with covariance matrix Γ and drift vector μ such that $\{X(t) - \mu t, \mathcal{F}_t, t \geq 0\}$ is a martingale under P, and*

(c) *L is a continuous $\{\mathcal{F}_t\}$-adapted d-dimensional process such that P-a.s. $L(0) = 0$, L is non-decreasing, and L_i increases only at times t when $Z_i(t) = 0$, $i = 1, \ldots, d$.*

Proof. Suppose that Z together $\{P_x, x \in S\}$ on $(C_S, \{\mathcal{M}_t\}, \mathcal{M})$ is an SRBM, associated with X and L as in (3.2) and (3.3). Then, $Z(t) \in S$, and under P_π, it is obvious that $Z(0) = X(0)$ has the distribution π. For $f \in C_b^2(S)$, by Itô's formula, it is clear that (3.13) is a $\{\mathcal{M}_t\}$-martingale. The conditions on L in (c) is equivalent to the condition in (3.14). Therefore, (Z, L) is a local time solution of the constrained martingale problem for $(G, D; \pi)$.

Conversely, suppose (Z, L) on a filtered probability space $(\Omega, \{\mathcal{F}_t\}, \mathcal{F}, P)$ is a local time solution of the constrained martingale problem for $(G, D; \pi)$. Define

$$\xi(t) \equiv Z(t) - Z(0) - RL(t) - \mu t = Z(t) - Z(0) - \sum_{i=1}^d v_i L_i(t) - \mu t,$$

we first show that ξ is a $\{\mathcal{F}_t\}$-Brownian motion with covariance matrix Γ and zero drift. For each integer $n > 0$, let

$$\sigma_n = \inf\{t \geq 0 : |\xi(t)| > n\}.$$

Then σ_n is a stopping time. Taking an $f \in C_K^2(\mathbb{R}^d)$ such that $f(x) = x_i$ on $\{x : |x| \le n\}$, since (3.13) is a martingale for the f, by the optional sampling theorem,

$$f(Z(t \wedge \sigma_n)) - \int_0^{t\wedge\sigma_n} Gf(Z(s))\, ds - \sum_{i=1}^{d} \int_0^{t\wedge\sigma_n} D_i f(X(s))\, dL_i(s) \tag{3.15}$$

is a continuous $\{\mathcal{F}_t\}$-martingale. Because $f(x) = x_i$, $Gf(x) = \mu_i$, and $D_j f(x) = v_j^i$ for $|x| \le n$, we see from (3.15) that $\xi_i(t \wedge \sigma_n)$ is a martingale, $i = 1, \ldots, d$. Since $\sigma_n \uparrow \infty$ as $n \to \infty$, ξ_i is a continuous local martingale. Similarly, by choosing $f \in C_K^2(\mathbb{R}^d)$ such that $f(x) = x_i x_j$ for $x \in \{x : |x| \le n\}$, (3.13) gives that

$$\begin{aligned} Z_i(t \wedge \sigma_n) Z_j(t \wedge \sigma_n) \;&-\; \Gamma_{ij}(t \wedge \sigma_n) - \int_0^{t\wedge\sigma_n} (\mu_i Z_j(s) + \mu_j Z_i(s))\, ds \\ &-\; \sum_{k=1}^{d} \int_0^{t\wedge\sigma_n} \left(v_k^i Z_j + v_k^i Z_i \right) dL_k(s) \end{aligned} \tag{3.16}$$

is a martingale. On the other hand, by Itö's formula, we have

$$\begin{aligned} Z_i(t \wedge \sigma_n) Z_j(t \wedge \sigma_n) \;&=\; Z_i(0) Z_j(0) + \int_0^{t\wedge\sigma_n} Z_i\, d\xi_j(s) + \int_0^{t\wedge\sigma_n} Z_j(s)\, d\xi_i(s) \\ &+\; \int_0^{t\wedge\sigma_n} (\mu_i Z_j(s) + \mu_j Z_i(s))\, ds + \langle \xi_i, \xi_j \rangle (t \wedge \sigma_n) \\ &+\; \sum_{k=1}^{d} \int_0^{t\wedge\sigma_n} \left(v_k^i Z_j + v_k^i Z_i \right) dL_k(s), \end{aligned} \tag{3.17}$$

where $\langle \xi_i, \xi_j \rangle(t)$ is the *quadratic variational process* of ξ_i and ξ_j. The first two stochastic integrals on the right hand side of (3.17) are martingales. From this and from (3.16) and (3.17), we have

$$\langle \xi_i, \xi_j \rangle (t \wedge \sigma_n) - \Gamma_{ij}(t \wedge \sigma_n)$$

is a martingale. Therefore, we have the quadratic variational process of ξ_i and ξ_j

$$\langle \xi_i, \xi_j \rangle (t) = \Gamma_{ij} t.$$

By Theorem 7.1 of [29, Chapter 2], we can find a d-dimensional *standard* $\{\mathcal{F}_t\}-$Brownian motion $B = (B(t))$ on the *same* probability space $(\Omega, \mathcal{F}, P)$ such that

$$\xi_i(t) = \sum_{k=1}^{d} (\Gamma^{\frac{1}{2}})_{ik} B_k(t), \quad i = 1, \cdots, d,$$

where $\Gamma^{\frac{1}{2}}$ is the square root of the positive definite covariance matrix Γ. Therefore, ξ is a d-dimensional Brownian motion starting from zero with covariance matrix Γ and drift zero.

By letting $X(t) \equiv Z(0) + \xi(t) + \mu t$, then X is a Brownian motion with initial distribution $PX(0)^{-1} = \pi$, covariance matrix Γ, and drift vector μ. Clearly, under P, (Z, X, L) satisfies the equations (3.1) to (3.3) with $X(0) = x$ replaced by $PX(0)^{-1} = \pi$ in (3.2). This proves our theorem. □

3.4 Existence of a Stationary Distribution

From Theorem 3.2, an SRBM Z is a strong Markov process. We ask when its stationary distribution exists, and if there is one, how to characterize such a stationary distribution.

Definition 3.7 *A probability measure π on S is called a stationary distribution of the SRBM Z if for every bounded Borel function f on S and every $t > 0$*

$$\int_S T_t f(x)\, \pi(dx) = \int_S f(x)\, \pi(dx),$$

where $\{T_t, t \geq 0\}$ is the semigroup as defined in Definition 3.5 associated with the strong Markov process Z.

In this section, we first establish a criterion for the existence of a stationary distribution via Liapunov functions. Then we show when $R^{-1} \geq 0$ and R is symmetric, Z has a stationary distribution if and only if $R^{-1}\mu < 0$. This case was not covered by Harrison and Williams [26], who considered only the case of a Minkowski reflection matrix.

3.4.1 Necessary Conditions

Let σ_i denote $(d-1)$-dimensional Lebesgue measure (surface measure) on the face F_i. The following proposition was proved in [26] when the reflection matrix R is Minkowski, and the proof can be generalized to the case when R is completely-$\mathcal{S}$ by using the following lemma.

Lemma 3.2 *Suppose that Z is an SRBM with local time L as in (3.3). Then for each $t \geq 0$,*

$$\sup_{x \in S} E_x L_i(t) < \infty.$$

Proof. R is a completely-$\mathcal{S}$ matrix, hence it is completely *saillante*. Therefore, the lemma is an immediate consequence of Lemma 1 in [2]. □

Proposition 3.2 *Any stationary distribution for an SRBM Z is unique. If π is the stationary distribution,*

(a) *π is equivalent to Lebesgue measure on S, and for each $x \in S$ and each $f \in C_b(S)$,*

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} E_x \left[f(Z(i)) \right] = \int_S f(z)\, d\pi(z);$$

(b) *there is a finite Borel measure ν_i on F_i such that $\nu_i \approx \sigma_i$ and for each bounded Borel function f on F_i and $t \geq 0$,*

$$E_\pi \left[\int_0^t f(Z(s))\, dL_i(s) \right] = t \int_{F_i} f\, d\nu_i, \quad (i = 1, 2, \ldots, d).$$

Proof. The proof of part (a) is essentially the same as the proof of Theorem 7.1 of [26], and with Lemma 3.2 replacing Lemma 8.4 in [26], the proof of part (b) can also be readily carried over from that of Theorem 8.1 of [26]. □

In terms of the primitive data (Γ, μ, R) of an SRBM, we have the following necessary conditions.

Theorem 3.4 *If Z has a stationary distribution, then*

(a) *the reflection matrix R is invertible, and*

(b) *$R^{-1}\mu < 0$ if $R^{-1} \geq 0$.*

Proof. Suppose that π is a stationary distribution of an SRBM Z. (a) Assume R is singular. Then there exists a non-trivial vector v such that $v'R = 0$, where "prime" is the transpose operator. For the SRBM Z, we have the semimartingale representation (3.1). Therefore

$$v'Z(t) = v'X(t) + v'RL(t) = v'X(t), \tag{3.18}$$

since $v'R = 0$. From Proposition 3.2 (a), Z is ergodic, and hence $v'Z$ is ergodic. On the other hand, $v'X$ is a one-dimensional $(v'\Gamma v, v'\mu)$-Brownian motion, which can not be ergodic, contradicting to (3.18). Thus, R can not be singular, this proves part (a).

(b) From part (a), R is invertible and from representation (3.1),

$$Z^*(t) \equiv R^{-1}Z(t) = R^{-1}X(t) + L(t), \quad t \geq 0.$$

Since $R^{-1} \geq 0$, we have $Z^*(t) \geq 0$. The rest of arguments can follow from the proof of Lemma 6.14 in [26]. □

3.4.2 A Sufficient Condition

When R is Minkowski, Harrison and Williams [26] proved that Z has a stationary distribution if and only if $R^{-1}\mu < 0$. Their proof relied heavily on the fact that R is Minkowski, in which case the SRBM Z has an alternative characterization as in the Appendix of [39]. By using the alternative characterization, they were able to show a transformed process of Z is stochastically monotone, and thus proved the above assertion. In the following, for a general completely-$\mathcal{S}$ reflection matrix R with $R^{-1} \geq 0$, we establish a sufficient condition for the existence of a stationary distribution via existence of a Liapunov function. If R is further assumed to be symmetric, we are able to construct such a Liapunov function when $R^{-1}\mu < 0$; thus Z has a stationary distribution in this case.

Let $C_0(S)$ be the set of $f \in C_b(S)$ such that $\lim_{x\to\infty} f(x) = 0$, i.e., for any $\epsilon > 0$, there exists an $R > 0$ such that for all $x \in \{|x| > R\} \cap S$, $|f(x)| < \epsilon$.

Definition 3.8 *An SRBM Z is said to be $C_0(S)$-Feller if for every $f \in C_0(S)$, $T_t f(x) \equiv E_x f(Z(t)) \in C_0(S)$ for each $t \geq 0$.*

Remark. Let $\{P(t,x,\cdot) \in \mathcal{P}(S), t \geq 0, x \in S\}$ be the transition probabilities of Z, that is, for each $A \in \mathcal{B}_S$, $(t,x) \to P(t,x,A)$ is a Borel measurable function and $T_t f(x) = \int_S f(y)P(t,x,dy)$ for each $f \in B(S)$. Since Z is Feller, it can be checked [8] that Z is $C_0(S)$-Feller if and only if for any compact subset $K \subset S$,

$$\lim_{x\to\infty} P(t,x,K) = 0. \tag{3.19}$$

A Markov process $Z = \{Z(t), t \geq 0\}$ taking values in S is said to be stochastically continuous if for every $f \in C_b(S)$, $t \to E_x f(Z(t))$ is a continuous function for each $x \in S$. The following proposition, tailored to the present setting, was proved in [8] for a Markov process taking values in any complete separable metric state space.

Proposition 3.3 *If Z is stochastically continuous, $C_0(S)$-Feller Markov process valued in S, then the following two statements are equivalent.*

(a) *For any $x \in S$ and any compact subset $K \subset S$,*

$$\lim_{T\to\infty} \frac{1}{T} \int_0^T P(t,x,K)\, dt = 0;$$

(b) *There exists no stationary distribution for the Markov process Z.* □

Lemma 3.3 *If $R^{-1} \geq 0$, then any SRBM Z with reflection matrix R is $C_0(S)$-Feller.*

Proof. From representation (3.1), we have

$$Z^* \equiv R^{-1}Z = R^{-1}X + L.$$

For any given $r > 0$, let

$$f(x) = \frac{r^2}{r^2 + |x|^2}, \quad x \in \mathbb{R}^d,$$

then $f \in C^2$ and we have

$$\frac{\partial f}{\partial x_i}(x) = -\frac{2x_i}{r^2} f^2(x), \tag{3.20}$$

$$\frac{\partial^2 f}{\partial x_i \partial x_j}(x) = \frac{8x_i x_j}{r^4} f^3(x) - \frac{2\delta_{ij}}{r^2} f^2(x), \tag{3.21}$$

where $\delta_{ij} = 0$ if $i \neq j$ and $\delta_{ij} = 1$ otherwise. Hence $f \in C_b^2(\mathbb{R}^d)$. For each $x \in S$, applying Itô's formula with f on the completion $(C_S, \tilde{\mathcal{M}}, P_x)$ of $(C_S, \mathcal{M}, P_x)$, we obtain P_x-a.s. for all $t \geq 0$

$$\begin{aligned} f(Z^*(t)) &= f(Z^*(0)) + \int_0^t G^* f(Z^*(s)) ds + \sum_{i=1}^d \int_0^t \frac{\partial f}{\partial x_i}(Z^*(s))\, d\xi_i(s) \\ &\quad + \sum_{i=1}^d \int_0^t \frac{\partial f}{\partial x_i}(Z^*(s))\, dL_i(s), \end{aligned} \tag{3.22}$$

where $G^* f$ is a second order elliptic operator associated with covariance matrix $R^{-1}\Gamma R'^{-1}$ and drift vector $R^{-1}\mu$ defined exactly the same way as Gf was defined in (3.10), and $\xi(t) \equiv R^{-1}(X(t) - \mu t)$. Since the derivatives of f are bounded, the third term on the right hand side of (3.22) is a martingale, therefore, by taking expectation, we have

$$\begin{aligned} E_x[f(Z^*(t))] &= f(x) + E_x\left[\int_0^t G^* f(Z^*(s))\, ds\right] \\ &\quad + E_x\left[\sum_{i=1}^d \int_0^t \frac{\partial f}{\partial x_i}(Z^*(s))\, dL_i(s)\right]. \end{aligned} \tag{3.23}$$

From (3.20) and (3.21), there exists a constant $\alpha > 0$ such that

$$|G^* f(x)| \leq \alpha f(x),$$

and the last term of (3.23) is equal to

$$E_x\left[-\frac{2}{r}\sum_{i=1}^d \int_0^t Z_i^*(s) f^2(Z(s))\, dL_i(s)\right],$$

which is non-positive since $Z^*(t) \geq 0$. Therefore, from (3.23), we have

$$E_x\left[f(Z^*(t))\right] \leq f(x) + \alpha \int_o^t E_x\left[f(Z^*(s))\right]\, ds.$$

Bellman's inequality gives

$$E_x\left[f(Z^*(t))\right] \leq f(x)e^{\alpha t},$$

and by generalized Chebyshev inequality, we have

$$P_x\{|Z^*(t)| \leq r\} \leq 2E_x\left[f(Z^*(t))\right] \leq 2f(x)e^{\alpha t} \to 0,$$

as $x \to \infty$. That is, (3.19) is true and from the Remark after Definition 3.8, Z^* is $C_0(S)$-Feller, which implies Z is $C_0(S)$-Feller. □

Now we are ready to prove the following sufficient condition for the existence of a stationary distribution. An f satisfying the conditions in the theorem is called a Liapunov function.

Theorem 3.5 *Assume $R^{-1} \geq 0$ and Z is a corresponding SRBM. Suppose that there is a non-negative $f \in C^2(S)$ such that $E_x\left[\int_0^t |\nabla f|^2(Z(s))\, ds\right] < \infty$ for each $x \in S$ and each $t \geq 0$, and for some $r > 0$*

$$Gf(x) \leq -1, \quad x \in \{|x| \geq r\} \cap S, \tag{3.24}$$

$$D_i f(x) \leq 0, \quad x \in F_i \quad (i = 1, 2 \ldots, d). \tag{3.25}$$

Then Z has a stationary distribution.

Proof. For the f, as before, applying Itô's formula as before to Z, and taking expectation with respect to P_x, we have

$$E_x\left[f(Z(t))\right] = f(x) + E_x\left[\int_0^t Gf(Z(s))\, ds\right] + \sum_{i=1}^d E_x\left[\int_0^t D_i f(Z(s))\, dL_i(s)\right]. \tag{3.26}$$

Because $D_i f(x) \leq 0$ on F_i $(i = 1, 2, \ldots, d)$, the last summation in (3.26) is non-positive. Since the right side of (3.26) is non-negative, we have

$$f(x) + \int_0^t Gf(Z(s))\, ds \geq 0. \tag{3.27}$$

Let $M = \sup_{x\in\{|x|\le r\}\cap S} Gf(x)$, noticing the condition (3.24), we have

$$\begin{aligned}(3.28)\quad E_x\left[\int_0^t Gf(Z(s))\,ds\right] &= E_x\left[\int_0^t 1_{\{|Z(s)|\le r\}} Gf(Z(s))\,ds\right] \\ &\quad + E_x\left[\int_0^t 1_{\{|Z(s)|>r\}} Gf(Z(s))\,ds\right] \\ &\le M\int_0^t P_x\{|Z(s)|\le r\}\,ds - \int_0^t P_x\{|Z(s)|>r\}\,ds \\ &= (M+1)\int_0^t P_x\{|Z(s)|\le r\}\,ds - t.\end{aligned}$$

From (3.27) and (3.28), we have

$$\int_0^t P_x\{|Z(s)|\le r\}\,ds \ge \frac{t}{M+1} - \frac{f(x)}{M+1}.$$

Therefore

$$\liminf_{t\to\infty}\frac{1}{t}\int_0^t P_x\{|Z(s)|\le r\}\,ds \ge \frac{1}{M+1} > 0.$$

Since Z is continuous, it is stochastically continuous. Also from Lemma (3.3), Z is $C_0(S)$-Feller. Hence, Proposition 3.3 asserts that there is a stationary distribution for Z. □

Corollary 3.3 *Assume $R^{-1} \ge 0$. If R is symmetric, then the corresponding SRBM has a stationary distribution if and only if*

$$\gamma \equiv -R^{-1}\mu > 0. \tag{3.29}$$

Proof. The necessity is given in Theorem 3.4. As to the sufficiency, it is easy to check that

$$f(x) = x'R^{-1}x$$

is a Liapunov function as in Theorem 3.5, and hence the sufficiency is immediately from Theorem 3.5. □

3.5 The Basic Adjoint Relationship (BAR)

In this section we will first derive a necessary condition, called the basic adjoint relationship (BAR), for the stationary distribution to satisfy. Then we will prove that (BAR) characterizes the stationary distribution, which was first conjectured in [26].

3.5.1 Necessity of (BAR)

The following proposition was first derived by Harrison and Williams [26] for the case where the reflection matrix R is Minkowski.

Proposition 3.4 *Suppose π is the stationary distribution for Z associated with boundary measures ν_i $(i = 1, 2, \ldots, d)$ defined as in Proposition (3.2) (b). Then for each $f \in C_b^2(S)$,*

$$\int_S Gf \, d\pi + \sum_{i=1}^{d} \int_{F_i} D_i f \, d\nu_i = 0. \qquad \text{(BAR)} \tag{3.30}$$

Proof. Applying Itô's formula, and taking expectation E_x, we have

$$E_x \left[f(Z(t)) \right] = f(x) + E_x \left[\int_0^t Gf(Z(s)) \, ds \right] + \sum_{i=1}^{d} E_x \left[\int_0^t D_i f(Z(s)) \, dL_i(s) \right].$$

Integrating both sides with respect to the stationary distribution π, we obtain

$$0 = t \int_S Gf \, d\pi + t \sum_{i=1}^{d} \int_{F_i} D_i f \, d\nu_i,$$

where part (b) of Proposition 3.2 was used to obtain the first integral term and Fubini's theorem for the second. Now (3.30) can be readily obtained. □

3.5.2 Sufficiency of (BAR)–A General Proof

Theorem 3.6 *Assume R is a completely-$\mathcal{S}$ matrix. Suppose that π_0 is a probability measure on S with support in $\mathcal{O}$, and π_1, …, π_d are positive finite measures with supports on $F_1, \ldots, F_d$ respectively. If they jointly satisfy the basic adjoint relationship (3.30), then π_0 is the stationary distribution for a (Γ, μ, R)-SRBM Z, and π_i is the boundary measure associated with the SRBM as in Proposition 3.2 (b), i.e.,*

$$E_{\pi_0} \left[\int_0^t f(Z(s)) \, dL_i(s) \right] = t \int_{F_i} f \, d\pi_i, \quad f \in B(F_i), (i = 1, 2, \ldots, d).$$

Remark. This theorem was first conjectured in [26], when the reflection matrix is Minkowski. The authors proved the conjecture is true when a certain skew symmetry condition on the data (Γ, R) is satisfied.

We have already given one proof of the theorem in Theorem 2.10 when the state space of an SRBM is two-dimensional rectangle. The key assumption in Theorem 2.10 is the

dimension being equal to two. There, we can uses Echeverria's theorem [10] or [11, Theorem 9.17 of Chapter 4] directly and the key to the proof is to prove $\mathcal{D}_0$ is dense in $C_0(S)$. However, it is not clear if $\mathcal{D}_0$ is dense in $C_0(S)$ when $d \geq 3$, and it seems hard to generalize the proof to the higher dimensional case. Also, recall that in Theorem 2.10 nothing was proved regarding the boundary measures π_i.

The new idea of proving Theorem 3.6 in general is to consider an SRBM as a solution of a *constrained martingale problem* as considered in Kurtz [32, 31], instead of the *martingale problem* considered in Theorem 2.10. Remember in Theorem 2.10, in order to stay within the framework of martingales, we had to select a relatively small domain $\mathcal{D}_0$ for the operator G. The smaller domain D_0 made the existence of a solution to the martingale problem a relatively easy task, but made the proof of uniqueness much harder. A solution of a constrained martingale problem is a pair of processes (Z, L), where Z takes values in S and L is the *local time* (or control process) of Z satisfying (3.3). In other words, by considering the constrained martingale problem, we keep track of all the boundary behavior of an SRBM. It was shown in Theorem 3.3 that an SRBM is equivalent to a solution to the constrained martingale. Therefore we can use the known uniqueness result of the SRBM, which enables us to provide a general complete proof of Theorem 3.6.

In the following, we first extend Echeverria-Weiss's theorem to the constrained martingale problem. This is basically a recapitulation of Theorem 4.1 of Kurtz [31].

Proposition 3.5 *Let π_0 be a probability measure on S with $\pi_0(\partial S) = 0$, and $\pi_1, \ldots, \pi_d$ be finite positive measures on S with the support of π_i contained in F_i, and suppose that*

$$\int_S Gf \, d\pi_0 + \sum_{i=1}^{d} \int_{F_i} D_i f \, d\pi_i = 0 \quad \text{for each } f \in C_K^2(\mathbb{R}^d). \tag{3.31}$$

Then there exists a solution (Z, L) on some probability space $(\Omega, \{\mathcal{F}_t\}, \mathcal{F}, P)$ of the constrained martingale problem for $(S, G, D; \pi_0)$ such that Z is stationary, and

$$E\left[\int_0^t 1_A(Z(s)) \, dL_i(s)\right] = t\pi_i(A), \qquad \text{for all } A \in \mathcal{B}(F_i).$$

To get a direct solution of a constrained martingale problem is sometimes difficult. The indirect method to get such a solution is usually by way of solving the *patchwork martingale problem* as discussed in Kurtz [32].

Definition 3.9 *For any $\pi \in \mathcal{P}(S)$, by a solution of the patchwork martingale problem for $(S, G, D; \pi)$ we mean a continuous process $(Z, L_0, L_1, \ldots, L_d)$ on some filtered probability*

space $(\Omega, \{\mathcal{F}_t\}, \mathcal{F}, P)$ *such that* $Z(t) \in S$ *for all* $t \geq 0$, $PZ(0)^{-1} = \pi$, P*-a.s.* $L_i(0) = 0$, $L_i(\cdot)$ *is non-decreasing,* $(i = 0, 1, \ldots, d)$, $\sum_{i=0}^{d} L_i(t) = t$, $L_0(\cdot)$ *increases only at times* t *when* $Z(t) \in \mathcal{O}$ *and* $L_i(\cdot)$ *increases only at times* t *when* $Z(t) \in F_i$ $(i = 1, \ldots, d)$ *and*

$$f(Z(t)) - \int_0^t Gf(Z(s))\, dL_0(s) - \sum_{i=1}^{d} \int_0^t D_i f(Z(s))\, dL_i(s)$$

is a $\{\mathcal{F}_t\}$*-martingale for all* $f \in C_K^2(\mathbb{R}^d)$.

Lemma 3.4 *Suppose* $(\pi_0, \pi_1, \ldots, \pi_d)$ *is as in Proposition 3.5. Then there exists a solution* $(Z, L_0, L_1 \ldots, L_d)$ *on some filtered probability space* $(\Omega, \{\mathcal{F}_t\}, \mathcal{F}, P)$ *of the patchwork martingale problem for* (S, G, D) *such that for each* $h > 0$,

$$(Z(\cdot), L_0(\cdot + h) - L_0(\cdot), \ldots, L_d(\cdot + h) - L_d(\cdot))$$

is a stationary process, $Z(t)$ *has distribution* $C^{-1} \sum_{i=0}^{d} \pi_i$ *where* $C = 1 + \sum_{i=1}^{d} \pi_i(F_i)$ *and* $E[L_i(t + h) - L_i(t)] = C^{-1} h \pi_i(F_i)$.

Proof. Define $Hf(x, u) = u_0 Gf(x) + \sum_{i=1}^{d} u_i D_i f(x)$ for $f \in C_K^2(\mathbb{R}^d)$ and $u = (u_0, \cdots, u_d) \in U$, the set of vectors with components 0 or 1 and $\sum_{i=0}^{d} u_i = 1$. It is clear that the following four conditions are satisfied.

1. $C_K^2(\mathbb{R}^d)$ is dense in $C_0(S)$,
2. For each $f \in C_K^2(\mathbb{R}^d)$ and $u \in U$, $Hf(\cdot, u) \in C_0(S)$,
3. For each $f \in C_K^2(\mathbb{R}^d)$,
$$\lim_{x \to \infty} \sup_{u \in U} Hf(x, u) = 0,$$
4. For each $u \in U$, $Hf(\cdot, u)$ satisfies the positive maximum principle, i.e., if $f(x) = \sup_z f(z) > 0$, then $Hf(x, u) \leq 0$.

Define $\nu \in \mathcal{P}(S \times U)$ so that

$$\int_{S \times U} h(x, u) \nu(dx \times du) = C^{-1} \left(\int_S h(x, e_0) \pi_0(dx) + \sum_{i=1}^{d} \int_{F_i} h(x, e_i) \pi_i(dx) \right).$$

Then $\int_{S \times U} Hf\, d\nu = 0$ for each $f \in C_K^2(\mathbb{R}^d)$, hence H and ν satisfy the conditions of Theorem 4.1 of Stockbridge [46]. Therefore there exists a stationary solution (Z, Λ) of the

controlled martingale problem for H; that is, (Z, Λ) is a stationary $S \times \mathcal{P}(U)$-valued process adapted to a filtration $\{\mathcal{F}_t\}$ on a probability space $(\Omega, \mathcal{F}, P)$ such that

$$f(Z(t)) - \int_0^t \int_U Hf(Z(s), u)\Lambda(s, du)\, ds \tag{3.32}$$
$$= f(Z(t)) - \int_0^t Gf(Z(s))\Lambda(s, \{e_0\})\, ds - \sum_{i=1}^{d} \int_0^t D_i f(Z(s))\Lambda(s, \{e_i\})\, ds$$

is a $\{\mathcal{F}_t\}$-martingale for each $f \in C_K^2(\mathbb{R}^d)$, and for $t \geq 0$,

$$E\left[1_A(Z(t))\Lambda(t, E)\right] = \nu(A \times E), \quad A \in \mathcal{B}_S, E \in \mathcal{B}_U. \tag{3.33}$$

Furthermore, the process Z can be taken as continuous. Defining

$$L_i(t) \equiv \int_0^t \Lambda(s, \{e_i\})\, ds, \quad (i = 0, 1, \ldots, d), \tag{3.34}$$

and noting that by (3.33),

$$\begin{aligned} E\left[\int_0^t 1_{\mathcal{O}}(Z(s))\, dL_0(s)\right] &= t\,\nu(\mathcal{O} \times \{e_0\}) = t\, C^{-1}\pi_0(\mathcal{O}) \\ E\left[\int_0^t \chi_{F_i}(Z(s))\, dL_i(s)\right] &= t\,\nu(F_i \times \{e_i\}) = t\, C^{-1}\pi_i(F_i) \quad (i = 1, 2, \ldots, d) \end{aligned}$$

so L_0 increase only when $Z(t) \in \mathcal{O}$ and L_i increases only when Z is in F_i $(i = 1, 2, \ldots, d)$ and $\sum_{i=0}^{d} L_i(t) = t$. Therefore $(Z, L_0, L_1, \ldots, L_d)$ is a solution of the patchwork martingale problem. □

Proof of Proposition 3.5. Consider a sequence of patchwork martingale problems with G fixed, but D_i replaced by nD_i. Then $(\pi_0, n^{-1}\pi_1, \ldots, n^{-1}\pi_d)$ satisfies (3.31) for the new family. Lemma 3.4 gives a sequence of processes $(Z^n, L_0^n, \ldots, L_d^n)$ satisfying the stationary conclusion of Lemma 3.4. Note that $E[L_0^n(t)] \to t$ and $nE[L_i^n(t)]$ is bounded in n for $i = 1, 2, \ldots, d$. This boundedness implies that the sequence of processes satisfies the Meyer–Zheng conditions (see Corollary 1.3 of Kurtz [30]), and it follows that there exists a limiting process (at least along a subsequence) which will be a stationary solution of the constrained martingale problem. □

Proof of Theorem 3.6. Because π_0 and $(\pi_1, \ldots, \pi_d)$ in Theorem 3.6 satisfy all the conditions in Proposition 3.5, there is a stationary local time solution Z of the constrained martingale problem for $(S, G, D; \pi_0)$, and π_i $(1, 2, \ldots, d)$ is the corresponding boundary measure. By Theorem 3.3, this solution Z is an SRBM, and the uniqueness of an SRBM asserts π_0 is the stationary distribution of the SRBM and π_i is related as in part (b) of Proposition 3.2. □

Chapter 4

Computing the Stationary Distribution of SRBM in an Orthant

4.1 Introduction

Let Z be a (Γ, μ, R)-SRBM whose stationary distribution π exists. It follows from Proposition 3.2 that π and its associated boundary measures ν_i are absolutely continuous with respect to Lebesgue measure dx on S and $d\sigma_i$ on F_i, respectively, $(i = 1, 2, \ldots, d)$. Denote $p_0 \equiv d\pi/dx$ on S, $p_i \equiv d\nu_i/d\sigma_i$ on F_i, and

$$p = (p_0; p_1, \ldots, p_d). \tag{4.1}$$

Although p contains both the stationary density p_0 and its associated boundary densities p_i, we simply call p the stationary density of the SRBM Z. In Section 11 of [26], the authors made two conjectures and raised one open problem.

(a) (Conjecture) The boundary density $p_i = \frac{1}{2}\Gamma_{ii} p_0|_{F_i}$, $i = 1, 2, \ldots, d$;

(b) (Conjecture) The basic adjoint relationship (3.30) characterizes the stationary distribution;

(c) (Problem) How to solve (3.30), which presumably means developing efficient numerical methods for computing important performance measures associated with the stationary density p, such as the means of the marginal distributions.

Conjecture (a) is not resolved in this dissertation. But for computation of the performance measures associated with the stationary density, whether the conjecture is true or not does not matter. In fact, the algorithm that we are going to propose will compute p_0 as well as the p_i's. As to Conjecture (b), it is resolved in Theorem 3.6. Problem (c) is the focus of this chapter.

With regard to Problem (c), for a driftless RBM in two dimensions the work of Harrison, Landau and Shepp [23] gives an analytical expression for the stationary distribution, and the availability of a package for computation of Schwartz-Christoffel transformations makes evaluation of the associated performance measures numerically feasible, cf. [52]. For the two-dimensional case with drift, Foddy [13] found analytical expressions for the stationary distributions for certain special domains, drifts, and directions of reflection, using Riemann-Hilbert techniques. In dimensions three and more, RBM's having stationary distributions of exponential form were identified in [27, 58] and these results were applied in [26, 28] to RBM's arising as approximations to open and closed queueing networks with homogeneous customer populations.

In this chapter we describe an approach to computation of stationary distributions p that seems to be widely applicable. Assuming the following Conjecture 4.1, we are able to provide a full proof of the convergence of the algorithm; all the numerical comparisons done thus far show that our algorithm gives reasonable accurate estimates and the convergence is relatively fast.

Conjecture 4.1 *Suppose that p_0 is an integrable Borel function in $\mathcal{O}$ such that $\int_S p_0\,dx = 1$ and $p_1, \ldots, p_d$ are integrable on $F_1, \ldots, F_d$ respectively. If they jointly satisfy the basic adjoint relationship*

$$\int_S (Gf \cdot p_0)\; dx + \sum_{i=1}^{d} \int_{F_i} (D_i f \cdot p_i)\; d\sigma_i = 0 \qquad \textit{for all } f \in C_b^2(S),$$

then p_i is non-negative $(i = 0, 1, \ldots, d)$.

If Brownian system models are to have an impact in the world of practical performance analysis, solving Problem (c) above is obviously crucial. In particular, practical methods are needed for determining stationary distributions, and it is very unlikely that general analytical solutions will ever be found. As a tool for analysis of queueing systems, the computer program described in this dissertation is obviously limited in scope, but our ultimate goal is to implement the same basic computational approach in a general routine

that can compete with software packages like PANACEA [38] and QNA [54] in the analysis of large, complicated networks.

Sections 4.2–4.3 focus on a description of a general method for computing the stationary density, and Section 4.4 describes a particular choice we made in order to implement the general method. (Readers will see that other choices are certainly possible). In Section 4.5, we consider a number of test problems, comparing the numerical results obtained with our algorithm against known exact results. Finally, Section 4.6 presents a number of concrete examples to show how our algorithm can be practically used for the performance analysis of queueing networks.

4.2 An Inner Product Version of (BAR)

In terms of the density function p, our basic adjoint relationship (3.30) becomes

$$\int_S (Gf \cdot p_0)\, dx + \sum_{i=1}^{d} \int_{F_i} (D_i f \cdot p_i)\, d\sigma_i = 0 \quad \text{for all } f \in C_b^2(S). \tag{4.2}$$

We first convert (4.2) into a compact form that will be used in the next section. Given an $f \in C_b^2(S)$, let

$$\mathcal{A}f \equiv (Gf; D_1 f, \ldots, D_d f) \tag{4.3}$$

and

$$d\lambda \equiv (dx; d\sigma_1, \ldots, d\sigma_d). \tag{4.4}$$

For a subset E of $\mathbb{R}^d$, let $\mathcal{B}(E)$ denote the set of functions which are $\mathcal{B}_E$-measurable. Let

$$\begin{aligned} L^i(S, d\lambda) \equiv \Bigg\{ & g = (g_0; g_1, \ldots, g_d) \in \mathcal{B}(S) \times \mathcal{B}(F_1) \times \cdots \times \mathcal{B}(F_d) : \\ & \int_S |g_0|^i\, dx + \sum_{j=1}^{d} \int_{F_j} |g_j|^i\, d\sigma_j < \infty \Bigg\}, \quad i = 1, 2, \end{aligned} \tag{4.5}$$

and for $g \in L^1(S, d\lambda)$, let

$$\int_S g d\lambda \equiv \int_S g_0\, dx + \sum_{i=1}^{d} \int_{F_i} g_i\, d\sigma_i.$$

For $g, h \in \mathcal{B}(S) \times \mathcal{B}(F_1) \times \cdots \times \mathcal{B}(F_d)$, we put $g \cdot h \equiv (g_0 h_0; g_1 h_1, \ldots, g_d h_d)$, and for $h > 0$, we put $g/h \equiv (g_0/h_0; g_1/h_1, \ldots, g_d/h_d)$. With this notation, the basic adjoint relationship

(4.2) can be rewritten as

$$\int_S (\mathcal{A}f \cdot p)\, d\lambda = 0, \qquad \text{for all } f \in C_b^2(S). \tag{4.6}$$

We start with the compact form of the basic adjoint relationship (4.6). Suppose $\mathcal{A}f$ and p were in $L^2(S, d\lambda)$ for all $f \in C_b^2(S)$; then the basic adjoint relationship (4.6) would amount to saying that p is orthogonal to $\mathcal{A}f$ for each $f \in C_b^2(S)$. Unfortunately, *there are* functions $f \in C_b^2(S)$ for which $\mathcal{A}f$ is not in $L^2(S, d\lambda)$, since the state space S is unbounded. Nevertheless, the above observation is the key to the algorithm that we are going to describe. In fact, based on the above observation, Dai and Harrison [9] developed an algorithm for computing the stationary density of an RBM in a two-dimensional rectangular state space; see Section 2.4 for details.

In order to carry over the algorithm in Chapter 2 to the case with unbounded state space, we need to introduce the concept of "reference measure". Let

$$q = (q_0; q_1, \ldots, q_d)\,,$$

where q_0 is a probability density in S and q_i is a positive integrable function over F_i (with respect to $(d-1)$-dimensional Lebesgue measure $d\sigma_i$). The function q will be called a *reference density*. In the following, we assume a reference density q has been chosen. We will come back the question of how to choose a reference density in Section 4.4. For the reference density q, we define the *reference measure*

$$d\eta \equiv q\, d\lambda = (q_0 dx; q_1 d\sigma_1, \ldots, q_d d\sigma_d) \tag{4.7}$$

where the measure $d\lambda$ is defined in (4.4). Similar to the definition of $L^i(S, d\lambda)$ and $\int_S g\, d\lambda$ for $g \in L^1(S, d\lambda)$, we can define $L^i(S, d\eta)$ and $\int_S g\, d\eta$ for $g \in L^1(S, d\eta)$. If we introduce a *new unknown* $r = p/q$, then our basic adjoint relationship (4.6) takes the form

$$\int_S (\mathcal{A}f \cdot r)\; d\eta = 0 \qquad \text{for all } f \in C_b^2(S). \tag{4.8}$$

4.3 Algorithm

In the following, we actually develop an algorithm to solve for this new unknown r. Of course, once one has r, one can get the stationary density p (interior density p_0 and boundary densities $p_1, \ldots, p_d$) from $p = r \cdot q$.

Associated with the measure η, we can define $L^2 \equiv L^2(S, d\eta)$ similarly as defining $L^2(S, d\lambda)$ in (4.5), taken with usual inner product $(\cdot, \cdot)$ and norm $||\cdot||$. Unless specified otherwise, all inner products and norms are taken in $L^2(S, d\eta)$. Because η is a finite measure and $\mathcal{A}f$ is bounded, we have $\mathcal{A}f \in L^2$ for each $f \in C_b^2(S)$. We define $H \subset L^2$ to be the *closed subspace*

$$H \equiv \text{ the closure of } \left\{\mathcal{A}f : f \in C_b^2(S)\right\}, \tag{4.9}$$

where the closure is taken in the usual L^2 norm. Let ϕ_0 be defined on S by

$$\phi_0 = (1; 0, \ldots, 0). \tag{4.10}$$

Proposition 4.1 *If the unknown function $r \in L^2$, then r is orthogonal to H and $(r, \phi_0) = 1$. Conversely, if there is a non-negative $r \in L^2$ such that r is orthogonal to H and $(r, \phi_0) = 1$, then $r \cdot q$ is the stationary density, where q is the chosen reference density.*

Proof. Assume $r \in L^2$. Since r satisfies the basic adjoint relationship (4.8), r is orthogonal to $\mathcal{A}f$ for every $f \in C_b^2(S)$. By taking a limit, we have that r is orthogonal to H. Also,

$$\begin{aligned}(r, \phi_0) &= \int_S r(x) \cdot \phi_0(x)\, d\eta = \int_S \frac{p}{q} \cdot \phi_0 \cdot q\, d\lambda \\ &= \int_S p \cdot \phi_0\, d\lambda = \int_S p_0\, dx = 1.\end{aligned}$$

The last equality holds because p is a probability density function in $\mathcal{O}$.

Conversely, assume there is a non-negative $r \in L^2$ such that r is orthogonal to H. Then r is orthogonal to $\mathcal{A}f$ for every $f \in C_b^2(S)$, or equivalently, r satisfies (4.8). Therefore, by definition, $r \cdot q$ satisfies (4.6). Since $r \cdot q$ is non-negative and $\int_S (r \cdot q)\, dx = 1$, it follows from Theorem 3.6 that $r \cdot q$ is the stationary density of the corresponding RBM. □

Proposition 4.2 *If $r \in L^2$, then the orthogonal complement $\tilde{\phi_0} \equiv \phi_0 - \bar{\phi_0}$ of ϕ_0 is non-zero in L^2 and is orthogonal to H, where $\bar{\phi_0}$ is the projection of ϕ_0 onto H defined by*

$$\bar{\phi_0} \equiv \text{argmin}_{\phi \in H} ||\phi_0 - \phi||. \tag{4.11}$$

Therefore, assuming Conjecture 4.1 to be true, we arrive at the following formula for the stationary density

$$p = \frac{1}{||\tilde{\phi_0}||^2} \tilde{\phi_0} \cdot q. \tag{4.12}$$

Proof. Assume $r \in L^2$. From Proposition 4.1, $(r, \phi_0) = 1$, and thus r is *not* orthogonal to ϕ_0. Also from Proposition 4.1, we know that r is orthogonal to H. Hence we conclude that ϕ_0 is *not* in H. Because H is closed and $\phi_0 \notin H$, the projection $\bar{\phi_0}$ defined in (4.11) of ϕ_0 is not equal to itself, which implies that $\tilde{\phi_0}$ is not zero in L^2. Of course, the orthogonal complement $\tilde{\phi_0}$ is orthogonal to H. Since

$$\begin{aligned}(\tilde{\phi_0}, \phi_0) &= \int_{\mathcal{O}} \tilde{\phi_0} \cdot q \, dx = \int_S \phi_0 \cdot \tilde{\phi_0} \, d\eta \\ &= (\phi_0, \phi_0 - \bar{\phi_0}) = (\phi_0 - \bar{\phi_0}, \phi_0 - \bar{\phi_0}) \\ &= ||\tilde{\phi_0}||^2 > 0,\end{aligned}$$

assuming Conjecture 4.1 to be true, then $\tilde{\phi_0}$ does not change sign and $r \equiv (\tilde{\phi_0}, \phi_0)^{-1} \tilde{\phi_0}$ satisfies all the conditions in Proposition 4.1. Therefore the assertion of Proposition 4.2 follows from Proposition 4.1. □

Proposition 4.3 *Let $\{H_n\}$ be a sequence of increasing subspaces of H, such that $H_n \uparrow H$. For each n, define ϕ_n to be the orthogonal complement of ϕ_0 onto H_n, i.e., $\phi_n = \phi_0 - \psi_n$ where*

$$\psi_n \equiv \text{argmin}_{\phi \in H_n} ||\phi_0 - \phi||. \tag{4.13}$$

If $r \in L^2$, and if we assume Conjecture 4.1 to be true, then

$$r_n \equiv \frac{\phi_n}{||\phi_n||^2} \to r \quad in \quad L^2, \quad as \; n \to \infty. \tag{4.14}$$

Furthermore, by setting $p_n = r_n \cdot q$, one has for all $f \in L^2$,

$$\int_S f \cdot p_n \, d\lambda \to \int_S f \cdot p \, d\lambda,$$

and if q is taken to be bounded, then $p_n \to p$ in $L^2(S, d\lambda)$ as $n \to \infty$.

Proof. Since $H_n \uparrow H$, $\phi_n \to \bar{\phi_0}$ as $n \to \infty$. Because $r \in L^2$, $\tilde{\phi_0}$ is non-zero in L^2. Therefore $||\phi_n||^2 \to ||\tilde{\phi_0}|| \neq 0$, and hence $r_n \equiv \phi_n / ||\phi_n||^2$ goes (in L^2) to r, which is $\tilde{\phi_0}/||\tilde{\phi_0}||^2$ under the assumption that Conjecture 4.1 is true. If $f \in L^2$, then

$$\begin{aligned}\left| \int_S f \cdot p_n \, d\lambda - \int_S f \cdot p \, d\lambda \right| &= \left| \int_S f \, (r_n - r) \, q \, d\lambda \right| \\ &= \left| \int_S f \, (r_n - r) \, d\eta \right| \leq ||f||^{1/2} ||r_n - r||^{1/2}.\end{aligned} \tag{4.15}$$

When q is bounded, then

$$\int_S |p_n - p|^2 \, d\lambda = \int_S |r_n - r|^2 q^2 \, d\lambda \leq \max_{x \in S} q(x) ||r_n - r||^2. \tag{4.16}$$

The rest of proof can be readily obtained from (4.15) and (4.16). □

Proposition 4.3 says that, when $r \in L^2$, we can calculate the corresponding stationary density p numerically by choosing appropriate finite dimensional subspaces H_n. However, when $r \notin L^2$, we can still define p_n via r_n as in (4.14). We conjecture that, in this case, p_n converges to p *weakly*, i.e.,

$$\int_S f \cdot p_n \, d\lambda \to \int_S f \cdot p \, d\lambda \quad \text{as } n \to \infty \text{ for all } f \in C_b(S).$$

4.4 Choosing a Reference Density and $\{H_n\}$

In this section we will choose a particular q to serve as the reference density function of the previous section. We will first define some quantities that are of interest in the queueing theoretic application of RBM. Let p be the stationary density of an RBM, and for $i = 1, 2, \ldots, d$ let

$$m_i \equiv \int_S (x_i \cdot p_0(x)) \, dx. \tag{4.17}$$

Then, $m \equiv (m_1, \ldots, m_d)'$ represents the long run mean position of the SRBM Z. For an SRBM arising from queueing networks, m_i corresponds to some performance measure of the corresponding queueing network (for example, the average waiting time or average queue length at station i). There are, of course, other quantities associated with p_0 that are of interest, such as the second moments or quantities like $\int_S \max(x_i, x_j) p_0(x) \, dx$ for $i \neq j$; see Nguyen [35]. Because our algorithm gives estimates for the density function itself, such extensions are trivial and we will only focus on the quantity m. Proposition 4.4 below indicates that, in calculating the steady-state mean vector m, it is enough to consider only "standard SRBM's".

Definition 4.1 *A semimartingale RBM with data (Γ, μ, R) is said to be standard if $\Gamma_{ii} = 1$ for $i = 1, 2, \ldots, d$, R is invertible and $\max_{1 \le i \le d} |(R^{-1}\mu)_i| = 1$.*

Proposition 4.4 *Suppose that Z with $\{P_x, x \in S\}$ is an SRBM with data (S, Γ, μ, R), that Z has a stationary distribution, and that the steady-state mean vector $m = (m_1, \ldots, m_d)'$ is defined via (4.17). Let $Z^* = \{Z^*(t), t \ge 0\}$ be defined by*

$$\begin{pmatrix} Z_1^*(t) \\ \vdots \\ Z_d^*(t) \end{pmatrix} = \begin{pmatrix} \frac{\gamma_{\max}}{\sqrt{\Gamma_{11}}} Z_1\left(\frac{t}{\gamma_{\max}^2}\right) \\ \vdots \\ \frac{\gamma_{\max}}{\sqrt{\Gamma_{dd}}} Z_d\left(\frac{t}{\gamma_{\max}^2}\right) \end{pmatrix}.$$

Then Z^* is a standard SRBM with data $(S, \Gamma^*, \mu^*, R^*)$, where

$$\Gamma^*_{ij} = \frac{\Gamma_{ij}}{\sqrt{\Gamma_{ii}}\sqrt{\Gamma_{jj}}}, \quad \mu^*_i = \frac{\mu_i}{\sqrt{\Gamma_{ii}}\gamma_{\max}}, \quad R^*_{ij} = \frac{R_{ij}}{R_{jj}}\frac{\sqrt{\Gamma_{jj}}}{\sqrt{\Gamma_{ii}}}, \quad 1 \le i \le d, 1 \le j \le d,$$

where $\gamma_{\max} = \max_{1\le i\le d} |\gamma_i|$ with

$$\gamma \equiv -R^{*-1} \begin{pmatrix} \mu_1/\sqrt{\Gamma_{11}} \\ \vdots \\ \mu_d/\sqrt{\Gamma_{dd}} \end{pmatrix}.$$

Moreover, Z^* has a stationary distribution and its steady-state mean vector m^* is related to m via

$$m_i = \frac{\sqrt{\Gamma_{ii}}}{\gamma_{\max}} m^*_i, \quad i = 1, 2, \ldots, d. \tag{4.18}$$

Proof. Suppose that Z with $\{P_x, x \in S\}$ is a (S, Γ, μ, R)-SRBM. Let m be the steady-state mean vector associated with the (S, Γ, μ, R)-SRBM. Let $\Gamma^* = (\Gamma^*_{ij})_{1\le i,j\le d}$ and $R^* = (R^*_{ij})_{1\le i,j\le d}$ be two $d \times d$ matrices defined by

$$\Gamma^*_{ij} = \frac{\Gamma_{ij}}{\sqrt{\Gamma_{ii}}\sqrt{\Gamma_{jj}}}, \quad R^*_{ij} = \frac{R_{ij}}{R_{jj}}\frac{\sqrt{\Gamma_{jj}}}{\sqrt{\Gamma_{ii}}}, \quad 1 \le i \le d, 1 \le j \le d.$$

Let $\mu^* = (\mu^*_1, \ldots, \mu^*_d)'$ where

$$\mu^*_i = \frac{\mu_i}{\sqrt{\Gamma_{ii}}\gamma_{\max}}, \quad 1 \le i \le d$$

and $\gamma_{\max} = \max_{1\le i\le d} |\gamma_i|$ with

$$\gamma \equiv -R^{*-1} \begin{pmatrix} \mu_1/\sqrt{\Gamma_{11}} \\ \vdots \\ \mu_d/\sqrt{\Gamma_{dd}} \end{pmatrix}.$$

Because Z is an (S, Γ, μ, R)-SRBM in the orthant $\mathbb{R}^d_+$. Then (3.1)–(3.3) hold. In particular, for each $x \in S$,

$$Z(t) = X(t) + RL(t), \quad t \ge 0, \quad P_x\text{-a.s.}$$

It follows then, P_x-almost surely, for each $t \ge 0$,

$$\begin{pmatrix} \frac{\gamma_{\max}}{\sqrt{\Gamma_{11}}} Z_1\left(\frac{t}{\gamma^2_{\max}}\right) \\ \vdots \\ \frac{\gamma_{\max}}{\sqrt{\Gamma_{dd}}} Z_d\left(\frac{t}{\gamma^2_{\max}}\right) \end{pmatrix} = \begin{pmatrix} \frac{\gamma_{\max}}{\sqrt{\Gamma_{11}}} X_1\left(\frac{t}{\gamma^2_{\max}}\right) \\ \vdots \\ \frac{\gamma_{\max}}{\sqrt{\Gamma_{dd}}} X_d\left(\frac{t}{\gamma^2_{\max}}\right) \end{pmatrix} + \tag{4.19}$$

$$+\begin{pmatrix} \frac{R_{11}}{R_{11}}\frac{\sqrt{\Gamma_{11}}}{\sqrt{\Gamma_{11}}} & \cdots & \frac{R_{1d}}{R_{dd}}\frac{\sqrt{\Gamma_{dd}}}{\sqrt{\Gamma_{11}}} \\ \vdots & \ddots & \vdots \\ \frac{R_{d1}}{R_{11}}\frac{\sqrt{\Gamma_{11}}}{\sqrt{\Gamma_{dd}}} & \cdots & \frac{R_{dd}}{R_{dd}}\frac{\sqrt{\Gamma_{dd}}}{\sqrt{\Gamma_{dd}}} \end{pmatrix} \begin{pmatrix} \frac{\gamma_{\max}}{\sqrt{\Gamma_{11}}} L_1\left(\frac{t}{\gamma_{\max}^2}\right) \\ \vdots \\ \frac{\gamma_{\max}}{\sqrt{\Gamma_{dd}}} L_d\left(\frac{t}{\gamma_{\max}^2}\right) \end{pmatrix}.$$

Let $Z^*(t)$ be the left hand side of (4.19), then $Z^* = \{Z^*(t), t \geq 0\}$ is an $(S, \Gamma^*, \mu^*, R^*)$-SRBM. By the definition of (Γ^*, μ^*, R^*), we have $\Gamma^*_{ii} = 1$, $R^*_{ii} = 1$ $(i = 1, \ldots, d)$ and $\max_{1 \leq i \leq d} |\gamma^*_i| = 1$ where

$$\gamma^* \equiv -R^{*-1}\mu^* = \frac{\gamma}{\gamma_{\max}}.$$

Therefore Z^* is a standard SRBM. Let $m^* = (m^*_1, \ldots, m^*_d)'$ be the steady-state mean of the SRBM Z^*. It follows that

$$m_i = \frac{\sqrt{\Gamma_{ii}}}{\gamma_{\max}} m^*_i, \quad i = 1, 2, \ldots, d,$$

and hence m can be computed in terms of m^*. □

When the SRBM is standard, all the associated data are properly scaled, and therefore the algorithm described in the next section is more stable.

Definition 4.2 *We say a stationary density for Z is of product form (or has a separable density) if the stationary density p_0 can be written as*

$$p_0(z) = \Pi_{k=1}^d p_0^k(z_k), \quad z = (z_1, \ldots, z_k) \in S, \tag{4.20}$$

where p_0', ..., p_0^d are all probability densities relative to Lebesgue measure on $\mathbb{R}_+$.

The following proposition is due to Harrison and Williams. (For a proof, see [26] Theorem 9.2.)

Proposition 4.5 *A standard SRBM Z has a product form stationary distribution if and only if (3.29) holds and the following condition holds:*

$$2\Gamma_{jk} = (R_{kj} + R_{jk}) \quad \textit{for} \quad j \neq k. \tag{4.21}$$

In this case, there is a constant C such that the density p is the exponential

$$z \to C \exp(-2\gamma \cdot z), \quad z = (z_1, \ldots, z_d) \in S, \tag{4.22}$$

i.e., p_0 is given by (4.22) and p_i is the restriction of this exponential to F_i $(i = 1, 2, \ldots, d)$, where γ is defined in (3.29).

Remark. Condition (4.21) holds if and only if $I - R$ is skew symmetric; Harrison and Williams [26] refer to (4.21) as a *skew symmetry* condition.

The above proposition asserts that the density p is of exponential form precisely when the *skew symmetry condition* (4.21) is satisfied. If we choose q to be the above exponential in (4.22), then $r \equiv p/q$ is identically one when (4.21) is satisfied. When the *skew symmetry condition* (4.21) is not satisfied, but is *almost* satisfied, we expect the density p is only slightly perturbated from the above exponential. That is, r is nearly equal to one. Therefore, we can think of r as some adjusting factor of how far the actual stationary density p is from the exponential solution. Based on these observations, we choose q to be the exponential in (4.22).

Corollary 4.1 *Fix the reference density function q to be the exponential in (4.22), and let $d\eta$ be as in (4.7). If all the assumptions in Proposition 4.3 are satisfied, then for each $f \in L^2$*

$$\int_S f(x)p_n(x)\,dx \to \int_S f(x)p(x)\,dx, \quad \text{as} \quad n \to \infty.$$

In particular, the approximating moment $m_i^n \equiv \int_S x_i p_n\,dx$ converges to m_i $(i = 1, 2, \ldots, d)$.

Proof. Since q is decaying exponentially, $f(x) \equiv x_i \in L^2 = L^2(S, d\eta)$. Therefore, the proof of the corollary is an immediate consequence of Proposition 4.3. □

Remark. Since the chosen q is smooth, our approximating function $p_n = r_n \cdot q$ is also smooth, whereas most of the time the stationary density p is singular. If we knew in advance the order of the singularities at the non-smooth parts of the boundary, we might be able to incorporate these singularities into the reference density function q, so that $r = p/q$ would be smooth. This would yield an algorithm which would converge faster. Unfortunately, there is no general result on the order of the singularities of the stationary density of an RBM in high dimensions.

The next proposition tells us how to choose a finite dimensional subspace H_n approximating H.

Proposition 4.6 *For each integer $n \geq 1$, let*

$$H_n = \{\mathcal{A}f : f \quad \text{is a polynomial of degree } \leq n\}. \tag{4.23}$$

Then $H_n \uparrow H$.

Proof. It is obvious that H_n is increasing. The proof of the convergence $H_n \to H$ is an immediate consequence of Proposition 7.1 and Remark 6.2 in the appendices of Ethier and Kurtz [11]. □

4.5 Numerical Comparisons

In this section, we will compare numerical results (QNET estimates) from our algorithm with some known analytical results for particular RBM's. The two cases discussed in this section are the only cases for which explicit analytical solutions are known, except for SRBM's with exponential stationary densities. Because the exponential solutions are incorporated in our algorithm, these are the only exact solutions available for checking the algorithm.

4.5.1 A Two-Dimensional RBM

Consider a two-dimensional RBM with covariance matrix $\Gamma = I$, drift vector of the form $\mu = (\mu_1, 0)'$ and reflection matrix

$$R = \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix}.$$

An RBM with these data (Γ, μ, R) arises as a Brownian approximation for a special type of tandem queue, cf. [21]. For this RBM, our stationary condition (3.29) reduces to $\mu_1 < 0$. This type of RBM was studied in Harrison [22]. There, after a transformation, the author was able to obtain a solution of product form for the stationary density in polar coordinates. The explicit form of the stationary density is

$$p(x) = Cr^{-1/2} e^{\mu_1 (r + x_1)} \cos(\theta/2), \quad x = (x_1, x_2) = (r\cos\theta, r\sin\theta), \tag{4.24}$$

where $C = \pi^{-1/2} (2|\mu_1|)^{3/2}$. Notice that $x = 0$ is the singular point of the density. The above density p is square integrable in $\mathcal{O}$ with respect to interior Lebesgue measure, but p is *not* square integrable over the boundary with respect to boundary Lebesgue measure. Therefore, p is not in L^2. By the scaling argument given in Proposition 4.4, it is enough to consider the case when $\mu_1 = -1$. It follows from (4.24) (cf. Greenberg [19]) that

$$m_1 = \frac{1}{2}, \qquad m_2 = \frac{3}{4}.$$

n	1	2	3	4
m_2	0.50000	0.83333	0.75000	0.75873
n	5	6	7	8
m_2	0.75133	0.75334	0.75225	0.75681

Table 4.1: Convergence of the Algorithm

Using our algorithm, taking $n = 5$ (n is the maximum degree of the polynomials we take in (4.23)), we have QNET estimates

$$m_1 = 0.50000, \qquad m_2 = 0.75133.$$

The QNET estimate of m_1 is exact as expected. If one takes the first station in the tandem queue in isolation, the first station will correspond to a one-dimensional RBM, whose stationary density is always of exponential form. It was rigorously proved in [22] that the one-dimensional marginal distribution in x_1 is indeed of exponential form. The above comparison shows that our algorithm can catch some marginal exponentials. Table 4.1 shows that if we require a one percent of accuracy, which is usually good enough in queueing network applications, the convergence is very fast, even for this very singular density.

4.5.2 Symmetric RBM's

An RBM is said to be *symmetric* if it is standard (cf. Definition 4.1) and its data (Γ, μ, R) are symmetric in the following sense: $\Gamma_{ji} = \Gamma_{ij} = \rho$ for $1 \le i < j \le d$, $\mu_i = -1$ for $1 \le i \le d$ and $R_{ji} = R_{ij} = -r$ for $1 \le i < j \le d$, where $r \ge 0$ and $r(d-1) < 1$. A symmetric RBM arises as a Brownian approximation of a symmetric *generalized Jakson network*. In such a network, each of the d station behaves exactly the same. Customers finishing service at one station will go to any one of the other $d-1$ stations with equal probability r and will leave the network with probability $1-(d-1)r$. For $d = 2$, the symmetric queueing network was used by Foschini to model a pair of communicating computers [14]. The author extensively studied the stationary density of the corresponding two-dimensional symmetric RBM.

Now because the data (Γ, μ, R) of a symmetric RBM is, in an obvious sense, invariant under permutation of the integer set $\{1, 2, \ldots, d\}$, it is clear that the stationary density $p_0(x)$ is symmetric, i.e.,

$$p_0(x_1, x_2, \ldots, x_d) = p_0(x_{\sigma(1)}, x_{\sigma(2)}, \ldots, x_{\sigma(d)}),$$

for any permutation σ on $\{1, 2, \ldots, d\}$. In particular, $\int_{F_j} x_i \cdot p_j \, d\sigma_j = \delta \equiv \int_{F_1} x_2 \cdot p_1 \, d\sigma_1$ for all $i \neq j$ and the marginal densities of p_0 are same, and hence

$$m_1 = m_2 = \cdots = m_d.$$

If we take $f = x_i^2$, then the basic adjoint relationship (3.30) gives

$$1 - 2m_1 - r \sum_{j=2}^{d} \int_{F_j} x_1 \cdot p_j \, d\sigma_j = 0. \tag{4.25}$$

Taking $f = x_1 x_2$, we have

$$\begin{aligned} \rho \;-\; & (m_1 + m_2) + \frac{1}{2} \int_{F_1} x_2 \cdot p_1 \, d\sigma_1 + \frac{1}{2} \int_{F_2} x_1 \cdot p_2 \, d\sigma_2 \\ -\; & \frac{1}{2} r \sum_{j=3}^{d} \int_{F_j} (x_1 + x_2) p_j \, d\sigma_j = 0. \end{aligned} \tag{4.26}$$

By symmetry, from (4.25) and (4.26) we get

$$1 - 2m_1 - \delta(d-1)r = 0, \tag{4.27}$$

$$\rho - 2m_1 + \delta - \delta(d-2)r = 0. \tag{4.28}$$

Solving these linear equations gives $\delta = (1 - \rho)/(1 + r)$ and

$$m_1 = \frac{1 - (d-2)r + (d-1)r\rho}{2(1+r)}. \tag{4.29}$$

Now we compare our numerical estimates of m_1 with the exact values of m_1 calculated from formula (4.29). When $d = 2$, the conditions on data (Γ, μ, R) yield $|\rho| < 1$ and $0 \leq r < 1$. Ranging $\rho \in \{-0.9, -0.5, 0.0, 0.5, 0.9\}$ and $r \in \{0.2, 0.4, 0.6, 0.8, 0.9, 0.95\}$, and taking $n = 3$, we obtain QNET estimates for m_1 (Table 4.2). Table 4.3 gives the relative errors between these QNET estimates and the exact values.

When $r = 1$ there is no corresponding SRBM. It is expected that when ρ is big (the skew symmetry condition (4.21) is far from being satisfied), the stationary density is very singular as $r \uparrow 1$. This phenomenon seems to be indicated in Table 4.3, where the performance of the algorithm degrades as r increases to one. When the dimension d is 3, then the restriction on the data gives $-1/2 < \rho < 1$ and $0 \leq r < 1/2$. Table 4.4 gives the relative errors between some QNET estimates and the exact values for m_1 in this case. When the dimension d is 4, then the restriction on the data gives $-1/3 < \rho < 1$ and $0 \leq r < 1/3$; relative errors between QNET estimates and exact values for m_1 are found in Table 4.5.

$r \setminus \rho$	-0.90	-0.50	0.00	0.50	0.90
0.20	0.341667	0.375000	0.416667	0.458333	0.491667
0.40	0.228571	0.285714	0.357143	0.428571	0.485714
0.60	0.143750	0.218750	0.312500	0.406250	0.481250
0.80	0.077778	0.166667	0.277778	0.388889	0.477778
0.90	0.050000	0.144737	0.263158	0.381579	0.476316
0.95	0.037179	0.134615	0.256410	0.378205	0.475641

Table 4.2: QNET Estimates for m_1 when $d = 2$ ($n = 3$)

$r \setminus \rho$	-0.90	-0.50	0.00	0.50	0.90
0.20	$9.75e{-16}$	$4.44e{-16}$	$-7.99e{-16}$	$-2.42e{-16}$	$3.39e{-16}$
0.40	$3.89e{-15}$	$-7.77e{-16}$	$-7.77e{-16}$	$5.18e{-16}$	$1.49e{-15}$
0.60	$4.63e{-15}$	$3.17e{-15}$	$8.88e{-16}$	$2.73e{-15}$	$-2.42e{-15}$
0.80	$2.32e{-15}$	$-4.33e{-15}$	$-1.20e{-15}$	$-2.63e{-14}$	$3.83e{-15}$
0.90	$-6.91e{-14}$	$2.45e{-14}$	$-2.08e{-13}$	$-1.81e{-12}$	$-7.46e{-13}$
0.95	$-9.59e{-14}$	$6.69e{-13}$	$5.44e{-12}$	$3.32e{-11}$	$1.04e{-10}$

Table 4.3: Relative Errors when $d = 2$ ($n = 3$)

$\rho \setminus r$	0.10	0.30	0.40	0.45
-0.40	$1.49e{-16}$	$2.04e{-15}$	$1.39e{-15}$	$-1.06e{-14}$
-0.20	$1.14e{-15}$	$-2.12e{-15}$	$-1.06e{-15}$	$-6.16e{-14}$
0.00	$-5.43e{-16}$	$-2.68e{-15}$	$-7.77e{-15}$	$-6.54e{-14}$
0.80	$-4.61e{-16}$	$-9.79e{-15}$	$8.65e{-14}$	$9.21e{-12}$
0.90	$-1.47e{-15}$	$-1.75e{-14}$	$3.66e{-13}$	$7.73e{-12}$

Table 4.4: Relative Errors when $d = 3$ ($n = 3$)

$\rho \setminus r$	0.10	0.20	0.30	0.33
-0.30	$-2.06e{-15}$	$4.60e{-15}$	$-2.01e{-14}$	$-2.23e{-13}$
-0.15	$-1.29e{-15}$	$7.18e{-15}$	$1.39e{-14}$	$-5.32e{-12}$
0.00	$9.16e{-16}$	$8.88e{-16}$	$-1.08e{-14}$	$-6.06e{-12}$
0.80	$-3.88e{-15}$	$7.40e{-16}$	$3.23e{-12}$	$4.60e{-10}$
0.90	$-5.14e{-15}$	$2.34e{-14}$	$4.78e{-12}$	$1.53e{-11}$

Table 4.5: Relative Errors when $d = 4$ ($n = 3$)

4.6 Queueing Network Applications

4.6.1 Queues in Series

Consider a network of d queues in series as pictured in Figure 4.1. The departure process from one service station forms the arrival process at the next service station. This serial queueing system has been studied by many authors, [15, 22, 45, 55, 49, 20] to name a few. Different authors studied this queueing system with different objectives. We will focus on the performance analysis of this network. To be more specific, we are interested in calculating the long-run average waiting time w_j that customers experience at station j. We will approximate the d-dimensional *workload process*, as described in Section 1.2 when $d = 2$, by a d-dimensional *semimartingale reflected Brownian motion* (SRBM) Z in the orthant $\mathbb{R}^d_+$. Then we will use the algorithm described in Sections 4.1 through 4.4 to solve for the stationary density of the corresponding SRBM numerically. From this, we will get back to the average waiting time in the original queueing system.

To find the long-run average waiting times even for this "simple" queueing system is difficult with general distributions (i.e., when they are not all exponential). The problem is that the arrival process to any queue beyond the first is not a renewal process; see Suresh and Whitt [49] and Berman and Westcott [1].

We characterize the network of queues in series by the $(2d + 1)$-tuple

$$(C_a^2, \rho_1, C_{s_1}^2, \ldots, \rho_d, C_{s_d}^2),$$

where d is the number of stations in the network, and as specified in Section 1.2, C_a^2 is the squared coefficient of variation for the interarrival time distribution, ρ_i is the mean service

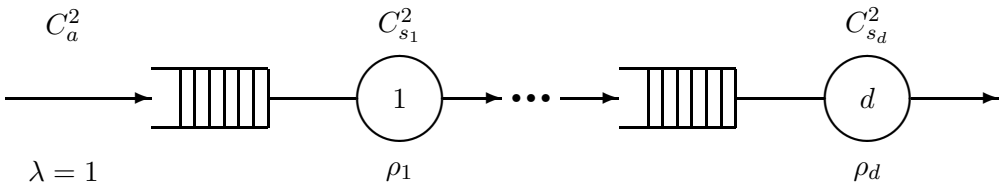

Figure 4.1: A Network of d Queues in Series

time at station i and $C_{s_i}^2$ is the squared coefficient of variation of the service time distribution at station i. We assume the arrival rate to station 1 is one so that ρ_i is the traffic intensity at station i. Using the method proposed in Harrison and Nguyen [24], as illustrated in Section 1.2, the d-dimensional current workload process $W(t) = (W_1(t), \ldots, W_d(t))'$ can be replaced (approximated) by a d-dimensional SRBM Z in the orthant $\mathbb{R}_+^d$ associated with data (Γ, μ, R). When $d = 2$, it has been derived in Section 1.2 that

$$\Gamma = \begin{pmatrix} \rho_1^2(C_a^2 + C_{s_1}^2) & -\rho_1\rho_2 C_{s_1}^2 \\ -\rho_1\rho_2 C_{s_1}^2 & \rho_2^2(C_{s_1}^2 + C_{s_2}^2) \end{pmatrix}, \quad \mu = \begin{pmatrix} \rho_1 - 1 \\ \rho_2/\rho_1 - 1 \end{pmatrix}, \quad R = \begin{pmatrix} 1 & 0 \\ -\rho_2/\rho_1 & 1, \end{pmatrix}.$$

Using the algorithm developed in an earlier part of this chapter, we can calculate the steady state mean $m = (m_1, m_2, \ldots, m_d)'$ of Z. It is suggested in [24] that the expected waiting times $w = (w_1, \ldots, w_d)'$ be approximated by the long-run average workload levels, and hence we have the estimates

$$w_j \doteq m_j, \quad j = 1, \ldots, d.$$

In the following, we will compare our QNET estimates of average waiting times for many system configurations and a variety of system parameters with simulation estimates as well as with estimates from other approximation schemes. Readers should notice that in using QNET to estimate w, there are two levels of errors that might occur. The first level comes from the model replacement ($w_j \doteq m_j$), and the second level of error comes from the numerical calculation of m_j.

Fraker's Experiment

Here we consider Fraker's [15] experiment. Fraker simulated eight different cases of eight single-server queues in series. In all cases, the external arrival process is Possion, and all service time distributions are Erlang. Each of four traffic intensities ($\rho = 0.3, 0.5, 0.7$ and 0.9) and each of four Erlang service-time distributions ($M = E_1, E_4, E_8$, and $D = E_\infty$) is assigned randomly to two of the eight nodes (M for exponential distributions or memoryless, and D for deterministic). Fraker's simulations consisted of three separate runs of 2500 customers each, with the first 500 being discarded to damp out the transient effects. Statistics were collected for six blocks of 1000 customers each. As pointed out by Whitt [55], the simulation runs were not long enough for good statistical accuracy, and no single estimate can be relied upon, but the total experiment clearly yields meaningful comparisons of approximations.

Whitt [55] compared his QNA approximations [54] with the first three cases of Fraker's simulation and approximation results. Currently, our QNET software is still under testing, and there is a limitation of the network size. We truncate Fraker's network by considering the first three stations as a sub-network of three queues in series. Table 4.6 compares various approximations of expected waiting at station j ($j = 1, 2, 3$) with Fraker's simulation results in the first three cases. All numbers except our QNET estimates are taken from Table 1–3 of [55]. The SIM column gives Fraker's simulation results, and the column under "Fraker" contains Fraker's approximation results. The QNA approximation results are listed in the QNA column, and our QNET estimates are in the QNET column. From Table 4.6, we can see that our QNET approximations, rooted in *heavy traffic* theory, perform very well, even under moderate traffic intensities. We remark that Fraker's approximations are especially designed for tandem queueing systems with Erlang service times, whereas our QNET method can handle a quite general class of networks. Table 4.7 shows the convergence of our algorithm for Case I, where $(m_1, \ldots, m_k)'$ is the steady-state mean of the corresponding k-dimensional SRBM that arises from the sub-network of the first k stations in series ($k = 2, 3, 4, 5$). From this table, we see that our algorithm is robust regardless of dimension.

Shimshak's Experiment

Shimshak [45] also developed approximations for expected waiting times of queues in series and compared them with simulations. He simulated two single-server queues in series with three different renewal arrival processes. In Shimshak's experiments I, III, and IV the interarrival time distributions are, respectively, exponential with $C_a^2 = 1$, hyperexponential (a mixture of two exponential distributions) with $C_a^2 = 4.0$, and Erlang (E_{10}) with $C_a^2 = 0.1$. (In experiment II the second queue has ten servers, so it will not be considered here.) Experiments III and IV are interesting additions to Fraker's experiment in the previous subsection because the external arrival process is non-Poisson.

These three cases were used by Whitt in [55] to validate his QNA software package. We follow his line of exposition except that we leave out two *first moment* approximation schemes which perform very poorly most of the time in these comparisons. The tandem queue contains four free variables. They are the two traffic intensities ($\rho = 0.6$ and 0.8) and two Erlang service time distributions (E_1 and E_{10}, $C_s^2 = 1.0$ or 0.1) at the two nodes. In each experiment, one of these four variables is held fixed. Hence, each experiment consists

Parameters			Waiting Times			
j	ρ_j	C_{s_j}	SIM	Fraker	QNA	QNET*
case I						
1	0.7	1/8	0.98	0.92	0.92	0.92
2	0.5	1	0.30	0.38	0.38	0.34
3	0.5	0	0.19	0.13	0.16	0.21
case II						
1	0.9	1	6.25	8.10	8.10	8.10
2	0.7	1/8	0.84	0.92	0.92	0.92
3	0.3	1/4	0.01	0.04	0.04	0.03
case III						
1	0.9	0	4.70	4.05	4.05	4.05
2	0.9	1/4	2.19	1.80	2.32	1.89
3	0.5	1	0.24	0.23	0.24	0.33

$*n = 5$

Table 4.6: Comparisons with Fraker's Experiments

m_1	m_2	m_3	m_4	m_5	n
0.918750	0.327284	0.194887	0.202654	0.015241	3
0.918750	0.331946	0.197427	0.215913	0.015034	4
0.918750	0.327474	0.196155	0.197442		3
0.918750	0.332758	0.198488	0.203457		4
0.918750	0.333974	0.212011			3
0.918750	0.335945	0.205642			4
0.918750	0.334451	0.204894			5
0.918750	0.327135				3
0.918750	0.330338				4
0.918750	0.330561				5

Table 4.7: Convergence of the Algorithm

of eight cases, containing all combinations of three variables, each with two possible values. In experiment I the service time distribution at node 1 is always E_{10}; in experiments III and IV the traffic intensity at node 1 is always 0.8.

Shimshak's simulations were obtained using GPSS and the regenerative method. The number of customers simulated in each case ranged from 15,000 to 25,000, depending on the traffic intensity. The statistical reliability is indicated by 95% confidence intervals. The simulation results and the approximations for expected *total* waiting times appear in Tables 4.8–4.10. The Fraker, Page and Marchal approximations are approximations devised by Shimshak using earlier methods. All the numbers except QNET estimates are taken from Table 6–8 of [55]. The half width of the 95% confidence interval is in parentheses below the simulation estimates. The estimated relative error appears below the approximation in parentheses. It is the approximation value minus the simulation value divided by the simulation value. The blanks in the Fraker column in Table 4.9 mean his approximation failed in these cases. We take $n = 6$ for our QNET calculations. In Table 4.8, except for Case 7, our QNET estimates give very good approximations of the total waiting times. In case 7, the squared coefficient of variation of the interarrivals is $C_a^2 = 1$, which is big compared to the squared coefficient of variation of the service times at both stations, $C_{s_1}^2 = C_{s_2}^2 = 0.1$. Case 7 in Table 4.8 shows that our QNET method fails to catch the high variability of the arrival process to the second queue, and therefore underestimates the average waiting time at the second station. Table 4.9 shows that our QNET approximations are uniformly better than QNA approximations, whereas in Table 4.10, our QNET approximations fail to get good approximations at all. Notice that in Table 4.10, for Case 2, Case 4, Case 6 and Case 8, because $C_a^2 = C_{s_1}^2 = 0.1$, there is a product form formula to calculate the corresponding steady-state mean m. In other words, the QNET calculation for m is *exact*, and all errors come from the model replacement. Presumably, this is due to the fact that traffic intensities are not high enough, and Table 4.10 confirms this. The rows with the largest errors have very small waiting times. This suggests that more validation of QNET is needed.

$C_a^2 = 1, \quad C_{s_1}^2 = 0.1$								
case	ρ	$C_{s_2}^2$	SIM	Fraker	Page	Marchal	QNA	QNET*
1	0.6		1.20	1.19	1.20	1.18	1.25	1.19
	0.6	1.0	(0.09)	(−0.01)	(0.00)	(−0.02)	(0.04)	(−0.01)
2	0.8		2.27	2.30	2.31	2.28	2.38	2.34
	0.6	1.0	(0.23)	(0.01)	(0.02)	(0.01)	(0.05)	(0.03)
3	0.6		0.78	0.77	0.84	0.84	0.84	0.70
	0.6	0.1	(0.06)	(−0.02)	(0.08)	(0.07)	(0.08)	(−0.10)
4	0.8		1.83	1.90	1.99	1.98	1.98	1.87
	0.6	0.1	(0.22)	(0.04)	(0.09)	(0.08)	(0.08)	(0.02)
5	0.6		3.41	3.07	3.10	3.06	3.21	3.26
	0.8	1.0	(0.43)	(−0.10)	(−0.09)	(−0.10)	(−0.06)	(−0.04)
6	0.8		4.33	3.85	3.70	3.84	4.07	4.23
	0.8	1.0	(0.60)	(−0.14)	(−0.10)	(−0.11)	(−0.06)	(−0.02)
7	0.6		1.93	1.60	1.73	1.72	1.77	0.86
	0.8	0.1	(0.27)	(−0.17)	(−0.10)	(−0.11)	(−0.08)	(−0.55)
8	0.8		2.48	2.43	2.58	2.57	2.63	2.34
	0.8	0.1	(0.29)	(−0.02)	(0.04)	(0.04)	(0.06)	(−0.06)
Average relative error				−0.05	−0.01	−0.02	+0.01	−0.09
Average absolute relative error				0.06	0.07	0.07	0.06	0.11

* $n = 6$ for QNET estimates.

Table 4.8: Comparisons with Shimshak's Experiment I

$C_a^2 = 4, \quad \rho_1 = 0.8$							
case	ρ_2	C_s^2	SIM	Fraker	Marchal	QNA	QNET*
1	0.6	1.0 1.0	9.08 (1.38)	10.30 (0.13)	10.39 (0.14)	9.39 (0.03)	9.03 (−0.01)
2	0.6	0.1 1.0	6.49 (0.73)		7.91 (0.22)	7.72 (0.19)	7.17 (0.10)
3	0.6	1.0 0.1	8.55 (1.20)	10.17 (0.19)	9.86 (0.15)	8.98 (0.05)	8.57 (0.00)
4	0.6	0.1 0.1	6.01 (0.73)		7.43 (0.24)	7.31 (0.22)	6.67 (0.10)
5	0.8	1.0 1.0	12.31 (2.26)	13.50 (0.10)	13.54 (0.10)	12.92 (0.05)	12.77 (0.04)
6	0.8	0.1 1.0	9.64 (1.33)		10.78 (0.12)	10.67 (0.11)	10.13 (0.05)
7	0.8	1.0 0.1	11.13 (1.37)	12.55 (−0.13)	11.90 (0.07)	11.49 (0.03)	10.95 (−0.02)
8	0.8	0.1 0.1	7.40 (0.95)		9.21 (0.24)	9.23 (0.25)	7.86 (0.06)
Average relative error				+0.14	+0.16	+0.12	+0.04
Average absolute relative error				0.14	0.16	0.12	0.05

* $n = 6$ for QNET estimates.

Table 4.9: Comparisons with Shimshak's Experiment III

$C_a^2 = 0.1, \quad \rho_1 = 0.8$							
case	ρ_2	C_s^2	SIM	Fraker	Marchal	QNA	QNET*
1	0.6	1.0	2.30	2.24	2.21	2.29	2.55
		1.0	(0.19)	(−0.03	(−0.04)	(0.00)	(0.11)
2	0.6	0.1	0.59	0.58	0.59	0.56	0.82
		1.0	(0.04)	(−0.02)	(0.00)	(−0.05)	(0.39)
3	0.6	1.0	1.95	1.81	1.84	1.89	2.18
		0.1	(0.45)	(−0.07)	(−0.06)	(−0.03)	(0.12)
4	0.6	0.1	0.25	0.27	0.35	0.20	0.41
		0.1	(0.02)	0.08	(0.40)	(−0.20)	(0.65)
5	0.8	1.0	3.84	4.27	4.26	4.21	4.24
		1.0	0.33	(0.11)	(0.11)	(0.10)	(0.10)
6	0.8	0.1	1.82	1.77	1.75	1.85	2.08
		1.0	(0.19)	(−0.03)	(−0.04)	(0.02)	(0.14)
7	0.8	1.0	2.68	2.79	2.88	2.77	2.98
		0.1	(0.52)	(0.04)	(0.07)	(0.03)	(0.11)
8	0.8	0.1	0.46	0.50	0.58	0.43	0.64
		0.1	(0.22)	(0.09)	(0.26)	(−0.06)	(0.39)
Average relative error				+0.02	+0.09	−0.02	+0.25
Average absolute relative error				0.06	0.12	0.06	0.25

* $n = 6$ for QNET estimates.

Table 4.10: Comparisons with Shimshak's Experiment IV

Suresh and Whitt's Experiment

Recently, Suresh and Whitt [49] did extensive simulations for a network of two queues in tandem. Their objective is to find a best order of two queues to minimize the average *total* waiting time, or equivalently to minimize the average sojourn time. They also compare their simulations with QNA approximations. For the purpose of this section, we are interested in comparing simulation results and approximation results for the expected waiting time at the *second* station only, because, in both QNA and QNET, the expected waiting time at the first station is approximated by Kingman's heavy traffic formula

$$w_1 = \rho_1 \left(\frac{\rho_1}{1-\rho_1} \right) \left(\frac{C_a^2 + C_{s_1}^2}{2} \right),$$

which has been tested intensively and found to be relatively good, see [54]. If the departure process from station 1 were a renewal process and its squared coefficient of variation C_d^2 could be calculated in some way, Kingman's formula could be used again to yield an estimate for the expected waiting at the second station. Unfortunately, the arrival process to the second queue (or the departure process from the first queue) is typically not renewal, see [49] and Berman and Westcott [1]. It is this difficulty which is challenging queueing performance analysts, and Whitt's QNA is trying to overcome the difficulty. The QNET method now provides another way for us to estimate the expected waiting at the second station. We should emphasize that QNET can do more in this respect, that is, it gives an approximation of the joint densities of the waiting times at the two stations.

Suresh and Whitt used the SIMAN simulation program to get the simulation estimates of the expected steady-state waiting times. In each case, they performed ten independent replications, using 30,000 arrivals in each replication, and they estimated 90% confidence intervals using t-statistics. An initial portion of each run (2000 customers) was discarded to allow the system to approach steady state. They considered various variability parameter triples $(C_a^2, C_{s_1}^2, C_{s_2}^2)$ for all combinations of the traffic intensities ρ_1 and ρ_2 in a representative range. For $C_{s_1}^2 \neq C_{s_2}^2$ they considered five variability triples $(C_a^2, C_{s_1}^2, C_{s_2}^2)$, namely, $(0.5, 0.5, 2.0)$, $(1.0, 0.5, 8.0)$, $(1.0, 2.0, 4.0)$, $(4.0, 0.5, 1.0)$ and $(4.0, 1.0, 4.0)$, ordered lexicographically; we will refer to them as Case 1, Case 2, Case 3, Case 4 and Case 5, respectively. For $C_{s_1}^2 = C_{s_2}^2$, they considered two variability triples $(C_a^2, C_{s_1}^2, C_{s_2}^2)$, namely, $(1.0, 0.5, 0.5)$ and $(1.0, 4.0, 4.0)$; we refer to them as Case 6 and Case 7. When $C_{s_1}^2 \neq C_{s_2}^2$, for each queue they considered four values of ρ_i: $0.3, 0.6, 0.8, 0.9$. Thus, associated with each variability triple are 32 cases of two queues in series:

$$(4 \text{ values of } \rho_1)\times\ (4 \text{ values of } \rho_2)\times\ (2 \text{ orders}) = 32 \text{ cases.}$$

When $C_{s_1}^2 = C_{s_2}^2$, they consider five values of ρ_i: $0.1, 0.2, 0.3, 0.6$ and 0.9. Thus for each triple there are 15 cases of two queues in series (they consider only $\rho_1 \neq \rho_2$, because otherwise the order does not matter).

When $C^2 = 0.5$, the E_2 distribution (Erlang of order 2, the convolution of two exponentials) is used in the simulation; when $C^2 = 1.0$, the exponential (M) distribution is used; when $C^2 > 1$, the H_2 distribution (hyperexponential, a mixture of two exponentials) with balanced means is used. Tables 4.11–4.17 give simulation estimates, Whitt's QNA estimates and QNET estimates of the average waiting time at the second station for all seven cases and two orders. In all tables, "Order I" refers to the given triple, and "Order II" is the reverse order. "D_i" denotes the difference between the approximation and the simulation estimate. This difference is the minimum of the absolute difference and the absolute relative difference, where

$$\text{absolute relative difference} = \frac{|\text{approximation} - \text{simulation estimate}|}{\text{simulation estimate}}.$$

"D_1" stands for QNET difference, and "D_2" stands for QNA difference. We take $n = 5$ for all QNET estimates. The estimated 90% confidence interval appears below the simulation estimate in each case. From Table 4.11-4.17, we see that except for Table 4.14 (Case 5), the QNET estimates give very good approximations most of the time. In situation (Order II of Case 2), the overall QNET estimates are not best, but still under heavy traffic ($\rho_1, \rho_2 \geq 0.8$), the QNET estimates are much better than the QNA approximations. This confirms the so called "heavy traffic theory", which asserts that QNET estimates are "good" under heavy traffic. These Tables clearly demonstrate that even under moderate traffic intensities, the QNET estimates are still good, and Table 4.16-4.17 show that QNET performs surprising well under low traffic intensities.

In Table 4.14, QNET performs very badly when the traffic intensities are *not* heavy or are *unbalanced* heavy. In fact, from Table 4.14, we see that in "Order I" (similarly in "Order II", when $\rho_1 \geq \rho_2$), station 1 is heavier loaded than station 2. The high variability of the arrival process to station 1 is smoothed out to a certain extent, and QNET is able to catch the correct variability for the arrival process to the second station. When $\rho_1 < \rho_2$, QNET performs badly. At this point, we do not know exactly why this happens. Intuitively, the first station is lightly loaded, and service time variability is small, therefore, customers need almost no waiting at station 1, get serviced with low variability and go to the second queue.

Therefore, the high variability of the arrival process to station 1 is carried over to the arrival process to station 2, see Reiman [40]. Somehow, QNET fails to catch this high variability of the arrival process to the second station and significantly underestimates the expected waiting time at the second station. This again demonstrate the need for more validation of the QNET method. Table 4.18 aggregates all balanced heavy traffic cases and gives an overall comparison of QNET estimates with QNA estimates under heavy traffic. It is clear from this table that QNET approximations are much better than QNA approximations under balanced heavy traffic conditions.

		$(C_a^2 = 0.50, \quad C_{s_1}^2 = 0.50, \quad C_{s_2}^2 = 2.00)$									
		Order I					Order II				
ρ_1	ρ_2	QNET	QNA	SIM	D_1	D_2	QNET	QNA	SIM	D_1	D_2
0.9	0.9	10.12	10.13	10.010 0.512	0.01	0.01	7.75	8.97	7.946 0.430	0.02	0.13
0.9	0.8	4.00	4.00	3.856 0.219	0.04	0.04	6.58	7.94	6.292 0.368	0.05	0.26
0.9	0.6	1.12	1.12	1.039 0.035	0.08	0.08	5.07	6.24	4.976 0.260	0.02	0.25
0.9	0.3	0.16	0.16	0.127 0.003	0.03	0.03	4.17	4.60	4.309 0.154	0.03	0.07
0.8	0.9	10.12	10.13	10.153 0.670	0.00	0.00	3.48	3.54	3.444 0.116	0.01	0.03
0.8	0.8	4.00	4.00	3.706 0.158	0.08	0.08	3.07	3.14	2.895 0.085	0.06	0.08
0.8	0.6	1.12	1.12	1.068 0.030	0.05	0.05	2.46	2.46	2.374 0.042	0.04	0.04
0.8	0.3	0.16	0.16	0.129 0.005	0.03	0.03	1.74	1.82	1.670 0.044	0.04	0.09
0.6	0.9	10.12	10.13	9.687 0.786	0.05	0.05	1.08	1.00	0.966 0.024	0.12	0.03
0.6	0.8	4.00	4.00	4.007 0.125	0.00	0.00	1.00	0.88	0.912 0.011	0.09	0.03
0.6	0.6	1.12	1.12	1.072 0.024	0.05	0.05	0.86	0.69	0.736 0.020	0.13	0.04
0.6	0.3	0.16	0.16	0.133 0.002	0.03	0.03	0.63	0.51	0.491 0.006	0.14	0.02
0.3	0.9	10.12	10.13	9.562 0.719	0.06	0.06	0.16	0.14	0.136 0.002	0.03	0.01
0.3	0.8	4.00	4.00	4.024 0.272	0.01	0.01	0.16	0.13	0.129 0.002	0.03	0.00
0.3	0.6	1.12	1.12	1.090 0.032	0.03	0.03	0.15	0.10	0.112 0.001	0.04	0.01
0.3	0.3	0.16	0.16	0.131 0.003	0.03	0.03	0.12	0.07	0.075 0.001	0.05	0.00
Average Difference					0.04	0.04				0.06	0.07

* $n = 5$ for QNET estimates

Table 4.11: Expected Waiting Times at the Second Queue: Case 1

		$(C_a^2 = 1.00, \quad C_{s_1}^2 = 0.50, \quad C_{s_2}^2 = 8.00)$									
		Order I					Order II				
ρ_1	ρ_2	QNET	QNA	SIM	D_1	D_2	QNET	QNA	SIM	D_1	D_2
0.9	0.9	36.09	34.81	37.04 4.41	0.03	0.06	24.32	29.04	24.05 1.76	0.01	0.21
0.9	0.8	14.05	13.75	13.35 1.26	0.05	0.03	18.36	24.22	18.12 0.93	0.01	0.34
0.9	0.6	3.89	3.87	3.80 0.24	0.02	0.02	10.54	16.28	12.63 0.80	0.17	0.29
0.9	0.3	0.55	0.55	0.52 0.02	0.03	0.03	6.60	8.63	7.55 0.38	0.13	0.14
0.8	0.9	36.42	35.15	37.58 5.26	0.03	0.06	11.38	11.47	10.74 0.45	0.06	0.07
0.8	0.8	14.28	13.89	13.78 1.57	0.04	0.01	9.65	9.57	8.95 0.47	0.08	0.07
0.8	0.6	3.94	3.91	3.84 0.25	0.03	0.02	6.29	6.43	6.33 0.27	0.01	0.02
0.8	0.3	0.55	0.56	0.49 0.02	0.06	0.06	2.94	3.41	3.36 0.10	0.12	0.02
0.6	0.9	36.45	35.72	34.61 3.73	0.05	0.03	3.68	3.23	2.86 0.14	0.29	0.13
0.6	0.8	14.40	14.11	13.16 1.50	0.09	0.07	3.31	2.69	2.57 0.11	0.29	0.05
0.6	0.6	4.02	3.97	3.96 0.24	0.03	0.02	2.72	1.81	2.07 0.04	0.31	0.13
0.6	0.3	0.56	0.57	0.52 0.02	0.04	0.05	1.38	0.96	1.11 0.03	0.24	0.14
0.3	0.9	36.45	36.27	31.12 4.44	0.17	0.17	0.55	0.46	0.24 0.01	0.31	0.22
0.3	0.8	14.40	14.34	13.33 0.60	0.08	0.07	0.54	0.38	0.25 0.00	0.29	0.13
0.3	0.6	4.05	4.03	4.10 0.25	0.01	0.02	0.50	0.26	0.27 0.01	0.23	0.01
0.3	0.3	0.57	0.58	0.55 0.02	0.02	0.03	0.39	0.14	0.20 0.00	0.18	0.06
Average Difference					0.05	0.05				0.16	0.12

$n = 5$ for QNET estimates

Table 4.12: Expected Waiting Times at the Second Queue: Case 2

		$(C_a^2 = 1.00, \quad C_{s_1}^2 = 2.00, \quad C_{s_2}^2 = 4.00)$									
		Order I					Order II				
ρ_1	ρ_2	QNET	QNA	SIM	D_1	D_2	QNET	QNA	SIM	D_1	D_2
0.9	0.9	22.03	23.53	23.026 2.776	0.06	0.02	19.13	21.99	18.447 1.520	0.04	0.19
0.9	0.8	9.07	9.30	8.887 0.544	0.02	0.05	16.23	19.93	16.562 1.427	0.02	0.20
0.9	0.6	2.64	2.61	2.829 0.116	0.07	0.08	13.00	16.52	14.538 1.192	0.11	0.14
0.9	0.3	0.39	0.37	0.402 0.008	0.02	0.03	12.22	13.24	12.967 1.738	0.06	0.02
0.8	0.9	20.92	22.84	23.698 2.050	0.12	0.04	8.45	8.69	8.981 0.426	0.06	0.03
0.8	0.8	8.70	9.02	8.686 0.378	0.00	0.04	7.57	7.87	7.747 0.376	0.02	0.02
0.8	0.6	2.57	2.54	2.577 0.079	0.00	0.01	5.95	6.53	6.438 0.343	0.08	0.01
0.8	0.3	0.38	0.36	0.399 0.004	0.02	0.04	4.88	5.23	4.896 0.146	0.00	0.07
0.6	0.9	20.31	21.71	21.969 2.683	0.07	0.01	2.61	2.44	2.618 0.067	0.00	0.07
0.6	0.8	8.13	8.58	8.131 0.311	0.00	0.05	2.43	2.21	2.352 0.079	0.03	0.06
0.6	0.6	2.45	2.41	2.485 0.086	0.02	0.03	2.13	1.84	2.012 0.063	0.06	0.09
0.6	0.3	0.37	0.34	0.368 0.015	0.00	0.02	1.52	1.47	1.566 0.044	0.03	0.06
0.3	0.9	20.25	20.61	21.480 3.439	0.06	0.04	0.39	0.35	0.345 0.009	0.04	0.00
0.3	0.8	8.01	8.14	8.408 0.472	0.05	0.03	0.38	0.32	0.345 0.008	0.04	0.03
0.3	0.6	2.26	2.29	2.337 0.081	0.03	0.02	0.36	0.26	0.321 0.009	0.04	0.06
0.3	0.3	0.35	0.33	0.340 0.007	0.01	0.01	0.30	0.21	0.248 0.005	0.06	0.04
Average Difference					0.03	0.03				0.04	0.07

* $n = 5$ for QNET estimates

Table 4.13: Expected Waiting Times at the Second Queue: Case 3

		$(C_a^2 = 4.00, \quad C_{s_1}^2 = 0.50, \quad C_{s_2}^2 = 1.00)$									
		Order I					Order II				
ρ_1	ρ_2	QNET	QNA	SIM	D_1	D_2	QNET	QNA	SIM	D_1	D_2
0.9	0.9	10.73	8.77	12.541 1.650	0.14	0.30	9.56	8.38	11.788 2.949	0.19	0.29
0.9	0.8	3.00	3.46	3.325 0.102	0.10	0.04	9.43	10.45	15.032 1.716	0.37	0.30
0.9	0.6	0.69	0.97	0.725 0.016	0.03	0.25	4.12	13.85	16.603 1.426	0.75	0.17
0.9	0.3	0.10	0.14	0.076 0.001	0.02	0.06	3.12	17.13	19.682 2.000	0.84	0.13
0.8	0.9	7.70	11.18	16.269 2.358	0.53	0.31	2.87	3.31	3.130 0.155	0.08	0.06
0.8	0.8	4.17	4.42	4.803 0.398	0.13	0.08	3.75	4.13	4.114 0.240	0.09	0.00
0.8	0.6	0.80	1.24	0.907 0.026	0.11	0.33	3.21	5.47	6.147 0.526	0.48	0.11
0.8	0.3	0.10	0.18	0.088 0.001	0.01	0.09	1.20	6.77	6.941 0.387	0.83	0.02
0.6	0.9	3.41	15.15	18.706 2.371	0.82	0.19	0.69	0.93	0.788 0.020	0.10	0.14
0.6	0.8	2.62	5.98	6.987 0.717	0.62	0.14	0.77	1.16	0.934 0.018	0.17	0.23
0.6	0.6	1.17	1.68	1.471 0.039	0.21	0.14	1.05	1.54	1.302 0.018	0.19	0.18
0.6	0.3	0.10	0.24	0.129 0.002	0.02	0.11	0.75	1.90	1.755 0.049	0.57	0.08
0.3	0.9	2.11	18.97	20.014 2.027	0.89	0.05	0.10	0.13	0.105 0.001	0.01	0.03
0.3	0.8	0.96	7.50	8.165 0.492	0.88	0.08	0.10	0.17	0.117 0.002	0.02	0.05
0.3	0.6	0.62	2.11	2.031 0.051	0.70	0.04	0.10	0.22	0.142 0.002	0.04	0.08
0.3	0.3	0.17	0.30	0.205 0.005	0.03	0.10	0.15	0.27	0.178 0.002	0.03	0.09
Average Difference					0.33	0.15				0.30	0.12

* $n = 5$ for QNET estimates

Table 4.14: Expected Waiting Times at the Second Queue: Case 4

		$(C_a^2 = 4.00, \quad C_{s_1}^2 = 1.00, \quad C_{s_2}^2 = 4.00)$									
		Order I					Order II				
ρ_1	ρ_2	QNET	QNA	SIM	D_1	D_2	QNET	QNA	SIM	D_1	D_2
0.9	0.9	26.11	22.56	25.776 2.604	0.01	0.12	20.25	20.25	20.453 1.104	0.01	0.01
0.9	0.8	9.14	8.91	10.105 0.854	0.10	0.12	20.25	20.25	25.409 2.870	0.20	0.20
0.9	0.6	2.32	2.51	2.652 0.130	0.13	0.05	20.25	20.25	23.394 2.348	0.13	0.13
0.9	0.3	0.32	0.36	0.357 0.012	0.04	0.00	20.25	20.25	21.580 2.733	0.06	0.06
0.8	0.9	29.08	24.62	33.220 5.395	0.12	0.26	8.00	8.00	9.201 0.534	0.13	0.13
0.8	0.8	10.30	9.73	11.360 0.723	0.09	0.14	8.00	8.00	8.455 0.535	0.05	0.05
0.8	0.6	2.51	2.74	2.868 0.188	0.13	0.05	8.00	8.00	8.643 0.441	0.07	0.07
0.8	0.3	0.32	0.39	0.399 0.009	0.08	0.01	8.00	8.00	7.991 0.327	0.00	0.00
0.6	0.9	31.58	28.03	31.279 3.484	0.01	0.10	2.25	2.25	2.298 0.063	0.00	0.00
0.6	0.8	11.89	11.07	12.942 1.016	0.08	0.14	2.25	2.25	2.443 0.099	0.08	0.08
0.6	0.6	2.90	3.11	3.389 0.150	0.14	0.08	2.25	2.25	2.523 0.045	0.11	0.11
0.6	0.3	0.34	0.44	0.444 0.011	0.10	0.00	2.25	2.25	2.316 0.054	0.03	0.03
0.3	0.9	32.35	31.31	27.840 1.955	0.16	0.12	0.32	0.32	0.269 0.005	0.05	0.05
0.3	0.8	12.73	12.37	13.667 1.304	0.07	0.10	0.32	0.32	0.286 0.005	0.04	0.04
0.3	0.6	3.48	3.48	3.611 0.226	0.04	0.04	0.32	0.32	0.330 0.009	0.01	0.01
0.3	0.3	0.41	0.50	0.528 0.020	0.11	0.03	0.32	0.32	0.328 0.006	0.01	0.01
Average Difference					0.065	0.086				0.063	0.063

* $n = 5$ for QNET estimates

Table 4.15: Expected Waiting Times at the Second Queue: Case 5

		$(C_a^2 = 1.00, \quad C_{s_1}^2 = 0.50, \quad C_{s_2}^2 = 0.50)$									
		Order I					Order II				
ρ_1	ρ_2	QNET	QNA	SIM	D_1	D_2	QNET	QNA	SIM	D_1	D_2
0.9	0.9	4.84	4.43	4.793 0.269	0.01	0.08	4.84	4.43	4.910 0.215	0.01	0.10
0.9	0.6	0.46	0.49	0.410 0.006	0.05	0.08	5.54	5.35	5.658 0.410	0.02	0.05
0.9	0.3	0.06	0.07	0.043 0.001	0.02	0.03	6.01	5.89	5.635 0.293	0.07	0.05
0.9	0.2	0.03	0.03	0.013 0.000	0.01	0.02	6.05	5.99	6.180 0.442	0.02	0.03
0.9	0.1	0.01	0.01	0.002 0.000	0.00	0.01	6.07	6.05	6.047 0.253	0.02	0.00
0.6	0.6	0.54	0.59	0.552 0.008	0.02	0.04	0.54	0.59	0.552 0.009	0.02	0.04
0.6	0.3	0.07	0.08	0.057 0.001	0.01	0.02	0.61	0.65	0.644 0.011	0.04	0.01
0.6	0.2	0.03	0.03	0.018 0.000	0.01	0.01	0.64	0.67	0.669 0.011	0.03	0.00
0.6	0.1	0.01	0.01	0.003 0.000	0.00	0.01	0.67	0.67	0.668 0.015	0.00	0.00
0.3	0.3	0.08	0.09	0.079 0.001	0.00	0.01	0.08	0.09	0.079 0.001	0.00	0.01
0.3	0.2	0.03	0.04	0.026 0.000	0.00	0.01	0.08	0.10	0.087 0.001	0.00	0.01
0.3	0.1	0.01	0.01	0.004 0.000	0.00	0.01	0.09	0.10	0.094 0.001	0.01	0.01
0.2	0.2	0.03	0.04	0.031 0.000	0.00	0.01	0.03	0.04	0.031 0.001	0.00	0.01
0.2	0.1	0.01	0.01	0.005 0.000	0.00	0.01	0.03	0.04	0.035 0.001	0.00	0.00
0.1	0.1	0.01	0.01	0.004 0.000	0.00	0.01	0.01	0.01	0.006 0.000	0.00	0.00
Average Difference					0.01	0.02				0.02	0.02

* $n = 5$ for QNET estimates

Table 4.16: Expected Waiting Times at the Second Queue: Case 6

		$(C_a^2 = 1.00, \quad C_{s_1}^2 = 4.00, \quad C_{s_2}^2 = 4.00)$									
		Order I					Order II				
ρ_1	ρ_2	QNET	QNA	SIM	D_1	D_2	QNET	QNA	SIM	D_1	D_2
0.9	0.9	26.43	30.09	25.505 2.955	0.04	0.18	26.43	30.09	28.072 2.921	0.06	0.07
0.9	0.6	3.47	3.34	3.405 0.144	0.02	0.02	20.57	24.62	22.857 4.542	0.10	0.08
0.9	0.3	0.51	0.48	0.530 0.019	0.02	0.05	20.27	21.34	20.511 1.629	0.01	0.04
0.9	0.2	0.20	0.19	0.199 0.005	0.00	0.01	20.26	20.74	19.132 2.606	0.06	0.08
0.9	0.1	0.04	0.04	0.046 0.001	0.00	0.00	20.25	20.37	20.067 1.768	0.01	0.02
0.6	0.6	2.94	2.74	2.874 0.103	0.02	0.05	2.94	2.74	2.918 0.116	0.01	0.06
0.6	0.3	0.48	0.39	0.457 0.013	0.02	0.07	2.32	2.37	2.494 0.114	0.07	0.05
0.6	0.2	0.20	0.15	0.188 0.004	0.01	0.04	2.27	2.30	2.231 0.088	0.02	0.03
0.6	0.1	0.04	0.03	0.045 0.002	0.00	0.01	2.25	2.26	2.271 0.135	0.01	0.00
0.3	0.3	0.42	0.34	0.378 0.011	0.04	0.04	0.42	0.34	0.371 0.009	0.05	0.03
0.3	0.2	0.17	0.13	0.153 0.006	0.02	0.02	0.39	0.33	0.356 0.006	0.03	0.03
0.3	0.1	0.04	0.03	0.039 0.002	0.00	0.01	0.34	0.32	0.346 0.012	0.01	0.02
0.2	0.2	0.16	0.13	0.141 0.003	0.02	0.01	0.16	0.13	0.145 0.006	0.02	0.02
0.2	0.1	0.04	0.03	0.034 0.002	0.01	0.01	0.15	0.13	0.134 0.003	0.01	0.01
0.1	0.1	0.04	0.03	0.034 0.002	0.00	0.01	0.04	0.03	0.033 0.003	0.00	0.01
Average Difference					0.015	0.035				0.031	0.037

* $n = 5$ for QNET estimates

Table 4.17: Expected Waiting Times at the Second Queue: Case 7

			Order I		Order II	
Case	ρ_1	ρ_2	QNET-D	QNA-D	QNET-D	QNA-D
1	0.9	0.9	0.01	0.01	0.02	0.13
	0.8	0.8	0.08	0.08	0.06	0.08
2	0.9	0.9	0.03	0.06	0.01	0.21
	0.8	0.8	0.04	0.01	0.08	0.07
3	0.9	0.9	0.06	0.02	0.04	0.19
	0.8	0.8	0.00	0.04	0.02	0.02
4	0.9	0.9	0.01	0.12	0.01	0.01
	0.8	0.8	0.09	0.14	0.05	0.05
5	0.9	0.9	0.14	0.30	0.20	0.29
	0.8	0.8	0.13	0.08	0.08	0.00
6	0.9	0.9	0.01	0.08	0.01	0.10
	0.6	0.6	0.02	0.04	0.02	0.04
7	0.9	0.9	0.04	0.18	0.06	0.07
	0.6	0.6	0.02	0.05	0.01	0.06
Average			0.05	0.09	0.05	0.10

Table 4.18: Overall Comparisons with QNA Approximations in Heavy Traffic

4.6.2 Analysis of a Multiclass Queueing Network

The two-station open queueing network pictured in Figure 4.2 has been suggested by Gelenbe and Pujolle [16] as a simplified model of a certain computer system. Server 1 represents a central processing unit (CPU) and server 2 a secondary memory. There are two classes of programs (jobs, or customers) flowing through the system, and they differ in their relative use of the CPU and the secondary memory. Jobs of class j ($j = 1, 2$) arrive at station 1 according to a Poisson process with rate α; and after completing service there they may either go on to station 2 (probability q_j) or leave the system (probability $1 - q_j$); each service at station 2 is followed by another service at station 1, after which the customer either return to station 2 or else leaves the system, again with probability q_j and $1 - q_j$, respectively. The service time distribution for class j customers at station i ($i, j = 1, 2$) is the same on every visit; its mean is τ_{ij} and its coefficient of variation (standard deviation divided by mean) is C_{ij}. Customers are served on a first-in-first-out basis, without regard to class, at each station. The specific numerical values that we will consider are such that class 1 makes heavier demands on the secondary memory but class 2 consumes more CPU time. Denoting by Q_i ($i = 1, 2$) the long-run average queue length at station i, including the customer being served there (if any), our goal is to estimate Q_1 and Q_2.

This open queueing network is within the class for which Harrison and Nguyen [24] have proposed an approximate Brownian model, but their initial focus is on the *current workload process* or *virtual waiting time process* $W(t) = (W_1(t), W_2(t))'$, rather than the queue length process; one may think of $W_i(t)$ as the time that a new arrival to station i at time t would have to wait before gaining access to the server. Harrison and Nguyen proposed that the process $W(t)$ be modeled or approximated by an RBM in the quadrant whose data (Γ, μ, R)

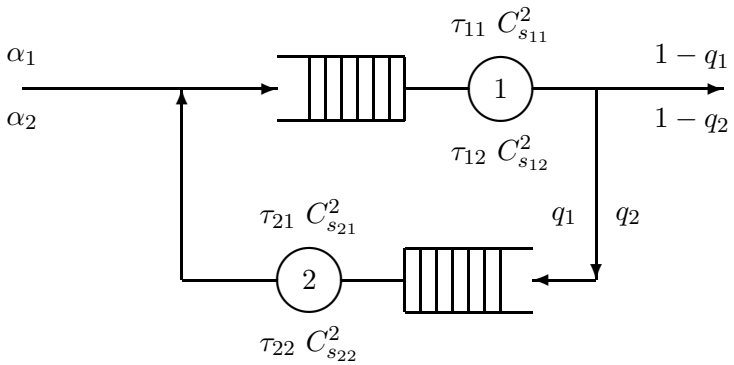

Figure 4.2: Model of an Interactive Computer

are derived from the parameters of the queueing system by certain formulas. Specializing those formulas to the case at hand one obtains

$$\mu = R(\rho - e), \quad \Gamma = TGT' \quad \text{and} \quad R = M^{-1}, \tag{4.30}$$

where e is the two-vector of ones, $\rho = (\rho_1, \rho_2)'$ is the vector of "traffic intensities"

$$\rho_1 = \frac{\alpha_1}{1-q_1}\tau_{11} + \frac{\alpha_2}{1-q_2}\tau_{12} \quad \text{and} \quad \rho_2 = \frac{\alpha_1 q_1}{1-q_1}\tau_{21} + \frac{\alpha_2 q_2}{1-q_2}\tau_{22},$$

and the matrices M, T and G are given by:

$$M = \begin{pmatrix} \frac{1}{\rho_1}F_{11} & \frac{1}{\rho_2}F_{12} \\ \\ \frac{1}{\rho_1}F_{21} & \frac{1}{\rho_2}F_{22} \end{pmatrix},$$

where

$$F = \begin{pmatrix} \frac{\alpha_1\tau_{11}}{(1-q_1)^2} + \frac{\alpha_2\tau_{12}}{(1-q_2)^2} & \frac{\alpha_1\tau_{11}q_1}{(1-q_1)^2} + \frac{\alpha_2\tau_{12}q_2}{(1-q_2)^2} \\ \\ \frac{\alpha_1\tau_{21}q_1}{(1-q_1)^2} + \frac{\alpha_2\tau_{22}q_2}{(1-q_2)^2} & \frac{\alpha_1\tau_{21}q_1}{(1-q_1)^2} + \frac{\alpha_2\tau_{22}q_2}{(1-q_2)^2} \end{pmatrix};$$

$$T = \begin{pmatrix} \frac{\tau_{11}}{1-q_1} & \frac{\tau_{11}}{1-q_1} & \frac{\tau_{12}}{1-q_2} & \frac{\tau_{12}}{1-q_2} \\ \\ \frac{\tau_{21}q_1}{1-q_1} & \frac{\tau_{21}}{1-q_1} & \frac{\tau_{22}q_2}{1-q_2} & \frac{\tau_{22}}{1-q_2} \end{pmatrix};$$

and

$$G = \begin{pmatrix} \alpha_1 + \frac{\alpha_1}{1-q_1}g_1 & -\frac{\alpha_1 q_1}{1-q-1}g_{12} & 0 & 0 \\ \\ -\frac{\alpha_1 q_1}{1-q-1}g_{12} & \alpha_1 q_1 + \frac{\alpha_1 q_1}{1-q_1}g_2 & 0 & 0 \\ \\ 0 & 0 & \alpha_2 + \frac{\alpha_2}{1-q_2}g_3 & -\frac{\alpha_1 q_2}{1-q_2}g_{34} \\ \\ 0 & 0 & -\frac{\alpha_1 q_2}{1-q_2}g_{34} & \alpha_2 q_2 + \frac{\alpha_2 q_2}{1-q_2}g_4 \end{pmatrix},$$

$\alpha_1 = 0.5,\ \alpha_2 = 0.25,\ q_1 = 0.5,\ q_2 = 0.2$								
	class 1				class 2			
	station 1		station 2		station 1		station 2	
case	mean	SCV	mean	SCV	mean	SCV	mean	SCV
1	0.5	1.0	0.5	2.0	1.0	0.0	1.0	1.0
2	0.5	0.2	0.5	2.0	1.0	0.0	1.0	1.0
3	0.5	1.0	1.0	1.0	0.5	1.0	1.0	1.0
4	0.5	3.0	0.5	2.0	0.5	0.0	1.0	1.0
5	0.5	3.0	0.5	1.0	0.5	0.0	1.0	0.2

Table 4.19: Parameters for the Multiclass Queueing Network

where

$$g_1 = \left(C_{s_{11}}^2 + q_1 C_{s_{21}}^2\right), \qquad g_2 = \left(C_{s_{21}}^2 + q_1 C_{s_{11}}^2\right), \qquad g_{12} = \left(C_{s_{11}}^2 + C_{s_{21}}^2\right),$$
$$g_3 = \left(C_{s_{12}}^2 + q_2 C_{s_{22}}^2\right), \qquad g_4 = \left(C_{s_{22}}^2 + q_2 C_{s_{12}}^2\right), \qquad g_{34} = \left(C_{s_{12}}^2 + C_{s_{22}}^2\right).$$

Let us denote by $m = (m_1, m_2)'$ the mean vector of the stationary distribution of the RBM whose data (Γ, μ, R) are computed via (4.30). In the approximation scheme of Harrison and Nguyen [24], which they call *the QNET method*, one approximates by m_i both the long-run average *virtual* waiting time and the long-run average *actual* waiting time at station i $(i = 1, 2)$. By Little's law $(L = \lambda W)$, we then have the following *QNET estimates* of the average queue length at the two stations:

$$Q_1 = \rho_1 + \left(\frac{\alpha_1}{1 - q_1} + \frac{\alpha_2}{1 - q_2}\right) m_1 \tag{4.31}$$

$$Q_2 = \rho_2 + \left(\frac{\alpha_1 q_1}{1 - q_1} + \frac{\alpha_2 q_2}{1 - q_2}\right) m_2 \tag{4.32}$$

Gelenbe and Pujolle [16] have simulated the performance of this simple queueing network in the five different cases described by Table 4.19,obtaining the results displayed in Table 4.20. All of the numerical results in the latter table except the QNET estimates are taken from Table 5.3 of [16]: the row labelled "SIM" gives simulation results, whereas the row labelled "TD" gives a "time division" approximation based on the classical theory of product-form queueing network, and that labelled "DC" gives a "diffusion approximation" that is essentially Whitt's [54] QNA scheme for two-moment analysis of system performance via "node decomposition". In essence, this last method uses a diffusion approximation to the

	case 1		case 2		case 3		case 4		case 5	
ρ	0.81	0.31	0.81	0.31	0.66	0.56	0.66	0.31	0.66	0.31
Q	Q_1	Q_2	Q_1	Q_2	Q_1	Q_2	Q_1	Q_2	Q_1	Q_2
SIM	3.62	0.49	3.15	0.45	1.92	1.24	2.35	0.51	2.37	0.50
TD	4.33	0.45	4.33	0.45	1.90	1.29	1.91	0.45	1.91	0.45
DC	3.66	0.50	2.91	0.49	1.90	1.29	2.77	0.53	2.77	0.50
QNET	3.83	0.50	3.40	0.49	1.90	1.29	2.48	0.54	2.41	0.48

Table 4.20: Mean Number of Customers for the Network Represented in Figure 4.2

queue length process of each station individually, after artificially decomposing the network into one-station subnetworks; the QNET method captures more subtle system interactions by considering the joint stationary distribution of an approximating *two-dimensional* diffusion process, the mean vector m of that stationary distribution being computed by means of the algorithm described in Chapter 4.

As Table 4.20 indicates, our QNET method gives very good approximations, somewhat better overall than either the TD or DC approximations. The network described by case 3 is in fact a "product form network", and for it all these approximation schemes give exact results.

Appendix A

Detailed Description of the Algorithm

In this appendix, we will give a detailed description of one version of the general algorithm described in Chapter 4. This version has been implemented in a software package, called BNET, to calculate the stationary density of a *semimartingale Reflected Brownian motion* (SRBM) in an *orthant* $\mathbb{R}_+^d$. BNET will be the kernel of a larger software package, tentatively called QNET, that can be used for performance analysis for a wide class of queueing network models. This implementation of BNET is by no means "optimal". However, the numerical results in Sections 4.5–4.6 show that it performs quite well.

The actual code of BNET for this implementation is in "C". In the first four sections, we will give a description of the algorithm that is independent of any particular language, and in Section 5 we will give the "C" code for most essential functions (routines) in BNET.

A.1 Input Parameters and Scaling

Inputs

The input parameters to BNET are (d, Γ, μ, R, n), where

a) d is a positive integer. It is the dimension of the state space of an SRBM;

b) Γ is a $d \times d$ positive definite matrix. It is the covariance matrix of an SRBM;

c) μ is a d-dimensional vector. It is the drift vector of an SRBM;

d) R is a $d \times d$ matrix with $\text{diag}(R) > 0$. It is the reflection matrix of an SRBM;

e) n is a positive integer. It is the maximum degree of polynomials that we used in (4.23).

In input mode, BNET will exit immediately and send a corresponding error message to a user, if one of the following happens:

i) d or n is not a positive integer;

ii) Γ is not symmetric or not positive definite;

iii) diag(R) $\not> 0$, SRBM does not exist;

iv) R is *not* invertible, stationary distribution does not exist;

v) $R^{-1}\mu \not< 0$, stationary distribution does not exist.

Scaling

After all the input parameters have passed the five tests above, we then scale parameters (Γ, μ, R) as follows.

$$\Gamma^*_{ij} = \frac{\Gamma_{ij}}{\sqrt{\Gamma_{ii}}\sqrt{\Gamma_{jj}}}, \quad R^*_{ij} = \frac{R_{ij}}{R_{jj}}\frac{\sqrt{\Gamma_{jj}}}{\sqrt{\Gamma_{ii}}}, \quad \mu^*_i = \frac{\mu_i}{\sqrt{\Gamma_{ii}}\gamma_{\max}},$$

where $\gamma_{\max}$ is the largest component of

$$\gamma \equiv -R^{*-1}\begin{pmatrix} \mu_1/\sqrt{\Gamma_{11}} \\ \vdots \\ \mu_d/\sqrt{\Gamma_{dd}} \end{pmatrix} > 0.$$

The idea of the scaling above is to obtain a (Γ^*, μ^*, R^*)-SRBM with $\Gamma^*_{ii} = 1$, $R^*_{ii} = 1$ and $\max_{1\le i\le d}\gamma^*_i = 1$, where

$$\gamma^* \equiv -R^{*-1}\mu^* = \frac{\gamma}{\gamma_{\max}}.$$

We believe that using this *normalized* set of data (Γ^*, μ^*, R^*) will make BNET more robust. The logic of the scaling was explained in Proposition 4.4. Suppose that m^* is the steady-state mean of a (Γ^*, μ^*, R^*)-SRBM. Then we have the steady state mean $m = (m_1, \ldots, m_d)'$ of the SRBM Z given by

$$m_i = \frac{\sqrt{\Gamma_{ii}}}{\gamma_{\max}} m^*_i, \quad i = 1, 2, \ldots, d. \tag{A.1}$$

The memory spaces of (Γ, μ, R) can be used to store the *converted* data (Γ^*, μ^*, R^*). The spaces of off-diagonal elements of Γ can be used to store the corresponding element of Γ^*. Since the diagonal elements of Γ^* are one, the spaces of diagonal elements of Γ still keep *old* values, i.e., Γ_{ii}'s. The old Γ_{ii} values are needed in (A.1). Similarly, the spaces of μ and R can be used to store the corresponding element of μ^* and R^*. In the mean time, we need additional spaces to keep the constant $\gamma_{\max}$, which is needed in (A.1).

From now on, $(d, \Gamma, \mu, R, n, \gamma_{\max})$ refers to *converted* data (with the asterisk omitted), instead of original input data. Notice that $\gamma_{\max}$ and the diagonal elements of Γ are not accessed until the final output routine, see (A.1).

A.2 Indexing

In light of the definition of H_n in (4.23), we need a systematic and efficient way to enumerate the elements of the following family of functions:

$$\left\{ x_1^{i_1} x_2^{i_2} \cdots x_d^{i_d} : 0 < i_1 + i_2 + \cdots + i_d \le n,\ i_j\text{'s are non-negative integers} \right\}.$$

It is equivalent to enumerate the elements of following subset of Z_+^d:

$$I(d, n) \equiv \{ (i_1, i_2, \ldots, i_d) : i_1 + i_2 + \cdots + i_d \le n,\ i_j\text{'s are non-negative integers} \}.$$

First, we have $\#I(d, n) = C_{n+d}^d \equiv (d+1)(d+2)\cdots(d+n)/n!$. To see this, suppose that there are n balls and d vertical bars. Mix these $d+n$ objects, and arrange them on a straight line. For each such arrangement, let i_1 be the number of balls to the left of the first bar, and i_2 be the number of balls between the first bar and the second bar. Similarly we can define i_3, ..., i_d. Obviously, such an $(i_1, \ldots, i_d) \in I(d, n)$, and each arrangement corresponds exactly one element in $I(d, n)$. The total number of such arrangements is C_{d+n}^d (Balls and bars are assumed to be indistinguishable). Therefore, we have shown $\#I(d, n) = C_{d+n}^d$. Since the dimensions of these sets I are used frequently, we can pre-compute them and store them in a $d \times (n+1)$ matrix C, i.e.,

$$C_{lk} \equiv \#I(l, k) = C_{l+k}^l = \frac{(l+1)(l+2)\cdots(l+k)}{k!}, \quad 1 \le l \le d, 0 \le k \le n.$$

Now for an $I = (i_1, i_2, \ldots, i_d) \in I(d, n)$, we are going to map it into an integer, denoted as $\mathbf{Index}(d, I)$, between 1 and C_{dn}. In the future, we are going to access this I through the

integer. The function **Index()** is defined as follows:

$$\begin{aligned}\textbf{Index}(d, \mathbf{I}) \equiv\ & \#I(d, i_1 + i_2 + \cdots + i_d - 1) + \\ & + \#I(d-1, i_2 + \cdots + i_d - 1) + \cdots + \\ & + \#I(1, i_d - 1) + 1.\end{aligned} \tag{A.2}$$

From this definition, we have $\textbf{Index}(d, (0, \ldots, 0)) = 1$ and $\textbf{Index}(d, (0, \ldots, 0, n)) = C_{dn}$.

Conversely, given an $1 \le i \le C_{dn}$, we need to find the unique $I = (i_1, i_2, \ldots, i_d) \in I(d, n)$ so that $i = \textbf{Index}(d, I)$. Such an I is going to be denoted as $\textbf{InverseIndex}(d, i)$, or simply I_i. We are going to add one more component for I_i, i.e.,

$$I_i = (i_0, i_1, \ldots, i_d), \quad \text{with} \quad i_0 = i_1 + \ldots + i_d.$$

Keeping track of i_0 sometimes makes our description easier as we will see shortly. Let k_d be the smallest k such that

$$\#I(d, k-1) < i \le \#I(d, k),$$

and k_{d-1} be the smallest k such that

$$\#I(d-1, k-1) < i - \#I(d, k_d - 1) \le \#I(d-1, k).$$

Inductively, we can define k_{d-2}, ..., and k_1, where k_1 is the smallest k such that

$$\#I(1, k-1) < i - \#I(d, k_d - 1) - \#I(d-1, k_{d-1} - 1) + \ldots + \#I(2, k_2 - 1) \le \#I(d-1, k).$$

We can use either *linear* search or *binary* search to locate above k. Having found (k_d, k_{d-1}, ..., k_1), we define

$$\begin{gathered} i_0 = k_d, \quad i_1 = k_d - k_{d-1}, \quad i_2 = k_{d-1} - k_{d-2} \\ \vdots \\ i_{d-1} = k_2 - k_1, \quad i_d = k_1. \end{gathered}$$

Then $I = (i_0, i_1, \ldots, i_d)$ has the desired property. Again, like C's, these I's can be pre-computed and stored in a $C_{dn} \times (d+1)$ matrix I, i.e., for $i = 1, \ldots, C_{dn}$,

$$I_i = (I_{i0}, I_{i1}, \ldots, I_{id}).$$

A.3 Basis and Inner Product

Natural Basis of H_n

Now we are able to enumerate the polynomials in H_n as $\mathcal{A}f_2$, ..., $\mathcal{A}f_{C_{nd}}$. The typical element of f_i's looks like

$$f_i(x_1, x_2, \ldots, x_d) = x_1^{I_{i1}} x_2^{I_{i2}} \cdots x_d^{I_{id}}.$$

For $2 \leq i \leq C_{dn}$, the first step is to store $\mathcal{A}f_i$. It has two parts, interior part which is equal to Gf_i and boundary part which is equal to $(D_1 f_i, \ldots, D_d f_i)$. Function Gf_i is a polynomial of degree $I_{i0} - 1$ with d variables and $D_l f_i$ is a polynomial of degree $I_{i0} - 1$ with $d - 1$ variable $(l = 1, \ldots, d)$. Following the enumeration scheme in the previous section, only the coefficients of these polynomials need to be stored. For each i, we enumerate the coefficients of Gf_i from $1, \ldots, C_{d(I_{i0}-1)}$. For $j = 1, 2, \ldots, d$,

a. if $I_{ij} \geq 1$,

$$Gf_i(I_{i1}, \ldots, I_{i(j-1)}, I_{ij} - 1, I_{i(j+1)}, \ldots, I_{id}) \equiv I_{ij}\mu_j;$$

b. if $I_{ij} \geq 2$, because $\Gamma_{jj} = 1$,

$$Gf_i(I_{i1}, \ldots, I_{i(j-1)}, I_{ij} - 2, I_{i(j+1)}, \ldots, I_{id}) \equiv \frac{1}{2} I_{ij}(I_{ij} - 1);$$

c. if $I_{ij} \geq 1$ & $I_{ik} \geq 1$, $k = j + 1, \ldots, d$,

$$Gf_i(I_{i1}, \ldots, , I_{ij} - 1, \ldots, I_{ik} - 1, \ldots, I_{id}) \equiv I_{ij} I_{ik} \Gamma_{jk};$$

d. for all the other cases of $(i_1, \ldots, i_d)$ with $i_1 + \cdots + i_d \leq I_{i0} - 1$,

$$Gf_i(i_1, \cdots, i_d) = 0.$$

Similarly, for $l = 1, 2, \ldots, d$, we enumerate the coefficients of $D_l f_i$ from $1, \ldots, C_{(d-1)(I_{i0}-1)}$. $D_l f_i$ has only $d - 1$ variable because $x_l = 0$. For $j = 1, 2, \ldots, d$,

e. if $I_{il} = 1$,

$$D_l f_i(I_{i1}, \ldots, \hat{I}_{il}, I_{id}) \equiv R_{ll};$$

f. if $I_{il} = 0$ & $I_{ij} \geq 1$,

$$D_l f_I(I_{i1}, \ldots, I_{ij} - 1, \ldots, \hat{I}_{il}, \ldots, I_{id}) \equiv I_{ij} R_{jl};$$

g. all the other cases

$$D_l f_i(i_1, \ldots, \hat{i_l}, \ldots, i_d) = 0,$$

where $\hat{i_l}$ means this component is missing in a vector. Therefore we can store $\mathcal{A}f_i$ in a sequence of records

$$\mathcal{A}f_i = \begin{cases} \mathcal{A}f_i \cdot \mathrm{itr} = Gf_i \\ \mathcal{A}f_i \cdot \mathrm{bd}[1 \ldots d] = (D_1 f_i, \ldots, D_d f_i), \end{cases} \qquad i = 2, 3, \ldots, C_{dn},$$

($\mathcal{A}f_1 = 0$ and is omitted).

Inner product

Let $f = (f \cdot \mathrm{itr}, f \cdot \mathrm{bd}_1, \ldots, f \cdot \mathrm{bd}_d)$ be a polynomial of degree k, i.e., $f \cdot \mathrm{itr}$ is a d variable polynomial of degree k and $f \cdot \mathrm{bd}_l$ is a $d-1$ variable polynomial of degree k. Let $g = (g \cdot \mathrm{itr}, g \cdot \mathrm{bd}_1, \ldots, g \cdot \mathrm{bd}_d)$ be another polynomial of degree m, the inner product $\langle f, k, g, m \rangle$ between f and g is defined as

$$\begin{aligned} \langle f, k, g, m \rangle &\equiv \int_S (f \cdot \mathrm{itr})(g \cdot \mathrm{itr}) \exp(-2\gamma \cdot x)\, dx + \frac{1}{2} \sum_{l=1}^{d} \int_{F_l} (f \cdot \mathrm{bd}_l)(g \cdot \mathrm{bd}_l) \exp(-2\gamma \cdot x)\, d\sigma_l \\ &= \sum_{i_1 + \ldots + i_d \le k} \sum_{j_1 + \ldots + j_d \le m} (f \cdot \mathrm{itr})(i_1, i_2, \ldots, i_d)(g \cdot \mathrm{itr})(j_1, j_2, \ldots, j_d) \times \\ &\quad \times \frac{(i_1 + j_1)!}{(2\gamma_1)^{i_1 + j_1 + 1}} \times \cdots \times \frac{(i_d + j_d)!}{(2\gamma_d)^{i_d + j_d + 1}} \\ &\quad + \frac{1}{2} \sum_{l=1}^{d} \sum_{i_1 + \ldots + \hat{i_l} + \ldots i_d \le k} \sum_{j_1 + \ldots + \hat{j_l} + \ldots j_d \le m} (f \cdot \mathrm{bd}_l)(i_1, \ldots, \hat{i_l}, \ldots i_d) \times \\ &\quad \times (g \cdot \mathrm{bd}_l)(j_1, \ldots, \hat{j_l}, \ldots, j_d) \frac{(i_1 + j_1)!}{(2\gamma_1)^{i_1 + j_1 + 1}} \times \cdots \ \hat{l} \text{ term } \cdots \times \frac{(i_d + j_d)!}{(2\gamma_d)^{i_d + j_d + 1}}. \end{aligned}$$

Because this inner product function is used frequently in BNET, reducing its run time will significantly improve the efficiency of BNET. Function call to calculate *weighting factor* $(i_l + j_l)/(2\gamma_l)^{i_1 + j_1 + 1}$ is expensive. Therefore we can also pre-compute them and store them as w, i.e.,

$$w[l][i] \equiv \frac{i!}{(2\gamma_l)^{i+1}}, \quad 1 \le l \le d, \quad 0 \le i \le 2n.$$

A.4 Stationary Density

There are several ways to find r_n defined in Proposition 4.3. In this section, we present two ways to do it. One is to use orthogonalization, the other is to solve a linear equation.

Orthogonalizing Basis Elements

Let $\psi_2 = \mathcal{A}f_2$, and for $i = 3, \ldots, C_{dn}$, let

$$\psi_i \equiv \mathcal{A}f_i - \sum_{k=2}^{i-1} \frac{\langle \mathcal{A}f_i, \psi_k \rangle}{\langle \psi_k, \psi_k \rangle} \psi_k.$$

Then $\{\psi_2, \psi_3, \ldots, \psi_{C_{dn}}\}$ is an orthogonal basis of H_n. Let

$$\psi_1 = (1, 1, \ldots, 1) \quad \text{and} \quad \phi_0 = (1, 0, \ldots, 0).$$

Then the orthogonal complement ϕ_n (see Proposition 4.3) of ψ_1 onto H_n is given by

$$\phi_n \equiv \psi_1 - \sum_{k=2}^{C_{dn}} \frac{\langle \psi_k, \psi_1 \rangle}{\langle \psi_k, \psi_k \rangle} \psi_k.$$

Hence

$$r_n \equiv \frac{\phi_n}{\langle \phi_n, \phi_0 \rangle}.$$

Letting

$$p_n = r_n \exp(-2\gamma \cdot x),$$

we have

$$\int_S (p_n \cdot \mathrm{itr})\, dx = 1, \quad m_j \doteq \int_S (x_i\, p_n \cdot \mathrm{itr})\, dx = \langle \tilde{x_i}, r_n \rangle,$$

where $\tilde{x_i}$ is the polynomial $(x_i, 0, \ldots, 0)$.

Solving Linear Equation

The orthogonal complement ϕ_n of ψ_1 on to H_n is equal to

$$\phi_n = \psi_1 - \bar{\psi_1},$$

where $\bar{\psi_1}$ is the projection of ψ_1 on to H_n. Because $\{\mathcal{A}f_i, 2 \leq i \leq C_{dn}\}$ is a basis for H_n, there exist constants $a_2, \ldots, a_{C_{dn}}$ such that

$$\bar{\psi_1} = \sum_{i=2}^{C_{dn}} a_i \mathcal{A}f_i.$$

Therefore

$$\phi_n = \psi_1 - \sum_{i=2}^{C_{dn}} a_i \mathcal{A}f_i.$$

By definition, $\langle \phi_n, \mathcal{A}f_i \rangle = 0$ $(i = 2, \ldots, C_{dn})$, and hence we obtain the following liner equation:

$$Aa = b, \tag{A.3}$$

where

$$A = (\langle \mathcal{A}f_i, \mathcal{A}f_j \rangle)_{2 \le i,j \le C_{dn}}, \quad a = (a_2, \ldots, a_{C_{dn}})', \quad b = (\langle \psi_1, \mathcal{A}f_2 \rangle, \ldots, \langle \psi_1, \mathcal{A}f_{C_{dn}} \rangle)'.$$

The matrix A is positive definite, therefore it has an inverse A^{-1}. Hence a is given by the unique solution $A^{-1}b$ of the linear equation (A.3). Once ϕ_n is found, the rest part of argument in getting stationary density is identical to that in the first part of this section.

A.5 "C" code for BNET

In this section, we present a "C" code of BNET. The `main()` function of BNET is defined as follows.

```
#include "bnet.h"
#include <malloc.h>
#include <math.h>
poly rn;
main()
{
  extern  void InputScaling();
  void         PreCompute(), Basis(), Density();
  poly         *Af;
  extern  void Output();

  InputScaling();
  PreCompute();
  Af = (poly *) malloc((unsigned) (c[d][n]+1) * sizeof(poly));
  if (!Af) Bneterror("Allocation Failure for Af in basis()");
  Basis(Af);
  Density(Af);
  Output();
}
```

The *header* file `"bnet.h"` will be described shortly. The data type `poly` will be defined in the file `"bnet.h"`. The function `InputScaling()` mainly deals with reading input parameters (d, Γ, μ, R, n) and scaling them as described in detail in Section A.1. The function

`Output()` gives whatever output that a user needs. These two functions are user dependent, hence we will not give their definitions here. We declared these two functions as *external* functions, because they are very likely located in a file different from the one that `main()` resides, or perhaps they are in a pre-compiled front–end module. The `malloc()` is a system function, allowing us to dynamically allocate memory space. A header file `"malloc.h"` or `"stdlib.h"`, depending on a particular system, should be included. In the following, we will concentrate on the implementation of three functions `PreCompute()`, `Basis()` and `Density()`. Even for these three functions, the definitions are not complete. But we do cover the most important parts of these routines. We also leave out such implementation details as checking errors, minimizing memory space usage and making a compact code. we believe the current code is easier for readers to read, without dealing with some unimportant coding details. To be definite, we assume that these three functions together with `main()` are in one file, say `"bnet.c"`. The following file `"bnet.h"` is included in the file `"bnet.c"`.

```
/* This is "bnet.h" file */

extern int     d;              /* dimension */
extern int     n;              /* maximum degree of polynomials */
extern double  *Gamma[];       /* covariance matrix */
extern double  mu[];           /* drift vector */
extern double  *R[];           /* reflection matrix */
extern double  gamma[];        /* = -R^{-1}mu */
extern int     *c[];           /* a d by (n+2) matrix holding "C"
                                  as described in Indexing section */
extern int     *I[];           /* as described in Indexing section */
extern int     *Ib[];          /* defined as I, but used over the
                                  boundary piece */
extern double  *w[];           /* a d by (2n+1) matrix holding
                                  weighting factors used in inner() */

typedef struct {
  double itr[];
  double *bd[];
} poly;
extern poly    rn;             /* the new unknown "rn", the density
                                  pn = rn exp(-2 gamma x) */
/* bnet utility functions */
extern int    *ivector();      /* allocate memory space to hold a vector */
extern double *dvector();      /* allocate memory space to hold a vector */
extern int    **imatrix();     /* allocate memory space to hold a matrix */
```

```
extern double **dmatrix();     /* allocate memory space to hold a matrix */
extern void   free_ivector();  /* free spaces allocated by ivector()     */
extern void   free_dvector();  /* free spaces allocated by dvector()     */
extern void   free_imatrix();  /* free spaces allocated by imatrix()     */
extern void   free_dmatrix();  /* free spaces allocated by dmatrix()     */
extern void   Bneterror();     /* bnet error handling function           */
```

The variables that are *declared* in `"bnet.h"` are *global* ones. Input parameters $(d, \Gamma, \mu, R, n, \gamma)$ will be defined and rescaled in the function `InputScaling()`, which we will not define here. Matrix $c[1 \dots d][-1 \dots n]$ is used to store those combinatorial numbers as defined in Section A.2. Matrix $I[2 \dots c_{dn}][0 \dots d]$ is used to store all the index for polynomials in the interior as described in Section A.2. Matrix $Ib[2 \dots c_{(d-1)n}][0 \dots (d-1)]$ is defined similarly to store all the index for polynomials over the boundary F_l $(l = 1, \dots, d)$, see function `PreCompute()`. Type `poly` is a new data type aiming to store a polynomial of arbitrary degree. A polynomial `f` with `poly` type has two parts, interior part `f.itr` and boundary part `f.bd`. We will *dynamically* allocate memory space for a polynomial with type `poly`. If a polynomial `f` is of degree $k \geq 0$, then `f.itr` will be an array ranging from 1 to c_{dk} and `f.bd` will be a $d \times c_{(d-1)k}$ matrix, whose lth row is an array ranging from 1 to $c_{(d-1)k}$ corresponding to the lth boundary piece of the polynomial. The last eight functions declared in `"bnet.h"` are utility functions. We illustrate the usage of functions `ivector()` and `free_ivector()`. The call `v=ivector(l, h)` will allocate memory spaces for a pointer `v` to hold a vector of integers $v[l \dots h]$, and the call `free_ivector(v, l, h)` frees all the space of `v` allocated by `ivector()`. The prefix `i` means that the relevant quantities are of `integer` type. The call of utility function `Bneterror()` will make BNET exit a system and print out a relevant warning message to a user. The usage of the rest of functions are similar, noting the prefix `d` means that the relevant quantities are of `double` type. These utility functions are not specified here. Interested readers are referred to [37] for details.

The following function `PreCompute()` computes and stores the combinatorial numbers, the indexes and weighting factors that we need later. The function `combi(l, k)`, which is not defined this document, returns 0 if $k = -1$ and C_{l+k}^{l} if $k \geq 0$.

```
int    *I[], *Ib[], *c[];
double *w[];

void PreCompute() {
  void       ComputeC();
  int      **ComputeIndex();
```

```
  double   **ComputeWeight();

  ComputeC();
  I = ComputeIndex(d);
  Ib = ComputeIndex(d-1);
  w = ComputeWeight();  /* w[l][i] = i!/(2 gamma_l)^{i+1} */
}

void ComputeC() {
  int i, l;
  extern int combi();

  c = imatrix(0, d, -1, n);    /* make  c[d+1][n+2] matrix */
  for (l=0; l<=d; l++) for (i=-1; i<= n; i++)
    c[l][i] = combi(l, i);     /* c[l][-1] = 0 */
}

int **ComputeIndex(dim)
     int dim;                  /* dim = d and d-1 */
{
  int  i, *II[], *InverseIndex();

  II = (int **)malloc((unsigned)c[dim][n]*sizeof(int *));
  for (i=1; i<=c[dim][n]; i++)   II[i] = InverseIndex(dim, i);
  return (II);
}

int *InverseIndex(dim, i)
     int dim, i;
{
  int j,  *k, *II;

  k = ivector(1, dim);
  for (j=dim; j>=1; j--) {
    k[j] = -1;
    while(i>c[j][k[j]]) k[j]++;   /* linear search */
    i -= c[j][k[j]-1];
   }
  II = ivector(0, d);
  II[0] = k[dim];
  for(j=1; j<dim; j++)  II[j] = k[dim-j+1] - k[dim-j];
  II[dim] = k[1];
  free_ivector(k, 1, dim);
```

```
  return (II);
}

void       ComputeWeight()
{
  int i, l;
  /*  extern double pow() is a math library function */
  long factorial();
  int twon;

  twon = 2 * n;
  w = dmatrix(1, d, 0, twon);
  for (l=1; l<=d ; l++) for (i=0; i<= twon; i++)
    w[l][i] = factorial(i)/pow(2*gamma[l], (double) i+1);
}
```

In the following routine `Basis()`, we find the natural basis $\mathcal{A}f_2, \ldots, \mathcal{A}f_{c_{dn}}$. Readers can refer to Section A.3 for more details. Function `Index()` is straightforwardly implemented according to the description in Section A.2. Function `initpoly()` is nothing but allocate enough space to hold a ith degree polynomial and initialize to be zero.

```
void  Basis(Af)
     poly *Af;
{
  int i, j, l, k;
  int *II, *IIb;
  poly initpoly();
  long Index();

  II = ivector(0, d);
  IIb= ivector(0, d-1);
  for (i=2; i<=c[d][n]; i++) {
    initpoly(A+i, I[i][0]-1);
    /* starts filling interior  polynomial */
    for (j=0; j<=d; j++) II[j] = I[i][j];
    for (j=1; j<=d; j++) {
      if (II[j]>=1) {
        II[0]--; II[j]--;
        Af[i].itr[Index(d, II)]=I[i][j]*mu[j];
        if (II[j]>=1) {
          II[0]--; II[j]--;
          Af[i].itr[Index(d,II)] = I[i][j]*(I[i][j]-1)/2;
```

```
          II[0]++; II[j]++;
        }
        II[0]++; II[j]++;
      }
    }
    for (j=1; j<=d; j++) for (l=j+1; l<=d; l++)
      if (II[j] >=1 && II[l] >=1) {
        II[0] -= 2; II[j]--; II[l]--;
        Af[i].itr[Index(d, II)] = I[i][j] * I[i][l] * Gamma[j][l];
        II[0] += 2; II[j]++; II[l]++;
      }
    /* starts filling  boundary polynomials */
    for (j=1; j<=d; j++) {
      if (II[j] == 1) {
        IIb[0] = II[0] - 1;
        for (l=1; l<j; l++) IIb[l] = II[l];
        for (l=j; l<d; l++) IIb[l] = II[l+1];
        Af[i].bd[j][Index(d-1, IIb)] = R[j][j];
      }
      if (II[j] == 0) {
        for (l=1; l<=d; l++) if (II[l] >= 1) {
          II[l]--;
          for (k=0; k<j; k++) IIb[k] = II[k];
          for (k=j; k<d; k++) IIb[k] = II[k+1];
          IIb[0]--; II[l]++;
          Af[i].bd[j][Index(d-1, IIb)] = II[l] * R[l][j];
        }
      }
    }
  }
  free_ivector(II, 0, d);
  free_ivector(IIb, 0, d-1);
}

void initpoly(f, k)
     poly *f;
     int k;
{
  int l, i;

  f->itr = dvector(1, c[d][k]);
  for (i=1; i<=c[d][k]; i++)
    (f->itr)[i] = 0.0;
```

```
  f->bd = dmatrix(1, d, 1, c[d-1][k]);
  for (l=1; l<=d; l++)
    for (i=1; i<=c[d-1][k]; i++)
      (f->bd)[l][i] = 0.0;
}

long Index(dim, II)
     int dim, *II;
{
  long tmp=1;
  int l, k=II[0]-1;

  for (l=dim; l>=1; l--) {
    tmp += c[l][k];
    k -= II[dim-l+1];
  }
  return (tmp);
}
```

Our next task is to define the function `Density()` to find the weighted density function r_n. In the definition, we need to calculate the *inner product* of two polynomials. Suppose that f and g are two polynomials with type `poly` of degree k and m, respectively. Then function `inner(f, t, g, m)` should return the inner product of f and g. It is implemented according the formula given in the "Inner Product" section. We comment that in calculating a term like $(i_1 + j_1)!/(2\gamma_1)^{i_1+j_1+1}$, one would like to use those pre-compute weighted factors $w[l][i]$. Apparently, storage for these quantities is not a problem, however, it saves tremendous number of functions calls to calculate $(i_1 + j_1)!$ and $(2\gamma_1)^{i_1+j_1+1}$ in real time. The function `half_linear(f, t, a, g, m, h)` performs $h = f + ag$ with $m \leq t$, where f and g are polynomials of degree k and m, respectively, and a is a real number.

```
void  Density(Af)
     poly *Af;
{
  int    i, j, l,  k;
  poly   phi_0;
  double tmp;
  extern double inner();
  extern void   half_linear();
  extern void   initpoly();
```

```
  extern void   orthogonalize();

  /* give space and initialize */
  initpoly(&rn, n-1);
  initpoly(&phi_0, 0);
  phi_0.itr[1] = 1.0;     /*phi_0 = (1; 0, 0, ..., 0) */
  rn.itr[1] = 1.0;        /*p = psi_1=(1; 1, 1, ..., 1) */
  for (l=1; l<=d; l++) {
    rn.bd[l][1] = 1.0;
    phi_0.bd[l][1] = 1.0;
  }                       /* set phi_0 = phi_1 */
                          /* end of initialization */

  orthogonalize(Af);
  for ( k=2; k<=c[d][n]; k++) {
    tmp = inner(Af[k], I[k][0]-1, Af[k], I[k][0]-1);
    if ( tmp ==0.0)
      Bneterror(" Can not be normalized when finding density");
    tmp = - inner(phi_0,  0 , Af[k], I[k][0]-1)/tmp;
    half_linear( rn, n-1, tmp, Af[k], I[k][0]-1, &rn);
  }
  /* normalize */
  for (l=1; l<=d; l++)    /* set phi_0 back to phi_0 */
    phi_0.bd[l][1] = 0.0;
  tmp = inner( phi_0, 0, rn, n-1);
  if ( tmp==0.0) Bneterror(" can not be noramlized into a density");
  tmp = 1/tmp;
  for ( i=1; i<=c[d][n-1]; i++)
    rn.itr[i] *= tmp;
  for ( j =1; j<=d; j++)
    for ( i=1; i<=c[d-1][n-1]; i++)
      rn.bd[j][i] *= tmp;
}


void orthogonalize(Af)
     poly *Af;
{
  int t, i;
  double tmp;
  extern double inner();
  extern void half_linear();
  extern void Bneterror();
```

```
  for ( t=3; t<=c[d][n]; t++) for ( i =2; i<t; i++) {
    tmp = inner(Af[i], I[i][0]-1, Af[i], I[i][0]-1);
    if (tmp ==0.0) Bneterror(" Can not orthogonalize, divisor zero ");
    tmp = -inner(Af[t], I[t][0]-1, Af[i], I[i][0]-1)/tmp;
    half_linear( Af[t], I[t][0]-1, tmp, Af[i], I[i][0]-1, &Af[t]);
  }
}


double inner(f, t,  g, m)
     poly f, g;
     int t, m;

{
  int i, j, l, k;
  double tmp=0.0;
  double prod;

  for ( i =1; i<=c[d][t]; i++) for ( j =1; j<=c[d][m]; j++) {
    prod = 1.0;
    for (l =1; l<=d;l++)
      prod *= w[l][I[i][l]+I[j][l]];
    tmp += f.itr[i]*g.itr[j]* prod;

  }
  for ( k =1; k<=d; k++) {
    for ( i =1; i<=c[d-1][t]; i++) for ( j =1; j<=c[d-1][m]; j++) {
      prod = 1.0;
      for (l =1; l<k;l++)
        prod *= w[l][Ib[i][l]+Ib[j][l]];
      for (l =k; l<d ;l++)
        prod *= w[l+1][Ib[i][l]+Ib[j][l]];
      tmp += 0.5 * f.bd[k][i] * g.bd[k][j] * prod;
    }
  }
  return (tmp);
}


void half_linear(f, t,  a,  g, m, h)
     /* h = f + a * g, deg(f)=t, deg(g) =m  m<= t */
     int t, m;
```

```
     poly f, g, *h;
     double a;
{
  int i, j;
  extern void Bneterror();

  if ( t<m) Bneterror(" the degree of first poly should be bigger ");
  for ( i=1; i<=c[d][m]; i++)
    h->itr[i] = f.itr[i] + a * g.itr[i];
  for ( j =1; j<=d; j++)   for ( i=1; i<=c[d-1][m]; i++)
    h->bd[j][i] = f.bd[j][i] + a * g.bd[j][i];
}
```

Bibliography

[1] M. Berman and M. Westcott. On queueing systems with renewal departure processes. *Adv. Appl. Prob.*, 15:657–673, 1983.

[2] A. Bernard and A. El Kharroubi. Régulations de processus dans le premier "orthant" de $\mathbb{R}^n$. *Stochastics*, to appear.

[3] P. Billingsley. *Convergence of Probability Measures.* Wiley, New York, 1968.

[4] P. Billingsley. *Probability and Measure.* Wiley, 2nd edition, 1986.

[5] J. Bramble and A. Schatz. Rayleigh–Ritz–Galerkin methods for Dirichlet's problem using subspaces without boundary conditions. *Comm. Pure Appl. Math.*, 23:653–675, 1970.

[6] K. L. Chung and R. J. Williams. *Introduction to Stochastic Integration.* Birkhäuser, Boston, 2nd edition, 1990.

[7] E. G. Coffman and M. I. Reiman. Diffusion approximations for computer/comunication systems. In G. Iazeolla, P. J. Courtois, and A. Hordijk, editors, *Mathematical Computer Performance and Reliability*, Amsterdam, 1984. North-Holland.

[8] J. G. Dai. The existence of invariant probability measures for a class of Feller processes. *Journal of Nanjing University: Natural Science Edition*, 21:245–252, 1985. (in Chinese with English abstract).

[9] J. G. Dai and J. M. Harrison. Steady-state analysis of RBM in a rectangle: numerical methods and a queueing application. *Annals of Applied Probability*, to appear in 1991.

[10] P. E. Echeverria. A criterion for invariant measures of Markov processes. *Z. Wahrsch. verw. Gebiete*, 61:1–16, 1982.

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

[12] C. Flores. Diffusion approximations for computer communications networks. In B. Gopinath, editor, *Computer Communications, Proc. Symp. Appl. Math.*, Providence, R. I., 1985. Amer. Math. Soc.

[13] M. Foddy. *Analysis of Brownian motion with drift, confined to a quadrant by oblique reflection*. PhD thesis, Department of Mathematics, Stanford University, 1983.

[14] G. J. Foschini. Equilibria for diffusion models of pairs of communicating computers–symmetric case. *IEEE Trans. Inform. Theory*, 28:273–284, 1982.

[15] J. R. Fraker. *Approximate techniques for the analysis of tandem queueing systems*. PhD thesis, Department of Industrial Engineering, Clemson University, 1971.

[16] E. Gelenbe and G. Pujolle. *Introduction to Queueing Networks*. Wiley, New York, 1987.

[17] D. Gilbard and N. S. Trudinger. *Elliptic Partial Differential Equations of Second Order*. Springer-Verlag, New York, 2nd edition, 1983.

[18] P. W. Glynn. Diffusion approximations. Technical Report 38, Department of Operations Research, Stanford University, 1989.

[19] B. S. Greenberg. *Queueing systems with returning customers and the order of tandem queues*. PhD thesis, University of California, Berkeley, 1986.

[20] B. S. Greenberg and R. W. Wolff. Optimal order of servers for tandem queues in light traffic. *Management Science*, 34:500–508, 1988.

[21] J. M. Harrison. The heavy traffic approximation for single server queues in series. *J. Appl. Prob.*, 10:613–629, 1973.

[22] J. M. Harrison. The diffusion approximation for tandem queues in heavy traffic. *Adv. Appl. Prob.*, 10:886–905, 1978.

[23] J. M. Harrison, H. Landau, and L. A. Shepp. The stationary distribution of reflected Brownian motion in a planar region. *Annals of Probability*, 13:744–757, 1985.

[24] J. M. Harrison and V. Nguyen. The QNET method for two-moment analysis of open queueing networks. *Queueing Systems: Theory and Applications*, 6:1–32, 1990.

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

[26] J. M. Harrison and R. J. Williams. Brownian models of open queueing networks with homogeneous customer populations. *Stochastics*, 22:77–115, 1987.

[27] J. M. Harrison and R. J. Williams. Multidimensional reflected Brownian motions having exponential stationary distributions. *Annals of Probability*, 15:115–137, 1987.

[28] J. M. Harrison, R. J. Williams, and H. Chen. Brownian models of closed queueing networks with homogeneous customer populations. *Stochastics*, 29:37–74, 1990.

[29] N. Ikeda and S. Watanabe. *Stochastic Differential Equations and Diffusion Processes.* North-Holland, New York, 1981.

[30] T. G. Kurtz. Random time change and convergence in distribution. to appear.

[31] T. G. Kurtz. A control formulation for constrained Markov processes. In *Proceedings of the AMS-SIAM Summer Seminar on Mathematics of Random Media*, Blacksburg, VA, May 29–June 9 1989. Virginia Polytechnic Institute and State University.

[32] T. G. Kurtz. Martingale problems for constrained markov processes. University of Maryland Distinguished Lecture Series on Stochastic Calculus, to appear.

[33] A. J. Lemoine. Networks of queues–a survey of weak convergence results. *Management Science*, 24:1175–1193, 1978.

[34] G. M. Lieberman. Oblique derivative problems in Lipschitz domains: II. discontinuous boundary data. *J. Reine Angew. Math.*, to appear.

[35] V. Nguyen. *Analysis of processing networks with parallel and sequential tasks.* PhD thesis, Department of Operations Research, Stanford University, 1990.

[36] W. P. Peterson. A heavy traffic limit theorem for networks of queues with multiple customer types. *Math. of Ops. Rsch.*, to appear.

[37] W. H. Press, B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling. *Numerical Recipes in C : The Art of Scientific Computing*. Cambridge University Press, New York, 1988.

[38] K. G. Ramakrishnan and D. Mitra. An overview of PANACEA, a software package for analyzing Markovian queueing networks. *Bell Sys. Tech. J.*, 61:2849–2872, 1982.

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

[40] M. I. Reiman. Asymptotocally exact decomposition approximations for queueing networks. Talk at the ORSA/TIMS National Meeting, Washington D. C., April 1988.

[41] M. I. Reiman and R. J. Williams. A boundary property of semimartingale reflecting Brownian motions. *Probability Theory and Related Fields*, 77:87–97, 1988 and 80:633, 1989.

[42] H. L. Royden. *Real analysis*. MacMillan, New York, 2nd edition, 1968.

[43] S. Serbin. *A computational investigation of least squares and other projection methods for the approximate solution of boundary value problems*. PhD thesis, Cornell University, Ithaca, New York, 1971.

[44] S. Serbin. Computational investigations of least–squares type methods for the approximate solution of boundary value problems. *Math. Comp.*, 29:777–793, 1975.

[45] D. G. Shimshak. A comparison of waiting time approximations in series queueing systems. *Naval Research Logistics Quarterly*, 26:499–509, 1979.

[46] R. H. Stockbridge. Time-average control of martingale problems: existence of a stationary solution. *Annals of Probability*, 18:190–205, 1990.

[47] D. W. Stroock and S. R. S. Varadhan. Difussion processes with boundary conditions. *Comm. Pure and Appl. Math.*, 24:147–225, 1971.

[48] D. W. Stroock and S. R. S. Varadhan. *Multidimensional Diffusion Processes*. Springer-Verlag, New York, 1979.

[49] S. Suresh and W. Whitt. Arranging queues in series: a simulation experiment. *Management Science*, submitted.

[50] L. M. Taylor and R. J. Williams. Existence and uniqueness of semimartingale reflecting Brownian motions on an orthant. in preparation.

[51] L. M. Taylor and R. J. Williams. Notes on SRBM. Private communication.

[52] L. N. Trefethen and R. J. Williams. Conformal mapping solution of Laplace's equation on a polygon with oblique derivative boundary conditions. *Journal of Computational and Applied Mathematics*, 14:227–249, 1986.

[53] S. R. S. Varadhan and R. J. Williams. Brownian motion in a wedge with oblique reflection. *Comm. Pure Appl. Math.*, 38:405–443, 1985.

[54] W. Whitt. The queueing network analyzer. *Bell Sys. Tech. J.*, 62:2817–2843, 1983.

[55] W. Whitt. Approximations for departure processes and queues in series. *Naval Research Logistics Quarterly*, 31:499–521, 1984.

[56] R. J. Williams. *Brownian motion in a wedge with oblique reflection at the boundary.* PhD thesis, Department of Mathematics, Stanford University, 1983.

[57] R. J. Williams. Reflected Brownian motion in a wedge: semimartingale property. *Z. Wahrsch. verw. Gebiete*, 69:161–176, 1985.

[58] R. J. Williams. Reflected Brownian motion with skew symmetric data in a polyhedral domain. *Probability Theory and Related Fields*, 75:459–485, 1987.