Matrix-Analytic Methods in Applied Probability

Leslie William Bright, B.Sc. (Hons.)

Thesis submitted for the degree of
Doctor of Philosophy
in
Applied Mathematics
at
The University of Adelaide
(Faculty of Mathematical and Computer Sciences)

Department of Applied Mathematics

SUB CRUCE LUMEN

March, 1996

Awarded 1996

Contents

Abstract

This thesis considers several questions in the field of matrix-analytic methods. These methods have been used extensively in applied probability to analyse a wide variety of problems. However, there still appears to be scope for further development both in the range of models that can be analysed, and in the type of analysis which can be employed.

The level dependent quasi-birth-and-death process (LDQBD) is an example of a matrix-analytic model. The LDQBD is a generalisation of the level independent quasi-birth-and-death process (QBD). QBDs are used frequently as a modelling tool when applying matrix-analytic methods but the LDQBD is used rather infrequently. In this thesis we develop theoretical results for LDQBDs as well as a number of algorithms that can be used to analyse LDQBDs. We model a number of systems using LDQBDs and provide numerical results.

The level independent M/G/1–type process and the level independent GI/M/1–type process are two other matrix-analytic models. Both of these models can be thought of as extensions of the QBD. Ramaswami developed a duality relationship for these two models. We give a new interpretation of Ramaswami's duality result and use this interpretation to develop an alternative duality relationship. We also present a duality result for level dependent M/G/1–type and GI/M/1–type processes which are generalisations of the level independent processes.

The thesis also considers quasi-stationary distributions for QBDs. We present an expression for the quasi-stationary distribution as well as algorithms for its numerical computation.

Signed Statement

This work contains no material which has been accepted for the award of any other degree or diploma in any university or other tertiary institution and, to the best of my knowledge and belief, contains no material previously published or written by another person, except where due reference has been made in the text.

I consent to this copy of my thesis, when deposited in the University Library, being available for loan and photocopying.

SIGNED: . DATE: 4/4/96.............

Acknowledgements

My sincere thanks go to Peter Taylor for all his efforts over the last three years. Peter's guidance, motivation and encouragement have been an enormous help in completing this thesis. My thanks also go to Nigel Bean for the many useful discussions we have had and the constructive suggestions he has made.

Financial support during my Ph.D. studies has come from The Australian Telecommunications and Electronics Research Board and an Australian Postgraduate Award. I hereby acknowledge these sources.

Finally I express my thanks to Shane Richards for his support as a fellow student, to my mum and to Alison.

Chapter 1

Introduction

Matrix-analytic methods form a subset of the field of algorithmic probability which in turn is a subset of the field of applied probability. Applied probability involves the modelling and analysis of stochastic systems. When we analyse a stochastic system we seek to gain an understanding of how the system behaves under various parameter values. To do this we usually attempt to obtain performance measures for the system. Frequently we say that we have obtained a "solution" for the system if we have obtained some form of expression (either implicit or explicit) for the performance measure we are interested in.

We define algorithmic probability as a methodology in which in both the modelling and analysis of the system, a primary aim of the probabilist is to obtain a solution of a form that leads naturally to an efficient algorithmic implementation. Matrix-analytic methods are an approach to algorithmic probability which involve the development of models with an underlying matrix structure. When we analyse these models using matrix-analytic methods, we aim to obtain algorithmically tractable solutions by exploiting the detailed matrix structure of the models. Matrix-analytic techniques have been used to analyse problems in areas such as queueing theory, manufacturing processes, inventory systems and telecommunications systems. Although matrix-analytic theory is at an advanced stage of its development, it is still developing at a rapid pace.

The seminal work in matrix-analytic methods was done by Evans [22] and Wallace [94]. They considered quasi-birth-and-death processes (hereby abbreviated to QBDs) which are a matrix extension of standard birth-death processes. In a QBD the state space is divided into subsets known as levels. From any level, the process can move up to the level immediately above the current level, down to the level immediately below the current level or to another state in the current level. An important feature of the QBD is that there is a homogeneous structure over the levels. That is, the behaviour of the process is essentially the same regardless of which level the process happens to be in. Both Evans and Wallace showed that the stationary distribution for a QBD can be expressed in the matrix-geometric form $\boldsymbol{x}_k = \boldsymbol{x}_{k-1}\boldsymbol{R}, k \geq 1$ where $\boldsymbol{x}_k$ gives the stationary probabilities for the states in level k. They also proposed methods for numerically computing the stationary distribution.

Neuts [68] extended the results of Evans and Wallace to the class of GI/M/1–type processes which are a matrix extension of the classical GI/M/1 queueing models. The state space for these models is also divided into levels and there is a homogeneous structure over these levels. The QBD is a special case of the GI/M/1–type process. Neuts focussed strongly on developing efficient numerical algorithms for the analysis of these processes. He also provided a probabilistic interpretation for the matrix-geometric form [71]. The book by Neuts [72] gives an extensive introduction to matrix-analytic methods and it has been a source of motivation for much research. In a later book, [73], Neuts considered M/G/1–type processes which are a matrix extension of the classical M/G/1 queueing models. The QBD can also be viewed as a special case of the M/G/1–type process.

Recently the models described above have been extended to the case where they have an inhomogeneous structure. That is, the behaviour of the process is dependent upon the particular level that the process is in. We refer to such processes as level dependent GI/M/1–type processes and level dependent M/G/1–type processes. A special case of both of these level dependent models is the level dependent QBD (hereby abbreviated to LDQBD). Ramaswami [80] gave a detailed presentation of the theory for these level dependent processes.

In this thesis we consider several areas of matrix-analytic theory. We begin in Chapter 2 with a detailed presentation of the general theory of LDQBDs. We use the work of Ramaswami [80] as a basis for our presentation. When analysing LDQBDs it is useful to determine the matrices $\boldsymbol{R}_k$ and $\boldsymbol{G}_k$. In Chapter 3 we present closed form expressions for these matrices and develop algorithms for numerically computing them. We give detailed probabilistic interpretations of our algorithms.

By evaluating the matrices $\boldsymbol{R}_k, k \geq 0$ for a LDQBD it is possible to compute its stationary distribution. In Chapter 4 we develop a general approach for computing the stationary distribution for a LDQBD and we also consider situations where the general approach can be simplified. In Chapter 5 we model a wide variety of systems as LDQBDs and we implement the algorithms developed in Chapter 4 to obtain some numerical results.

In Chapter 6 we consider level dependent GI/M/1–type and M/G/1–type processes. In this section we present a number of theoretical results and in particular, a number of duality type results. Ramaswami [79] developed a duality result for (level independent) GI/M/1–type and M/G/1–type processes and Asmussen and Ramaswami [7] gave a probabilistic interpretation of this duality result. By considering level dependent processes we develop an alternative dual process to Ramaswami's dual. Our duality result can be given a probabilistic interpretation in terms of the time reverse process and we relate this interpretation to that given by Asmussen and Ramaswami.

In Chapter 7 we consider quasi-stationary and limiting-conditional distributions for QBDs. To compute the quasi-stationary distribution for a QBD one needs to evaluate the decay parameter α and a matrix $\boldsymbol{R}(-\alpha)$. Algorithms for computing both these quantities are presented. We give a discussion of the numerical implementation of these algorithms.

At the end of the thesis we present our conclusions and give some suggestions for further research.

Chapter 2

Level Dependent Quasi-Birth-and-Death Processes

2.1 Introduction

In this chapter we present the elementary theory for level dependent quasi-birth-and-death processes (hereby abbreviated to LDQBDs). Our presentation is based on that given in Ramaswami [80]. Ramaswami follows the trend started by Neuts (see [72]) of avoiding the use of the Perron-Frobenius theory of non negative matrices in favour of using probabilistic arguments. Ramaswami's presentation differs from previous presentations by its use of the theory of terminating renewal processes. The beauty of Ramaswami's presentation is that it starts by considering the stationary probabilities and derives these directly by appealing to results for terminating renewal processes. This is in contrast to traditional approaches that begin by considering taboo first passage probabilities and the non linear matrix equations that they satisfy.

We begin this chapter by defining LDQBDs. We then present some results from the theory of terminating renewal processes that will be needed in the sequel. Next we consider the form of the stationary probabilities and we show how this leads to the consideration of certain first passage times. We then show that these passage times satisfy a set of non linear matrix equations. Following this we introduce a set of first passage probabilities that are useful when analysing LDQBDs. At the end of the chapter we present conditions under which a LDQBD is recurrent and transient.

2.2 Level dependent quasi-birth-and-death processes

2.2.1 Discrete time LDQBDs

Consider a discrete time two-dimensional Markov chain $X(n)$ on the state space $\mathcal{S} = \{(k,j) : k \geq 0, 1 \leq j \leq M(k)\}$ with transition matrix of the block partitioned form

$$P = \begin{pmatrix} A_1^{(0)} & A_0^{(0)} & 0 & 0 & \cdots \\ A_2^{(1)} & A_1^{(1)} & A_0^{(1)} & 0 & \cdots \\ 0 & A_2^{(2)} & A_1^{(2)} & A_0^{(2)} & \cdots \\ 0 & 0 & A_2^{(3)} & A_1^{(3)} & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix} \tag{2.2.1}$$

where $A_2^{(k)}, k \geq 1$, $A_1^{(k)}, k \geq 0$, $A_0^{(k)}, k \geq 0$ are matrices of order $M(k) \times M(k-1)$, $M(k) \times M(k)$, and $M(k) \times M(k+1)$ respectively. A Markov chain with this structure is referred to as a level dependent quasi-birth-and-death process (LDQBD).

In the state (k,j), k is referred to as the *level* of the state and j is referred to as the *phase* of the state. We refer to the set of states $\{(k,j) : 1 \leq j \leq M(k)\}$ as *level* k and we denote this set by $\boldsymbol{k}$. From the structure of $\boldsymbol{P}$ we see that from level k, the process can either move up to level $k+1$, down to level $k-1$ or stay in level k. The matrices $A_0^{(k)}$, $A_1^{(k)}$ and $A_2^{(k)}$ respectively give the probabilities for these transitions. Due to this property we say that $X(n)$ is skip free on the levels.

2.2.2 Continuous time LDQBDs

The continuous time analogue of a discrete time LDQBD can be defined as follows. A continuous time LDQBD is a continuous time two-dimensional Markov process $X(t)$ on the state space $\mathcal{S} = \{(k,j) : k \geq 0, 1 \leq j \leq M(k)\}$ with q-matrix of the block partitioned form

$$Q = \begin{pmatrix} Q_1^{(0)} & Q_0^{(0)} & 0 & 0 & \cdots \\ Q_2^{(1)} & Q_1^{(1)} & Q_0^{(1)} & 0 & \cdots \\ 0 & Q_2^{(2)} & Q_1^{(2)} & Q_0^{(2)} & \cdots \\ 0 & 0 & Q_2^{(3)} & Q_1^{(3)} & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix} \tag{2.2.2}$$

where $Q_2^{(k)}, k \geq 1$, $Q_1^{(k)}, k \geq 0$, $Q_0^{(k)}, k \geq 0$ are matrices of order $M(k) \times M(k-1)$, $M(k) \times M(k)$, and $M(k) \times M(k+1)$ respectively. The definitions and comments made above about discrete time LDQBDs all carry over to continuous time.

2.2.3 Assumptions

For LDQBDs in both discrete and continuous time we will assume that the process is irreducible and aperiodic on a subset of $\mathcal{S}$ of the form $\{(k,j) : 0 \leq k \leq K, 1 \leq j \leq M(k)\}$ for some value K such that $0 \leq K \leq \infty$. The case where K is finite allows us to consider a LDQBD on a finite state space. To avoid unnecessary complications we also assume that $M(k)$ is finite for all $k \geq 0$. However, we note that most of the results we present can be extended to the case where $M(k)$ is infinite. For more details we refer to Ramaswami and Taylor [84] and Ramaswami [80].

In the case of a discrete time LDQBD, we assume that $\boldsymbol{P}$ is stochastic unless otherwise stated. For continuous time LDQBDs we assume that $\boldsymbol{Q}$ is stable and conservative. That is, $-q_{ii} < +\infty\ \forall i$ and $\sum_{j \in \mathcal{S}} q_{ij} = 0\ \forall i$.

2.2.4 Level independent quasi-birth-and-death processes

The special case of a discrete time LDQBD where $M(k) = M, k \geq 0$, $\boldsymbol{A}_0^{(k)} = \boldsymbol{A}_0, k \geq 0$, $\boldsymbol{A}_1^{(k)} = \boldsymbol{A}_1, k \geq 1$, and $\boldsymbol{A}_2^{(k)} = \boldsymbol{A}_2, k \geq 1$ is called a discrete time level independent quasi-birth-and-death process. A continuous time level independent quasi-birth-and-death process is defined in a similar way. We will hereby refer to such processes by the abbreviation QBD. We will see throughout the chapter that many of the results for LDQBDs simplify greatly for the case of a QBD.

Evans [22] was the first to consider processes with a QBD structure. He showed that the stationary distribution $\boldsymbol{x} = (\boldsymbol{x}_0, \boldsymbol{x}_1, \dots)$ satisfies the relationship

$$\boldsymbol{x}_k = \boldsymbol{x}_0 \boldsymbol{R}^k \qquad k \geq 0 \tag{2.2.3}$$

where $\boldsymbol{R}$ is a non negative solution to a matrix-quadratic equation. Wallace [94] also derived the relationship (2.2.3). Both Wallace and Evans suggested algorithmic methods for calculating the matrix $\boldsymbol{R}$. In particular, Wallace developed computer routines for doing this. It is also of interest to mention that Wallace is responsible for the name quasi-birth-and-death process. Neuts [68] showed that the relationship (2.2.3) also holds for GI/M/1–type processes which are an extension of QBDs. He also gave a probabilistic interpretation of the matrix $\boldsymbol{R}$ (see [71]).

2.2.5 LDQBDs that are level independent above a certain level

Another special case of a discrete time LDQBD is the case where for some value of K, $\boldsymbol{A}_0^{(k)} = \boldsymbol{A}_0$, $\boldsymbol{A}_1^{(k)} = \boldsymbol{A}_1$, and $\boldsymbol{A}_2^{(k)} = \boldsymbol{A}_2$, $\forall k \geq K$. In this case we have a discrete time LDQBD that is level independent for levels K and above. Similar processes exist in continuous time.

2.3 Some theory for terminating renewal processes

In this section we present some results from renewal theory and in particular results for terminating renewal processes. We refer to Feller [25] and Karlin and Taylor [38] as references on renewal theory.

In renewal processes we consider independently identically distributed random variables $T_1, T_2, \dots$ with common distribution function $L(.)$. If we stipulate that $L(0) = 0$, the random variables T_i are positive. The random variables T_i are known as the inter-event times. Define S_0 to be a non negative random variable and define the random variables S_n, $n \geq 1$ by $S_n = S_0 + T_1 + T_2 + \dots + T_n$. The variables S_n are known as the renewal epochs and the sequence $\{S_n, n \geq 0\}$ is called the renewal process. The renewal process is said to be pure if $S_0 = 0$ and delayed otherwise.

Define $U(t)$ to be the expected number of renewal epochs in the interval $[0,t]$ with the origin counting as a renewal epoch. It is convenient to think of U as generating a measure on $[0,\infty)$ where an interval $I=(a,b]$ has measure $U\{I\}=U(b)-U(a)$. The renewal process is said to be *terminating* if $L(\infty)<1$ in which case the expected number of renewals before termination is given by $U(\infty)=\frac{1}{1-L(\infty)}$. A standard result for terminating renewal processes (see [25, chapter 11]) states that if $z(x)$ has a limit as $x\to\infty$ and $z(x)=0\ \forall x\leq 0$ then

$$\int_{y=0}^{t} z(t-y)U\{dy\}\to z(\infty)U(\infty)=\frac{z(\infty)}{1-L(\infty)} \qquad \text{as } t\to\infty. \tag{2.3.1}$$

When considering discrete time LDQBDs we will consider renewal processes where $L(.)$ is a discrete distribution with span equal to one. In this case equation (2.3.1) becomes

$$\sum_{y=0}^{t} z(t-y)U\{(y-1,y]\}\to z(\infty)U(\infty)=\frac{z(\infty)}{1-L(\infty)} \qquad \text{as } t\to\infty. \tag{2.3.2}$$

2.4 The stationary probabilities

In this section we consider the form of the stationary probabilities for LDQBDs. We will use the following notation;

$P_{k,j}(A)$ = Probability of the event A given that the process starts in the state (k,j),

${}_mP_{k,j}(A)$ = Probability that the event A occurs and the process does not visit level m in the intermediate steps, given that the process starts in the state (k,j).

We make the following definitions for both discrete and continuous time LDQBDs. Let $\boldsymbol{P}(\boldsymbol{\ell},\boldsymbol{k};n)$ be a $M(\ell)\times M(k)$ matrix whose (i,j)th element is $P_{\ell,i}[X(n)=(k,j)]$ and let ${}_m\boldsymbol{P}(\boldsymbol{\ell},\boldsymbol{k};n)$ be a $M(\ell)\times M(k)$ matrix whose (i,j)th element is ${}_mP_{\ell,i}[X(n)=(k,j)]$. The (i,j)th element of $\boldsymbol{P}(\boldsymbol{\ell},\boldsymbol{k};n)$ gives the probability that, starting in the state (ℓ,i), the process is in state (k,j) at time n. The (i,j)th element of ${}_m\boldsymbol{P}(\boldsymbol{\ell},\boldsymbol{k};n)$ gives the probability that, starting in the state (ℓ,i), the process is in state (k,j) at time n and has not visited level m in between. Let $x(k,j)=\lim_{n\to\infty}P_{\ell,i}[X(n)=(k,j)]$. It is known from the theory of Markov chains that $x(k,j)$ exists and by irreducibility is independent of the initial state (ℓ, i). The quantities $x(k,j)$ are known as the stationary probabilities of the Markov chain. We define the vector $\boldsymbol{x}_k$ by $\boldsymbol{x}_k = (x(k,1), x(k,2), \ldots, x(k, M(k))), k \geq 0$ and we put $\boldsymbol{x} = (\boldsymbol{x}_0, \boldsymbol{x}_1, \ldots)$.

2.4.1 Discrete time LDQBDs

We obtain an expression for $[\boldsymbol{P}(\boldsymbol{0}, \boldsymbol{k}; n)]_{ij}$ as follows. Starting in $(0,i)$, for the process to be in (k,j) at time n, it must visit level $k-1$ one or more times in between. It is clear that no matter how many visits are made to level $k-1$ before reaching (k,j), the process must make a final visit to level $k-1$ at some time τ such that $0 \leq \tau < n$. That is, there is a time $0 \leq \tau < n$ such that the process is in level $k-1$ at time τ and in the period $(\tau, n]$, the process does not visit level $k-1$ again. By conditioning on this time τ, we can write

$$[\boldsymbol{P}(\boldsymbol{0}, \boldsymbol{k}; n)]_{ij} = \sum_{h=1}^{M(k-1)} \sum_{\nu=1}^{M(k)} \sum_{\tau=0}^{n-1} [\boldsymbol{P}(\boldsymbol{0}, \boldsymbol{k}-\boldsymbol{1}; n-\tau-1)]_{ih} [\boldsymbol{A}_0^{(k-1)}]_{h\nu} [{}_{\boldsymbol{k-1}}\boldsymbol{P}(\boldsymbol{k}, \boldsymbol{k}; \tau)]_{\nu j}. \tag{2.4.1}$$

The following theorem is adapted from Theorem 3.1 in Ramaswami [80].

Theorem 2.4.1 *For a discrete time LDQBD that is irreducible and aperiodic,*

$$\boldsymbol{x}_k = \boldsymbol{x}_{k-1} \boldsymbol{R}_{k-1} \qquad k \geq 1 \tag{2.4.2}$$

where $[\boldsymbol{R}_{k-1}]_{ij}$ is the expected total number of visits to the state (k,j), before returning to level $k-1$, given the chain starts in the state $(k-1, i)$.

Proof: Taking limits as $n \to \infty$ of both sides of equation (2.4.1) gives

$$x(k,j) = \sum_{h=1}^{M(k-1)} \sum_{\nu=1}^{M(k)} [\boldsymbol{A}_0^{(k-1)}]_{h\nu} \lim_{n\to\infty} \sum_{\tau=0}^{n-1} [\boldsymbol{P}(\boldsymbol{0}, \boldsymbol{k}-\boldsymbol{1}; n-\tau-1)]_{ih} [{}_{\boldsymbol{k-1}}\boldsymbol{P}(\boldsymbol{k}, \boldsymbol{k}; \tau)]_{\nu j} \tag{2.4.3}$$

where taking the limit inside the first two sums is justified because the sums are finite. Now $[{}_{\boldsymbol{k-1}}\boldsymbol{P}(\boldsymbol{k}, \boldsymbol{k}; \tau)]_{\nu j}$ gives the expected number of renewals in the interval $(\tau-1, \tau]$ for the delayed renewal process consisting of visits to the state (k,j), avoiding level $k-1$, given the process started in the state (k,ν). This is because the number of renewals in $(\tau-1, \tau]$ is either one with probability $[{}_{\boldsymbol{k-1}}\boldsymbol{P}(\boldsymbol{k}, \boldsymbol{k}; \tau)]_{\nu j}$ or zero. Since the chain is irreducible, this renewal process is terminating and hence we can apply equation (2.3.2) to give

$$x(k,j) = \sum_{h=1}^{M(k-1)} \sum_{\nu=1}^{M(k)} x(k-1,h)[\boldsymbol{A}_0^{(k-1)}]_{h\nu} \sum_{\tau=0}^{\infty} [{}_{k-1}\boldsymbol{P}(\boldsymbol{k},\boldsymbol{k};\tau)]_{\nu j} \tag{2.4.4}$$

where $\sum_{\tau=0}^{\infty}[{}_{k-1}\boldsymbol{P}(\boldsymbol{k},\boldsymbol{k};\tau)]_{\nu j}$ is the expected number of visits to the state (k,j) before the process reaches level $k-1$ for the first time, given the process starts in the state (k,ν).

Writing equation (2.4.4) in matrix form

$$\boldsymbol{x}_k = \boldsymbol{x}_{k-1}\boldsymbol{A}_0^{(k-1)} \sum_{\tau=0}^{\infty} {}_{k-1}\boldsymbol{P}(\boldsymbol{k},\boldsymbol{k};\tau) \tag{2.4.5}$$

and defining $\boldsymbol{R}_{k-1}$ by

$$\boldsymbol{R}_{k-1} = \boldsymbol{A}_0^{(k-1)} \sum_{\tau=0}^{\infty} {}_{k-1}\boldsymbol{P}(\boldsymbol{k},\boldsymbol{k};\tau) \qquad k \geq 1, \tag{2.4.6}$$

ensures that $[\boldsymbol{R}_{k-1}]_{ij}$ has the necessary interpretation and $\boldsymbol{x}_k = \boldsymbol{x}_{k-1}\boldsymbol{R}_{k-1}$ as required.

□

For the special case of a QBD, Theorem 2.4.1 reduces to the following.

Corollary 2.4.1 *For a discrete time QBD that is irreducible and aperiodic,*

$$\boldsymbol{x}_k = \boldsymbol{x}_{k-1}\boldsymbol{R} \qquad k \geq 1 \tag{2.4.7}$$

where for all $k \geq 1$, $[\boldsymbol{R}]_{ij}$ *is the expected total number of visits to the state* (k,j), *before returning to level* $k-1$, *given the chain starts in the state* $(k-1,i)$.

An interesting question to consider is when will the matrices $\boldsymbol{R}_k$ defined in Theorem 2.4.1 be finite? In the case of a QBD, Neuts [72, Lemma 1.2.1] used Theorem 3 in Chung [14] to show that if the QBD is positive recurrent then the matrix $\boldsymbol{R}$ is finite. Neuts' primary interest was in positive recurrent processes and hence he did not consider the finiteness of $\boldsymbol{R}$ for transient or null recurrent QBDs. As it happens, Theorem 3 in [14] does not require that the QBD be positive recurrent. Hence the theorem can be used to prove that $\boldsymbol{R}$ is also finite when the QBD is transient or null recurrent. Furthermore, the same Theorem 3 can be used to prove that the matrices $\boldsymbol{R}_k$ are finite for both recurrent and transient LDQBDs. An alternative proof that $\boldsymbol{R}_k$ is finite for all LDQBDs was provided by Ramaswami and Taylor [84]. They showed that finiteness follows by a direct application of Corollary 1 of Lemma 5.4 in Seneta [89]. Hence we can state the following theorem.

Theorem 2.4.2 *For a discrete time irreducible and aperiodic LDQBD the matrices* $\{\boldsymbol{R}_k, k \geq 0\}$ *are finite.*

2.4.2 Continuous time LDQBDs

Analogous results to those obtained for discrete time LDQBDs can be obtained for continuous time LDQBDs. The derivations are similar to those used in discrete time.

The continuous time analogue of equation (2.4.1) is

$$[\boldsymbol{P}(0,k;t)]_{ij} = \sum_{h=1}^{M(k-1)} \sum_{\nu=1}^{M(k)} \int_{\tau=0}^{t} [\boldsymbol{P}(0,\boldsymbol{k}-1;t-\tau)]_{ih}[\boldsymbol{Q}_0^{(k-1)}]_{h\nu}[{}_{\boldsymbol{k}-1}\boldsymbol{P}(\boldsymbol{k},\boldsymbol{k};\tau)]_{\nu j}\, d\tau. \tag{2.4.8}$$

The following is the continuous time analogue of Theorem 2.4.1.

Theorem 2.4.3 *For a stable continuous time LDQBD that is irreducible and aperiodic,*

$$\boldsymbol{x}_k = \boldsymbol{x}_{k-1}\boldsymbol{R}_{k-1} \qquad k \geq 1 \tag{2.4.9}$$

where $[\boldsymbol{R}_{k-1}]_{ij}$ *is the expected sojourn time in the state* (k, j)*, in units of the mean sojourn time in the state* $(k-1, i)$*, before returning to level* $k-1$*, given the chain starts in the state* $(k-1, i)$*.*

Proof: Taking limits as $t \to \infty$ of both sides of equation (2.4.8) gives

$$x(k,j) = \sum_{h=1}^{M(k-1)} \sum_{\nu=1}^{M(k)} [\boldsymbol{Q}_0^{(k-1)}]_{h\nu} \lim_{t\to\infty} \int_{\tau=0}^{t} [\boldsymbol{P}(0,\boldsymbol{k}-1;t-\tau)]_{ih}[{}_{\boldsymbol{k}-1}\boldsymbol{P}(\boldsymbol{k},\boldsymbol{k};\tau)]_{\nu j}\, d\tau \tag{2.4.10}$$

where taking the limit inside the first two sums is justified because the sums are finite. Now $[{}_{\boldsymbol{k}-1}\boldsymbol{P}(\boldsymbol{k},\boldsymbol{k};\tau)]_{\nu j}\, d\tau$ gives the expected number of renewals in the interval $(\tau, \tau + d\tau]$ for the delayed renewal process consisting of visits to the state (k, j), avoiding level $k-1$, given the process started in state (k, ν). Since the chain is irreducible, this renewal process is terminating and hence we can apply equation (2.3.1) to give

$$x(k,j) = \sum_{h=1}^{M(k-1)} \sum_{\nu=1}^{M(k)} x(k-1,h)[\boldsymbol{Q}_0^{(k-1)}]_{h\nu} \int_{\tau=0}^{\infty} {}_{k-1}[\boldsymbol{P}(\boldsymbol{k},\boldsymbol{k};\tau)]_{\nu j}\, d\tau. \tag{2.4.11}$$

Again $\int_{\tau=0}^{\infty}[{}_{\boldsymbol{k}-1}\boldsymbol{P}(\boldsymbol{k},\boldsymbol{k};\tau)]_{\nu j}\, d\tau$ is the expected total sojourn time in the state (k, j) before the process reaches level $k-1$ for the first time, given the process starts in the state (k, ν).

If we write equation (2.4.11) in the matrix form

$$\boldsymbol{x}_k = \boldsymbol{x}_{k-1} \boldsymbol{Q}_0^{(k-1)} \int_{\tau=0}^{\infty} {}_{k-1}\boldsymbol{P}(\boldsymbol{k}, \boldsymbol{k}; \tau) d\tau \tag{2.4.12}$$

then defining $\boldsymbol{R}_{k-1}$ by

$$\boldsymbol{R}_{k-1} = \boldsymbol{Q}_0^{(k-1)} \int_{\tau=0}^{\infty} {}_{k-1}\boldsymbol{P}(\boldsymbol{k}, \boldsymbol{k}; \tau) d\tau \qquad k \geq 1, \tag{2.4.13}$$

gives us the result provided $[\boldsymbol{R}_{k-1}]_{ij}$ has the required interpretation.

$[\boldsymbol{R}_{k-1}]_{ij}$ can be written as

$$\begin{aligned} [\boldsymbol{R}_{k-1}]_{ij} &= \sum_{\ell=1}^{M(k)} [\boldsymbol{Q}_0^{(k-1)}]_{i\ell} \int_{\tau=0}^{\infty} [{}_{k-1}\boldsymbol{P}(\boldsymbol{k}, \boldsymbol{k}; \tau)]_{\ell j} \, d\tau & (2.4.14) \\ &= \sum_{\ell=1}^{M(k)} \frac{[\boldsymbol{Q}_0^{(k-1)}]_{i\ell}}{-[\boldsymbol{Q}_1^{(k-1)}]_{ii}} (-[\boldsymbol{Q}_1^{(k-1)}]_{ii}) \int_{\tau=0}^{\infty} [{}_{k-1}\boldsymbol{P}(\boldsymbol{k}, \boldsymbol{k}; \tau)]_{\ell j} \, d\tau. & (2.4.15) \end{aligned}$$

Now $\dfrac{[\boldsymbol{Q}_0^{(k-1)}]_{i\ell}}{-[\boldsymbol{Q}_1^{(k-1)}]_{ii}}$ is the probability that the process first moves into state (k, ℓ) given it started in state $(k-1, i)$ and $(-[\boldsymbol{Q}_1^{(k-1)}]_{ii}) \int_0^{\infty} [{}_{k-1}\boldsymbol{P}(\boldsymbol{k}, \boldsymbol{k}; \tau)]_{\ell j} \, d\tau$ is the expected sojourn time in the state (k, j), in units of the mean sojourn time in state $(k-1, i)$, before the process reaches level $k-1$ for the first time, given the process starts in the state (k, ℓ). Hence we see that $[\boldsymbol{R}_{k-1}]_{ij}$ has the required interpretation. □

For the special case of a QBD, Theorem 2.4.3 reduces to the following.

Corollary 2.4.2 *For a stable continuous time LDQBD that is irreducible and aperiodic,*

$$\boldsymbol{x}_k = \boldsymbol{x}_{k-1} \boldsymbol{R} \qquad k \geq 1 \tag{2.4.16}$$

where for all $k \geq 1$, $[\boldsymbol{R}]_{ij}$ is the expected sojourn time in the state (k, j), in units of the mean sojourn time in the state $(k-1, i)$, before returning to level $k-1$, given the chain starts in the state $(k-1, i)$.

We now show that the matrices $\boldsymbol{R}_k$ defined in Theorem 2.4.3 are finite. To do this we consider the jump chain. The jump chain is a discrete time LDQBD obtained from the continuous time LDQBD by observing it only at time points where there is a change of state. The jump chain is governed by the matrices

$$\begin{aligned} \boldsymbol{A}_0^{(k)} &= \Phi_k^{-1} \boldsymbol{Q}_0^{(k)}, & (2.4.17) \\ \boldsymbol{A}_1^{(k)} &= \Phi_k^{-1} \boldsymbol{Q}_1^{(k)} + \boldsymbol{I}, & (2.4.18) \\ \boldsymbol{A}_2^{(k)} &= \Phi_k^{-1} \boldsymbol{Q}_2^{(k)} & (2.4.19) \end{aligned}$$

where Φ_k = -diag($[Q_1^{(k)}]_{ii}$). The jump chain is a very useful tool when analysing continuous time processes. It enables us to derive results for continuous time processes from corresponding results for discrete time processes. This is useful since it is often much easier to analyse discrete time processes since they generally have a clearer physical interpretation. We will use the jump chain frequently throughout this thesis.

Let $\{R_k^J, k \geq 0\}$ be the family of R_k matrices for the jump chain. From the probabilistic interpretations of the matrices R_k for discrete and continuous time processes we have the relationship

$$R_k = \Phi_k R_k^J \Phi_{k+1}^{-1} \qquad k \geq 0 \tag{2.4.20}$$

or equivalently

$$[R_k]_{ij} = \frac{[R_k^J]_{ij}(-[Q_1^{(k+1)}]_{jj})^{-1}}{(-[Q_1^{(k)}]_{ii})^{-1}} \qquad k \geq 0. \tag{2.4.21}$$

To see this, note that $(-[Q_1^{(k+1)}]_{jj})^{-1}$ is the mean sojourn time in the state $(k+1, j)$ per visit to $(k+1, j)$ and $(-[Q_1^{(k)}]_{ii})^{-1}$ is the mean sojourn time in the state (k, i) per visit to (k, i). Recall that the (i, j)th element of R_k^J is the expected number of visits by the jump chain to the state $(k+1, j)$, before returning to level k, given the process started in state (k, i). Multiplying $[R_k^J]_{ij}$ by $(-[Q_1^{(k+1)}]_{jj})^{-1}$ gives the expected total sojourn time by the continuous time process in the state $(k+1, j)$ before returning to level k, given the process started in state (k, i). Dividing $[R_k^J]_{ij}(-[Q_1^{(k+1)}]_{jj})^{-1}$ by $(-[Q_1^{(k)}]_{ii})^{-1}$ converts the above into units of the mean sojourn time in state (k, i). This gives us that the right hand side of equation (2.4.21) is equal to the expected sojourn time in the state $(k+1, j)$, in units of the mean sojourn time in the state (k, i), before returning to level k, given the chain starts in the state (k, i). We now have the following theorem.

Theorem 2.4.4 *For a stable continuous time LDQBD that is irreducible and aperiodic, the matrices* $\{R_k, k \geq 0\}$ *are finite.*

Proof: The result follows directly from equation (2.4.20) and Theorem 2.4.2. □

2.5 The matrices $\{\boldsymbol{R}_k, k \geq 0\}$

In the previous section we saw that the matrices $\{\boldsymbol{R}_k, k \geq 0\}$ arise naturally when considering the stationary probabilities. In this section we provide further results relating to these matrices.

Firstly we present the following lemma which appears in a more general setting in Ramaswami [80].

Lemma 2.5.1 *1. For a discrete time LDQBD*

$$\boldsymbol{A}_0^{(k-1)} \sum_{\tau=0}^{\infty} {}_{k-1}\boldsymbol{P}(\boldsymbol{k}, \boldsymbol{k}+\boldsymbol{\ell}; \tau) = \boldsymbol{R}_{k-1}\boldsymbol{R}_k \ldots \boldsymbol{R}_{k+\ell-1} \qquad k \geq 1,\ \ell \geq 0. \quad (2.5.1)$$

2. For a continuous time LDQBD

$$\boldsymbol{Q}_0^{(k-1)} \int_{t=0}^{\infty} {}_{k-1}\boldsymbol{P}(\boldsymbol{k}, \boldsymbol{k}+\boldsymbol{\ell}; t)dt = \boldsymbol{R}_{k-1}\boldsymbol{R}_k \ldots \boldsymbol{R}_{k+\ell-1} \qquad k \geq 1,\ \ell \geq 0. \tag{2.5.2}$$

Proof: Firstly consider the discrete time result. For $\ell = 0$ the result follows from equation (2.4.6). The result for general ℓ can be shown using induction. Assume the result holds for some ℓ. By using similar conditioning arguments to those used to derive equation (2.4.1) we can write

$$\begin{aligned}
&\boldsymbol{A}_0^{(k-1)} \sum_{\tau=0}^{\infty} {}_{k-1}\boldsymbol{P}(\boldsymbol{k}, \boldsymbol{k}+\boldsymbol{\ell}+\boldsymbol{1}; \tau) \\
&= \boldsymbol{A}_0^{(k-1)} \sum_{\tau=0}^{\infty} \sum_{n=0}^{\tau-1} {}_{k-1}\boldsymbol{P}(\boldsymbol{k}, \boldsymbol{k}+\boldsymbol{\ell}; \tau-n-1)\boldsymbol{A}_0^{(k+\ell)} {}_{k+\ell}\boldsymbol{P}(\boldsymbol{k}+\boldsymbol{\ell}+\boldsymbol{1}, \boldsymbol{k}+\boldsymbol{\ell}+\boldsymbol{1}; n).
\end{aligned}$$

Swapping the order of the summations gives

$$\begin{aligned}
&\boldsymbol{A}_0^{(k-1)} \sum_{\tau=0}^{\infty} {}_{k-1}\boldsymbol{P}(\boldsymbol{k}, \boldsymbol{k}+\boldsymbol{\ell}+\boldsymbol{1}; \tau) \\
&= \boldsymbol{A}_0^{(k-1)} \sum_{n=0}^{\infty} \sum_{\tau=0}^{\infty} {}_{k-1}\boldsymbol{P}(\boldsymbol{k}, \boldsymbol{k}+\boldsymbol{\ell}; \tau)\boldsymbol{A}_0^{(k+\ell)} {}_{k+\ell}\boldsymbol{P}(\boldsymbol{k}+\boldsymbol{\ell}+\boldsymbol{1}, \boldsymbol{k}+\boldsymbol{\ell}+\boldsymbol{1}; n) \\
&= \boldsymbol{A}_0^{(k-1)} \sum_{\tau=0}^{\infty} {}_{k-1}\boldsymbol{P}(\boldsymbol{k}, \boldsymbol{k}+\boldsymbol{\ell}; \tau)\boldsymbol{A}_0^{(k+\ell)} \sum_{n=0}^{\infty} {}_{k+\ell}\boldsymbol{P}(\boldsymbol{k}+\boldsymbol{\ell}+\boldsymbol{1}, \boldsymbol{k}+\boldsymbol{\ell}+\boldsymbol{1}; n) \\
&= \boldsymbol{R}_{k-1} \ldots \boldsymbol{R}_{k+\ell-1}\boldsymbol{R}_{k+\ell}
\end{aligned}$$

where the last equality results from the inductive assumption and equation (2.4.6). The proof of the continuous time result is done in a similar fashion. □

The following theorem gives equations satisfied by the matrices $\{\boldsymbol{R}_k, k \geq 0\}$ in discrete and continuous time.

Theorem 2.5.1 *1. For a discrete time LDQBD the matrices $\{\boldsymbol{R}_k, k \geq 0\}$ satisfy the equations*

$$\boldsymbol{R}_k = \boldsymbol{A}_0^{(k)} + \boldsymbol{R}_k \boldsymbol{A}_1^{(k+1)} + \boldsymbol{R}_k \boldsymbol{R}_{k+1} \boldsymbol{A}_2^{(k+2)} \qquad k \geq 0. \tag{2.5.3}$$

2. For a continuous time LDQBD the matrices $\{\boldsymbol{R}_k, k \geq 0\}$ satisfy the equations

$$\boldsymbol{0} = \boldsymbol{Q}_0^{(k)} + \boldsymbol{R}_k \boldsymbol{Q}_1^{(k+1)} + \boldsymbol{R}_k \boldsymbol{R}_{k+1} \boldsymbol{Q}_2^{(k+2)} \qquad k \geq 0. \tag{2.5.4}$$

Proof: From equation (2.4.6) we have for a discrete time LDQBD

$$\begin{aligned} \boldsymbol{R}_k &= \boldsymbol{A}_0^{(k)} \sum_{\tau=0}^{\infty} {}_k\boldsymbol{P}(\boldsymbol{k}+\boldsymbol{1}, \boldsymbol{k}+\boldsymbol{1}; \tau) \\ &= \boldsymbol{A}_0^{(k)} + \boldsymbol{A}_0^{(k)} \sum_{\tau=0}^{\infty} {}_k\boldsymbol{P}(\boldsymbol{k}+\boldsymbol{1}, \boldsymbol{k}+\boldsymbol{1}; \tau+1) \\ &= \boldsymbol{A}_0^{(k)} + \boldsymbol{A}_0^{(k)} \sum_{\tau=0}^{\infty} \Big({}_k\boldsymbol{P}(\boldsymbol{k}+\boldsymbol{1}, \boldsymbol{k}+\boldsymbol{1}; \tau) \boldsymbol{A}_1^{(k+1)} \\ &\qquad\qquad + {}_k\boldsymbol{P}(\boldsymbol{k}+\boldsymbol{1}, \boldsymbol{k}+\boldsymbol{2}; \tau) \boldsymbol{A}_2^{(k+2)} \Big) \\ &= \boldsymbol{A}_0^{(k)} + \boldsymbol{R}_k \boldsymbol{A}_1^{(k+1)} + \boldsymbol{R}_k \boldsymbol{R}_{k+1} \boldsymbol{A}_2^{(k+2)} \end{aligned}$$

where the last equality follows from Lemma 2.5.1.

In order to prove the result in continuous time we note that if u is the time that the process makes its final visit to level $k+1$ before time τ, then by conditioning on u we can write

$${}_k\boldsymbol{P}(\boldsymbol{k}+\boldsymbol{1}, \boldsymbol{k}+\boldsymbol{1}; \tau) = e^{\boldsymbol{Q}_1^{(k+1)}\tau} + \int_{u=0}^{\tau} {}_k\boldsymbol{P}(\boldsymbol{k}+\boldsymbol{1}, \boldsymbol{k}+\boldsymbol{2}; u) \boldsymbol{Q}_2^{(k+2)} e^{\boldsymbol{Q}_1^{(k+1)}(\tau-u)}\, du. \tag{2.5.5}$$

The result is obtained by multiplying both sides of (2.5.5) by $\boldsymbol{Q}_0^{(k)}$, integrating from τ equals zero to infinity and applying the results in Lemma 2.5.1. □

2.6 The matrices $\{\boldsymbol{U}_k, k \geq 1\}$

In the previous section we showed that the matrices $\{\boldsymbol{R}_k, k \geq 0\}$ satisfy equations (2.5.3) and (2.5.4) in discrete and continuous time respectively. However, these equations may not necessarily have a unique solution. In this section we consider which particular solutions to equations (2.5.3) and (2.5.4) are equal to the matrices $\boldsymbol{R}_k$. In order to do this we introduce the family of matrices $\{\boldsymbol{U}_k, k \geq 1\}$.

2.6.1 Discrete time LDQBDs

Define the matrices $\{\boldsymbol{U}_k, k \geq 1\}$ by

$$\boldsymbol{U}_k = \boldsymbol{A}_1^{(k)} + \boldsymbol{R}_k \boldsymbol{A}_2^{(k+1)} \qquad k \geq 1. \tag{2.6.1}$$

Suppose the LDQBD starts in level k and we observe the process only at time points when it is in level k, before it visits level $k-1$ for the first time. If we call this partially observed process C_k, then $\boldsymbol{U}_k$ is the transition matrix for C_k. This is easily derived by conditioning on the time of the last visit to level $k+1$ before the process returns to level k. Note that the processes $C_k, k \geq 1$ are transient processes because the LDQBD is irreducible. The following lemma gives an expression for $\boldsymbol{R}_k$ in terms of $\boldsymbol{U}_k$.

Lemma 2.6.1 *The matrices $\boldsymbol{R}_k$ are related to the matrices $\boldsymbol{U}_k$ by*

$$\boldsymbol{R}_{k-1} = \boldsymbol{A}_0^{(k-1)}(\boldsymbol{I} - \boldsymbol{U}_k)^{-1} \qquad k \geq 1. \tag{2.6.2}$$

Proof: For the case of a QBD, equation (2.6.2) becomes $\boldsymbol{R} = \boldsymbol{A}_0(\boldsymbol{I} - \boldsymbol{U})^{-1}$ and this has been proved by Latouche [55]. From equation (2.4.6) we have

$$\boldsymbol{R}_{k-1} = \boldsymbol{A}_0^{(k-1)} \sum_{\tau=0}^{\infty} {}_{k-1}\boldsymbol{P}(\boldsymbol{k}, \boldsymbol{k}; \tau)$$

where $\sum_{\tau=0}^{\infty}[{}_{k-1}\boldsymbol{P}(\boldsymbol{k}, \boldsymbol{k}; \tau)]_{ij}$ is the expected number of visits to the state (k, j), before the process reaches level $k-1$ for the first time, given the process starts in the state (k, i). In terms of the process C_k, $\sum_{\tau=0}^{\infty}[{}_{k-1}\boldsymbol{P}(\boldsymbol{k}, \boldsymbol{k}; \tau)]_{ij}$ gives the expected number of visits by C_k to the state (k, j), before absorption occurs, given C_k starts in the state (k, i). From Theorems 3.2.1 and 3.2.6 in Kemeny and Snell [41], since the LDQBD is irreducible, we have that $\sum_{\tau=0}^{\infty}[{}_{k-1}\boldsymbol{P}(\boldsymbol{k}, \boldsymbol{k}; \tau)]_{ij} = [(\boldsymbol{I} - \boldsymbol{U}_k)^{-1}]_{ij}$ and hence the result follows. ☐

The following theorem is a special case of Theorem 3.4 in Ramaswami [80].

Theorem 2.6.1 *The families of matrices $\{\boldsymbol{R}_k, k \geq 0\}$ and $\{\boldsymbol{U}_k, k \geq 1\}$ are the minimal non negative solutions to the equations (2.6.1) and (2.6.2).*

Proof: Define the sequences $\boldsymbol{U}_k(n), n \geq 0$ and $\boldsymbol{R}_k(n), n \geq 0$ as follows;

$$\boldsymbol{U}_k(0) = \boldsymbol{A}_1^{(k)}, \tag{2.6.3}$$

$$\boldsymbol{R}_k(n) = \boldsymbol{A}_0^{(k)}(\boldsymbol{I} - \boldsymbol{U}_k(n))^{-1} \qquad n \geq 0, \tag{2.6.4}$$

$$\boldsymbol{U}_k(n) = \boldsymbol{A}_1^{(k)} + \boldsymbol{R}_k(n-1)\boldsymbol{A}_2^{(k+1)} \qquad n \geq 1. \tag{2.6.5}$$

Using induction it can be shown that the (i,j)th element of $\boldsymbol{U}_k(n)$ and $\boldsymbol{R}_k(n)$ have the following interpretations. The (i,j)th element of $\boldsymbol{U}_k(n)$ gives the probability that the next visit to level k is through phase j given that the process started in state (k,i) and the process has not visited level $k-1$ or any of the states above level $k+n$ in between. The (i,j)th element of $\boldsymbol{R}_k(n)$ gives the expected number of visits to $(k+1,j)$ before returning to level k, given that the process started in state (k,i) and the process has not visited any of the states above level $k+1+n$ in between. Given this interpretation of $\boldsymbol{U}_k(n)$ and $\boldsymbol{R}_k(n)$, it is clear that $\boldsymbol{U}_k(n) \to \boldsymbol{U}_k$ and $\boldsymbol{R}_k(n) \to \boldsymbol{R}_k$ as $n \to \infty$. Now suppose that $\bar{\boldsymbol{U}}_k$ and $\bar{\boldsymbol{R}}_k$ are non negative solutions of the equations (2.6.1) and (2.6.2). Again using induction it can be shown that $\bar{\boldsymbol{U}}_k \geq \boldsymbol{U}_k(n)$ and $\bar{\boldsymbol{R}}_k \geq \boldsymbol{R}_k(n)$ for all $n \geq 0$. Taking the limit as $n \to \infty$ we have $\bar{\boldsymbol{U}}_k \geq \boldsymbol{U}_k$ and $\bar{\boldsymbol{R}}_k \geq \boldsymbol{R}_k$. These last inequalities imply that $\boldsymbol{U}_k$ and $\boldsymbol{R}_k$ must be the minimal non negative solutions to equations (2.6.1) and (2.6.2). ☐

The following corollary states which particular solutions to equations (2.5.3) are equal to the matrices $\boldsymbol{R}_k$.

Corollary 2.6.1 *The family of matrices* $\{\boldsymbol{R}_k, k \geq 0\}$ *is the minimal non negative solution to the equations (2.5.3).*

Proof: Define the sequence $\bar{\boldsymbol{R}}_k(n), n \geq 0$ as follows;

$$\bar{\boldsymbol{R}}_k(0) = \boldsymbol{A}_0^{(k)}, \tag{2.6.6}$$

$$\bar{\boldsymbol{R}}_k(n+1) = \boldsymbol{A}_0^{(k)} + \bar{\boldsymbol{R}}_k(n)\boldsymbol{A}_1^{(k+1)} + \bar{\boldsymbol{R}}_k(n)\bar{\boldsymbol{R}}_{k+1}(n)\boldsymbol{A}_2^{(k+2)} \qquad n \geq 0. \tag{2.6.7}$$

It can be shown via induction that $\bar{\boldsymbol{R}}_k(n)$ is monotonically increasing and $\bar{\boldsymbol{R}}_k(n) \leq \boldsymbol{R}_k \ \forall n$. Hence via the dominated convergence theorem, $\bar{\boldsymbol{R}}_k(n) \to \boldsymbol{R}_k^*$ where $\boldsymbol{R}_k^*$ satisfies equation (2.5.3) and is called the minimal non negative solution to the equation. For $n \geq 0$ define $\bar{\boldsymbol{U}}_k(n)$ by

$$\bar{\boldsymbol{U}}_k(n) = \boldsymbol{A}_1^{(k)} + \bar{\boldsymbol{R}}_k(n)\boldsymbol{A}_2^{(k+1)} \qquad n \geq 0. \tag{2.6.8}$$

It can be shown via induction that $\bar{\boldsymbol{R}}_k(n) \geq \boldsymbol{R}_k(n)$ and $\bar{\boldsymbol{U}}_k(n) \geq \boldsymbol{U}_k(n) \ \forall n \geq 0$ where $\boldsymbol{R}_k(n)$ and $\boldsymbol{U}_k(n)$ are defined by equations (2.6.3)-(2.6.5). Taking limits as $n \to \infty$ of $\bar{\boldsymbol{R}}_k(n) \geq \boldsymbol{R}_k(n)$ gives $\boldsymbol{R}_k^* \geq \boldsymbol{R}_k$. Since $\boldsymbol{R}_k^*$ is the minimal non negative solution to equation (2.5.3), we also have $\boldsymbol{R}_k^* \leq \boldsymbol{R}_k$ which gives $\boldsymbol{R}_k^* = \boldsymbol{R}_k$. ☐

2.6.2 Continuous time LDQBDs

Here we define the matrices $\{U_k, k \geq 1\}$ by

$$U_k = Q_1^{(k)} + R_k Q_2^{(k+1)} \qquad k \geq 1. \tag{2.6.9}$$

Defining the partially observed processes C_k as we did in discrete time, it can be shown that U_k is the q-matrix for C_k. The following lemma is the continuous time analogue of Lemma 2.6.1. Note that Latouche [55] proved the result for the case of a QBD.

Lemma 2.6.2 *The matrices R_k are related to the matrices U_k by*

$$R_{k-1} = Q_0^{(k-1)}(-U_k)^{-1} \qquad k \geq 1. \tag{2.6.10}$$

Proof: From equation (2.4.13) we have

$$R_{k-1} = Q_0^{(k-1)} \int_{\tau=0}^{\infty} {}_{k-1}P(k, k; \tau) d\tau \tag{2.6.11}$$

where $[{}_{k-1}P(k, k; \tau)]_{ij}$ can be interpreted as the probability that the process C_k is in phase j at time τ, given that C_k started in phase i. From Markov process theory, this is equal to $[e^{U_k \tau}]_{ij}$. Hence we have

$$\begin{aligned} R_{k-1} &= Q_0^{(k-1)} \int_{\tau=0}^{\infty} e^{U_k \tau} d\tau && (2.6.12) \\ &= Q_0^{(k-1)}(-U_k)^{-1} && (2.6.13) \end{aligned}$$

where the last equality follows from results in Graham [31]. ☐

The following theorem and corollary give the continuous time analogues of Theorem 2.6.1 and Corollary 2.6.1. The proofs follow the same lines as those of Theorem 2.6.1 and Corollary 2.6.1. The details are omitted.

Theorem 2.6.2 *The families of matrices $\{R_k, k \geq 0\}$ and $\{U_k, k \geq 1\}$ are the minimal non negative solutions to the equations (2.6.9) and (2.6.10).*

Corollary 2.6.2 *The family of matrices $\{R_k, k \geq 0\}$ is the minimal non negative solution to the equations (2.5.4).*

2.7 The matrices $\{G_k, k \geq 1\}$

In this section we introduce a third family of matrices, $\{G_k, k \geq 1\}$ that arise when studying LDQBDs. The matrices G_k are defined as follows for both discrete and continuous time LDQBDs. The (i,j)th element of the matrix G_k is the probability that, starting in the state (k,i), the process eventually reaches level $k-1$ and does so through the state $(k-1,j)$. We prove that the matrices G_k are the minimal non negative solutions to a family of non linear matrix equations and show how the matrices G_k are related to the matrices U_k and R_k.

2.7.1 Discrete time LDQBDs

The following theorem gives the relationship between the matrices G_k and U_k.

Theorem 2.7.1 *The families of matrices* $\{G_k, k \geq 1\}$ *and* $\{U_k, k \geq 1\}$ *satisfy the following equations;*

$$\begin{aligned} G_k &= (I - U_k)^{-1} A_2^{(k)}, & (2.7.1) \\ U_k &= A_1^{(k)} + A_0^{(k)} G_{k+1}. & (2.7.2) \end{aligned}$$

Proof: Latouche [55] proved the results for the case of a QBD. We generalise Latouche's methods to the case of a LDQBD.

In the proof of Lemma 2.6.1 we showed that the (i,j)th element of $(I - U_k)^{-1}$ is the expected number of visits to the state (k,j), before the process reaches level $k-1$ for the first time, given the process starts in state (k,i). Given this result and the probabilistic interpretation of G_k, equation (2.7.1) follows from Theorem 3.3.7 in Kemeny and Snell [41].

In Section 2.6.1 we showed that $[U_k]_{ij}$ is the probability that the next visit to level k is through phase j, given that the process started in state (k,i) and the process has not visited level $k-1$ in between. Hence equation (2.7.2) follows directly by conditioning on the state of the process after the first step. □

The following theorem gives the relationship between the matrices G_k and R_k.

Theorem 2.7.2 *The families of matrices $\{G_k, k \geq 1\}$ and $\{R_k, k \geq 0\}$ satisfy the following equations;*

$$\begin{aligned} G_k &= (I - A_1^{(k)} - R_k A_2^{(k+1)})^{-1} A_2^{(k)}, & (2.7.3) \\ R_k &= A_0^{(k)} (I - A_1^{(k+1)} - A_0^{(k+1)} G_{k+2})^{-1}. & (2.7.4) \end{aligned}$$

Proof: Equation (2.7.3) follows by substituting equation (2.6.1) into (2.7.1) and equation (2.7.4) follows by substituting equation (2.7.2) into (2.6.2). □

The following theorem shows that the matrices $\{G_k, k \geq 1\}$ satisfy a family of non linear matrix equations.

Theorem 2.7.3 *The matrices $\{G_k, k \geq 1\}$ satisfy the equations*

$$G_k = A_2^{(k)} + A_1^{(k)} G_k + A_0^{(k)} G_{k+1} G_k \qquad k \geq 1. \tag{2.7.5}$$

Proof: The result can be proved directly by conditioning on the state after the first step or algebraically as follows. Substituting equation (2.7.2) into equation (2.7.1) gives

$$G_k = (I - A_1^{(k)} - A_0^{(k)} G_{k+1})^{-1} A_2^{(k)} \tag{2.7.6}$$

which upon pre-multiplying both sides by $I - A_1^{(k)} - A_0^{(k)} G_{k+1}$ and re-arranging gives (2.7.5). □

The following theorem and corollary can be proved using the same methods as those used in the proofs of Theorem 2.6.1 and Corollary 2.6.1. We omit the details.

Theorem 2.7.4 *The families of matrices $\{G_k, k \geq 1\}$ and $\{U_k, k \geq 1\}$ are the minimal non negative solutions to the equations (2.7.1) and (2.7.2).*

Corollary 2.7.1 *The family of matrices $\{G_k, k \geq 1\}$ is the minimal non negative solution to the equations (2.7.5).*

2.7.2 Continuous time LDQBDs

We now present the continuous time analogues of the results presented in the previous section. The following theorem gives the relationship between the matrices G_k and U_k. Note that Latouche [55] proved the result for the case of a QBD.

Theorem 2.7.5 *The families of matrices* $\{G_k, k \geq 1\}$ *and* $\{U_k, k \geq 1\}$ *satisfy the following equations;*

$$\begin{aligned} G_k &= (-U_k)^{-1}Q_2^{(k)}, & (2.7.7) \\ U_k &= Q_1^{(k)} + Q_0^{(k)}G_{k+1}. & (2.7.8) \end{aligned}$$

Proof: In order for the process to make a first passage from level k to level $k-1$, it must make a number of return visits to level k (possibly zero), avoiding level $k-1$, and then make a direct transition from level k to level $k-1$. Hence we have

$$G_k = \left(\int_{\tau=0}^{\infty} {}_{k-1}P(k, k; \tau)d\tau\right)Q_2^{(k)}. \tag{2.7.9}$$

We saw in the proof of Lemma 2.6.2 that ${}_{k-1}P(k, k; \tau) = e^{U_k\tau}$ so we have

$$\begin{aligned} G_k &= \left(\int_{\tau=0}^{\infty} e^{U_k\tau}d\tau\right)Q_2^{(k)} & (2.7.10) \\ &= (-U_k)^{-1}Q_2^{(k)}. & (2.7.11) \end{aligned}$$

Equation (2.7.8) follows from the probabilistic interpretations of U_k and G_k. ☐

The following theorem gives the relationship between the matrices G_k and R_k.

Theorem 2.7.6 *The families of matrices* $\{G_k, k \geq 1\}$ *and* $\{R_k, k \geq 0\}$ *satisfy the following equations;*

$$\begin{aligned} G_k &= (-Q_1^{(k)} - R_kQ_2^{(k+1)})^{-1}Q_2^{(k)}, & (2.7.12) \\ R_k &= Q_0^{(k)}(-Q_1^{(k+1)} - Q_0^{(k+1)}G_{k+2})^{-1}. & (2.7.13) \end{aligned}$$

Proof: Equation (2.7.12) follows by substituting equation (2.6.9) into (2.7.7) and equation (2.7.13) follows by substituting equation (2.7.8) into (2.6.10). ☐

Theorem 2.7.7 *The matrices* $\{G_k, k \geq 1\}$ *satisfy the equations*

$$0 = Q_2^{(k)} + Q_1^{(k)}G_k + Q_0^{(k)}G_{k+1}G_k \qquad k \geq 1. \tag{2.7.14}$$

Proof: Substituting equation(2.7.8) into equation (2.7.7) gives

$$G_k = (-Q_1^{(k)} - Q_0^{(k)}G_{k+1})^{-1}Q_2^{(k)} \tag{2.7.15}$$

which upon pre-multiplying both sides by $-Q_1^{(k)} - Q_0^{(k)}G_{k+1}$ and re-arranging gives (2.7.14). ☐

The following theorem and corollary can be proved using the same methods as those used in the proofs of Theorem 2.6.1 and Corollary 2.6.1.

Theorem 2.7.8 *The families of matrices* $\{\boldsymbol{G}_k, k \geq 1\}$ *and* $\{\boldsymbol{U}_k, k \geq 1\}$ *are the minimal non negative solutions to the equations (2.7.7) and (2.7.8).*

Corollary 2.7.2 *The family of matrices* $\{\boldsymbol{G}_k, k \geq 1\}$ *is the minimal non negative solution to the equations (2.7.14).*

2.8 Ergodicity conditions

In this section we consider conditions under which LDQBDs are recurrent, (positive or null) and transient. The following theorem gives a necessary and sufficient condition for a LDQBD to be recurrent.

Theorem 2.8.1 *An irreducible and aperiodic LDQBD (either discrete or continuous time) is recurrent if and only if the Markov chain (process) governed by the transition matrix (generator matrix)* $\boldsymbol{A}_1^{(0)} + \boldsymbol{R}_0\boldsymbol{A}_2^{(1)}$ $(\boldsymbol{Q}_1^{(0)} + \boldsymbol{R}_0\boldsymbol{Q}_2^{(1)})$ *is positive recurrent.*

Proof: Since the proofs for discrete and continuous time are similar we consider only the discrete time case. If we start the Markov chain in level 0 and observe it only at time points when it is in level 0, then the transition probability matrix for this partially observed process is $\boldsymbol{A}_1^{(0)} + \boldsymbol{R}_0\boldsymbol{A}_2^{(1)}$. This partially observed process is an irreducible Markov chain on a finite state space so if it is positive recurrent then $\boldsymbol{A}_1^{(0)} + \boldsymbol{R}_0\boldsymbol{A}_2^{(1)}$ is stochastic. Using equations (2.6.2) and (2.7.1) we can show that $\boldsymbol{R}_0\boldsymbol{A}_2^{(1)} = \boldsymbol{A}_0^{(0)}\boldsymbol{G}_1$ and so since $\boldsymbol{A}_1^{(0)} + \boldsymbol{R}_0\boldsymbol{A}_2^{(1)}$ is stochastic, we have

$$\begin{aligned}
(\boldsymbol{A}_1^{(0)} + \boldsymbol{R}_0\boldsymbol{A}_2^{(1)})\boldsymbol{e} &= \boldsymbol{e} \\
\Rightarrow (\boldsymbol{A}_1^{(0)} + \boldsymbol{A}_0^{(0)}\boldsymbol{G}_1)\boldsymbol{e} &= \boldsymbol{e} \\
\Rightarrow \boldsymbol{A}_0^{(0)}\boldsymbol{G}_1\boldsymbol{e} &= \boldsymbol{A}_0^{(0)}\boldsymbol{e}
\end{aligned}$$

and hence $\boldsymbol{G}_1$ is stochastic. Now given the probabilistic interpretation of $\boldsymbol{G}_k$, since the LDQBD is irreducible, $\boldsymbol{G}_k$ must be stochastic for all $k \geq 1$ and hence the LDQBD must be recurrent.

Now suppose the LDQBD is recurrent. This implies that $\boldsymbol{G}_1$ is stochastic and hence $\boldsymbol{A}_1^{(0)} + \boldsymbol{R}_0\boldsymbol{A}_2^{(1)} = \boldsymbol{A}_1^{(0)} + \boldsymbol{A}_0^{(0)}\boldsymbol{G}_1$ is stochastic. Therefore since the chain with transition matrix $\boldsymbol{A}_1^{(0)} + \boldsymbol{R}_0\boldsymbol{A}_2^{(1)}$ is an irreducible Markov chain on a finite state space, the chain must be positive recurrent. □

We now give conditions for a recurrent LDQBD to be positive recurrent.

Theorem 2.8.2 *An irreducible, aperiodic and recurrent LDQBD (either discrete or continuous time) is positive recurrent if and only if*

$$\sum_{n=0}^{\infty} \boldsymbol{R}_0 \cdots \boldsymbol{R}_n \boldsymbol{e} < \infty. \tag{2.8.1}$$

Proof: We provide the proof for discrete time processes only. Firstly suppose that $\sum_{n=0}^{\infty} \boldsymbol{R}_0 \cdots \boldsymbol{R}_n \boldsymbol{e} < \infty$. From Theorem 2.8.1 the chain with transition matrix $\boldsymbol{A}_1^{(0)} + \boldsymbol{R}_0\boldsymbol{A}_2^{(1)}$ is positive recurrent. Let $\boldsymbol{y}_0$ be the stationary probability vector for this chain. If we define $\boldsymbol{x}_k, k \geq 0$ by $\boldsymbol{x}_0 = [1 + \boldsymbol{y}_0 \sum_{k=0}^{\infty} \prod_{\ell=0}^{k-1} \boldsymbol{R}_\ell \boldsymbol{e}]^{-1} \boldsymbol{y}_0$ and $\boldsymbol{x}_k = \boldsymbol{x}_{k-1}\boldsymbol{R}_{k-1}, k \geq 1$ then $\boldsymbol{x}_k, k \geq 0$ satisfy the global balance equations for the LDQBD and sum to one. Furthermore, since $\boldsymbol{y}_0 > \boldsymbol{0}$, $\boldsymbol{x}_k > \boldsymbol{0}\ \forall k \geq 0$. Therefore the LDQBD is positive recurrent since we have shown that the stationary probabilities are positive.

Now suppose the LDQBD is positive recurrent. This implies that $\sum_{k=0}^{\infty} \boldsymbol{x}_k \boldsymbol{e} = 1$ which implies $\boldsymbol{x}_0 \sum_{k=0}^{\infty} \prod_{\ell=0}^{k-1} \boldsymbol{R}_\ell \boldsymbol{e} = 1$ since $\boldsymbol{x}_k = \boldsymbol{x}_{k-1}\boldsymbol{R}_{k-1}, k \geq 1$. Hence (2.8.1) must hold. □

For the case of a QBD we have the following corollaries to Theorems 2.8.1 and 2.8.2. The proofs are obvious.

Corollary 2.8.1 *An irreducible and aperiodic QBD (either discrete or continuous time) is recurrent if and only if the Markov chain (process) governed by the transition matrix (generator matrix)* $\boldsymbol{A}_1^{(0)} + \boldsymbol{R}\boldsymbol{A}_2 \quad (\boldsymbol{Q}_1^{(0)} + \boldsymbol{R}\boldsymbol{Q}_2)$ *is positive recurrent.*

Corollary 2.8.2 *An irreducible, aperiodic and recurrent QBD (either discrete or continuous time) is positive recurrent if and only if* $(\boldsymbol{I} - \boldsymbol{R})^{-1}$ *exists and has finite row sums.*

Combining the results of Theorems 2.8.1 and 2.8.2 we can state that an irreducible and aperiodic LDQBD is positive recurrent if and only if the Markov chain (process) with transition matrix (generator matrix) $\boldsymbol{A}_1^{(0)} + \boldsymbol{R}_0\boldsymbol{A}_2^{(1)} \quad (\boldsymbol{Q}_1^{(0)} + \boldsymbol{R}_0\boldsymbol{Q}_2^{(1)})$ is positive recurrent and condition (2.8.1) holds. We note that Ramaswami [80, Theorem 3.7] gave this result whereas we have separated his result into two parts.

We now make some comments on Theorems 2.8.1 and 2.8.2. The Markov chain with transition matrix $\boldsymbol{A}_1^{(0)} + \boldsymbol{R}_0\boldsymbol{A}_2^{(1)}$ has a finite state space, and it is irreducible, so the only way it isn't positive recurrent is if $\boldsymbol{A}_1^{(0)} + \boldsymbol{R}_0\boldsymbol{A}_2^{(1)}$ is strictly substochastic. Hence from Theorem 2.8.1, if we can compute $\boldsymbol{R}_0$ then we can determine if the LDQBD is recurrent. Similarly, a continuous time process is transient if and only if $\boldsymbol{Q}_1^{(0)} + \boldsymbol{R}_0\boldsymbol{Q}_2^{(1)}$ is strictly non-conservative. Of course we can only find $\boldsymbol{R}_0$ and check that $\boldsymbol{A}_1^{(0)} + \boldsymbol{R}_0\boldsymbol{A}_2^{(1)}$ is stochastic to within some degree of accuracy. Hence any efforts to verify these conditions numerically must be applied with due care. It is clearly not possible to verify condition (2.8.1) numerically so other methods are needed to prove positive recurrence.

It seems that it is not possible to develop easily verifiable conditions for a LDQBD to be positive recurrent. This means that each problem must be analysed separately. However, for most problems, methods can be derived to show whether or not the process is positive recurrent. In the case of a discrete time QBD, Neuts [72] derived conditions under which the QBD is positive recurrent. We give this result in the following theorem together with the corresponding result for continuous time QBDs.

Theorem 2.8.3 *1. Consider a discrete time QBD where* $\boldsymbol{A} = \boldsymbol{A}_0 + \boldsymbol{A}_1 + \boldsymbol{A}_2$ *is irreducible and stochastic. Let* $\boldsymbol{\pi}$ *be such that* $\boldsymbol{\pi A} = \boldsymbol{\pi}$ *and* $\boldsymbol{\pi e} = 1$. *Then*

(a) if $\boldsymbol{\pi A}_0\boldsymbol{e} < \boldsymbol{\pi A}_2\boldsymbol{e}$ *then the LDQBD is positive recurrent,*

(b) if $\boldsymbol{\pi A}_0\boldsymbol{e} = \boldsymbol{\pi A}_2\boldsymbol{e}$ *then the LDQBD is null recurrent,*

(c) if $\boldsymbol{\pi A}_0\boldsymbol{e} > \boldsymbol{\pi A}_2\boldsymbol{e}$ *then the LDQBD is transient.*

2. Consider a continuous time QBD where $\hat{\boldsymbol{Q}} = \boldsymbol{Q}_0 + \boldsymbol{Q}_1 + \boldsymbol{Q}_2$ *is an irreducible q-matrix. Let* $\boldsymbol{\pi}$ *be such that* $\boldsymbol{\pi}\hat{\boldsymbol{Q}} = \boldsymbol{0}$ *and* $\boldsymbol{\pi e} = 1$. *Then*

(a) if $\boldsymbol{\pi Q}_0\boldsymbol{e} < \boldsymbol{\pi Q}_2\boldsymbol{e}$ *then the LDQBD is positive recurrent,*

(b) if $\boldsymbol{\pi Q}_0\boldsymbol{e} = \boldsymbol{\pi Q}_2\boldsymbol{e}$ *then the LDQBD is null recurrent,*

(c) if $\boldsymbol{\pi Q}_0\boldsymbol{e} > \boldsymbol{\pi Q}_2\boldsymbol{e}$ *then the LDQBD is transient.*

Proof: We refer to Neuts [72, Chapter 1] for the proof of part *1*. To prove part *2* we consider the jump chain which we introduced in Section 2.4.2. The jump chain is a discrete time QBD governed by the matrices $\boldsymbol{A}_1^{(0)} = \boldsymbol{\Phi}_0^{-1}\boldsymbol{Q}_1^{(0)} + \boldsymbol{I}$, $\boldsymbol{A}_i = \boldsymbol{\Phi}^{-1}\boldsymbol{Q}_i + \delta_{i1}\boldsymbol{I}$ $i = 0, 1, 2$ where $\boldsymbol{\Phi}_0 = -\mathrm{diag}([\boldsymbol{Q}_1^{(0)}]_{ii})$, $\boldsymbol{\Phi} = -\mathrm{diag}([\boldsymbol{Q}_1]_{ii})$ and δ_{ij} is the Kronecker delta. Now since

$\pi\hat{Q} = \mathbf{0}$, we have that

$$\begin{aligned}
&\pi\boldsymbol{\Phi}A \\
&= \pi(\boldsymbol{\Phi}A_0 + \boldsymbol{\Phi}A_1 + \boldsymbol{\Phi}A_2) \\
&= \pi(Q_0 + Q_1 + \boldsymbol{\Phi} + Q_2) \\
&= \pi\hat{Q} + \pi\boldsymbol{\Phi} \\
&= \pi\boldsymbol{\Phi}.
\end{aligned}$$

Hence if we set $\bar{\pi} = \pi\boldsymbol{\Phi}/(\pi\boldsymbol{\Phi}e)$ then $\bar{\pi}$ satisfies $\bar{\pi}A = \bar{\pi}$ and $\bar{\pi}e = 1$. Using part *1* and this expression for $\bar{\pi}$ we can show that the jump chain is positive recurrent, null recurrent or transient as $\pi Q_0 e < \pi Q_2 e$, $\pi Q_0 e = \pi Q_2 e$ or $\pi Q_0 e > \pi Q_2 e$. Now the proof is complete if we can show that positive recurrence, null recurrence or transience for the jump chain implies positive recurrence, null recurrence or transience respectively of the continuous time process. Proposition 3.2 in Anderson [5] states that if the jump chain is transient, then the continuous time process is also transient. Now suppose that $\bar{m} = (\bar{m}_0, \bar{m}_1, \ldots)$ is an invariant measure for the jump chain. A standard result from Markov chain theory gives us that $m = (\bar{m}_0\boldsymbol{\Phi}_0^{-1}, \bar{m}_1\boldsymbol{\Phi}^{-1}, \bar{m}_2\boldsymbol{\Phi}^{-1}, \ldots)$ is an invariant measure for the continuous time process. If we let H be the maximum row sum of $\boldsymbol{\Phi}^{-1}$ then if the jump chain is positive recurrent we have

$$\begin{aligned}
&\sum_{i=0}^{\infty} \bar{m}_i e < \infty \\
\Rightarrow &\sum_{i=0}^{\infty} \bar{m}_i H e < \infty \\
\Rightarrow &\sum_{i=0}^{\infty} \bar{m}_i \boldsymbol{\Phi}^{-1} e < \infty
\end{aligned}$$

since $\boldsymbol{\Phi}^{-1}e \leq He$. Hence we have that $\sum_{i=0}^{\infty} m_i e < \infty$ and therefore the continuous time process is also positive recurrent. Similarly, if the jump chain is null recurrent then the continuous time process will be null recurrent. This completes the proof. □

It is easy to adapt Theorem 2.8.3 to the case of a LDQBD that is eventually level independent. We hence have the following theorem.

Theorem 2.8.4 *1. Consider a discrete time LDQBD where* $A_0^{(k)} = A_0$, $A_1^{(k)} = A_1$, *and* $A_2^{(k)} = A_2$, $\forall k \geq K$ *and* $A = A_0 + A_1 + A_2$ *is irreducible and stochastic. Let* π *be such that* $\pi A = \pi$ *and* $\pi e = 1$. *Then*

(a) *if* $\boldsymbol{\pi} A_0 \mathbf{e} < \boldsymbol{\pi} A_2 \mathbf{e}$ *then the LDQBD is positive recurrent,*

(b) *if* $\boldsymbol{\pi} A_0 \mathbf{e} = \boldsymbol{\pi} A_2 \mathbf{e}$ *then the LDQBD is null recurrent,*

(c) *if* $\boldsymbol{\pi} A_0 \mathbf{e} > \boldsymbol{\pi} A_2 \mathbf{e}$ *then the LDQBD is transient.*

2. *Consider a continuous time QBD where* $Q_0^{(k)} = Q_0$, $Q_1^{(k)} = Q_1$, *and* $Q_2^{(k)} = Q_2$, $\forall k \geq K$ *and* $\hat{Q} = Q_0 + Q_1 + Q_2$ *is an irreducible q-matrix. Let* $\boldsymbol{\pi}$ *be such that* $\boldsymbol{\pi}\hat{Q} = \mathbf{0}$ *and* $\boldsymbol{\pi}\mathbf{e} = 1$*. Then*

(a) *if* $\boldsymbol{\pi} Q_0 \mathbf{e} < \boldsymbol{\pi} Q_2 \mathbf{e}$ *then the LDQBD is positive recurrent,*

(b) *if* $\boldsymbol{\pi} Q_0 \mathbf{e} = \boldsymbol{\pi} Q_2 \mathbf{e}$ *then the LDQBD is null recurrent,*

(c) *if* $\boldsymbol{\pi} Q_0 \mathbf{e} > \boldsymbol{\pi} Q_2 \mathbf{e}$ *then the LDQBD is transient.*

Chapter 3

Calculating the matrices $\boldsymbol{G}_k$ and $\boldsymbol{R}_k$

3.1 Introduction

In this chapter we present algorithms for numerically computing the matrices $\boldsymbol{G}_k$ and $\boldsymbol{R}_k$. These algorithms will be used in Chapter 4 when we develop algorithms for computing the stationary distribution of a LDQBD. The algorithm for computing $\boldsymbol{R}_k$ was presented in Bright and Taylor [12] but it is given in greater detail here.

We begin the chapter by presenting closed form expressions for $\boldsymbol{G}_k$ and $\boldsymbol{R}_k$. In the case of discrete time processes the expressions we give are taken from Ramaswami and Taylor [84]. We derive expressions for continuous time processes by considering the jump chain which we introduced in Section 2.4.2.

Having derived explicit expressions for $\boldsymbol{G}_k$ and $\boldsymbol{R}_k$ we show how these expressions can be used to compute the matrices numerically. Since the expressions for both $\boldsymbol{G}_k$ and $\boldsymbol{R}_k$ are in terms of an infinite sum, we provide appropriate methods for truncating these sums. The terms in both the infinite sums are defined by recursive relationships and we discuss an efficient way of implementing these relationships.

Since the algorithms we present are for LDQBDs, it would seem that they are also suitable for QBDs. However, the extra structure possessed by QBDs means that there are more efficient methods available for their analysis. In the case of a QBD, Latouche and Ramaswami [56] have developed very efficient methods for evaluating $\boldsymbol{G}$ and $\boldsymbol{R}$ and we give a summary of these in Sections 3.5 and 3.6. We also note that an algorithm developed by Bini and Meini [10] can be used to compute $\boldsymbol{G}$ and $\boldsymbol{R}$. This algorithm is similar in structure to Latouche and Ramaswami's algorithm yet it has the added feature that it can also be applied to the more general M/G/1–type process. We conclude the chapter with some discussion on the algorithms presented. In particular, we give probabilistic interpretations of the algorithms.

3.2 An expression for the matrix G_k

In this section we derive closed form expressions for the matrices G_k in discrete and continuous time. The results we present for discrete time processes are taken from Ramaswami and Taylor [84].

3.2.1 Discrete time LDQBDs

In order to derive closed form expressions for the matrices $\{G_k, k \geq 1\}$ we define two additional families of matrices, namely $\{U_k^\ell, k \geq 1, 0 \leq \ell \leq \left\lfloor \frac{\log k}{\log 2} \right\rfloor\}$ and $\{D_k^\ell, k \geq 1, 0 \leq \ell \leq \left\lfloor \frac{\log k}{\log 2} \right\rfloor\}$. To define these matrices we first define the following stopping time.

Definition 3.2.1 $\gamma(\boldsymbol{k}) = \inf\{n \geq 0 : X(n) \in \boldsymbol{k}\}$.

In the following definition $P(A)$ represents the probability of event A.

Definition 3.2.2 *The matrix* U_k^ℓ *has dimensions* $M(k) \times M(k+2^\ell)$ *and is defined for* $k \geq 1$ *and* $0 \leq \ell \leq \left\lfloor \frac{\log k}{\log 2} \right\rfloor$ *by*

$$\left[U_k^\ell\right]_{ij} = P\left(\gamma(\boldsymbol{k}+\boldsymbol{2^\ell}) < \gamma(\boldsymbol{k}-\boldsymbol{2^\ell}) \text{ and } X(\gamma(\boldsymbol{k}+\boldsymbol{2^\ell})) = (k+2^\ell, j) \mid X(0) = (k,i)\right) \tag{3.2.1}$$

and the matrix D_k^ℓ *has dimensions* $M(k) \times M(k-2^\ell)$ *and is defined for* $k \geq 1$ *and* $0 \leq \ell \leq \left\lfloor \frac{\log k}{\log 2} \right\rfloor$ *by*

$$\left[D_k^\ell\right]_{ij} = P\left(\gamma(\boldsymbol{k}-\boldsymbol{2^\ell}) < \gamma(\boldsymbol{k}+\boldsymbol{2^\ell}) \text{ and } X(\gamma(\boldsymbol{k}-\boldsymbol{2^\ell})) = (k-2^\ell, j) \mid X(0) = (k,i)\right). \tag{3.2.2}$$

It is clear from the definitions why the quantities can only be defined for a finite number of ℓ values. From now on we will write expressions like (3.2.2) in the abbreviated matrix form

$$D_k^\ell = P\left(\gamma(\boldsymbol{k}-\boldsymbol{2^\ell}) < \gamma(\boldsymbol{k}+\boldsymbol{2^\ell}) \text{ and } X(\gamma(\boldsymbol{k}-\boldsymbol{2^\ell})) \mid X(0) \in \boldsymbol{k}\right) \tag{3.2.3}$$

as in Ramaswami and Taylor [84].

The following theorem appears in [84] as Lemma 2.1. There the authors also give the result for the case where $M(k) = \infty$.

Theorem 3.2.1 *$\mathbf{U}^\ell_k$ and $\mathbf{D}^\ell_k$ defined above are given by the following recursive equations.*

$$\mathbf{U}^0_k = (\mathbf{I} - \mathbf{A}^{(k)}_1)^{-1}\mathbf{A}^{(k)}_0, \tag{3.2.4}$$

$$\mathbf{D}^0_k = (\mathbf{I} - \mathbf{A}^{(k)}_1)^{-1}\mathbf{A}^{(k)}_2, \tag{3.2.5}$$

$$\mathbf{U}^{\ell+1}_k = \left[\mathbf{I} - \mathbf{U}^\ell_k\mathbf{D}^\ell_{k+2^\ell} - \mathbf{D}^\ell_k\mathbf{U}^\ell_{k-2^\ell}\right]^{-1}\mathbf{U}^\ell_k\mathbf{U}^\ell_{k+2^\ell}, \tag{3.2.6}$$

$$\mathbf{D}^{\ell+1}_k = \left[\mathbf{I} - \mathbf{U}^\ell_k\mathbf{D}^\ell_{k+2^\ell} - \mathbf{D}^\ell_k\mathbf{U}^\ell_{k-2^\ell}\right]^{-1}\mathbf{D}^\ell_k\mathbf{D}^\ell_{k-2^\ell}. \tag{3.2.7}$$

Proof: The expressions for $\mathbf{U}^0_k$ and $\mathbf{D}^0_k$ are easily shown. By using simple probabilistic arguments it can be shown that

$$\mathbf{U}^{\ell+1}_k = (\mathbf{U}^\ell_k\mathbf{D}^\ell_{k+2^\ell} + \mathbf{D}^\ell_k\mathbf{U}^\ell_{k-2^\ell})\mathbf{U}^{\ell+1}_k + \mathbf{U}^\ell_k\mathbf{U}^\ell_{k+2^\ell}. \tag{3.2.8}$$

From the definitions of $\mathbf{U}^\ell_k$ and $\mathbf{D}^\ell_k$ the matrix $\mathbf{U}^\ell_k\mathbf{D}^\ell_{k+2^\ell} + \mathbf{D}^\ell_k\mathbf{U}^\ell_{k-2^\ell}$ gives the probabilities that the process returns to level k after having visited level $k+2^\ell$ or $k-2^\ell$ but not having visited levels $k+2^{\ell+1}$ or $k-2^{\ell+1}$. Since the process is irreducible, this matrix must be sub-stochastic and again irreducibility implies the inverse $[\mathbf{I} - \mathbf{U}^\ell_k\mathbf{D}^\ell_{k+2^\ell} - \mathbf{D}^\ell_k\mathbf{U}^\ell_{k-2^\ell}]^{-1}$ exists. Hence equation (3.2.8) can be rewritten as equation (3.2.6). A similar argument can be used to derive equation (3.2.7). ☐

The following theorem which appears as Theorem 2.1 in Ramaswami and Taylor [84], gives an explicit expression for the matrix G_k for a discrete time LDQBD. The theorem is a generalisation of the corresponding result for a QBD given in Latouche and Ramaswami [56] .

Theorem 3.2.2 *For a discrete time LDQBD the matrices $\{G_k, k \geq 1\}$ are explicitly given by*

$$G_k = \sum_{\ell=0}^{\infty}\left[\prod_{i=0}^{\ell-1}\mathbf{U}^i_{k-1+2^i}\right]\mathbf{D}^\ell_{k-1+2^\ell} \tag{3.2.9}$$

where $\mathbf{U}^\ell_k$ and $\mathbf{D}^\ell_k$ are defined by equations (3.2.4)-(3.2.7).

Proof: We firstly show that

$$\prod_{i=0}^{\ell}\mathbf{U}^i_{k-1+2^i} = P\Big(\gamma(\mathbf{k}-1+2^{\ell+1}) < \gamma(\mathbf{k}-1) \text{ and } X(\gamma(\mathbf{k}-1+2^{\ell+1})) \mid X(0) \in \mathbf{k}\Big) \tag{3.2.10}$$

and

$$\prod_{i=0}^{\ell} U^{i}_{k-1+2^i} D^{\ell+1}_{k-1+2^{\ell+1}} = P\Big(\gamma(\mathbf{k-1+2^{\ell+1}}) < \gamma(\mathbf{k-1}) < \gamma(\mathbf{k-1+2^{\ell+2}})$$
$$\text{and } X(\gamma(\mathbf{k-1})) \mid X(0) \in \mathbf{k}\Big). \tag{3.2.11}$$

We do this by induction. Equation (3.2.10) clearly holds for $\ell = 0$. Make the assumption that it holds for a general ℓ.

$$P\Big(\gamma(\mathbf{k-1+2^{\ell+2}}) < \gamma(\mathbf{k-1}) \text{ and } X(\gamma(\mathbf{k-1+2^{\ell+2}})) \mid X(0) \in \mathbf{k}\Big)$$

$$= P\Big(\gamma(\mathbf{k-1+2^{\ell+1}}) < \gamma(\mathbf{k-1}) \text{ and } X(\gamma(\mathbf{k-1+2^{\ell+1}})) \mid X(0) \in \mathbf{k}\Big)$$

$$P\Big(\gamma(\mathbf{k-1+2^{\ell+2}}) < \gamma(\mathbf{k-1}) \text{ and } X(\gamma(\mathbf{k-1+2^{\ell+2}})) \mid$$
$$\gamma(\mathbf{k-1+2^{\ell+1}}) < \gamma(\mathbf{k-1}) \text{ and } X(\mathbf{k-1+2^{\ell+1}})\Big)$$

$$= \left[\prod_{i=0}^{\ell} U^{i}_{k-1+2^i}\right] P\Big(\gamma(\mathbf{k-1+2^{\ell+2}}) < \gamma(\mathbf{k-1})$$
$$\text{and } X(\gamma(\mathbf{k-1+2^{\ell+2}})) \mid X(0) \in \mathbf{k-1+2^{\ell+1}}\Big)$$

$$= \left[\prod_{i=0}^{\ell} U^{i}_{k-1+2^i}\right] U^{\ell+1}_{k-1+2^{\ell+1}}$$

which is (3.2.10) for the $\ell + 1$ case. Expression (3.2.11) is proved similarly.

Now from the definition of $\boldsymbol{G}_k$ we can write $\boldsymbol{G}_k$ in the form

$$\boldsymbol{G}_k = P\left(\gamma(\mathbf{k-1}) < \infty \text{ and } X(\gamma(\mathbf{k-1})) \mid X(0) \in \mathbf{k}\right). \tag{3.2.12}$$

Hence we need to determine when $\gamma(\mathbf{k-1})$ is finite, given $X(0) \in \mathbf{k}$. Since $X(n)$ is skip free on the levels we have that, if $X(0) \in \mathbf{k}$, then $\gamma(\mathbf{k-1+2^{\ell+1}}) > \gamma(\mathbf{k-1+2^{\ell}}) \geq 2^{\ell}-1$. This implies that $\lim_{\ell\to\infty} \gamma(\mathbf{k-1+2^{\ell}}) = \infty$. We also have that $\gamma(\mathbf{k-1+2^{\ell}}) < \infty$ a.s. for all $\ell \geq 0$ because the process is irreducible. Hence we can say that $\gamma(\mathbf{k-1})$ is finite if and only if its value is in one of the intervals $(\gamma(\mathbf{k-1+2^{\ell}}), \gamma(\mathbf{k-1+2^{\ell+1}}))$

$\ell \geq 0$. Hence

$$
\begin{aligned}
G_k = \sum_{\ell=0}^{\infty} P\Big(\gamma(\boldsymbol{k}-\mathbf{1}+\mathbf{2}^{\ell}) < \gamma(\boldsymbol{k}-\mathbf{1}) < \gamma(\boldsymbol{k}-\mathbf{1}+\mathbf{2}^{\ell+1}) \\
\text{and } X(\gamma(\boldsymbol{k}-\mathbf{1})) \mid X(0) \in \boldsymbol{k}\Big) \\
= \sum_{\ell=0}^{\infty} \prod_{i=0}^{\ell-1} U^{i}_{k-1+2^{i}} D^{\ell}_{k-1+2^{\ell}}
\end{aligned}
$$

by (3.2.11). □

3.2.2 Continuous time LDQBDs

To obtain an expression for G_k for a continuous time LDQBD we consider the jump chain. As mentioned in Section 2.4.2, the jump chain is a discrete time LDQBD obtained from the continuous time LDQBD by observing it only at time points where there is a change of state. The jump chain is governed by the matrices

$$
\begin{aligned}
A_0^{(k)} &= \Phi_k^{-1} Q_0^{(k)}, && (3.2.13)\\
A_1^{(k)} &= \Phi_k^{-1} Q_1^{(k)} + I, && (3.2.14)\\
A_2^{(k)} &= \Phi_k^{-1} Q_2^{(k)} && (3.2.15)
\end{aligned}
$$

where Φ_k = -diag($[Q_1^{(k)}]_{ii}$). The continuous time version of Theorem 3.2.2 is as follows.

Theorem 3.2.3 *For a continuous time LDQBD the matrices* $\{G_k, k \geq 1\}$ *are explicitly given by*

$$
G_k = \sum_{\ell=0}^{\infty} \left[\prod_{i=0}^{\ell-1} U^{i}_{k-1+2^{i}} \right] D^{\ell}_{k-1+2^{\ell}} \tag{3.2.16}
$$

where U_k^{ℓ} *and* D_k^{ℓ} *are defined recursively by*

$$
\begin{aligned}
U_k^0 &= (-Q_1^{(k)})^{-1} Q_0^{(k)}, && (3.2.17)\\
D_k^0 &= (-Q_1^{(k)})^{-1} Q_2^{(k)}, && (3.2.18)\\
U_k^{\ell+1} &= \left[I - U_k^{\ell} D^{\ell}_{k+2^{\ell}} - D_k^{\ell} U^{\ell}_{k-2^{\ell}} \right]^{-1} U_k^{\ell} U^{\ell}_{k+2^{\ell}}, && (3.2.19)\\
D_k^{\ell+1} &= \left[I - U_k^{\ell} D^{\ell}_{k+2^{\ell}} - D_k^{\ell} U^{\ell}_{k-2^{\ell}} \right]^{-1} D_k^{\ell} D^{\ell}_{k-2^{\ell}}. && (3.2.20)
\end{aligned}
$$

Proof: By the definition of G_k it is clear that G_k for a continuous time LDQBD is equal to G_k for its jump chain. Hence by Theorem 3.2.2 G_k is given by equation (3.2.9) where

U_k^ℓ and D_k^ℓ are defined by equations (3.2.4)-(3.2.7). When the values for $A_0^{(k)}$, $A_1^{(k)}$ and $A_2^{(k)}$ given by equations (3.2.13)-(3.2.15) are substituted, equations (3.2.4)-(3.2.7) reduce to equations (3.2.17)-(3.2.20). □

3.3 An expression for the matrix R_k

In this section we derive closed form expressions for the matrices R_k in discrete and continuous time. Our derivation in discrete time is based on a duality result that will be presented in Chapter 6. The result in continuous time is obtained by considering the jump chain.

3.3.1 Discrete time LDQBDs

In order to obtain an explicit expression for R_k, we appeal to a duality result that will be presented in Chapter 6. The result there is presented for M/G/1–type and GI/M/1–type processes which can be thought of as extensions of LDQBDs. The results in Chapter 6 are presented only for continuous time processes although analogous results hold for discrete time processes. We state the result here for the special case of a discrete time LDQBD. For more details we refer to Section 6.12 of Chapter 6.

Consider a discrete time LDQBD with transition matrix given by equation (2.2.1) and invariant measure m. Define $\Delta_k = \text{diag}(m_k)$ and put

$$\bar{A}_0^{(k)} = \Delta_k^{-1}\left(A_2^{(k+1)}\right)'\Delta_{k+1} \qquad k \geq 0, \qquad (3.3.1)$$

$$\bar{A}_1^{(k)} = \Delta_k^{-1}\left(A_1^{(k)}\right)'\Delta_k \qquad k \geq 0, \qquad (3.3.2)$$

$$\bar{A}_2^{(k)} = \Delta_k^{-1}\left(A_0^{(k-1)}\right)'\Delta_{k-1} \qquad k \geq 1. \qquad (3.3.3)$$

It is easy to show that the matrices $\bar{A}_0^{(k)}$, $\bar{A}_1^{(k)}$ and $\bar{A}_2^{(k)}$ define an irreducible, aperiodic LDQBD. It can be shown using methods similar to those used in the proof of Theorem 6.12.1 that, if $\bar{G}_k$ is the minimal non negative solution to

$$\bar{G}_k = \bar{A}_2^{(k)} + \bar{A}_1^{(k)}\bar{G}_k + \bar{A}_0^{(k)}\bar{G}_{k+1}\bar{G}_k \qquad (3.3.4)$$

then R_k is given by

$$R_k = \Delta_k^{-1}\bar{G}'_{k+1}\Delta_{k+1} \qquad k \geq 0. \qquad (3.3.5)$$

In Chapter 6 we give a probabilistic interpretation of the continuous time analogue of equation (3.3.5). This interpretation is easily extended to discrete time processes.

In order to be able to construct the matrices $\bar{A}_i^{(k)}, i = 0, 1, 2$, an invariant measure $\boldsymbol{m}$ must exist. From Markov chain theory we know that if the LDQBD is recurrent then an invariant measure exists. Hence the only situation where an invariant measure may fail to exist is when the LDQBD is transient. Latouche and Taylor [58] showed that it is always possible to construct a matrix-geometric invariant measure for a transient QBD. Furthermore, they give a simple method for doing so. They also showed that under certain conditions it is possible to construct a matrix-product form invariant measure for transient LDQBDs. If we are not concerned with the form of the invariant measure, it can be shown by applying Theorem 2 in Harris [36] that an invariant measure exists for all transient LDQBDs. Theorem 2 of [36] is stated in Chapter 6 as Lemma 6.2.2 and can be applied by noting that for all states (k, j), the only states (ℓ, i) for which $P_{(\ell,i),(k,j)} > 0$ are the states in levels $k-1, k$ and $k+1$. Since the set of states consisting of the states in levels $k-1$, k and $k+1$ is a finite set of states, Theorem 2 of [36] implies that an invariant measure exists for all transient discrete time irreducible LDQBDs. Hence we can always construct the matrices $\bar{A}_i^{(k)}$, $i = 0, 1, 2$ defined by equations (3.3.1)-(3.3.3).

The following theorem presents the same result as that given in Theorem 2.2 of Ramaswami and Taylor [84]. However, since Ramaswami and Taylor were particularly interested in QBDs where $M(k) = \infty \; \forall k \geq 0$, they needed to make the assumption that an invariant measure exists. Since we have assumed that $M(k) < \infty \; \forall k \geq 0$, an invariant measure will always exist. We prove the following theorem using the expression (3.3.5) whereas Ramaswami and Taylor [84] proved their Theorem 2.2 using a relationship similar to, but not equivalent to (3.3.5).

Theorem 3.3.1 *For a discrete time LDQBD the matrices* $\{\boldsymbol{R}_k, k \geq 0\}$ *are explicitly given by*

$$\boldsymbol{R}_k = \sum_{\ell=0}^{\infty} \widetilde{\boldsymbol{U}}_k^{\ell} \left[\prod_{i=0}^{\ell-1} \widetilde{\boldsymbol{D}}_{k+2^{\ell-i}}^{\ell-1-i} \right] \tag{3.3.6}$$

where $\widetilde{\boldsymbol{U}}_k^{\ell}$ *and* $\widetilde{\boldsymbol{D}}_k^{\ell}$ *are* $M(k) \times M(k+2^{\ell})$ *and* $M(k) \times M(k-2^{\ell})$ *matrices respectively* *and are defined recursively by the equations*

$$\begin{aligned}
\widetilde{U}_k^0 &= A_0^{(k)}(I - A_1^{(k+1)})^{-1}, && (3.3.7)\\
\widetilde{D}_k^0 &= A_2^{(k)}(I - A_1^{(k-1)})^{-1}, && (3.3.8)\\
\widetilde{U}_k^{\ell+1} &= \widetilde{U}_k^{\ell}\widetilde{U}_{k+2^\ell}^{\ell}\left[I - \widetilde{U}_{k+2^{\ell+1}}^{\ell}\widetilde{D}_{k+3.2^\ell}^{\ell} - \widetilde{D}_{k+2^{\ell+1}}^{\ell}\widetilde{U}_{k+2^\ell}^{\ell}\right]^{-1}, && (3.3.9)\\
\widetilde{D}_k^{\ell+1} &= \widetilde{D}_k^{\ell}\widetilde{D}_{k-2^\ell}^{\ell}\left[I - \widetilde{U}_{k-2^{\ell+1}}^{\ell}\widetilde{D}_{k-2^\ell}^{\ell} - \widetilde{D}_{k-2^{\ell+1}}^{\ell}\widetilde{U}_{k-3.2^\ell}^{\ell}\right]^{-1}. && (3.3.10)
\end{aligned}$$

Proof: From equation (3.3.5) we have

$$\begin{aligned}
R_k &= \Delta_k^{-1}\bar{G}_{k+1}'\Delta_{k+1} && (3.3.11)\\
&= \Delta_k^{-1}\left(\sum_{\ell=0}^{\infty}\left[\prod_{i=0}^{\ell-1}\bar{U}_{k+2^i}^{i}\right]\bar{D}_{k+2^\ell}^{\ell}\right)'\Delta_{k+1} && (3.3.12)\\
&= \Delta_k^{-1}\sum_{\ell=0}^{\infty}(\bar{D}_{k+2^\ell}^{\ell})'\left[\prod_{i=0}^{\ell-1}(\bar{U}_{k+2^{\ell-1-i}}^{\ell-1-i})'\right]\Delta_{k+1} && (3.3.13)\\
&= \sum_{\ell=0}^{\infty}\Delta_k^{-1}(\bar{D}_{k+2^\ell}^{\ell})'\Delta_{k+2^\ell}\left[\prod_{i=0}^{\ell-1}\Delta_{k+2^{\ell-i}}^{-1}(\bar{U}_{k+2^{\ell-1-i}}^{\ell-1-i})'\Delta_{k+2^{\ell-1-i}}\right] && (3.3.14)
\end{aligned}$$

so provided we can show

$$\widetilde{U}_k^{\ell} = \Delta_k^{-1}(\bar{D}_{k+2^\ell}^{\ell})'\Delta_{k+2^\ell} \qquad \ell \geq 0, \tag{3.3.15}$$

and

$$\widetilde{D}_k^{\ell} = \Delta_k^{-1}(\bar{U}_{k-2^\ell}^{\ell})'\Delta_{k-2^\ell} \qquad \ell \geq 0, \tag{3.3.16}$$

where $\bar{U}_k^{\ell}$ and $\bar{D}_k^{\ell}$ are given by equations (3.2.4)-(3.2.7) with $A_0^{(k)}$, $A_1^{(k)}$ and $A_2^{(k)}$ replaced by $\bar{A}_0^{(k)}$, $\bar{A}_1^{(k)}$ and $\bar{A}_2^{(k)}$ respectively, then we have the result. This can be shown using induction. ☐

3.3.2 Continuous time LDQBDs

To obtain a closed from expression for the matrix R_k for continuous time LDQBDs we use the jump chain as we did when deriving an expression for the matrix G_k in continuous time. We showed in Chapter 2, Section 2.4.2 that if R_k^J is the R_k matrix for the jump chain then

$$R_k = \Phi_k R_k^J \Phi_{k+1}^{-1}. \tag{3.3.17}$$

We now prove the following theorem.

Theorem 3.3.2 *For a continuous time LDQBD the matrices $\{\boldsymbol{R}_k, k \geq 0\}$ are explicitly given by*

$$\boldsymbol{R}_k = \sum_{\ell=0}^{\infty} \widetilde{\boldsymbol{U}}_k^{\ell} \prod_{i=0}^{\ell-1} \widetilde{\boldsymbol{D}}_{k+2^{\ell-i}}^{\ell-1-i} \qquad k \geq 0 \tag{3.3.18}$$

where $\widetilde{\boldsymbol{U}}_k^{\ell}$ and $\widetilde{\boldsymbol{D}}_k^{\ell}$ are $M(k) \times M(k+2^{\ell})$ and $M(k) \times M(k-2^{\ell})$ matrices respectively and are defined recursively by

$$\widetilde{\boldsymbol{U}}_k^{0} = \boldsymbol{Q}_0^{(k)}(-\boldsymbol{Q}_1^{(k+1)})^{-1}, \tag{3.3.19}$$

$$\widetilde{\boldsymbol{D}}_k^{0} = \boldsymbol{Q}_2^{(k)}(-\boldsymbol{Q}_1^{(k-1)})^{-1}, \tag{3.3.20}$$

$$\widetilde{\boldsymbol{U}}_k^{\ell+1} = \widetilde{\boldsymbol{U}}_k^{\ell}\widetilde{\boldsymbol{U}}_{k+2^{\ell}}^{\ell}\left[\boldsymbol{I} - \widetilde{\boldsymbol{U}}_{k+2^{\ell+1}}^{\ell}\widetilde{\boldsymbol{D}}_{k+3.2^{\ell}}^{\ell} - \widetilde{\boldsymbol{D}}_{k+2^{\ell+1}}^{\ell}\widetilde{\boldsymbol{U}}_{k+2^{\ell}}^{\ell}\right]^{-1}, \tag{3.3.21}$$

$$\widetilde{\boldsymbol{D}}_k^{\ell+1} = \widetilde{\boldsymbol{D}}_k^{\ell}\widetilde{\boldsymbol{D}}_{k-2^{\ell}}^{\ell}\left[\boldsymbol{I} - \widetilde{\boldsymbol{U}}_{k-2^{\ell+1}}^{\ell}\widetilde{\boldsymbol{D}}_{k-2^{\ell}}^{\ell} - \widetilde{\boldsymbol{D}}_{k-2^{\ell+1}}^{\ell}\widetilde{\boldsymbol{U}}_{k-3.2^{\ell}}^{\ell}\right]^{-1}. \tag{3.3.22}$$

Proof: From equations (3.3.17) and (3.3.6) we have

$$\boldsymbol{R}_k = \boldsymbol{\Phi}_k \sum_{\ell=0}^{\infty} \bar{\boldsymbol{U}}_k^{\ell} \left[\prod_{i=0}^{\ell-1} \bar{\boldsymbol{D}}_{k+2^{\ell-i}}^{\ell-1-i}\right] \boldsymbol{\Phi}_{k+1}^{-1} \tag{3.3.23}$$

where $\bar{\boldsymbol{U}}_k^{\ell}$ and $\bar{\boldsymbol{D}}_k^{\ell}$ are given by equations (3.3.7)-(3.3.10) with $\boldsymbol{A}_0^{(k)}$, $\boldsymbol{A}_1^{(k)}$ and $\boldsymbol{A}_2^{(k)}$ given by equations (3.2.13)-(3.2.15). Equation (3.3.23) is equivalent to

$$\boldsymbol{R}_k = \sum_{\ell=0}^{\infty} \boldsymbol{\Phi}_k \bar{\boldsymbol{U}}_k^{\ell} \boldsymbol{\Phi}_{k+2^{\ell}}^{-1} \left[\prod_{i=0}^{\ell-1} \boldsymbol{\Phi}_{k+2^{\ell-i}}^{-1} \bar{\boldsymbol{D}}_{k+2^{\ell-i}}^{\ell-1-i} \boldsymbol{\Phi}_{k+2^{\ell-1-i}}^{-1}\right] \tag{3.3.24}$$

and hence the proof is completed by using induction to show that

$$\widetilde{\boldsymbol{U}}_k^{\ell} = \boldsymbol{\Phi}_k \bar{\boldsymbol{U}}_k^{\ell} \boldsymbol{\Phi}_{k+2^{\ell}}^{-1} \tag{3.3.25}$$

and

$$\widetilde{\boldsymbol{D}}_k^{\ell} = \boldsymbol{\Phi}_k \bar{\boldsymbol{D}}_k^{\ell} \boldsymbol{\Phi}_{k-2^{\ell}}^{-1}. \tag{3.3.26}$$

□

Theorem 3.3.2 was given in Lemma 1 of Bright and Taylor [12] but there it was assumed that the LDQBD is positive recurrent. We have shown here that the positive recurrence condition can be removed.

3.4 Efficiently computing G_k and R_k

In this section we develop algorithms for numerically computing the matrices G_k and R_k. These algorithms are based on the closed form expressions for G_k and R_k that were given in Sections 3.2 and 3.3. Since these expressions are in terms of infinite sums, in order to evaluate G_k and R_k via these expressions, we must truncate the infinite sums. The other issue to consider when computing G_k and R_k via the expressions in Sections 3.2 and 3.3 is how to implement efficiently the recursive relationships defining the matrices U_k^ℓ and D_k^ℓ and $\widetilde{U}_k^\ell$ and $\widetilde{D}_k^\ell$. At the beginning of this section we consider how to truncate the infinite sums and then we consider the question of efficient implementation of the various recursive relationships.

3.4.1 Truncation rules for G_k

Define $G_k(\ell)$ to be the sum of the first $\ell+1$ terms in the infinite sum in (3.2.9) or (3.2.16). When calculating a numerical value for G_k we take $G_k = G_k(L)$ for some appropriately chosen value of L. That is, we truncate the sum at $\ell = L$. A simple way of determining a value of L is stated in the following truncation rule.

Truncation rule 3.4.1 *Take $G_k = G_k(L)$ where L is the smallest value of ℓ such that $[G_k(\ell) - G_k(\ell-1)]_{max} < \epsilon$ for some tolerance ϵ where $[M]_{max}$ is the maximum entry of the matrix M.*

Truncation rule 3.4.1 simply says that we should stop computing extra terms once we compute a term that is sufficiently small. Note however that the rule does not guarantee that $[G_k(\ell) - G_k(\ell-1)]_{max} < \epsilon$ holds $\forall \ell \geq L$ which is what we want to hold. In the case where $X(n)$ is positive recurrent, it is known that G_k is stochastic which leads to the following more precise truncation rule.

Truncation rule 3.4.2 *If $X(n)$ is positive recurrent, take $G_k = G_k(L)$ where L is the smallest value of ℓ such that $\|G_k(\ell)\mathbf{e} - \mathbf{e}\|_\infty < \epsilon$ for some tolerance ϵ.*

3.4.2 Truncation rules for R_k

Define $R_k(\ell)$ to be the sum of the first $\ell + 1$ terms in the infinite sum in (3.3.6) or (3.3.18). As we did in the case of G_k, when calculating a numerical value for R_k we take

$R_k = R_k(L)$ for some appropriately chosen value of L. A simple way of determining a value of L is stated in the following truncation rule which is similar to Truncation rule 3.4.1.

Truncation rule 3.4.3 *Take* $R_k = R_k(L)$ *where* L *is the smallest value of* ℓ *such that* $[R_k(\ell) - R_k(\ell - 1)]_{max} < \epsilon$ *for some tolerance* ϵ.

Truncation rule 3.4.3 has the same problem as Truncation rule 3.4.1. That is, it does not guarantee that $[R_k(\ell) - R_k(\ell - 1)]_{max} < \epsilon$ holds $\forall \ell > L$. However, in many situations it is possible, by considering the structure of the LDQBD, to show that $[R_k(\ell) - R_k(\ell - 1)]_{max} < \epsilon$ holds $\forall \ell > L$ when L is chosen via Truncation rule 3.4.3. As was the case when calculating G_k, it is possible to derive a more precise truncation rule provided the LDQBD is positive recurrent. This can be done by extending Lemma 1.2.3 of Neuts [72] to the case of a LDQBD, as we now show.

From equation (2.6.2) we have for a discrete time LDQBD,

$$\begin{aligned} R_k A_2^{(k+1)} &= A_0^{(k)}(I - U_{k+1})^{-1} A_2^{(k+1)} & (3.4.1) \\ &= A_0^{(k)} G_{k+1} & (3.4.2) \end{aligned}$$

where we obtain the last equality from equation (2.7.1). Now in the case of a positive recurrent LDQBD, G_k is stochastic so we have

$$R_k A_2^{(k+1)} e = A_0^{(k)} e. \qquad (3.4.3)$$

Similarly, for a continuous time LDQBD we can show that

$$R_k Q_2^{(k+1)} e = Q_0^{(k)} e. \qquad (3.4.4)$$

So in the case of a positive recurrent LDQBD we have the following truncation rule.

Truncation rule 3.4.4 *Take* $R_k = R_k(L)$ *where* L *is the smallest value of* ℓ *such that for some tolerance* ϵ, $\|R_k(\ell) A_2^{(k+1)} e - A_0^{(k)} e\|_\infty < \epsilon$ *for a discrete time LDQBD or* $\|R_k(\ell) Q_2^{(k+1)} e - Q_0^{(k)} e\|_\infty < \epsilon$ *for a continuous time LDQBD.*

3.4.3 Truncation rules for QBDs

Obviously the truncation rules in the previous two sections also apply to QBDs. However, in the case of a QBD, it is possible to develop an alternative truncation rule for evaluating R. This truncation rule can be stated as follows.

Truncation rule 3.4.5 *Take* $\boldsymbol{R} = \boldsymbol{R}(L)$ *where* L *is the smallest value of* ℓ *such that* $\|\boldsymbol{a}\boldsymbol{R}(\ell) - \alpha\boldsymbol{a}\|_\infty < \epsilon$ *for some tolerance* ϵ *where* α *is the spectral radius of* $\boldsymbol{R}$ *and* $\boldsymbol{a}$ *is the corresponding left eigenvector.*

In order to apply Truncation rule 3.4.5 we need to calculate both α and $\boldsymbol{a}$. We discuss methods for doing this later in the chapter.

Note that Truncation rule 3.4.5 can in theory be extended to the case of a LDQBD provided the matrix $\boldsymbol{R}_k$ is square. However, this would not provide a practical means of evaluating $\boldsymbol{R}_k$ since it seems to be difficult to compute the eigenvalue and corresponding left eigenvector of $\boldsymbol{R}_k$ efficiently.

It is interesting to note the similarity between Truncation rules 3.4.5 and 3.4.2. In Truncation rule 3.4.2 $\boldsymbol{G}_k$ is stochastic and hence its spectral radius is equal to one and the corresponding right eigenvector is equal to $\boldsymbol{e}$. So while in Truncation rule 3.4.5 we calculate the spectral radius of $\boldsymbol{R}$ and the corresponding left eigenvector, in Truncation rule 3.4.2 we calculate the known spectral radius of $\boldsymbol{G}$ and the corresponding right eigenvector.

3.4.4 Algorithms for computing G_k

For both discrete and continuous time processes we approximate $\boldsymbol{G}_k$ by

$$\boldsymbol{G}_k(L) = \sum_{\ell=0}^{L} \left[\prod_{i=0}^{\ell-1} \boldsymbol{U}^i_{k-1+2^i} \right] \boldsymbol{D}^\ell_{k-1+2^\ell} \tag{3.4.5}$$

where $\boldsymbol{U}^\ell_k$ and $\boldsymbol{D}^\ell_k$ are defined by equations (3.2.4)-(3.2.7) and (3.2.17)-(3.2.20) in discrete and continuous time respectively. Since the expressions for $\boldsymbol{U}^\ell_k$ and $\boldsymbol{D}^\ell_k$ are the same in discrete and continuous time except for the definitions of $\boldsymbol{U}^0_k$ and $\boldsymbol{D}^0_k$ it is sufficient to consider only discrete time processes. When considering continuous time processes we simply need to calculate $\boldsymbol{U}^0_k$ and $\boldsymbol{D}^0_k$ according to equations (3.2.17)-(3.2.18) rather than (3.2.4)-(3.2.5). In the rest of this section we will assume that $\boldsymbol{U}^\ell_k$ and $\boldsymbol{D}^\ell_k$ are defined by equations (3.2.4)-(3.2.7).

An elementary algorithm for computing $\boldsymbol{G}_k$ is as follows.

Algorithm 3.4.1 (Computing G_k)

```
    $\ell = 0$
```

```
$\prod = I$

$G_k(0) = D_k^0$

do
    $\ell = \ell + 1$

    $U = U_{k-1+2^{\ell-1}}^{\ell-1}$

    $D = D_{k-1+2^\ell}^{\ell}$

    $\prod = \prod \times U$

    $G_k(\ell) = G_k(\ell - 1) + \prod \times D$

until ( $G_k(\ell)$ obeys the truncation rule )

$L = \ell$

$G_k = G_k(L)$
```

It is clear from equations (3.2.4)-(3.2.7) that the steps that involve the most work in Algorithm 3.4.1 are those that calculate U and D. Therefore, in order to develop an efficient algorithm we must develop an efficient way of calculating both these terms. It can be seen from equations (3.2.4)-(3.2.7) that to calculate $U_{k-1+2^\ell}^{\ell}$ and $D_{k-1+2^\ell}^{\ell}$ (for fixed ℓ) we must calculate the inverse $[I - U_{k-1+2^\ell}^{\ell-1} D_{k-1+3.2^{\ell-1}}^{\ell-1} - D_{k-1+2^\ell}^{\ell-1} U_{k-1+2^{\ell-1}}]^{-1}$. In Algorithm 3.4.1 this inverse is calculated in successive iterations of the ***do*** loop. At one iteration it is computed in order to calculate $D_{k-1+2^\ell}^{\ell}$ and in the following iteration it is computed again in order to calculate $U_{k-1+2^\ell}^{\ell}$. Algorithm 3.4.1 can be altered as follows to avoid calculating this inverse twice.

Algorithm 3.4.2 (Computing G_k)

```
$\ell = 0$
$U = U^0_k,\ D = D^0_k$
$\prod = I$
$G_k(0) = D$
do
    $\ell = \ell + 1$
    $\prod = \prod \times U$
    $U = U^{\ell}_{k-1+2^{\ell}},\ D = D^{\ell}_{k-1+2^{\ell}}$
    $G_k(\ell) = G_k(\ell - 1) + \prod \times D$
until ( $G_k(\ell)$ obeys the truncation rule )
$L = \ell$
$G_k = G_k(L)$
```

Note that in the new algorithm we must update the value of $\prod$ before we calculate the new values of U and D. In Algorithm 3.4.2 we calculate $U^L_{k-1+2^L}$ but do not use it. However, this only requires additional matrix multiplications because we already have the necessary inverse from calculating $D^L_{k-1+2^L}$

Since the matrices $U = U^{\ell}_{k-1+2^{\ell}}$ and $D = D^{\ell}_{k-1+2^{\ell}}$ are calculated together, we refer to this pair of matrices as the UD–pair $UD(\ell, k - 1 + 2^{\ell})$. To implement the above algorithm efficiently we must develop an efficient method of computing the necessary UD–pairs.

At the end of Algorithm 3.4.2 we have calculated the UD–pairs $UD(\ell, k - 1 + 2^{\ell})$ for

$\ell = 0, 1, \ldots, L$ and fixed k. From equations (3.2.4)-(3.2.7) it can be seen that to compute $UD(\ell, k-1+2^\ell)$ we need to calculate other UD-pairs $UD(n,m)$ for $n < \ell$. We consider a way to compute these quantities so that no UD-pair is computed more than once. This can be done by storing various UD-pairs. To determine which UD-pairs to calculate and which to store, it is useful to represent the recursive relationships expressed by equations (3.2.6) and (3.2.7) by a diagram. Figure 3.4.1 shows that to calculate $UD(\ell, k-1+2^\ell)$, we need to know the UD-pairs $UD(\ell-1, k-1+2^{\ell-1})$, $UD(\ell-1, k-1+2.2^{\ell-1})$, and $UD(\ell-1, k-1+3.2^{\ell-1})$. By using Figure 3.4.1 repeatedly, we see that to calculate $UD(2, k+3)$ we need first to calculate $UD(1, k+m)$ for $m = 1, 3, 5$ and $UD(0, k+m)$ for $m = 0, 1, \ldots, 6$. Similarly, to calculate $UD(3, k+7)$ we need first to calculate $UD(2, k+m)$ for $m = 3, 7, 11$, $UD(1, k+m)$ for $m = 1, 3, 5, \ldots, 13$ and $UD(0, k+m)$ for $m = 0, 1, 2, \ldots, 14$. These last two statements are represented by Figure 3.4.2. To calculate

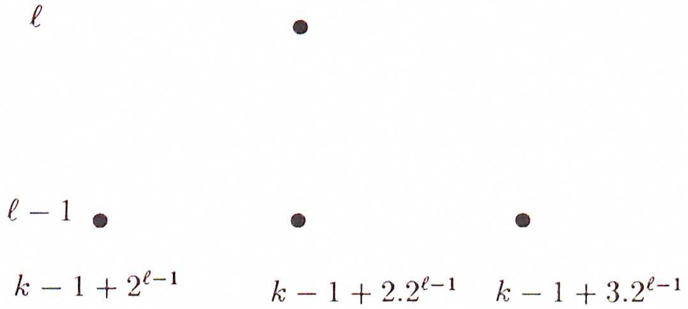

Figure 3.4.1: Recursive relationship between UD-pairs.

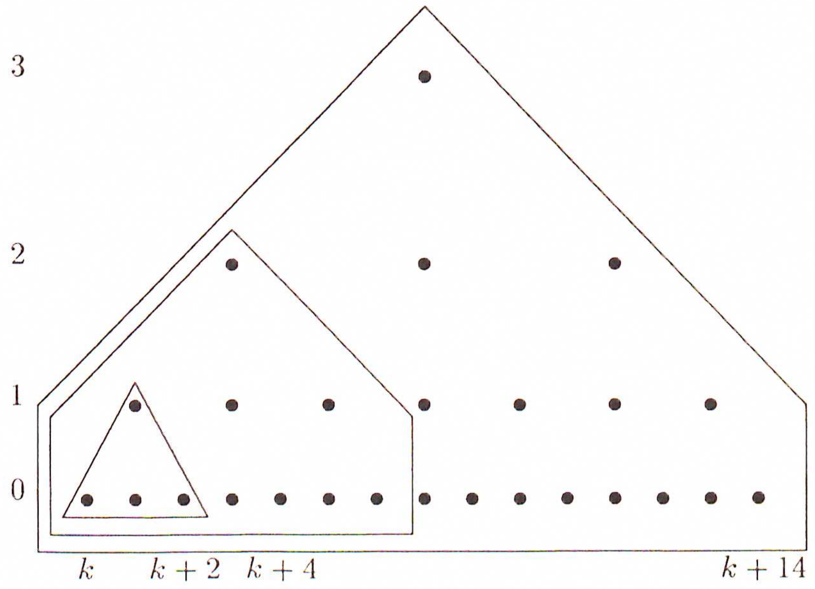

Figure 3.4.2: Dependencies between UD-pairs

$UD(1, k+1)$ we need to calculate all the UD–pairs in the triangle in Figure 3.4.2. To calculate $UD(2, k+3)$ we need to calculate all the UD–pairs in the smaller pentagon in Figure 3.4.2 and to calculate $UD(3, k+7)$, we need to calculate all the UD–pairs in the larger pentagon. Figure 3.4.2 can be extended to any arbitrary value of ℓ. From Figure 3.4.2 we see that if we calculate $UD(\ell, k-1+2^\ell)$ for $\ell = 0, 1, 2, 3$ by directly implementing equations (3.2.6)-(3.2.7) then we will be computing some UD–pairs over and over again. For example we would calculate the pairs $UD(0, k)$, $UD(0, k+1)$ and $UD(0, k+2)$ once when computing $UD(1, k+1)$, again when computing $UD(2, k+3)$ and a third time when computing $UD(3, k+7)$. However, it is possible to calculate UD–pairs in an order that will result in not computing any UD–pair more than once. The UD–pairs needed for Algorithm 3.4.2 can be calculated (and used) in the following order.

Algorithm 3.4.3 (Evaluating UD–pairs for Algorithm 3.4.2)

```
$\ell = 0$

compute $UD(0, k)$, use it in Algorithm 3.4.2 and store it

do
  for $i = 0$ to $\ell$
    for $j = k - 1 + 2^{\ell+1}$ to $k - 1 + 2^{\ell+2} - 2^i$, in steps of size $2^i$
      compute $UD(i, j)$ and store it

  $\ell = \ell + 1$

  compute $UD(\ell, k - 1 + 2^\ell)$, use it in Algorithm 3.4.2 and store it

  remove all $UD$–pairs from storage except the pairs
  $UD(j, k - 1 + 2^{\ell+1} - 2^j)$, $j = 0, 1, \ldots, \ell$

until (Algorithm 3.4.2 stops)
```

3.4.5 Algorithms for computing R_k

One way of computing the matrix R_k is first to compute the matrix G_{k+2} and then use either equation (2.7.4) for discrete time processes or equation (2.7.13) for continuous time processes. However, as mentioned earlier, R_k can also be computed by using the expression (3.3.6) or (3.3.18). In this section we present an algorithm for doing this. The form of the algorithm is very similar to that given in the previous section for computing G_k. This is to be expected since the expressions for R_k have a similar form to the expressions for G_k.

As was the case for G_k, when computing R_k the same approach can be used for both discrete and continuous time processes. The only difference is that $\widetilde{U}^{\ell}_{k}$ and $\widetilde{D}^{\ell}_{k}$ are defined by equations (3.3.7)-(3.3.10) in discrete time and by equations (3.3.19)-(3.3.22) in continuous time.

By considering equation (3.3.6) we can develop the following algorithm for computing R_k. It is similar to Algorithm 3.4.2 for computing G_k.

Algorithm 3.4.4 (Computing R_k)

```
$\ell = 0$

$\widetilde{U} = \widetilde{U}^{0}_{k}$, $\widetilde{D} = \widetilde{D}^{0}_{k+2}$

$\prod = I$

$R_k(0) = \widetilde{U}$

do
    $\ell = \ell + 1$

    $\prod = \widetilde{D} \times \prod$

    $\widetilde{U} = \widetilde{U}^{\ell}_{k}$, $\widetilde{D} = \widetilde{D}^{\ell}_{k+2^{\ell+1}}$
```

```
        $R_k(\ell) = R_k(\ell - 1) + \widetilde{U} \times \Pi$

    until ( $R_k(\ell)$ obeys the truncation rule )

    $L = \ell$

    $R_k = R_k(L)$
```

In Algorithm 3.4.4 we calculate $\widetilde{U}^{\ell}_{k}$ and $\widetilde{D}^{\ell}_{k+2^{\ell+1}}$ at the same step. This is because both quantities involve the same inverse and so it is sensible to compute them together to avoid calculating the same inverse twice. This can be seen by considering equations (3.3.7)-(3.3.10). We refer to $\widetilde{U}^{\ell}_{k}$ and $\widetilde{D}^{\ell}_{k+2^{\ell+1}}$ as the $\widetilde{UD}$–pair $\widetilde{UD}(\ell, k)$. At the end of Algorithm 3.4.4 we have calculated the $\widetilde{UD}$–pairs $\widetilde{UD}(\ell, k)$ for $\ell = 0, 1, \ldots, L$ and fixed k. To do this we need to calculate various other $\widetilde{UD}$–pairs $\widetilde{UD}(n, m)$ for $n < \ell$. As we did for the UD–pairs in the case of G_k, we develop a way to compute these quantities so that $\widetilde{UD}$–pairs are not computed more than once. The key to doing this is to store various $\widetilde{UD}$–pairs. To determine which $\widetilde{UD}$–pairs to store note that $\widetilde{U}^{\ell}_{k}$ and $\widetilde{U}^{\ell}_{k+2^{\ell+1}}$ are given by

$$\begin{aligned}
\widetilde{U}^{\ell}_{k} &= \widetilde{U}^{\ell-1}_{k} \widetilde{U}^{\ell-1}_{k+2^{\ell-1}} \left[I - \widetilde{U}^{\ell-1}_{k+2.2^{\ell-1}} \widetilde{D}^{\ell-1}_{k+3.2^{\ell-1}} - \widetilde{D}^{\ell-1}_{k+2.2^{\ell-1}} \widetilde{U}^{\ell-1}_{k+2^{\ell-1}} \right]^{-1}, \qquad (3.4.6) \\
\widetilde{D}^{\ell}_{k+2^{\ell+1}} &= \widetilde{D}^{\ell-1}_{k+2^{\ell+1}} \widetilde{D}^{\ell-1}_{k+3.2^{\ell-1}} \left[I - \widetilde{U}^{\ell-1}_{k+2.2^{\ell-1}} \widetilde{D}^{\ell-1}_{k+3.2^{\ell-1}} - \widetilde{D}^{\ell-1}_{k+2.2^{\ell-1}} \widetilde{U}^{\ell-1}_{k+2^{\ell-1}} \right]^{-1}. \qquad (3.4.7)
\end{aligned}$$

Equations (3.4.6) and (3.4.7) are represented by Figure 3.4.3. From Figure 3.4.3 we see that to calculate $\widetilde{UD}(\ell, k)$, we need to know the $\widetilde{UD}$–pairs $\widetilde{UD}(\ell - 1, k)$, $\widetilde{UD}(\ell - 1, k + 2^{\ell-1})$, and $\widetilde{UD}(\ell - 1, k + 2.2^{\ell-1})$.

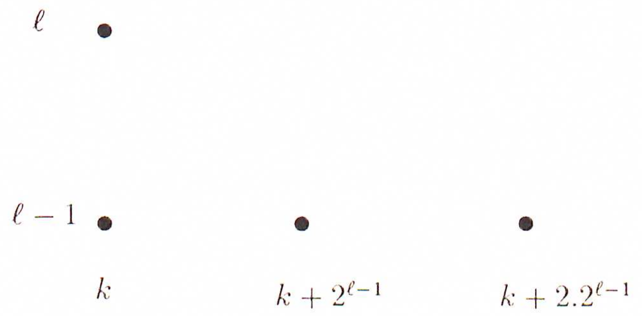

Figure 3.4.3: Recursive relationship between $\widetilde{UD}$–pairs.

By appealing to Figure 3.4.3 it can be seen that to calculate $\widetilde{UD}(2, k)$ we need first to calculate $\widetilde{UD}(1,k+m)$ for $m = 0,2,4$ and $\widetilde{UD}(0,k+m)$ for $m = 0,1,\dots,6$. Similarly, to calculate $\widetilde{UD}(3,k)$ we need first to calculate $\widetilde{UD}(2,k+m)$ for $m = 0,4,8$, $\widetilde{UD}(1,k+m)$ for $m = 0,2,4,\dots,12$ and $\widetilde{UD}(0,k+m)$ for $m = 0,1,2,\dots,14$. The last two statements are represented by Figure 3.4.4. To calculate $\widetilde{UD}(1,k)$ we need to calculate all the $\widetilde{UD}$–pairs

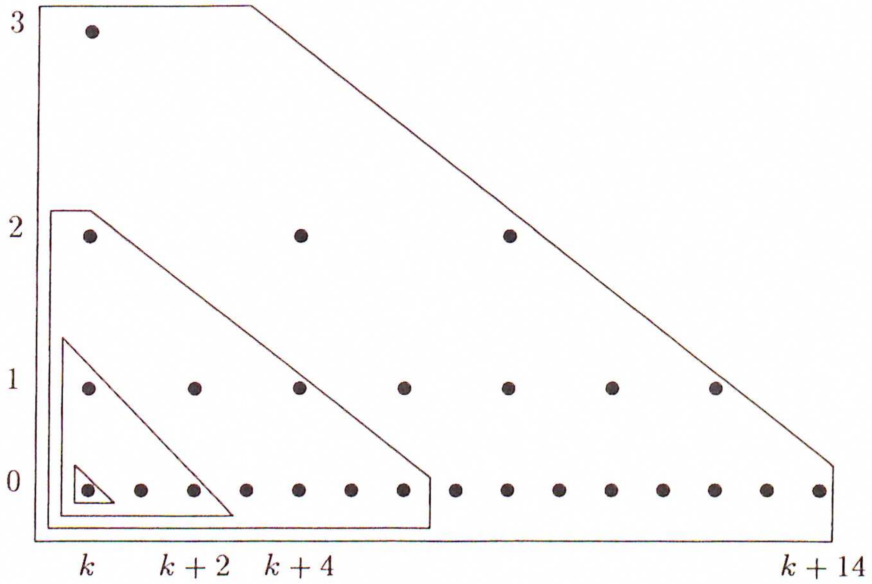

Figure 3.4.4: Dependencies between $\widetilde{UD}$–pairs

in the second smallest triangle in Figure 3.4.4. To calculate $\widetilde{UD}(2,k)$ we need to calculate all the $\widetilde{UD}$–pairs in the smaller pentagon and to calculate $\widetilde{UD}(3,k)$ we need to calculate all the $\widetilde{UD}$–pairs in the larger pentagon. It is clear how Figure 3.4.4 can be extended to any arbitrary value of ℓ. Figure 3.4.4 shows that if we calculate $\widetilde{UD}(\ell,k)$ for $\ell = 0,1,2,3$ by directly implementing equations (3.3.7)-(3.3.10) then we will be computing some $\widetilde{UD}$–pairs over and over again. However, this can be avoided by calculating the $\widetilde{UD}$–pairs in a particular order. The following algorithm outlines the order in which the $\widetilde{UD}$–pairs should be calculated (and used) in Algorithm 3.4.4.

Algorithm 3.4.5 (Evaluating $\widetilde{UD}$–pairs for Algorithm 3.4.4)

```
$\ell = 0$

compute $\widetilde{UD}(0,k)$, use it in Algorithm 3.4.4 and store it

do
```

```
    for $i = 0$ to $\ell$
        for $j = k + (2^{\ell-i+1} - 1)2^i$ to $k + 2(2^{\ell-i+1} - 1)2^i$ in steps of size $2^i$
            compute $\widetilde{UD}(i, j)$ and store it

    $\ell = \ell + 1$

    compute $\widetilde{UD}(\ell, k)$, use it in Algorithm 3.4.4 and store it

    remove all $\widetilde{UD}$-pairs from storage except the pairs
    $\widetilde{UD}(j, k + (2^{\ell-j} - 1)2^{j+1})$, $j = 0, 1, \ldots, \ell$

until (Algorithm 3.4.4 stops)
```

3.5 Expressions for the matrices G and R for QBDs

In the case of a QBD, the expressions for G_k and R_k derived in Sections 3.2 and 3.3 can be simplified by exploiting the homogeneity of the process. When this is done, the expressions simplify to those expressions given in Latouche and Ramaswami [56].

Consider a discrete time QBD. Due to the homogeneity of QBDs we have that $G_k = G_1 \; \forall k \geq 1$. From Theorem 3.2.2 we have

$$G_1 = \sum_{\ell=0}^{\infty} \left[\prod_{i=0}^{\ell-1} U^i_{2^i} \right] D^\ell_{2^\ell} \tag{3.5.1}$$

where $U^\ell_{2^\ell}$ and $D^\ell_{2^\ell}$ are given by

$$\begin{aligned}
U^0_1 &= (I - A_1)^{-1} A_0, && (3.5.2)\\
D^0_1 &= (I - A_1)^{-1} A_2, && (3.5.3)\\
U^{\ell+1}_{2^{\ell+1}} &= \left[I - U^\ell_{2^{\ell+1}} D^\ell_{3.2^\ell} - D^\ell_{2^{\ell+1}} U^\ell_{2^\ell} \right]^{-1} U^\ell_{2^{\ell+1}} U^\ell_{3.2^\ell}, && (3.5.4)\\
D^{\ell+1}_{2^{\ell+1}} &= \left[I - U^\ell_{2^{\ell+1}} D^\ell_{3.2^\ell} - D^\ell_{2^{\ell+1}} U^\ell_{2^\ell} \right]^{-1} D^\ell_{2^{\ell+1}} D^\ell_{2^\ell}. && (3.5.5)
\end{aligned}$$

Now it is clear that due to the homogeneity of QBDs, U^ℓ_k and D^ℓ_k will be independent of k. Hence we can set $U_k^\ell = \hat{U}^\ell$ $\forall k$ and then the above equations become

$$\begin{aligned}
\hat{U}^0 &= (I - A_1)^{-1} A_0, && (3.5.6)\\
\hat{D}^0 &= (I - A_1)^{-1} A_2, && (3.5.7)\\
\hat{U}^{\ell+1} &= \left[I - \hat{U}^\ell \hat{D}^\ell - \hat{D}^\ell \hat{U}^\ell\right]^{-1} (\hat{U}^\ell)^2, && (3.5.8)\\
\hat{D}^{\ell+1} &= \left[I - \hat{U}^\ell \hat{D}^\ell - \hat{D}^\ell \hat{U}^\ell\right]^{-1} (\hat{D}^\ell)^2 && (3.5.9)
\end{aligned}$$

Note that equations (3.5.6)-(3.5.9) are equivalent to equations (16)-(17) in Latouche and Ramaswami [56]. We can now state the following theorem which appears as Theorem 4.1 in [56].

Theorem 3.5.1 *For a discrete time QBD the matrix G is given explicitly by*

$$G = \sum_{\ell=0}^{\infty} \left[\prod_{i=0}^{\ell-1} \hat{U}^i\right] \hat{D}^\ell \tag{3.5.10}$$

where $\hat{U}^\ell$ and $\hat{D}^\ell$ are defined recursively for $\ell \geq 0$ by equations (3.5.6)-(3.5.9).

Latouche and Ramaswami gave a probabilistic proof of Theorem 3.5.1. We used a generalised version of their proof to prove Theorem 3.2.2.

An expression for G in continuous time can be obtained in a similar way to give the following theorem which appears as Theorem 7.1 in [56].

Theorem 3.5.2 *For a continuous time QBD the matrix G is given explicitly by*

$$G = \sum_{\ell=0}^{\infty} \left[\prod_{i=0}^{\ell-1} \hat{U}^i\right] \hat{D}^\ell \tag{3.5.11}$$

where

$$\begin{aligned}
\hat{U}^0 &= (-Q_1)^{-1} Q_0, && (3.5.12)\\
\hat{D}^0 &= (-Q_1)^{-1} Q_2, && (3.5.13)
\end{aligned}$$

and $\hat{U}^\ell$ and $\hat{D}^\ell$ are defined recursively for $\ell \geq 0$ by equations (3.5.8)-(3.5.9).

For completeness we present expressions for R in discrete and continuous time. Their derivations are obvious. The result for discrete time is presented in Theorem 6.1 of [56].

Theorem 3.5.3 *For a discrete time QBD the matrix $\boldsymbol{R}$ is given explicitly by*

$$\boldsymbol{R} = \sum_{\ell=0}^{\infty} \widetilde{\boldsymbol{U}}^{\ell} \left(\prod_{i=0}^{\ell-1} \widetilde{\boldsymbol{D}}^{\ell-1-i} \right) \tag{3.5.14}$$

where $\widetilde{\boldsymbol{U}}^{\ell}$ and $\widetilde{\boldsymbol{D}}^{\ell}$ are defined recursively for $\ell \geq 0$ by

$$\begin{aligned} \widetilde{\boldsymbol{U}}^{0} &= \boldsymbol{A}_0(\boldsymbol{I} - \boldsymbol{A}_1)^{-1}, && (3.5.15) \\ \widetilde{\boldsymbol{D}}^{0} &= \boldsymbol{A}_2(\boldsymbol{I} - \boldsymbol{A}_1)^{-1}, && (3.5.16) \\ \widetilde{\boldsymbol{U}}^{\ell+1} &= (\widetilde{\boldsymbol{U}}^{\ell})^2 \left[\boldsymbol{I} - \widetilde{\boldsymbol{U}}^{\ell}\widetilde{\boldsymbol{D}}^{\ell} - \widetilde{\boldsymbol{D}}^{\ell}\widetilde{\boldsymbol{U}}^{\ell}\right]^{-1}, && (3.5.17) \\ \widetilde{\boldsymbol{D}}^{\ell+1} &= (\widetilde{\boldsymbol{D}}^{\ell})^2 \left[\boldsymbol{I} - \widetilde{\boldsymbol{U}}^{\ell}\widetilde{\boldsymbol{D}}^{\ell} - \widetilde{\boldsymbol{D}}^{\ell}\widetilde{\boldsymbol{U}}^{\ell}\right]^{-1}. && (3.5.18) \end{aligned}$$

Theorem 3.5.4 *For a continuous time QBD the matrix $\boldsymbol{R}$ is given explicitly by*

$$\boldsymbol{R} = \sum_{\ell=0}^{\infty} \widetilde{\boldsymbol{U}}^{\ell} \left(\prod_{i=0}^{\ell-1} \widetilde{\boldsymbol{D}}^{\ell-1-i} \right) \tag{3.5.19}$$

where

$$\begin{aligned} \widetilde{\boldsymbol{U}}^{0} &= \boldsymbol{Q}_0(-\boldsymbol{Q}_1)^{-1}, && (3.5.20) \\ \widetilde{\boldsymbol{D}}^{0} &= \boldsymbol{Q}_2(-\boldsymbol{Q}_1)^{-1}, && (3.5.21) \end{aligned}$$

and $\widetilde{\boldsymbol{U}}^{\ell}$ and $\widetilde{\boldsymbol{D}}^{\ell}$ are defined recursively for $\ell \geq 0$ by equations (3.5.17)-(3.5.18).

3.6 Computing G and R for QBDs

It is clear that the algorithms presented so far for computing $\boldsymbol{G}_k$ and $\boldsymbol{R}_k$ for LDQBDs also apply to QBDs. However, the simpler analytic expressions for $\boldsymbol{G}$ and $\boldsymbol{R}$ in the QBD case mean that we can develop much more efficient algorithms. If the algorithms for LDQBDs are used for a QBD they treat the process as if it is a LDQBD. That is, they completely disregard the homogeneity of the process which leads to the unnecessary storage of many matrices. Hence the algorithms presented in the previous sections are not appropriate for QBDs. In this section we give a detailed presentation of algorithms developed by Latouche and Ramaswami [56] for computing the matrices $\boldsymbol{G}$ and $\boldsymbol{R}$ for QBDs. Clearly these algorithms should also be used when analysing LDQBDs that are level independent above a certain level. Bini and Meini [10] have also developed an algorithm that can be used to compute $\boldsymbol{G}$ and $\boldsymbol{R}$ for QBDs. The structure of their algorithm is similar to the algorithm of Latouche and Ramaswami. However, it has the additional feature that it can be used in the analysis of M/G/1-type processes. We give a brief description of their algorithm at the end of this section.

3.6.1 Computing G

The following algorithm for evaluating G for a QBD was presented in Figure 2 of [56].

Algorithm 3.6.1 (Computing G)

```
$\ell = 0$

$U = \hat{U}^0$, $D = \hat{D}^0$

$\prod = I$

$G(0) = D$

do
    $\ell = \ell + 1$

    $\prod = \prod \times U$

    $U = \hat{U}^\ell$, $D = \hat{D}^\ell$

    $G(\ell) = G(\ell - 1) + \prod \times D$

until ( $G(\ell)$ obeys the truncation rule )

$L = \ell$

$G = G(L)$
```

The above algorithm is more efficient than Algorithm 3.4.2 because it is much easier to compute the quantities U and D and it does not require the storage of many matrices as does Algorithm 3.4.2.

3.6.2 Computing R

The form of equations (3.5.14) and (3.5.19) suggest the following algorithm for computing the matrix R for a QBD.

Algorithm 3.6.2 (Computing R)

```
$\ell = 0$

$U = \widetilde{U}^0$, $D = \widetilde{D}^0$

$\Pi = I$

$R(0) = U$

do
    $\ell = \ell + 1$

    $\Pi = D \times \Pi$

    $U = \widetilde{U}^\ell$, $D = \widetilde{D}^\ell$

    $R(\ell) = R(\ell - 1) + U \times \Pi$

until ( $R(\ell)$ obeys the truncation rule )

$L = \ell$

$R = R(L)$
```

To apply Truncation rule 3.4.5 in Algorithm 3.6.2 we need to know both α and $\boldsymbol{a}$ which, as mentioned in Section 3.4.3 are the spectral radius of $\boldsymbol{R}$ and the corresponding left eigenvector. We now outline a simple method given in Neuts [72] for calculating these quantities in discrete time.

For $0 \le z \le 1$, let $\chi(z)$ be the maximal eigenvalue of

$$A^*(z) = A_0 + zA_1 + z^2A_2. \tag{3.6.1}$$

First consider finding the maximum eigenvalue, α of $\boldsymbol{R}$. Neuts [72, Lemma 1.3.2] showed that α is the smallest positive solution to the equation $z = \chi(z)$ for $0 < z \le 1$. If the QBD is not positive recurrent then $z = 1$ is the only solution to this equation. If the QBD is positive recurrent, this equation has a second solution as well as the solution $z = 1$. For positive recurrent QBDs it can also be shown that $\chi'(1-) > 1$ and hence it is easy to show that if z_0 is such that $\chi(z_0) > z_0$ then $\alpha > z_0$ and if z_0 is such that $\chi(z_0) < z_0$ then $\alpha < z_0$. Hence α can be found using a bisection search. Using arguments similar to those used in the proof of Lemma 1.3.2 of Neuts [72], it can be shown that $\boldsymbol{a}$ is the left eigenvector of $A^*(\alpha)$ corresponding to the eigenvalue α. Hence it is a simple matter to find $\boldsymbol{a}$ once α is known.

To find α and $\boldsymbol{a}$ for a continuous time QBD, we consider the uniformised chain. A continuous time Markov process with q-matrix $\boldsymbol{Q}$ is uniform if $q = \sup_i -q_{ii} < \infty$ and then its uniformised chain is defined to be the discrete time Markov chain with transition probability matrix equal to $\frac{1}{r}\boldsymbol{Q} + \boldsymbol{I}$ for $r \ge q$. Continuous time QBDs are uniform due to their homogeneity. The uniformised chain for a continuous time QBD has $A_0 = \frac{1}{r}Q_0$, $A_1 = \frac{1}{r}Q_1 + I$ and $A_2 = \frac{1}{r}Q_2$. Now α and $\boldsymbol{a}$ for a continuous time QBD can be found using similar methods to those used above for discrete time processes. If we define $Q^*(z)$ for $0 \le z \le 1$ by

$$Q^*(z) = Q_0 + zQ_1 + z^2Q_2,$$

then $Q^*(z)$ has positive off diagonal elements and negative diagonal elements. For such matrices there exists analogues of the Perron-Frobenius results for non negative matrices. For more details see Theorem 2.6 in Seneta [89]. Using Theorem 2.6 from [89] we can define $\bar{\chi}(z)$ to be the maximal eigenvalue (that is, the eigenvalue with maximum real part) of $Q^*(z)$.

Now $A^*(z)$ for the uniformised chain is given by

$$\begin{aligned} A^*(z) &= \frac{1}{r}Q_0 + z(\frac{1}{r}Q_1 + I) + z^2\frac{1}{r}Q_2 \\ &= \frac{1}{r}Q^*(z) + zI \end{aligned}$$

and hence we have the relationship

$$\chi(z) = \frac{1}{r}\bar{\chi}(z) + z. \tag{3.6.2}$$

It can be shown that the R matrix for the uniformised chain is equal to the R matrix for the continuous time QBD. Hence from the results for discrete time QBDs we have that α is the smallest positive solution to $z = \chi(z)$ for $0 < z \leq 1$. From equation (3.6.2) this is equivalent to finding the smallest positive solution to $\bar{\chi}(z) = 0$ for $0 < z \leq 1$. As before, α can be found by using a bisection search. If z_0 is such that $\bar{\chi}(z_0) > 0$ then $\alpha > z_0$ and if z_0 is such that $\bar{\chi}(z_0) < 0$ then $\alpha < z_0$. It can be shown that $\boldsymbol{a}$ is the left eigenvector of $Q^*(\alpha)$ corresponding to the eigenvalue α. Hence once we know α, we can easily compute $\boldsymbol{a}$.

3.6.3 LDQBDs that are level independent above a certain level

For a LDQBD that is level independent for level K and above, we have $G_k = G \; \forall k \geq K$ and $R_k = R \; \forall k \geq K$. To compute G and R we can use Algorithms 3.6.1 and 3.6.2 respectively. To compute G_k, $1 \leq k < K$ and R_k, $0 \leq k < K$ we may use Algorithm 3.4.2 and Algorithm 3.4.4 but these algorithms do not exploit the level independent structure for levels K and above. From equations (2.6.2) and (2.6.1) we have for a discrete time QBD that

$$\begin{aligned} R_{K-1} &= A_0^{(K-1)}(I - U_K)^{-1} && (3.6.3) \\ &= A_0^{(K-1)}(I - A_1^{(K)} - R_K A_2^{(K+1)})^{-1} && (3.6.4) \\ &= A_0^{(K-1)}(I - A_1 - RA_2)^{-1} && (3.6.5) \end{aligned}$$

which suggests that a way to determine R_{K-1} would be to first compute R (using Algorithm 3.6.2) and then to compute R_{K-1} using equation (3.6.5). Since equations (2.6.2) and (2.6.1) say that for any $0 \leq k \leq K$,

$$R_{k-1} = A_0^{(k-1)}(I - A_1^{(k)} - R_k A_2^{(k+1)})^{-1}, \tag{3.6.6}$$

a possible method of finding $R_k, 0 \leq k < K$ for a discrete time QBD is to compute R using Algorithm 3.6.2 and then use equation (3.6.6) to compute $R_{K-1}, R_{K-2}, \ldots, R_0$.

Equations (2.7.1) and (2.7.2) state that for a discrete time QBD, for any $1 \leq k \leq K$,

$$G_k = (I - A_1^{(k)} - A_0^{(k)} G_{k+1})^{-1} A_2^{(k)}. \tag{3.6.7}$$

Hence a possible method for finding $G_k, 1 \leq k < K$ for a discrete time QBD is to compute G using Algorithm 3.6.1 and then to use equation (3.6.7) to compute $G_{K-1}, G_{K-2}, \ldots, G_1$.

The above arguments can be summarised by the following algorithms.

Algorithm 3.6.3 (Computing $\{R_k, 0 \leq k \leq K\}$)

```
Compute $R_K$ using Algorithm 3.6.2

for $k = K - 1$ downto 0
    $R_k = A_0^{(k)}(I - A_1^{(k+1)} - R_{k+1} A_2^{(k+2)})^{-1}$
```

Algorithm 3.6.4 (Computing $\{G_k, 1 \leq k \leq K\}$)

```
Compute $G_K$ using Algorithm 3.6.1

for $k = K - 1$ downto 0
    $G_k = (I - A_1^{(k)} - A_0^{(k)} G_{k+1})^{-1} A_2^{(k)}$
```

Similar arguments to those used above can be applied to continuous time QBDs. We omit the details.

3.6.4 An alternative algorithm

Recently Bini and Meini [10] developed an algorithm based on FFT-techniques that can be used to find the matrix G for a QBD. Their algorithm also holds for the more general M/G/1-type process which we will consider in Chapter 6. In the special case of a QBD, the algorithm of Bini and Meini is similar in both its structure and efficiency to the algorithm of Latouche and Ramaswami. We do not give a presentation of Bini and Meini's algorithm since to appreciate its structure fully we should present it in the M/G/1–type context. We note that although in the QBD case its efficiency is similar to the algorithm of Latouche and Ramaswami, in the case of a M/G/1–type process it is the most efficient algorithm available.

3.7 Analysis of the algorithms

In this section we discuss certain aspects of the algorithms presented in this chapter. Algorithms 3.6.1 and 3.6.2 have been discussed in detail by Latouche and Ramaswami [56]. We give a discussion of the more general Algorithms 3.4.2 and 3.4.4. In particular, we provide physical interpretations of these algorithms for both discrete and continuous time LDQBDs. We obtain these interpretations by generalising the interpretations of Latouche and Ramaswami's algorithms.

3.7.1 Algorithms for computing G_k

When Algorithm 3.4.2 stops, we have calculated the UD–pairs $UD(\ell, k - 1 + 2^\ell)$ for $\ell = 0, 1, \ldots, L$ and fixed k. If these UD–pairs are computed according to Algorithm 3.4.3 then it can be shown that a total of $4.2^L - L - 3$ inverses must be calculated. If the UD–pairs were calculated by directly implementing the recursive equations (3.2.6) and (3.2.7) then a total of approximately $(3/8).5^L$ inverses must be calculated. Note that with Algorithm 3.4.3, although we have significantly reduced the number of inverses that we need to calculate, the total number of inverses still increases exponentially with L. However, for many practical examples only a small value of L is needed. This can be seen by appealing to the physical interpretations of the explicit expressions for G_k. We firstly consider the physical interpretation for discrete time LDQBDs and then we consider continuous time LDQBDs.

Discrete time LDQBDs

We showed in the proof of Theorem 3.2.2 that $\left[\prod_{i=0}^{\ell-1} U_{k-1+2^i}^i\right] D_{k-1+2^\ell}^\ell$ gives the probability that the process visits level $k-1$ after it visits level $k - 1 + 2^\ell$ and before it visits level $k - 1 + 2^{\ell+1}$, given the process starts in level k at time $t = 0$. That is, this collects the probability mass of all sample paths that go at least as high as level $k - 1 + 2^\ell$ but no higher than level $k-1+2^{\ell+1}$, before their first return to level $k-1$. Therefore, when we truncate the infinite sum (3.2.9) at $\ell = L$, we only consider sample paths that remain below level $k-1+2^{L+1}$. Provided L is large enough, these sample paths will contain the bulk of the probability mass and the remaining sample paths will have very low probability masses and we can therefore ignore them. Obviously it is possible to construct processes for which a very large value of L is needed. However, even for relatively small L (say five or six), we consider paths which move up 64 or 128 levels before returning. For many physical processes this accounts for most of the probability mass. Hence for most processes only a moderate value of L will be required and therefore the number of inverses that we need to compute will not be too large.

Continuous time LDQBDs

We saw in the proof of Theorem 3.2.3 that the matrix G_k for a continuous time LDQBD is equal to the matrix G_k for its jump chain. Hence we can say that $\left[\prod_{i=0}^{\ell-1} U^i_{k-1+2^i}\right] D^\ell_{k-1+2^\ell}$ has the same physical interpretation as it did in the discrete time case. All the comments for discrete time processes made above carry over to continuous time processes.

3.7.2 Algorithms for computing R_k

Algorithm 3.4.4 has a similar interpretation to Algorithm 3.4.2. When Algorithm 3.4.4 stops, we have calculated the $\widetilde{UD}$-pairs $\widetilde{UD}(\ell, k)$ for $\ell = 0, 1, \ldots, L$ and fixed k. If these $\widetilde{UD}$-pairs are computed according to Algorithm 3.4.5 then a total of $4.2^L - L - 3$ inverses must be calculated whereas if they are calculated by directly implementing the recursive equations (3.3.9) and (3.3.10) then a total of approximately $(3/8).5^L$ inverses must be calculated. Hence it is important that the $\widetilde{UD}$-pairs be computed in an efficient manner. We now give a physical interpretation of equations (3.3.6) and (3.3.18).

Discrete time LDQBDs

In order to give a probabilistic interpretation of equation (3.3.6) we give an alternative expression for R_k. To do this we define $\delta(k) = \inf\{n > 0 : X(n) \in \boldsymbol{k}\}$ and we let $\mathbb{I}(\cdot)$ be an indicator function. From equation (2.4.6) we have that

$$R_k = A_0^{(k)} \sum_{\tau=0}^{\infty} {}_kP(\mathbf{k}+\mathbf{1}, \mathbf{k}+\mathbf{1}; \tau) \tag{3.7.1}$$

$$= A_0^{(k)} \sum_{\tau=0}^{\infty} E\left[\mathbb{I}(X(\tau) \in \mathbf{k}+\mathbf{1}) \mid X(0) \in \mathbf{k}+\mathbf{1}, X(s) \notin \mathbf{k}\ 0 < s < \tau\right] \tag{3.7.2}$$

$$= A_0^{(k)} E\left[\sum_{\tau=0}^{\delta(k)} \mathbb{I}(X(\tau) \in \mathbf{k}+\mathbf{1}) \mid X(0) \in \mathbf{k}+\mathbf{1}\right]. \tag{3.7.3}$$

Using this notation, it can be shown that

$$\widetilde{U}_k^{\ell} = A_0^{(k)} E\left[\sum_{\tau=\delta(k)\wedge\delta(k+2^{\ell})}^{\delta(k)\wedge\delta(k+2^{\ell+1})} \mathbb{I}(X(\tau) \in \mathbf{k}+\mathbf{2}^{\ell}) \mid X(0) \in \mathbf{k}+\mathbf{1}\right], \tag{3.7.4}$$

$$\widetilde{D}_k^{\ell} = A_2^{(k)} E\left[\sum_{\tau=\delta(k)\wedge\delta(k-2^{\ell})}^{\delta(k)\wedge\delta(k-2^{\ell+1})} \mathbb{I}(X(\tau) \in \mathbf{k}-\mathbf{2}^{\ell}) \mid X(0) \in \mathbf{k}-\mathbf{1}\right], \tag{3.7.5}$$

and after some manipulation, it can be shown that the ℓth term of the sum in equation (3.3.6) has the interpretation

$$A_0^{(k)} E\left[\sum_{\tau=\delta(k)\wedge\delta(k+2^{\ell})}^{\delta(k)\wedge\delta(k+2^{\ell+1})} \mathbb{I}(X(\tau) \in \mathbf{k}+\mathbf{1}) \mid X(0) \in \mathbf{k}+\mathbf{1}\right]. \tag{3.7.6}$$

That is, the ℓth term only considers sample paths that have gone at least as high as level $k + 2^{\ell}$ but have not reached level $k + 2^{\ell+1}$. Note the similarity between these sample paths and those included in the ℓth term of the infinite sum expression for G_k. As we mentioned when interpreting the expressions for G_k, for most practical processes only a moderate value of L is required to include the sample paths that contain most of the probability mass.

Continuous time LDQBDs

Using similar arguments to those used for discrete time processes we can write R_k as

$$R_k = Q_0^{(k)} E\left[\int_{t=0}^{\delta(k)} \mathbb{I}(X(t) \in \mathbf{k}+\mathbf{1})dt \mid X(0) \in \mathbf{k}+\mathbf{1}\right]. \tag{3.7.7}$$

To obtain a physical interpretation of equation (3.3.18) we consider the jump chain. In the proof of Theorem 3.3.2 we showed that

$$\widetilde{\boldsymbol{U}}_k^\ell = \boldsymbol{\Phi}_k \bar{\boldsymbol{U}}_k^\ell \boldsymbol{\Phi}_{k+2^\ell}^{-1} \tag{3.7.8}$$

and

$$\widetilde{\boldsymbol{D}}_k^\ell = \boldsymbol{\Phi}_k \bar{\boldsymbol{D}}_k^\ell \boldsymbol{\Phi}_{k-2^\ell}^{-1} \tag{3.7.9}$$

where $\bar{\boldsymbol{U}}_k^\ell$ and $\bar{\boldsymbol{D}}_k^\ell$ relate to the jump chain and are given by equations (3.3.7)-(3.3.10) with $\boldsymbol{A}_0^{(k)}$, $\boldsymbol{A}_1^{(k)}$ and $\boldsymbol{A}_2^{(k)}$ given by equations (3.2.13)-(3.2.15).

If we denote the jump chain by $X^J(t)$ then from equations (3.7.4) and (3.7.8) and since $\Phi_k \boldsymbol{A}_0^{(k)} = \boldsymbol{Q}_0^{(k)}$, we have that

$$\widetilde{\boldsymbol{U}}_k^\ell = \boldsymbol{Q}_0^{(k)} E\left[\sum_{\tau=\delta(k)\wedge\delta(k+2^\ell)}^{\delta(k)\wedge\delta(k+2^{\ell+1})} \mathbb{I}(X^J(\tau) \in \boldsymbol{k}+\boldsymbol{2}^\ell) \mid X^J(0) \in \boldsymbol{k}+\boldsymbol{1}\right] \Phi_{k+2^\ell}^{-1}. \tag{3.7.10}$$

Now the expectation in equation (3.7.10) calculates the expected number of visits to level $k+2^\ell$ by the jump chain, before the jump chain reaches level k or level $k+2^{\ell+1}$, given the jump chain starts in level $k+1$. Hence the product $E[\cdot]\Phi_{k+2^\ell}^{-1}$ in equation (3.7.10) gives the "expected number of visits by the jump chain to level $k+2^\ell$ (subject to the restrictions described above), multiplied by the expected time spent in level $k+2^\ell$ per visit." Clearly this is equivalent to observing the continuous time process and calculating the expected sojourn time spent in level $k+2^\ell$. Therefore we can write

$$\widetilde{\boldsymbol{U}}_k^\ell = \boldsymbol{Q}_0^{(k)} E\left[\int_{t=\delta(k)\wedge\delta(k+2^\ell)}^{\delta(k)\wedge\delta(k+2^{\ell+1})} \mathbb{I}(X(t) \in \boldsymbol{k}+\boldsymbol{2}^\ell)dt \mid X(0) \in \boldsymbol{k}+\boldsymbol{1}\right]. \tag{3.7.11}$$

Similarly we can show that

$$\widetilde{\boldsymbol{D}}_k^\ell = \boldsymbol{Q}_2^{(k)} E\left[\int_{t=\delta(k)\wedge\delta(k-2^\ell)}^{\delta(k)\wedge\delta(k-2^{\ell+1})} \mathbb{I}(X(t) \in \boldsymbol{k}-\boldsymbol{2}^\ell)dt \mid X(0) \in \boldsymbol{k}-\boldsymbol{1}\right]. \tag{3.7.12}$$

Equation (3.3.17) states that $\boldsymbol{R}_k = \boldsymbol{\Phi}_k \boldsymbol{R}_k^J \boldsymbol{\Phi}_{k+1}^{-1}$ where $\boldsymbol{R}_k^J$ is the $\boldsymbol{R}_k$ matrix for the jump chain. Hence by using equation (3.7.6), we can show that the ℓth term of (3.3.18) has the interpretation

$$\boldsymbol{Q}_0^{(k)} E\left[\int_{t=\delta(k)\wedge\delta(k+2^\ell)}^{\delta(k)\wedge\delta(k+2^{\ell+1})} \mathbb{I}(X(t) \in \boldsymbol{k}+\boldsymbol{1})dt \mid X(0) \in \boldsymbol{k}+\boldsymbol{1}\right]. \tag{3.7.13}$$

Hence the ℓth term of (3.3.18) also considers only sample paths that have gone at least as high as level $k+2^\ell$ but have not reached level $k+2^{\ell+1}$.

3.7.3 Storage

Storage is also an issue when using Algorithms 3.4.2 and 3.4.4. The various UD–pairs and $\widetilde{UD}$–pairs needed for these algorithms are evaluated using Algorithms 3.4.3 and 3.4.5 respectively. Both these algorithms require the storage of a large number of matrices. We will show that Algorithm 3.4.3 can be modified so as to reduce the storage space required. The discussion we will give can be easily adapted to Algorithm 3.4.5.

Consider the ***do*** loop of Algorithm 3.4.3 for $\ell = 1$. From the previous iteration of the loop we will have the UD–pairs $UD(0, k+2)$ and $UD(1, k+1)$ in storage. In the loop we proceed in the following order.

1. Compute and store $UD(0, k+3)$, $UD(0, k+4)$, $UD(0, k+5)$. $UD(0, k+6)$,
2. Compute and store $UD(1, k+3)$, $UD(1, k+5)$.
3. Compute and store $UD(2, k+3)$.
4. Remove all UD–pairs except $UD(0, k+6)$, $UD(1, k+5)$, $UD(2, k+3)$.

Note that after step 1 we have $UD(0, k+\ell)$ $\ell = 2, \dots, 6$ and $UD(1, k+1)$ in storage. After step 2 we have added $UD(1, k+3)$ and $UD(1, k+5)$ to the storage space. It is not necessary to have all these UD–pairs in storage at the same time. To see this, note that steps 1 to 4 can be performed as follows.

1. Compute and store $UD(0, k+3)$, $UD(0, k+4)$.
2. Use $UD(0, k+\ell)$ $\ell = 2, 3, 4$ to compute $UD(1, k+3)$.
3. Remove $UD(0, k+2)$, $UD(0, k+3)$ from storage and keep $UD(0, k+4)$.
4. Compute and store $UD(0, k+5)$, $UD(0, k+6)$.
5. Use $UD(0, k+\ell)$ $\ell = 4, 5, 6$ to compute $UD(1, k+5)$.
6. Remove $UD(0, k+4)$, $UD(0, k+5)$ from storage and keep $UD(0, k+6)$.
7. Use $UD(1, k+\ell)$ $\ell = 1, 3, 5$ to compute $UD(2, k+3)$.
8. Remove $UD(1, k+1)$, $UD(1, k+3)$ from storage and keep $UD(0, k+6)$, $UD(1, k+5)$, $UD(2, k+3)$.

By following the second sequence of eight steps we will always have less UD–pairs in storage than we would if we were to follow the first sequence of four steps. By extending the above argument to larger values of ℓ, the storage space required by Algorithm 3.4.3 can be reduced significantly.

Obviously even by implementing the algorithms so as to minimise the amount of storage required, it is still possible for some processes to require an excessive amount of storage. Hence the algorithms we have presented may not be applicable to some processes. However, they are amenable to a large range of problems. In Chapter 4 we will discuss how we can use the algorithms as part of a procedure for calculating the stationary distribution of LDQBDs, and in Chapter 5 we apply this procedure to a variety of examples.

Chapter 4

Calculating the stationary distribution for LDQBDs

4.1 Introduction

In this chapter we present algorithms for computing the stationary distribution for a LDQBD. The majority of what we present here has appeared in two previous works by the author (see Bright and Taylor, [12] and [13]). However, we give a more detailed presentation here.

In Chapter 5 we will apply the algorithms presented in this chapter to a variety of problems. Since the examples we consider there are continuous time processes, we present the results in this chapter for continuous time processes. All the results we present have analogues in discrete time and we state these, without proofs, at the end of the chapter.

QBDs have been used to model systems such as queues in a random environment, high speed communication systems, database systems and multiprogramming systems. Nelson [66] computed the stationary distribution for a parallel processing system and used it to show that the number of jobs concurrently in execution grows linearly with the load of the system. Latouche [54] modelled a multiprogramming-multiprocessor computer system as a QBD and provided an algorithmic technique for computing the stationary probabilities. Zukerman [98] modelled a hybrid switching system as a QBD and used the stationary distribution to compute the average delay in the system. Neuts [69, 67], and Zukerman and Kirton [99] also modelled systems as QBDs and computed the stationary probabilities. Since LDQBDs are a generalisation of level independent processes, they enable us to model a wider class of problems. For example, the $PH/M/1$ queue (i.e the single server queue with a phase-type arrival process) can be analysed using a QBD whereas the $PH/M/\infty$ queue can be analysed using a LDQBD. Similarly, LDQBDs enable us to model infinite server queues in a random environment rather than just single server queues. Various types of retrial queues can be modelled as LDQBDs as can a number of bivariate queueing models. LDQBDs also enable us to model more complex communication systems and multiprogramming systems. In Chapter 5 we will discuss in more detail the variety of problems that can be modelled by LDQBDs.

The stationary distributions of QBDs have been dealt with widely in the literature. Evans [22] and Wallace [94] were the first to show that the stationary distribution for a QBD has the so called matrix-geometric form $\boldsymbol{x}_k = \boldsymbol{x}_{k-1}\boldsymbol{R}$. They also proposed various methods for computing numerical values of the stationary distribution. Neuts [68] showed that the stationary distribution for GI/M/1-type processes also has the matrix-geometric form and he gave a probabilistic interpretation of the matrix $\boldsymbol{R}$, [71]. Following Neuts' work, many contributions have been made to the literature. Recently Latouche and Ramaswami [56] developed a logarithmic reduction algorithm that can be used to calculate the stationary distribution. The computational complexity of this algorithm is significantly less than that for previously existing algorithms. An algorithm developed by Bini and Meini [10] which is also applicable to the more general M/G/1-type process can also be used to compute the stationary distribution for a QBD. The computational complexity and structure of this algorithm is similar to the algorithm of Latouche and Ramaswami.

The special case where the LDQBD has a finite state space has also received significant attention in the literature. This occurs when for some K, $\boldsymbol{Q}_0^{(K)} = 0$ and $\boldsymbol{Q}_0^{(k)} = \boldsymbol{Q}_1^{(k)} = \boldsymbol{Q}_2^{(k)} = \boldsymbol{0}\ \forall k \geq K+1$. Hajek [34] initially considered QBDs on a finite state space and recently Naoumov [65] developed an algorithm that appears to be the most efficient currently available. Gaver, Jacobs and Latouche [29] developed a general method for LDQBDs on a finite state space and Ye and Li [97] have developed a computationally efficient algorithm for the case where the generator is piecewise level independent. In the level dependent case on an infinite state space, as far as we know, the problem of calculating the stationary distribution has not been addressed. The algorithms we present for LDQBDs are based on the algorithms presented in Chapter 3 for computing the matrices $\boldsymbol{R}_k$. As mentioned in Chapter 3, the algorithms presented there are extensions of Latouche and Ramaswami's logarithmic reduction algorithm.

We begin the chapter by giving an explicit expression for the stationary distribution of a continuous time LDQBD. Following this we propose a general method for computing the stationary distribution for continuous time processes and we then consider a number of special cases where a more sophisticated approach may be adopted. At the end of the chapter we give a statement of the corresponding results for discrete time processes.

4.2 The stationary distribution of a continuous time LDQBD

In Chapter 2 we showed that the stationary distribution $\boldsymbol{x} = (\boldsymbol{x}_0, \boldsymbol{x}_1, \dots)$ for a positive recurrent continuous time LDQBD has the form $\boldsymbol{x}_k = \boldsymbol{x}_{k-1}\boldsymbol{R}_{k-1}, k \geq 1$. For the rest of the chapter we will assume that $X(t)$ is positive recurrent. In Chapter 3 we discussed how to calculate the matrices $\boldsymbol{R}_k$ but we have not yet discussed how to find the vector $\boldsymbol{x}_0$. The following theorem states how to find $\boldsymbol{x}_0$ for a positive recurrent continuous time LDQBD and gives a closed form expression for the stationary distribution $\boldsymbol{x}_k, k \geq 0$.

Theorem 4.2.1 *If $X(t)$ is a continuous time positive recurrent LDQBD then its stationary distribution $\boldsymbol{x}$ is given by*

$$\boldsymbol{x}_k = \boldsymbol{x}_0 \prod_{\ell=0}^{k-1} \boldsymbol{R}_\ell \qquad k \geq 0 \tag{4.2.1}$$

where the matrices $\boldsymbol{R}_k$ are defined in Chapter 2 and $\boldsymbol{x}_0$ is a positive solution to

$$\boldsymbol{x}_0(\boldsymbol{Q}_1^{(0)} + \boldsymbol{R}_0\boldsymbol{Q}_2^{(1)}) = \boldsymbol{0} \tag{4.2.2}$$

such that

$$\boldsymbol{x}_0 \sum_{k=0}^{\infty} \left[\prod_{\ell=0}^{k-1} \boldsymbol{R}_\ell \right] \boldsymbol{e} = 1. \tag{4.2.3}$$

Proof: Equation (4.2.1) follows directly from equation (2.4.9). For a continuous time LDQBD we know from the global balance equations that $\boldsymbol{x}_0$ and $\boldsymbol{x}_1$ satisfy

$$\boldsymbol{x}_0\boldsymbol{Q}_1^{(0)} + \boldsymbol{x}_1\boldsymbol{Q}_2^{(1)} = \boldsymbol{0}$$

which upon substituting $\boldsymbol{x}_1 = \boldsymbol{x}_0\boldsymbol{R}_0$ becomes

$$\boldsymbol{x}_0(\boldsymbol{Q}_1^{(0)} + \boldsymbol{R}_0\boldsymbol{Q}_2^{(1)}) = \boldsymbol{0}. \tag{4.2.4}$$

Theorem 2.8.1 guarantees that equation (4.2.4) has a positive solution such that (4.2.3) holds. Equation (4.2.3) guarantees that $\sum_{k=0}^{\infty}\boldsymbol{x}_k\boldsymbol{e} = 1$ so $\boldsymbol{x}_k$ must be the stationary distribution. □

We can only calculate the infinite sum in (4.2.3) if we know $\{\boldsymbol{R}_k, k \geq 0\}$. This is rarely the case since the $\boldsymbol{R}_k$ matrices can usually be found only numerically. Hence we must calculate a truncated sum. If we define $[\boldsymbol{x}_k(K^*)]_j$, $0 \leq k \leq K^*$ to be the stationary probability that $X(t)$ is in the state (k, j) conditional on $X(t)$ being in the set $\{(k,j)|0 \leq k \leq K^*, 1 \leq j \leq M(k)\}$ then it is clear that $\boldsymbol{x}_k(K^*)$, $0 \leq k \leq K^*$ is given by

$$\boldsymbol{x}_k(K^*) = \boldsymbol{x}_0(K^*)\prod_{\ell=0}^{k-1}\boldsymbol{R}_\ell \qquad 0 \leq k \leq K^*, \tag{4.2.5}$$

where $\boldsymbol{x}_0(K^*)$ satisfies (4.2.2) and

$$\boldsymbol{x}_0(K^*)\sum_{k=0}^{K^*}\left[\prod_{\ell=0}^{k-1}\boldsymbol{R}_\ell\right]\boldsymbol{e} = 1. \tag{4.2.6}$$

Note that $\boldsymbol{x}_k(K^*)$ is not the invariant measure of a truncated process, rather it is the invariant measure of $X(t)$ for level k, normalised over the states in and below level K^*. We discuss this in more detail in Section 4.3.

It is obvious that $\forall\, K^* \geq 0$, $\boldsymbol{x}_k(K^*)$ is an upper bound for $\boldsymbol{x}_k$. So we have

$$\boldsymbol{x}_k \leq \boldsymbol{x}_k(K^*) \qquad \forall\, K^* \geq 0 \tag{4.2.7}$$

and

$$\boldsymbol{x}_k = \lim_{K^*\to\infty}\boldsymbol{x}_k(K^*) \qquad \forall k \geq 0 \tag{4.2.8}$$

where the inequality and the limit hold elementwise. Provided we take K^* large enough, we hope to have $\boldsymbol{x}_k(K^*) \approx \boldsymbol{x}_k$. The question of finding a sufficiently large value of K^* will be considered in Section 4.4.

4.3 Calculating $\boldsymbol{x}_k(K^*)$ for a given value of K^*

Assume K^* is given. The following theorem gives a starting point for considering this situation.

Theorem 4.3.1 *The vector* $\boldsymbol{x}(K^*) = (\boldsymbol{x}_0(K^*), \boldsymbol{x}_1(K^*), \ldots, \boldsymbol{x}_{K^*}(K^*))$ *where* $\boldsymbol{x}_k(K^*)$ *is given by (4.2.5) is the stationary distribution for the irreducible Markov process on the state space* $\{(k,j) : 0 \leq k \leq K^*, 1 \leq j \leq M(k)\}$ *with q-matrix* $\boldsymbol{Q}(K^*)$ *given by*

$$\boldsymbol{Q}(K^*) = \begin{pmatrix} \boldsymbol{Q}_1^{(0)} & \boldsymbol{Q}_0^{(0)} & & & \\ \boldsymbol{Q}_2^{(1)} & \boldsymbol{Q}_1^{(1)} & \boldsymbol{Q}_0^{(1)} & & \\ & \ddots & \ddots & \ddots & \\ & & \boldsymbol{Q}_2^{(K^*-1)} & \boldsymbol{Q}_1^{(K^*-1)} & \boldsymbol{Q}_0^{(K^*-1)} \\ & & & \boldsymbol{Q}_2^{(K^*)} & \boldsymbol{U}_{K^*} \end{pmatrix} \qquad (4.3.1)$$

where $\boldsymbol{U}_{K^*}$ *is given by equation (2.6.9).*

Hence $\boldsymbol{x}(K^*)$ *is the unique solution to the system of equations*

$$\boldsymbol{x}(K^*)\boldsymbol{Q}(K^*) = \boldsymbol{0} \qquad (4.3.2)$$

subject to

$$\sum_{k=0}^{K^*} \boldsymbol{x}_k(K^*)\boldsymbol{e} = 1. \qquad (4.3.3)$$

Proof: The process with q-matrix $\boldsymbol{Q}(K^*)$ has an interpretation as the original LDQBD on the infinite state space observed only at time points where it is in levels less than or equal to K^*. To see this recall from Chapter 2 that, $\boldsymbol{U}_{K^*}$ is the q-matrix for the original LDQBD observed only at time points where it is in level K^*, before it visits level $K^* - 1$ for the first time.

From Markov chain theory we have that the unique solution to equation (4.3.2) subject to (4.3.3) is the stationary distribution for the Markov process with q-matrix $\boldsymbol{Q}(K^*)$. The proof is completed by using equation (2.5.4) to verify that $\boldsymbol{x}_k(K^*)$ given by (4.2.5) satisfies equation (4.3.2). □

It is clear that $\boldsymbol{Q}(K^*)$ has negative diagonal elements and non negative off diagonal elements. Furthermore, $\boldsymbol{Q}(K^*)$ is an irreducible matrix. To show that $\boldsymbol{Q}(K^*)$ is a q-matrix, we need to show that $\boldsymbol{Q}(K^*)\boldsymbol{e} = \boldsymbol{0}$. This is clearly the case provided $(\boldsymbol{Q}_2^{(K^*)} +$

$U_{K^*})e = \mathbf{0}$. Now provided the LDQBD is positive recurrent we have that $G_{K^*+1}e = e$ so from equation (2.7.8) we have

$$\begin{aligned}
&(Q_2^{(K^*)} + U_{K^*})e && (4.3.4)\\
&= (Q_2^{(K^*)} + Q_1^{(K^*)} + Q_0^{(K^*)}G_{K^*+1})e && (4.3.5)\\
&= \mathbf{0}. && (4.3.6)
\end{aligned}$$

From equation (2.6.9) we see that to calculate U_{K^*} we need to know R_{K^*}. Hence if we calculate R_{K^*} using Algorithm 3.4.4 then from Theorem 4.3.1 we can calculate $x(K^*)$ by solving (4.3.2) subject to (4.3.3). There are a number of numerical methods available to do this. The Gauss-Seidel scheme is an often used iterative approach.

In Bright and Taylor [12], [13] the authors computed $x(K^*)$ by computing $\{R_k, 0 \le k \le K^* - 1\}$ and then using equation (4.2.5). This method is equivalent to the method proposed by Gaver, Jacobs and Latouche in [29]. The algorithm presented there is essentially an implementation of block Gaussian elimination (or block LU decomposition). Gaver, Jacobs and Latouche also gave a probabilistic interpretation of their algorithm.

In the case where $Q(K^*)$ has the form

$$Q(K^*) = \begin{pmatrix}
Q_1^{(0)} & Q_0^{(0)} & & & & & \\
Q_2(1) & Q_1(1) & Q_0(1) & & & & \\
\ddots & \ddots & \ddots & & & & \\
 & Q_2(1) & Q_1(1) & Q_0(1) & & & \\
 & & \ddots & \ddots & \ddots & & \\
 & & & Q_2(N) & Q_1(N) & Q_0(N) & \\
 & & & \ddots & \ddots & \ddots & \\
 & & & & Q_2(N) & Q_1(N) & Q_0(N)\\
 & & & & & Q_2^{(K^*)} & Q_1^{(K^*)}
\end{pmatrix} \qquad (4.3.7)$$

the q-matrix $Q(K^*)$ is piecewise level independent. That is, the levels $1, 2, \ldots, K^* - 1$ can be divided into N groups of levels such that the behaviour of the process is level independent within each of these groups. In this case the folding algorithm of Ye and Li [97] exploits the piecewise level independence of $\boldsymbol{Q}(K^*)$. Their algorithm in the piecewise level independent case is more efficient than that of Gaver, Jacobs and Latouche [29]. In the case where the transition rates $\boldsymbol{Q}_i^{(k)}, i = 0, 1, 2$ change at every level k, the algorithms of Ye and Li and Gaver, Jacobs and Latouche are equivalent.

There are a number of additional methods for solving $\boldsymbol{x}\boldsymbol{Q}(K^*) = \boldsymbol{0}$ for the case where $\boldsymbol{Q}(K^*)$ has the form

$$\boldsymbol{Q}(K^*) = \begin{pmatrix} \boldsymbol{Q}_1^{(0)} & \boldsymbol{Q}_0^{(0)} & & & \\ \boldsymbol{Q}_2^{(1)} & \boldsymbol{Q}_1 & \boldsymbol{Q}_0 & & \\ & \ddots & \ddots & \ddots & \\ & & \boldsymbol{Q}_2 & \boldsymbol{Q}_1 & \boldsymbol{Q}_0^{(K^*-1)} \\ & & & \boldsymbol{Q}_2^{(K^*)} & \boldsymbol{Q}_1^{(K^*)} \end{pmatrix}. \qquad (4.3.8)$$

Here the matrix $\boldsymbol{Q}(K^*)$ essentially represents a QBD on a finite state space with some level dependence at the boundaries. Hajek [34] initially considered q-matrices of the form (4.3.8) and showed that $\boldsymbol{x}(K^*)$ can be written as the linear combination of two matrix-geometric terms. Recently Naoumov [65] refined Hajek's expression for $\boldsymbol{x}(K^*)$. To compute the matrix-geometric terms in Naoumov's expression for $\boldsymbol{x}(K^*)$ requires the solution of two matrix quadratic equations. A possible way of doing this is to use the algorithms of Latouche and Ramaswami [56] or Bini and Meini [10]. However, Naoumov has developed an alternative algorithm for doing this. Although his algorithm is similar in structure to both the algorithm of Latouche and Ramaswami and that of Bini and Meini, his algorithm is slightly more efficient.

In the rest of this chapter we will compute $\boldsymbol{x}_k(K^*)$ as we did in [12] and [13]. That is, we will compute $\{\boldsymbol{R}_k, 0 \leq k \leq K^* - 1\}$ and then use equation (4.2.5). Note however that we do not claim that this is the most efficient method. Indeed, although this method is numerically stable for a wide range of parameter values, in some cases it is subject to underflow and overflow errors. In such cases alternative methods may be more appropriate.

In the case where K^* is sufficiently large, $\boldsymbol{x}(K^*)$ is an approximation to the stationary distribution for the LDQBD with q-matrix (2.2.2). This is because

$$\sum_{k=0}^{K^*} \left[\prod_{\ell=0}^{k-1} \boldsymbol{R}_\ell \right] \boldsymbol{e} \approx \sum_{k=0}^{\infty} \left[\prod_{\ell=0}^{k-1} \boldsymbol{R}_\ell \right] \boldsymbol{e} \tag{4.3.9}$$

and so $\boldsymbol{x}_0(K^*)$ calculated via (4.2.5) is approximately the same as $\boldsymbol{x}_0$ calculated via (4.2.1).

An alternative method for computing the stationary distribution for a LDQBD is to truncate the LDQBD to a finite state space. Neuts and Rao [74] used this method to compute the stationary distribution for a multiserver retrial queue. They also proposed another more efficient method for approximating the stationary distribution for the particular class of model they considered. We now describe how to apply the truncation method and the differences between it and our approach. Until now the truncation method has been the only available method of analysis.

Given a LDQBD with q-matrix (2.2.2), an approximation $\hat{\boldsymbol{x}}$ to the true stationary distribution $\boldsymbol{x}$ can be obtained by solving the equation

$$\hat{\boldsymbol{x}} \hat{\boldsymbol{Q}} = \boldsymbol{0} \tag{4.3.10}$$

subject to $\hat{\boldsymbol{x}}\boldsymbol{e} = 1$ where $\hat{\boldsymbol{Q}}$ has the form

$$\hat{\boldsymbol{Q}} = \begin{pmatrix} \boldsymbol{Q}_1^{(0)} & \boldsymbol{Q}_0^{(0)} & & & \\ \boldsymbol{Q}_2^{(1)} & \boldsymbol{Q}_1^{(1)} & \boldsymbol{Q}_0^{(1)} & & \\ & \ddots & \ddots & \ddots & \\ & & \boldsymbol{Q}_2^{(\hat{K}-1)} & \boldsymbol{Q}_1^{(\hat{K}-1)} & \boldsymbol{Q}_0^{(\hat{K}-1)} \\ & & & \boldsymbol{Q}_2^{(\hat{K})} & \hat{\boldsymbol{Q}}_1^{(\hat{K})} \end{pmatrix} \tag{4.3.11}$$

with $\hat{\boldsymbol{Q}}_1^{(\hat{K})}$ given by

$$\hat{\boldsymbol{Q}}_1^{(\hat{K})} = \boldsymbol{Q}_1^{(\hat{K})} + \boldsymbol{Q}_0^{(\hat{K})} \tag{4.3.12}$$

or

$$\hat{Q}_1^{(\hat{K})} = Q_1^{(\hat{K})} + \mathrm{diag}(Q_0^{(\hat{K})} e). \tag{4.3.13}$$

The value of $\hat{K}$ in the above must be chosen so that the amount of probability mass that is lost due to truncation is minimal. In the first truncation method where $\hat{Q}_1^{(\hat{K})}$ is given by equation (4.3.12), the transition rate from phase i in level $\hat{K}$ to phase j in level $\hat{K}+1$ in the LDQBD on the infinite state space, is redirected to phase j in level $\hat{K}$. In the second truncation method where $\hat{Q}_1^{(\hat{K})}$ is given by equation (4.3.13), the transition rate from phase i in level $\hat{K}$ to phase j in level $\hat{K}+1$ in the LDQBD on the infinite state space, is set to zero.

Both the truncation methods are significantly different to the process with q-matrix $Q(K^*)$ given by (4.3.1). The process truncated according to (4.3.12) assumes that as soon as the original LDQBD reaches level K^*+1, it immediately returns to level K^*, without changing phases. The process truncated according to (4.3.13) assumes that the original LDQBD never reaches the set $\{(k,j) : k > K^*, 1 \leq j \leq M(k)\}$. In contrast, the process with q-matrix $Q(K^*)$ given by (4.3.1) is a process on a finite state space, but it includes all relevant information about the original LDQBD on the infinite state space.

To calculate $x_k(K^*), 0 \leq k \leq K^*$ using equation (4.2.5) we need to know $R_k, 0 \leq k \leq K^*-1$. We saw in Chapter 3 that we can compute these matrices by using Algorithm 3.4.4. However, to calculate $\{R_k, 0 \leq k \leq K^*-1\}$, it is not necessary to use Algorithm 3.4.4 repeatedly. From equations (2.6.10) and (2.6.9) we have

$$\begin{aligned} R_k &= Q_0^{(k)}(-U_{k+1})^{-1} && (4.3.14)\\ &= Q_0^{(k)}(-Q_1^{(k+1)} - R_{k+1}Q_2^{(k+2)})^{-1}. && (4.3.15)\end{aligned}$$

So a possible method for calculating $\{R_k, 0 \leq k \leq K^*-1\}$ is to use Algorithm 3.4.4 to calculate R_{K^*-1} and then use equation (4.3.15) to calculate $R_{K^*-2}, R_{K^*-3}, \ldots, R_0$. In order for the recursion (4.3.15) to be stable we require that the inverses contain non negative elements. The following theorem states that this is indeed the case.

Theorem 4.3.2 *For continuous time LDQBDs the matrix $(-U_k)^{-1}$ has non negative elements.*

Proof: In the proof of Lemma 2.6.2 we showed that $[(-U_k)^{-1}]_{ij}$ is the expected total sojourn time in state (k,j) before the process reaches level $k-1$ for the first time, given the process starts in the state (k, i). Hence $(-\boldsymbol{U}_k)^{-1}$ has non negative elements. □

An algorithm for calculating $\{\boldsymbol{x}_k(K^*),\ 0 \le k \le K^*\}$ for a given value of K^* is as follows.

Algorithm 4.3.1 (Calculating $\boldsymbol{x}_k(K^*), 0 \le k \le K^*$ given K^*)

```
Calculate $\boldsymbol{R}_{K^*-1}$ using Algorithm 3.4.4

Recursively calculate $\boldsymbol{R}_{K^*-2}, \boldsymbol{R}_{K^*-3}, \ldots, \boldsymbol{R}_0$

Solve $\boldsymbol{x}_0(K^*)(\boldsymbol{Q}_1^{(0)} + \boldsymbol{R}_0\boldsymbol{Q}_2^{(1)}) = 0$
subject to $\boldsymbol{x}_0(K^*)\boldsymbol{e} = 1$

for $k = 1$ to $K^*$ do
    $\boldsymbol{x}_k(K^*) = \boldsymbol{x}_{k-1}(K^*)\boldsymbol{R}_{k-1}$

    Normalise $\boldsymbol{x}_0(K^*), \boldsymbol{x}_1(K^*), \ldots, \boldsymbol{x}_k(K^*)$
    subject to $\sum_{m=0}^{k} \boldsymbol{x}_m(K^*)\boldsymbol{e} = 1$
```

4.4 Choosing a value for K^*

When calculating $\{\boldsymbol{x}_k(K^*),\ 0 \le k \le K^*\}$ we want to find a level K^* such that the stationary probability of being in a state above level K^* is approximately zero. In some special cases it is possible to calculate K^* *a priori*. In this section we firstly discuss a general method for calculating K^* and then we look at some special cases.

4.4.1 A general method

If $X(t)$ is positive recurrent then $\boldsymbol{x}_k(K^*), k \ge 0$ is an upper bound for the stationary distribution. By choosing K^* very large, $\boldsymbol{x}_k(K^*)$ will hopefully give a close approximation to the stationary distribution. However, as K^* increases, so does the computational effort required to calculate $\boldsymbol{x}_k(K^*)$. To see this, recall from Theorem 4.3.1 that $\boldsymbol{x}_k(K^*)$ is the stationary distribution for the Markov process on the finite state space $\{(k, j) : 0 \le k \le K^*, 1 \le j \le M(k)\}$ with q-matrix given by equation (4.3.1). Hence, as K^* increases, the finite state space increases in size and hence does the computational effort required to compute $\boldsymbol{x}_k(K^*)$. Therefore it is desirable to determine a value of K^* that is sufficiently large to obtain a close approximation but not too large so as to require an excessive amount of computation.

A simple approach to finding a value of K^* consists of trying various values of K^* until we obtain one that is satisfactory. It is certainly possible to do better for some LDQBDs and we will consider such cases in later sections. However, for most processes of physical interest, the following algorithm will halt at a suitable value of K^*.

Algorithm 4.4.1 (A general method for calculating $\boldsymbol{x}_k(K^*), 0 \leq k \leq K^*$)

```
Set $K^*_{old}$ equal to an initial guess for $K^*$

Calculate $\boldsymbol{x}_k(K^*_{old}), 0 \leq k \leq K^*_{old}$ using Algorithm 4.3.1

do
    Choose $K^*_{new}$ such that $K^*_{new} > K^*_{old}$

    for $k = 0$ to $K^*_{old}$ do
        $\boldsymbol{x}_k(K^*_{new}) = \boldsymbol{x}_k(K^*_{old})$

    for $k = K^*_{old} + 1$ to $K^*_{new}$ do
        $\boldsymbol{x}_k(K^*_{new}) = \boldsymbol{x}_{k-1}(K^*_{new})\boldsymbol{R}_{k-1}$

        Normalise $\boldsymbol{x}_0(K^*_{new}), \boldsymbol{x}_1(K^*_{new}), ..., \boldsymbol{x}_k(K^*_{new})$
        subject to $\sum_{m=0}^{k} \boldsymbol{x}_m(K^*)\boldsymbol{e} = 1$

until $\left( \max_{k:0 \leq k \leq K^*_{old}} (|| \boldsymbol{x}_k(K^*_{new}) - \boldsymbol{x}_k(K^*_{old}) ||_{max}) < \epsilon \right)$
```

The stopping criterion in Algorithm 4.4.1 states that if the changes in the values of the vectors $\boldsymbol{x}_k(K^*), 0 \leq k \leq K^*_{old}$ are sufficiently small, then the current value of K^* is sufficient. We hope that by choosing K^* in this fashion that we will find a value of K^*

such that $\boldsymbol{x}_k \approx \mathbf{0}\ \forall k > K^*$. It is certainly possible to construct LDQBDs for which this is not the case. However, for most LDQBDs that arise in practice, the method will work.

We note that the matrices $\boldsymbol{Q}_0^{(k)}$, $\boldsymbol{Q}_1^{(k)}$ and $\boldsymbol{Q}_2^{(k)}$ must be expressible in some parametric form. Indeed this is necessary if a process is to be analysed using a computer. It would be very difficult, if not impossible, even to input into a computer the parameters of a process involving a large number of unrelated $\boldsymbol{Q}_i^{(k)}$ matrices. By considering the parametric form of the matrices $\boldsymbol{Q}_i^{(k)}$, it should be possible to tell if the value of K^* found by Algorithm 4.4.1 is accurate. For the above reasons, careful consideration must be given to the structure of the LDQBD when using this method to find K^*.

4.4.2 LDQBDs on a finite state space

If we put $\boldsymbol{Q}_0^{(K)} = \mathbf{0}$ and $\boldsymbol{Q}_2^{(k)} = \boldsymbol{Q}_1^{(k)} = \boldsymbol{Q}_0^{(k)} = \mathbf{0}\ \forall\, k \geq K+1$ then $X(t)$ is irreducible on the state space $\{(k,j) : 0 \leq k \leq K, 1 \leq j \leq M(k)\}$ so we have a LDQBD on a finite state space. Therefore if $\boldsymbol{x} = (\boldsymbol{x}_0, ..., \boldsymbol{x}_K)$ is the stationary distribution of $X(t)$ then $\boldsymbol{x}_k = \boldsymbol{x}_k(K),\ 0 \leq k \leq K$. That is, we have $K^* = K$.

From Algorithm 4.3.1, our first step in finding $\boldsymbol{x}_k, 0 \leq k \leq K$ is to find $\boldsymbol{R}_{K-1}$. From the physical interpretation of $\boldsymbol{R}_k$, it is clear that $\boldsymbol{R}_K = 0$ and hence from equation (4.3.15) we have that $\boldsymbol{R}_{K-1} = \boldsymbol{Q}_0^{(K-1)}(-\boldsymbol{Q}_1^{(K)})^{-1}$. An algorithm for computing $\boldsymbol{x}_k(K^*)$ is as follows.

Algorithm 4.4.2 (Finite state space)

```
$\boldsymbol{R}_{K-1} = \boldsymbol{Q}_0^{(K-1)}(-\boldsymbol{Q}_1^{(K)})^{-1}$

Recursively calculate $\boldsymbol{R}_{K-2}, \boldsymbol{R}_{K-3}, ..., \boldsymbol{R}_0$

Solve $\boldsymbol{x}_0(\boldsymbol{Q}_1^{(0)} + \boldsymbol{R}_0\boldsymbol{Q}_2^{(1)}) = 0$
subject to $\boldsymbol{x}_0\boldsymbol{e} = 1$.

for $k = 1$ to $K$ do
    $\boldsymbol{x}_k = \boldsymbol{x}_{k-1}\boldsymbol{R}_{k-1}$
```

Normalise $\boldsymbol{x}_0, \boldsymbol{x}_1, ..., \boldsymbol{x}_k$ *such that* $\sum_{m=0}^{k} \boldsymbol{x}_m \boldsymbol{e} = 1$

In the case of a finite state space, our algorithm is equivalent to the continuous time algorithm presented by Gaver, Jacobs and Latouche [29]. Gaver, Jacobs and Latouche considered the matrices $\bar{\boldsymbol{C}}_k$, $0 \leq k \leq K$ defined by $\bar{\boldsymbol{C}}_K = \boldsymbol{Q}_1^{(K)}$ and $\bar{\boldsymbol{C}}_k = \boldsymbol{Q}_1^{(k)} + \boldsymbol{Q}_0^{(k)}(-\bar{\boldsymbol{C}}_{k+1})^{-1}\boldsymbol{Q}_2^{(k+1)}$, $K-1 \geq k \geq 0$. They show that the stationary distribution is given by

$$\boldsymbol{x}_k = \boldsymbol{x}_0 \prod_{\ell=0}^{k-1} \boldsymbol{Q}_0^{(\ell)}(-\bar{\boldsymbol{C}}_{\ell+1})^{-1} \qquad 0 \leq k \leq K \tag{4.4.1}$$

where $\boldsymbol{x}_0$ is the solution to

$$\boldsymbol{x}_0 \bar{\boldsymbol{C}}_0 = \boldsymbol{0} \tag{4.4.2}$$

such that

$$\boldsymbol{x}_0 \sum_{k=0}^{K} \left[\prod_{\ell=0}^{k-1} \boldsymbol{Q}_0^{(\ell)}(-\bar{\boldsymbol{C}}_{\ell+1})^{-1} \right] \boldsymbol{e} = 1. \tag{4.4.3}$$

It can be shown that the matrix $\bar{\boldsymbol{C}}_k$ is equal to the matrix $\boldsymbol{U}_k, 0 \leq k \leq K$. To see this recall that

$$\boldsymbol{U}_k = \boldsymbol{Q}_1^{(k)} + \boldsymbol{R}_k \boldsymbol{Q}_2^{(k+1)}. \tag{4.4.4}$$

Since $\boldsymbol{R}_K = \boldsymbol{0}$, we have that

$$\begin{aligned} \boldsymbol{U}_K &= \boldsymbol{Q}_1^{(K)} && (4.4.5) \\ &= \bar{\boldsymbol{C}}_K. && (4.4.6) \end{aligned}$$

Now assume that $\bar{\boldsymbol{C}}_k = \boldsymbol{U}_k$ holds for arbitrary $0 < k \leq K$ and consider $\bar{\boldsymbol{C}}_{k-1}$.

$$\begin{aligned} \bar{\boldsymbol{C}}_{k-1} &= \boldsymbol{Q}_1^{(k-1)} + \boldsymbol{Q}_0^{(k-1)}(-\bar{\boldsymbol{C}}_k)^{-1}\boldsymbol{Q}_2^{(k)} && (4.4.7) \\ &= \boldsymbol{Q}_1^{(k-1)} + \boldsymbol{Q}_0^{(k-1)}(-\boldsymbol{U}_k)^{-1}\boldsymbol{Q}_2^{(k)} && (4.4.8) \\ &= \boldsymbol{Q}_1^{(k-1)} + \boldsymbol{R}_{k-1}\boldsymbol{Q}_2^{(k)} && (4.4.9) \\ &= \boldsymbol{U}_{k-1} && (4.4.10) \end{aligned}$$

where the second last equality follows from equation (2.6.10).

Given this expression for $\bar{\boldsymbol{C}}_k$, we have that

$$\begin{aligned} &\boldsymbol{Q}_0^{(k)}(-\bar{\boldsymbol{C}}_{k+1})^{-1} && (4.4.11)\\ &= \boldsymbol{Q}_0^{(k)}(-\boldsymbol{U}_{k+1})^{-1} && (4.4.12)\\ &= \boldsymbol{R}_k && (4.4.13)\end{aligned}$$

from equation (4.3.15). So equations (4.4.1) and (4.4.2) reduce to equations (4.2.5) and (4.2.2) and hence the two approaches are equivalent.

Gaver, Jacobs and Latouche showed that $\boldsymbol{Q}_0^{(k)}(-\bar{\boldsymbol{C}}_{k+1})^{-1}$ is the counterpart to Neut's rate matrix $\boldsymbol{R}$ for an infinite level independent QBD. Hence their solution is in some way related to Neut's matrix-geometric solutions. A consequence of our approach is that it shows directly that the case of a finite state space LDQBD fits directly into the matrix-analytic structure of infinite LDQBDs.

4.4.3 LDQBDs that are level independent above a certain level

Neuts [72] considered LDQBDs that are level independent above a certain level. The approach there also uses $\boldsymbol{x}_k = \boldsymbol{x}_K(\boldsymbol{R}_K)^{k-K}$ to find $\boldsymbol{x}_k$, $k > K$ which is what we obtain from equation (4.2.1). To find $\boldsymbol{x}_k$, $0 \le k \le K$ for a continuous time process, Neuts solved

$$\begin{aligned} \boldsymbol{x}_0\boldsymbol{Q}_1^{(0)} + \boldsymbol{x}_1\boldsymbol{Q}_2^{(1)} &= 0, && && (4.4.14)\\ \boldsymbol{x}_i\boldsymbol{Q}_0^{(i)} + \boldsymbol{x}_{i+1}\boldsymbol{Q}_1^{(i+1)} + \boldsymbol{x}_{i+2}\boldsymbol{Q}_2^{(i+2)} &= 0 && 0 \le i \le K-2, && (4.4.15)\\ \boldsymbol{x}_{K-1}\boldsymbol{Q}_0^{(K-1)} + \boldsymbol{x}_K\boldsymbol{Q}_1^{(K)} + \boldsymbol{x}_K\boldsymbol{R}_K\boldsymbol{Q}_2^{(K)} &= 0, && && (4.4.16)\\ \sum_{i=0}^{K-1} \boldsymbol{x}_i\boldsymbol{e} + \boldsymbol{x}_K(\boldsymbol{I} - \boldsymbol{R}_K)^{-1}\boldsymbol{e} &= 1. && && (4.4.17)\end{aligned}$$

That is, Neuts solved the global balance equations for the process. This does not highlight the matrix-analytic structure of $\boldsymbol{x}_k$ for $k \le K$. A nice property of our method is that it shows that $\boldsymbol{x}_k$ has a matrix-analytic form for all $0 \le k \le K$.

In this particular case, it is possible to determine if $X(t)$ is positive recurrent and to determine the stationary distribution exactly. As mentioned in Chapter 2, Section 2.8, $X(t)$ is positive recurrent if and only if $\boldsymbol{\pi}\boldsymbol{Q}_0\boldsymbol{e} < \boldsymbol{\pi}\boldsymbol{Q}_2\boldsymbol{e}$ where $\hat{\boldsymbol{Q}} = \boldsymbol{Q}_0 + \boldsymbol{Q}_1 + \boldsymbol{Q}_2$ and $\boldsymbol{\pi}$

is the solution to $\boldsymbol{\pi}\hat{\boldsymbol{Q}} = \mathbf{0}$ such that $\boldsymbol{\pi e} = 1$. If $X(t)$ is positive recurrent then

$$\boldsymbol{x}_0 \sum_{k=0}^{\infty} \left[\prod_{\ell=0}^{k-1} \boldsymbol{R}_\ell \right] \boldsymbol{e} \tag{4.4.18}$$

$$= \boldsymbol{x}_0 \left[\sum_{k=0}^{K} \prod_{\ell=0}^{k-1} \boldsymbol{R}_\ell + \prod_{\ell=0}^{K} \boldsymbol{R}_m (I - \boldsymbol{R}_K)^{-1} \right] \boldsymbol{e} \tag{4.4.19}$$

$$< \infty \tag{4.4.20}$$

and the stationary distribution can be calculated exactly using equation (4.2.1). This is one of the few cases where this is possible. A similar argument can be applied in discrete time. Since $X(t)$ is a level independent QBD for levels K and above, we should calculate $\boldsymbol{R}_K$ via Algorithm 3.6.2. An algorithm for this case is given below.

Algorithm 4.4.3

```
Calculate $\boldsymbol{R}_K$ via Algorithm 3.6.2

Recursively calculate $\boldsymbol{R}_{K-1}, \boldsymbol{R}_{K-2}, ..., \boldsymbol{R}_0$

Solve $\boldsymbol{x}_0(\boldsymbol{Q}_1^{(0)} + \boldsymbol{R}_0\boldsymbol{Q}_2^{(1)}) = 0$
subject to $\boldsymbol{x}_0 \left[\sum_{k=0}^{K} \prod_{m=0}^{k-1} \boldsymbol{R}_m + \prod_{m=0}^{K} \boldsymbol{R}_m(\boldsymbol{I} - \boldsymbol{R}_K)^{-1}\right] \boldsymbol{e} = 1$

Calculate $\boldsymbol{x}_k, k \geq 0$ according to equation (4.2.1)
```

4.4.4 Choosing K^* by constructing a dominating process

In this section we consider LDQBDs for which the following condition holds.

Condition 1 : $\forall\, k, i\ \exists\, j$ such that $[\boldsymbol{Q}_2^{(k)}]_{ij} > 0$.

If Condition 1 holds, then at any phase at any level, $X(t)$ moves down to a phase in the level below according to some positive rate. Note that this class of processes is a non-trivial class. For example, infinite server queues and some bivariate queueing models we will consider in Chapter 5 have this property when modelled as LDQBDs.

In choosing a value for K^* we want

$$\sum_{k=K^*+1}^{\infty} \ell_k < \epsilon \tag{4.4.21}$$

where $\ell_k = \boldsymbol{x}_k \boldsymbol{e}$ is the marginal stationary probability of being in level k. Suppose we consider a birth-death process $\bar{Y}(t)$ on the state space $\{0, 1, 2, \dots\}$ with transition rates $\bar{q}(k, j)$ and stationary distribution $\bar{\ell}_k$. If we can construct $\bar{Y}(t)$ such that

$$\sum_{k=K^*+1}^{\infty} \ell_k \leq \sum_{k=K^*+1}^{\infty} \bar{\ell}_k \tag{4.4.22}$$

then a possible method of choosing K^* is to take K^* to be the smallest value of K such that

$$\sum_{k=K+1}^{\infty} \bar{\ell}_k < \epsilon. \tag{4.4.23}$$

If we take

$$\bar{q}(k, k+1) = [Q_0^{(k)} \boldsymbol{e}]_{max} \qquad k \geq 0 \tag{4.4.24}$$

$$\bar{q}(k, k-1) = [Q_0^{(k)} \boldsymbol{e}]_{min} \qquad k \geq 1 \tag{4.4.25}$$

where $[\boldsymbol{v}]_{min}$ and $[\boldsymbol{v}]_{max}$ are the minimum and maximum entries of the vector $\boldsymbol{v}$ respectively, then provided the stationary distribution $\bar{\ell}_k$ exists, we expect that (4.4.22) will hold since intuitively $\bar{Y}(t)$ will have more probability mass in its tail than the original LDQBD. Since $\bar{Y}(t)$ is a birth-death process, if the stationary distribution exists, we have

$$\sum_{k=K+1}^{\infty} \bar{\ell}_k = \sum_{k=K+1}^{\infty} \bar{\ell}_0 \prod_{j=0}^{k-1} \frac{\bar{q}(j, j+1)}{\bar{q}(j+1, j)}. \tag{4.4.26}$$

A sufficient condition for the stationary distribution to exist is $\bar{q}(j, j+1)/\bar{q}(j+1, j) \leq \alpha < 1 \ \forall j \geq J$ for some J since then the terms in the infinite sum will be bounded above by a geometric series. In this case we have

$$\begin{aligned}
\sum_{k=K^*+1}^{\infty} \bar{\ell}_k &= \bar{\ell}_{K^*+1} + \sum_{k=K^*+2}^{\infty} \bar{\ell}_{K^*+1} \prod_{j=0}^{k-K^*-2} \frac{\bar{q}(K^*+1+j, K^*+2+j)}{\bar{q}(K^*+2+j, K^*+1+j)} \\
&\leq \epsilon + \epsilon \sum_{k=K^*+2}^{\infty} \alpha^{k-K^*-1} \\
&= \epsilon \left(\frac{1}{1-\alpha} \right).
\end{aligned}$$

So in this case we have an upper bound on $\sum_{k=K+2}^{\infty} \bar{\ell}_k$ and the smaller α is, the smaller the bound will be. Intuitively we expect that

$$\sum_{k=K+1}^{\infty} \ell_k \leq \frac{\epsilon}{1-\alpha}$$

so we can say $x_k(K^*)$ is approximately equal to the stationary distribution of being in level k. In other cases more detailed analysis of $X(t)$ may be necessary to determine if this choice of K^* is sufficiently large for $\boldsymbol{x}_k(K^*)$ to approximate $\boldsymbol{x}_k$. At the very least, $\boldsymbol{x}_k(K^*)$ can always be interpreted as an upper bound for the stationary distribution.

As mentioned previously, the matrices $\boldsymbol{Q}_i^{(k)}$ must be expressible in some parametric form. If this parametric form is simple, it may be possible to obtain an analytical expression for $\sum_{k=K+1}^{\infty} \bar{\ell}_k$ and it is then straightforward to find a value for K^*. When we don't have an analytical expression for $\bar{\ell}_k$, we can still compute the quantity $\{\bar{\ell}_k(\bar{K}), 0 \leq k \leq \bar{K}\}$ given by

$$\bar{\ell}_k(\bar{K}) = \bar{\ell}_0(\bar{K}) \prod_{j=1}^{k-1} \frac{\bar{q}(j, j+1)}{\bar{q}(j+1, j)} \qquad 0 \leq k \leq \bar{K}, \tag{4.4.27}$$

where $\bar{\ell}_0(\bar{K})$ is found using $\sum_{k=0}^{\bar{K}} \bar{\ell}_k(\bar{K}) = 1$. It is clear that for any $\bar{K} \geq 0$, we have $\bar{\ell}_k(\bar{K}) \geq \bar{\ell}_k \ \forall\ 0 \leq k \leq \bar{K}$ and $\bar{\ell}_k(\infty) = \bar{\ell}_k$. We can then choose K^* to be the smallest value of K such that $\bar{\ell}_K(K) < \epsilon$.

The reasoning behind this choice of K^* is as follows. Firstly, if $\bar{\ell}_{K^*}(K^*) < \epsilon$ then we expect that $\ell_{K^*} < \epsilon$ so the stationary probability that $X(t)$ is in level K^* is small (according to some tolerance ϵ). We want K^* such that the stationary probability that $X(t)$ is in level $K^* + 1$ or above is small. So when we choose K^* in this way, we assume that if the stationary probability that $X(t)$ is in level K^* is small then the stationary probability that $X(t)$ is in level K^*+1 or above is also small. We hope that this will be the case but there will be situations where this will fail to hold. Hence careful consideration must be given to the structure of the matrices $\boldsymbol{Q}_i^{(k)}$ before we apply this method.

The discussion we have presented so far gives some justification for finding K^* by constructing the birth-death process $\bar{Y}(t)$. However, it is certainly desirable to give a formal mathematical justification for this choice of K^*. It seems that a direct approach is difficult and hence we choose to consider stochastically dominating processes. In order to do this we present some results on stochastic orderings and stochastic domination. We illustrate some of the concepts by considering a birth-death process.

Some theory of stochastic orderings

The key definitions and theorems that we state here are taken from Massey [63]. We also cite Stoyan [91] and Smeitink [90] as further references on stochastic orderings and their applications.

Let $\preceq$ indicate a *partial order* and $\prec$ indicate a *quasi-order*. For a set S, $\preceq$ is a partial order on S if

(i) $s \preceq s \ \forall s \in S$;

(ii) $s \preceq t$ and $t \preceq s$ imply $s = t$;

(iii) $s \preceq t$ and $t \preceq u$ imply $s \preceq u$,

and $\prec$ is a quasi-order on S if

(i) $s \prec s$ is false $\ \forall s \in S$;

(ii) $s \prec t$ and $t \prec u$ imply $s \prec u$.

Ross [88] states that given a quasi-order $\prec$, a partial order $\preceq$ can be defined by

$$x \preceq y \text{ if and only if } x \prec y \text{ or } x = y. \tag{4.4.28}$$

Consider the set $\mathcal{S} = \{0, 1, 2, \dots\}$ which is the state space for a birth-death process. We can define a partial order $\preceq$ on $\mathcal{S}$ by $k \preceq j$ if $k \leq j$. If we define the quasi order $\prec$ by $k \prec j$ if $k < j$ then we can define a partial order $\preceq$ according to (4.4.28) and this partial order will be equivalent to the partial order defined by $k \preceq j$ if $k \leq j$.

If S is the state space of a stochastic process and is endowed with a partial order, then for any subset Γ of S, define

$$\Gamma^{\uparrow} = \{y | x \preceq y \text{ for some } x \in \Gamma\}.$$

Γ is then said to be an *increasing set* if $\Gamma = \Gamma^{\uparrow}$. For the partial order $\preceq$ on $\mathcal{S} = \{0, 1, 2, \dots\}$ defined by $k \preceq j$ if $k \leq j$, the increasing sets are of the form $\{k, k+1, k+2, \dots\}$.

The following definition is equivalent to a definition given in Massey [63]. The definition is obtained from that of Massey by noting that any monotone function can be written as a linear combination of indicator functions of increasing sets.

Definition 4.4.1 *Let $\mathbf{y}$ and $\bar{\mathbf{y}}$ be probability density functions on a partially ordered set S. Then $\bar{\mathbf{y}}$ stochastically dominates $\mathbf{y}$ (written $\mathbf{y} \leq_s \bar{\mathbf{y}}$) if for all monotone functions $f(.) : S \to R$*

$$\sum_{n \in S} f(n)[\mathbf{y}]_n \leq \sum_{n \in S} f(n)[\bar{\mathbf{y}}]_n. \tag{4.4.29}$$

Definition 4.4.2 *The process $\bar{X}(t)$ with state space S and density function $\bar{\boldsymbol{x}}(t)$ stochastically dominates the process $X(t)$ with state space S and density function $\boldsymbol{x}(t)$ if*

$$\begin{aligned} \boldsymbol{x}(0) &\leq_s \bar{\boldsymbol{x}}(0) \\ \Rightarrow \boldsymbol{x}(t) &\leq_s \bar{\boldsymbol{x}}(t) \qquad \forall t \geq 0. \end{aligned} \tag{4.4.30}$$

If $\bar{X}(t)$ stochastically dominates $X(t)$ then equation (4.4.30) holds $\forall t \geq 0$ so we can take the limit as $t \to \infty$ of both sides of the inequality and obtain

$$\boldsymbol{x} \leq_s \bar{\boldsymbol{x}} \tag{4.4.31}$$

provided the stationary distributions $\boldsymbol{x}$ and $\bar{\boldsymbol{x}}$ exist.

The following theorem is based on Theorem 5.3 in Massey [63].

Lemma 4.4.1 *Let $X_1(t)$ and $X_2(t)$ be uniform Markov processes on the state space S with transition rates $q_1(i,j)$ and $q_2(i,j)$ respectively (i.e for $m = 1,2$, $\sum_j q_m(i,j)$ is bounded $\forall\, j$). If $\preceq$ is a partial order on S then $X_2(t)$ stochastically dominates $X_1(t)$ if and only if for all $x \preceq y$ in S and all increasing sets Γ, the following hold*

$$\text{if } x, y \in \Gamma, \qquad \sum_{z \notin \Gamma} q_1(x,z) \geq \sum_{z \notin \Gamma} q_2(y,z) \tag{4.4.32}$$

and

$$\text{if } x, y \notin \Gamma, \qquad \sum_{z \in \Gamma} q_1(x,z) \leq \sum_{z \in \Gamma} q_2(y,z). \tag{4.4.33}$$

Brandt and Last [11] extended Lemma 4.4.1 by removing the requirement that $X_1(t)$ and $X_2(t)$ be uniform. For convenience we state this as a lemma.

Lemma 4.4.2 *Lemma 4.4.1 holds for non-uniform $X_1(t)$ and $X_2(t)$.*

To illustrate the use of Lemma 4.4.2 we let $X_1(t)$ and $X_2(t)$ be birth-death processes on $S = \{0, 1, \dots\}$. Defining $\preceq$ by $k \preceq j$ if $k \leq j$, the increasing sets are of the form $\{k, k+1, k+2, \dots\}$. By considering conditions (4.4.32) and (4.4.33), it is clear that for any birth-death process $X_1(t)$, provided we choose $q_2(k,j)$ such that

$$q_2(k, k+1) \geq q_1(k, k+1) \quad k \geq 0 \tag{4.4.34}$$

and

$$q_2(k,k-1) \leq q_1(k,k-1) \quad k \geq 1, \tag{4.4.35}$$

then $X_2(t)$ will stochastically dominate $X_1(t)$.

We now return to the question of finding K^*. Firstly we observe that the marginal probabilities ℓ_k for the LDQBD $X(t)$ can be thought of as the stationary probabilities for a birth-death process. Call this birth-death process $Y(t)$. One way of showing that the condition (4.4.22) holds is to show that $\bar{Y}(t)$ stochastically dominates $Y(t)$. A possible way to do this is to use Lemma 4.4.2 but this would require knowing the transition rates for $Y(t)$. Since we don't know these transition rates, we instead consider the original LDQBD $X(t)$. We attempt to construct a LDQBD $\bar{X}(t)$ that stochastically dominates $X(t)$ and such that its marginal probabilities are given by $\bar{\ell}_k$. From the results for stochastically dominating processes, we can then conclude that (4.4.22) holds.

We note that if we can show that $\bar{X}(t)$ stochastically dominates $X(t)$ then we are showing a much stronger result than we actually need. This is because stochastic domination is essentially a time dependent result whereas the condition (4.4.22) is a stationary result. However, it is still worthwhile investigating if $\bar{X}(t)$ does stochastically dominate $X(t)$.

It is simple to show that the probabilities $\bar{\ell}_k$ are the marginal stationary probabilities for the LDQBD $\bar{X}(t)$ with q-matrix defined by

$$\bar{Q} = \begin{pmatrix} \bar{Q}_1^{(0)} & \bar{Q}_0^{(0)} & 0 & 0 & \cdots \\ \bar{Q}_2^{(1)} & \bar{Q}_1^{(1)} & \bar{Q}_0^{(1)} & 0 & \cdots \\ 0 & \bar{Q}_2^{(2)} & \bar{Q}_1^{(2)} & \bar{Q}_0^{(2)} & \cdots \\ 0 & 0 & \bar{Q}_2^{(3)} & \bar{Q}_1^{(3)} & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix} \tag{4.4.36}$$

where

$$[\bar{Q}_0^{(k)}]_{ij} = \frac{[Q_0^{(k)}e]_{max}}{M(k+1)} \quad k \geq 0, \quad 1 \leq i \leq M(k), \quad 1 \leq j \leq M(k+1), \tag{4.4.37}$$

$$[\bar{Q}_2^{(k)}]_{ij} = \frac{[Q_2^{(k)}e]_{min}}{M(k-1)} \quad k \geq 1, \quad 1 \leq i \leq M(k), \quad 1 \leq j \leq M(k-1), \tag{4.4.38}$$

$$[\bar{Q}_1^{(k)}]_{ij} = [Q_1^{(k)}]_{ij} \qquad k \geq 0, \quad 1 \leq i,j \leq M(k), \quad i \neq j. \tag{4.4.39}$$

In order to determine if $\bar{X}(t)$ stochastically dominates $X(t)$, we use Lemma 4.4.2. In order to do this we need to define a partial order $\preceq$. The obvious partial order to consider is defined by $(i,j) \preceq (k,\ell)$ if $i \leq k,\ j \leq \ell$. This leads to increasing sets of the form

$$\{(k,j) : k \geq \bar{k}, j \geq \bar{j}\} \tag{4.4.40}$$

for some $\bar{k}$ and $\bar{j}$. It is easy to show that for these increasing sets the conditions (4.4.32) and (4.4.33) in Lemma 4.4.2 do not hold. We therefore need to choose another partial order. Since the state space for a LDQBD is partitioned into levels, a natural set of increasing sets to consider are sets of the form

$$\{(k,j) : k \geq \bar{k}\}. \tag{4.4.41}$$

That is, each increasing set consists of all states in and above a certain level $\bar{k}$. The fact that LDQBDs are skip free on the levels should assist us in showing conditions (4.4.32) and (4.4.33) hold.

In order to define a partial order such that the increasing sets are of the form (4.4.41) we define the relation $\prec$ by $(i,j) \prec (k,l)$ if and only if $i < k$. It is easy to show that $\prec$ is a quasi-order. Define the relation $\preceq$ according to (4.4.28).

Now according to this partial order, the increasing sets are of the form

$$\Phi_{i,I} = \{(k,j) | k \geq i, 1 \leq j \leq M(k)\} \cup \{(i-1,j) | j \in I\} \tag{4.4.42}$$

where $I \subseteq \{1, 2, \ldots M(i-1)\}$ is some non-empty set of indices. That is, the increasing sets contain all the states in levels i and above and at least one state in level $i-1$. Let $\Phi(i) = \bigcup_I \Phi_{i,I}$ and define Θ by

$$\Theta = \bigcup_{i=1}^{\infty} \Phi(i).$$

Thus $\Phi(i)$ is the set of all increasing sets Γ such that $(\ell, j) \in \Gamma\ \forall\, \ell \geq i,\ 1 \leq j \leq M(\ell)$, and $(i-1, j) \in \Gamma$ for at least one $1 \leq j \leq M(i-1)$ and Θ is the set of all increasing sets. Note that $\Phi(i) \cap \Phi(j) = \phi\ \ \forall\ i \neq j$ and that the cardinality of Θ is infinite but the cardinality of $\Phi(i)$ is finite $\forall\ i \geq 1$.

We now attempt to show that conditions (4.4.32) and (4.4.33) hold with $q_1(x, z)$ given by the matrices $Q_0^{(k)}$, $Q_1^{(k)}$ and $Q_2^{(k)}$ for the original LDQBD and $q_2(y, z)$ given by the matrices $\bar{Q}_0^{(k)}$, $\bar{Q}_1^{(k)}$ and $\bar{Q}_2^{(k)}$ of equations (4.4.37)-(4.4.39).

Firstly consider the case where $\Gamma \in \Phi(i), i > 1$ and $x = y = (i-1, j) \in \Gamma$. We have

$$\sum_{z \notin \Gamma} q_1(x, z) = \sum_{\substack{\ell s.t \\ (i-1,\ell) \notin \Gamma}} [Q_1^{(i-1)}]_{j\ell} + [Q_2^{(i-1)} \mathbf{e}]_j \tag{4.4.43}$$

and

$$\begin{aligned} \sum_{z \notin \Gamma} q_2(y, z) &= \sum_{\substack{\ell s.t \\ (i-1,\ell) \notin \Gamma}} [\bar{Q}_1^{(i-1)}]_{j\ell} + [\bar{Q}_2^{(i-1)} \mathbf{e}]_j && (4.4.44) \\ &= \sum_{\substack{\ell s.t \\ (i-1,\ell) \notin \Gamma}} [Q_1^{(i-1)}]_{j\ell} + [Q_2^{(i-1)} \mathbf{e}]_{min} && (4.4.45) \\ &\leq \sum_{z \notin \Gamma} q_1(x, z) && (4.4.46) \end{aligned}$$

and hence condition (4.4.32) holds for this case.

Now consider where $\Gamma \in \Phi(i),\ i > 1$ and $x = (i-1, j) \in \Gamma,\ y = (i, k) \in \Gamma$. We have

$$\sum_{z \notin \Gamma} q_1(x, z) = \sum_{\substack{\ell s.t \\ (i-1,\ell) \notin \Gamma}} [Q_1^{(i-1)}]_{j\ell} + [Q_2^{(i-1)} \mathbf{e}]_j \tag{4.4.47}$$

and

$$\begin{aligned} \sum_{z \notin \Gamma} q_2(y, z) &= \sum_{\substack{\ell s.t \\ (i-1,\ell) \notin \Gamma}} [\bar{Q}_2^{(i)}]_{k\ell} && (4.4.48) \\ &\leq [Q_2^{(i)} \mathbf{e}]_{min}. && (4.4.49) \end{aligned}$$

So in general (4.4.33) does not hold and hence we cannot show that $\bar{X}(t)$ stochastically dominates $X(t)$. We note that $\bar{X}(t)$ may still stochastically dominate $X(t)$ but we have not managed to prove the result via Lemma 4.4.2. Other attempts at proving the result may be made by defining different partial orderings and appealing to Lemma 4.4.2.

However, it seems that the partial ordering we defined is the most natural ordering to consider so we feel that further attempts will be unsuccessful.

Although we have seen that $\bar{X}(t)$ does not stochastically dominate $X(t)$, the discussion above shows that $\bar{X}(t)$ is close to satisfying the conditions (4.4.32) and (4.4.33). In the following theorem we define a process which is similar to $\bar{X}(t)$ defined above and we prove that it stochastically dominates $X(t)$.

Theorem 4.4.1 *Let $X(t)$ be a continuous time LDQBD which satisfies Condition 1. Moreover, let $\bar{X}(t)$ be a continuous time LDQBD on the same state space as $X(t)$ and with q-matrix of the form given by (4.4.36) where*

$$[\bar{\boldsymbol{Q}}_0^{(0)}]_{ij} = [\boldsymbol{Q}_0^{(0)}]_{ij}, \tag{4.4.50}$$

$$[\bar{\boldsymbol{Q}}_0^{(k)}]_{ij} = max\left[\frac{[\boldsymbol{Q}_0^{(k-1)}\boldsymbol{e}]_{max}}{M(k+1)}, [\boldsymbol{Q}_0^{(k)}]_{ij}\right] \qquad k \geq 1, \tag{4.4.51}$$

$$[\bar{\boldsymbol{Q}}_2^{(1)}]_{ij} = 0, \tag{4.4.52}$$

$$[\bar{\boldsymbol{Q}}_2^{(k)}]_{ij} = min\left[\frac{[\boldsymbol{Q}_2^{(k-1)}\boldsymbol{e}]_{min}}{M(k-1)}, [\boldsymbol{Q}_2^{(k)}]_{ij}\right] \qquad k \geq 2, \tag{4.4.53}$$

$$[\bar{\boldsymbol{Q}}_1^{(k)}]_{ij} = [\boldsymbol{Q}_1^{(k)}]_{ij} \qquad k \geq 0, \quad i \neq j. \tag{4.4.54}$$

Then $\bar{X}(t)$ stochastically dominates $X(t)$.

Proof: We prove the result using Lemma 4.4.2. As we did previously we define the relation $\prec$ by $(i,j) \prec (k,l)$ if and only if $i < k$ and define the relation $\preceq$ according to (4.4.28).

We have seen that the increasing sets are given by

$$\Theta = \bigcup_{i=1}^{\infty} \Phi(i)$$

where $\Phi(i) = \cup_I \Phi_{i,I}$ with $\Phi_{i,I}$ defined by (4.4.42).

In order to prove the theorem, we need to show that conditions (4.4.32) and (4.4.33) hold in Lemma 4.4.2 with $q_1(x,z)$ given by the matrices $\boldsymbol{Q}_0^{(k)}$, $\boldsymbol{Q}_1^{(k)}$ and $\boldsymbol{Q}_2^{(k)}$ for the original LDQBD and $q_2(y,z)$ given by the matrices $\bar{Q}_0^{(k)}$, $\bar{Q}_1^{(k)}$ and $\bar{Q}_2^{(k)}$ of equations (4.4.50)-(4.4.51).

First look at the case $x \prec y \in \Gamma$. Take $\Gamma \in \Phi(i),\ i > 1$ and $x = (i-1,j)$, $y = (i,k)$. We have

$$\sum_{z \notin \Gamma} q_1(x,z) = \sum_{\substack{\ell s.t \\ (i-1,\ell)\notin\Gamma}} [Q_1^{(i-1)}]_{j\ell} + [Q_2^{(i-1)}e]_j \tag{4.4.55}$$

and

$$\sum_{z \notin \Gamma} q_2(y,z) = \sum_{\substack{\ell s.t \\ (i-1,\ell)\notin\Gamma}} [\bar{Q}_2^{(i)}]_{k\ell} \tag{4.4.56}$$

$$\leq [Q_2^{(i-1)}e]_{min} \tag{4.4.57}$$

so it is clear that (4.4.32) holds. Note that with $\Gamma \in \Phi(i),\ \ i > 1$, (4.4.32) holds trivially for all other $x \prec y \in \Gamma$. Now consider $\Gamma \in \Phi(1)$ and $x = (0,j)$, $y = (1,k)$. We have

$$\sum_{z \notin \Gamma} q_1(x,z) = \sum_{\substack{\ell s.t \\ (0,\ell)\notin\Gamma}} [Q_1^{(0)}]_{j\ell} \tag{4.4.58}$$

and

$$\sum_{z \notin \Gamma} q_2(y,z) = \sum_{\substack{\ell s.t \\ (0,\ell)\notin\Gamma}} [\bar{Q}_2^{(1)}]_{k\ell} \tag{4.4.59}$$

$$= 0 \tag{4.4.60}$$

so (4.4.32) holds. For all other $x \prec y \in \Gamma \in \Phi(1)$, (4.4.32) holds trivially.

Now consider $x = y \in \Gamma$. For $\Gamma \in \Phi(i), i > 1$ the only non-trivial cases of x that we must consider are $x = (i,j)$ and $x = (i-1,j)$. For $x = (i,j)$ we have

$$\sum_{z \notin \Gamma} q_1(x,z) = \sum_{\substack{\ell s.t \\ (i-1,\ell)\notin\Gamma}} [Q_2^{(i)}]_{j\ell} \tag{4.4.61}$$

and

$$\sum_{z \notin \Gamma} q_2(y,z) = \sum_{\substack{\ell s.t \\ (i-1,\ell)\notin\Gamma}} [\bar{Q}_2^{(i)}]_{j\ell} \tag{4.4.62}$$

$$\leq \sum_{z \notin \Gamma} q_1(x,z) \tag{4.4.63}$$

from (4.4.53). For $x = (i-1, j)$ we have

$$\sum_{z \notin \Gamma} q_1(x, z) = \sum_{\substack{\ell s.t \\ (i-1,\ell) \notin \Gamma}} [Q_1^{(i-1)}]_{j\ell} + [Q_2^{(i-1)} e]_j \tag{4.4.64}$$

and

$$\sum_{z \notin \Gamma} q_2(y, z) = \sum_{\substack{\ell s.t \\ (i-1,\ell) \notin \Gamma}} [\bar{Q}_1^{(i-1)}]_{j\ell} + [\bar{Q}_2^{(i-1)} e]_j \tag{4.4.65}$$

$$\leq \sum_{z \notin \Gamma} q_1(x, z) \tag{4.4.66}$$

from (4.4.54) and (4.4.53). For $\Gamma \in \Phi(1)$ we need only consider $x = (1, j)$ and $x = (0, j)$. For $x = (1, j)$ we have

$$\sum_{z \notin \Gamma} q_1(x, z) = \sum_{\substack{\ell s.t \\ (0,\ell) \notin \Gamma}} [Q_2^{(1)}]_{j\ell} \tag{4.4.67}$$

and

$$\sum_{z \notin \Gamma} q_2(y, z) = \sum_{\substack{\ell s.t \\ (0,\ell) \notin \Gamma}} [\bar{Q}_2^{(1)}]_{j\ell} \tag{4.4.68}$$

$$= 0 \tag{4.4.69}$$

and for $x = (0, j)$ we have

$$\sum_{z \notin \Gamma} q_1(x, z) = \sum_{\substack{\ell s.t \\ (0,\ell) \notin \Gamma}} [Q_1^{(0)}]_{j\ell} \tag{4.4.70}$$

and

$$\sum_{z \notin \Gamma} q_2(y, z) = \sum_{\substack{\ell s.t \\ (0,\ell) \notin \Gamma}} [\bar{Q}_1^{(0)}]_{j\ell} \tag{4.4.71}$$

$$= \sum_{\substack{\ell s.t \\ (0,\ell) \notin \Gamma}} [Q_1^{(0)}]_{j\ell} \tag{4.4.72}$$

so (4.4.32) holds in both cases.

The proof that (4.4.33) holds for all $x \preceq y \in S$ and all increasing sets Γ is similar to the above. This completes the proof. □

We now make some comments about Theorem 4.4.1. We refer to the process $\bar{X}(t)$ defined in Theorem 4.4.1 as the *dominating process*. Note that the dominating process is an irreducible LDQBD on the state space $\{(k,j) : k \geq 1, 1 \leq j \leq M(k)\}$. In general $\bar{X}(t)$ has a complex structure and therefore cannot be used to find a value of K^*. However, in the case where for all k we have

$$\frac{[Q_0^{(k-1)}e]_{max}}{M(k+1)} \geq [Q_0^{(k)}]_{ij} \qquad \forall i,j, \tag{4.4.73}$$

and

$$\frac{[Q_2^{(k-1)}e]_{min}}{M(k-1)} \leq [Q_2^{(k)}]_{ij} \qquad \forall i,j, \tag{4.4.74}$$

the dominating process has the property that its marginal distribution over the levels is easy to calculate. It is then possible to use methods similar to those we have described above to calculate a value of K^* using this marginal distribution.

In the situation where (4.4.73) and (4.4.74) hold, the dominating process is similar in structure to the LDQBD with matrices $Q_i^{(k)}$ defined by (4.4.37)-(4.4.39). Although we have shown that the later LDQBD does not in general stochastically dominate the original LDQBD, the fact that it is similar to the dominating process suggests that it can come close to stochastically dominating the original LDQBD. We note that conditions (4.4.73) and (4.4.74) may not hold even if $Q_i^{(k)}$ have relatively simple expressions. For example, if $Q_0^{(k)} = \text{diag}(\boldsymbol{\lambda}), \forall k \geq 0$ and $M(k) = M, \forall k \geq 0$ then condition (4.4.73) becomes $\lambda_{max}/M > \lambda_i$ which does not hold at least for i such that $\lambda_i = \lambda_{max}$.

When we presented the stochastic domination theory, we mentioned that for any given birth-death process it is always possible to construct a second birth-death process that stochastically dominates the first. This result is also true for any given LDQBD. If we replace equations (4.4.51) and (4.4.53) by

$$[\bar{Q}_0^{(k)}]_{ij} = \lambda_k \tag{4.4.75}$$

and

$$[\bar{Q}_2^{(k)}]_{ij} = \mu_k \tag{4.4.76}$$

respectively where

$$\lambda_k > max\left[\frac{[Q_0^{(k-1)}e]_{max}}{M(k+1)}, [Q_0^{(k)}]_{ij}\right] \tag{4.4.77}$$

and

$$\mu_k < min\left[\frac{[Q_2^{(k-1)}e]_{min}}{M(k-1)}, [Q_2^{(k)}]_{ij}\right] \tag{4.4.78}$$

then it is clear that the resulting LDQBD will also stochastically dominate $X(t)$. Furthermore, when they exist, the marginal stationary probabilities will be easy to calculate for this process. Of course it will often be difficult to choose λ_k and μ_k so that conditions (4.4.77) and (4.4.78) hold and so that the resulting LDQBD is also positive recurrent. Even if it is possible to do so, the resulting value of K^* will generally be much larger than necessary. Hence this is not a practical way to select a value for K^*.

We summarise the above discussion as follows. When Condition 1 holds, a suitable method for computing K^* is to consider the birth-death process $\bar{Y}(t)$ with transition rates given by (4.4.24) and (4.4.25) and provided $\bar{Y}(t)$ is positive recurrent, we can choose K^* to be the smallest value of K such that equation (4.4.23) holds. We will illustrate this method of finding K^* in the next chapter when we consider some numerical examples. We state the following algorithm for the case when Condition 1 holds.

Algorithm 4.4.4 (Calculating $\boldsymbol{x}_k(K^*), 0 \leq k \leq K^*$ when Condition 1 holds)

```
Construct the birth-death process with transition rates
given by (4.4.24) and (4.4.25)

if ( the birth-death process is positive recurrent )
    Set $K^*$ equal to the minimum value of $K$ such that $\sum_{k=K+1}^{\infty} \bar{\ell}_k < \epsilon$
else
    choose $K^*$ using some other method

Proceed with Algorithm 3.4.4 using the current value of $K^*$
```

Before concluding this section we note that the arguments we have presented here also apply to processes for which Condition 1 holds only for states in and above a certain level $\hat{K}$. This weaker form of Condition 1 is as follows.

Condition 2 : $\forall k \geq \hat{K}$ and $\forall i \; \exists j \; s.t \, [Q_2^{(k)}]_{ij} > 0$.

The extensions of our arguments to the case where Condition 2 holds are clear. We omit the details.

4.5 The stationary distribution of a discrete time LDQBD

As mentioned at the beginning of the chapter all the results presented for continuous time LDQBDs have discrete time analogues. In this section we state these discrete time results without proof. Their proofs are similar to those for the continuous time results.

The discrete time equivalent of Theorem 4.2.1 is as follows.

Theorem 4.5.1 *If $X(t)$ is a discrete time positive recurrent LDQBD then its stationary distribution $\boldsymbol{x}$ is given by*

$$\boldsymbol{x}_k = \boldsymbol{x}_0 \prod_{\ell=0}^{k-1} \boldsymbol{R}_\ell \qquad k \geq 0 \tag{4.5.1}$$

where the matrices $\boldsymbol{R}_k$ are defined in Chapter 2 and $\boldsymbol{x}_0$ is a positive solution to

$$\boldsymbol{x}_0(\boldsymbol{A}_1^{(0)} + \boldsymbol{R}_0 \boldsymbol{A}_2^{(1)}) = \boldsymbol{x}_0 \tag{4.5.2}$$

such that

$$\boldsymbol{x}_0 \sum_{k=0}^{\infty} \left[\prod_{\ell=0}^{k-1} \boldsymbol{R}_\ell \right] \boldsymbol{e} = 1. \tag{4.5.3}$$

If we define $(\boldsymbol{x}_k(K^*))_j$, $0 \leq k \leq K^*$ by

$$\boldsymbol{x}_k(K^*) = \boldsymbol{x}_0(K^*) \prod_{\ell=0}^{k-1} \boldsymbol{R}_\ell \qquad 0 \leq k \leq K^* \tag{4.5.4}$$

where $\boldsymbol{x}_0(K^*)$ satisfies (4.5.2) and

$$\boldsymbol{x}_0(K^*) \sum_{k=0}^{K^*} \left[\prod_{m=0}^{k-1} \boldsymbol{R}_m \right] \boldsymbol{e} = 1 \tag{4.5.5}$$

then $\boldsymbol{x}_k(K^*)$ has the same interpretation as it did for continuous time processes. Hence by taking K^* large enough, we hope to have $\boldsymbol{x}_k(K^*) \approx \boldsymbol{x}_k$.

The discrete time equivalent of Theorem 4.3.1 is as follows.

Theorem 4.5.2 *The vector* $\boldsymbol{x}(K^*) = (\boldsymbol{x}_0(K^*), \boldsymbol{x}_1(K^*), \ldots, \boldsymbol{x}_{K^*}(K^*))$ *where* $\boldsymbol{x}_k(K^*)$ *is given by (4.5.4) is the stationary distribution for the irreducible Markov chain on the state space* $\{(k,j) : 0 \le k \le K^*, 1 \le j \le M(k)\}$ *with transition matrix* $\boldsymbol{P}(K^*)$ *given by*

$$\boldsymbol{P}(K^*) = \begin{pmatrix} \boldsymbol{A}_1^{(0)} & \boldsymbol{A}_0^{(0)} & & & & \\ \boldsymbol{A}_2^{(1)} & \boldsymbol{A}_1^{(1)} & \boldsymbol{A}_0^{(1)} & & & \\ & \ddots & \ddots & \ddots & & \\ & & \boldsymbol{A}_2^{(K^*-1)} & \boldsymbol{A}_1^{(K^*-1)} & \boldsymbol{A}_0^{(K^*-1)} \\ & & & \boldsymbol{A}_2^{(K^*)} & \boldsymbol{U}_{K^*} \end{pmatrix} \tag{4.5.6}$$

where $\boldsymbol{U}_{K^*}$ *is given by equation (2.6.1).*

Hence $\boldsymbol{x}(K^*)$ *is the unique solution to the system of equations*

$$\boldsymbol{x}(K^*)\boldsymbol{P}(K^*) = \boldsymbol{x}(K^*) \tag{4.5.7}$$

subject to

$$\sum_{k=0}^{K^*} \boldsymbol{x}_k(K^*)\boldsymbol{e} = 1. \tag{4.5.8}$$

The methods we described for solving $\boldsymbol{x}(K^*)\boldsymbol{Q}(K^*) = \boldsymbol{0}$ for continuous time processes can all be adapted to the case of discrete time processes. Again, the method we propose is to compute $\boldsymbol{R}_0, \ldots, \boldsymbol{R}_{K^*-1}$ and then to use equation (4.5.4).

From equations (2.6.1) and (2.6.2) we have

$$\begin{aligned} \boldsymbol{R}_k &= \boldsymbol{A}_0^{(k)}(\boldsymbol{I} - \boldsymbol{U}_{k+1})^{-1} && (4.5.9) \\ &= \boldsymbol{A}_0^{(k)}(\boldsymbol{I} - \boldsymbol{A}_1^{(k+1)} - \boldsymbol{R}_{k+1}\boldsymbol{A}_2^{(k+2)})^{-1}. && (4.5.10) \end{aligned}$$

Therefore we can calculate $\boldsymbol{R}_0, \ldots, \boldsymbol{R}_{K^*-1}$ by using Algorithm 3.4.4 to calculate $\boldsymbol{R}_{K^*-1}$ and then using equation (4.5.10) to calculate $\boldsymbol{R}_{K^*-2}, \boldsymbol{R}_{K^*-3}, \ldots, \boldsymbol{R}_0$. As was the case for continuous time processes, the recursion in equation (4.5.10) is stable since it can be shown that the inverse contains non negative elements.

Hence the discrete time version of Algorithm 4.3.1 is as follows.

Algorithm 4.5.1 (Calculating $\boldsymbol{x}_k(K^*), 0 \leq k \leq K^*$ given K^*)

```
Calculate $\boldsymbol{R}_{K^*-1}$ using Algorithm 3.4.4

Recursively calculate $\boldsymbol{R}_{K^*-2}, \boldsymbol{R}_{K^*-3}, ..., \boldsymbol{R}_0$

Solve $\boldsymbol{x}_0(K^*)(\boldsymbol{A}_1^{(0)} + \boldsymbol{R}_0\boldsymbol{A}_2^{(1)}) = \boldsymbol{x}_0(K^*)$
subject to $\boldsymbol{x}_0(K^*)\boldsymbol{e} = 1$

for $k = 1$ to $K^*$ do
    $\boldsymbol{x}_k(K^*) = \boldsymbol{x}_{k-1}(K^*)\boldsymbol{R}_{k-1}$

    Normalise $\boldsymbol{x}_0(K^*), \boldsymbol{x}_1(K^*), ..., \boldsymbol{x}_k(K^*)$ s.t. $\sum_{m=0}^{k} \boldsymbol{x}_m(K^*)\boldsymbol{e} = 1$
```

Note that the only difference between Algorithm 4.5.1 and Algorithm 4.3.1 is that in the latter we solve $\boldsymbol{x}_0(K^*)(\boldsymbol{Q}_1^{(0)} + \boldsymbol{R}_0\boldsymbol{Q}_2^{(1)}) = \boldsymbol{0}$ whereas in the former we solve $\boldsymbol{x}_0(K^*)(\boldsymbol{A}_1^{(0)} + \boldsymbol{R}_0\boldsymbol{A}_2^{(1)}) = \boldsymbol{x}_0(K^*)$. This is the only change that is needed in order to apply Algorithms 4.4.1, 4.4.2 and 4.4.3 to the relevant discrete time processes.

The theory of stochastic orderings and stochastic dominance carries over to discrete time processes. In order to apply the results in Section 4.4.4 to discrete time processes, it is simply a matter of replacing the transition rates with the corresponding transition probabilities. So for example, Condition 1 is replaced by $\forall\, k, i\ \exists\, j$ such that $[\boldsymbol{A}_2^{(k)}]_{ij} > 0$. With these changes, Algorithm 4.4.4 can be applied to discrete time processes.

Chapter 5

Applications

5.1 Introduction

In this chapter we present a variety of queueing systems that can be modelled as LDQBDs. We firstly consider a number of retrial queues. We then consider some queues with an infinite server capacity and finally we consider some bivariate queueing models. We model all the systems we consider as continuous time LDQBDs. We do this because the systems we consider are better represented by continuous time processes than discrete time processes. As we mentioned in Chapter 4, there is only a small difference between the analysis of continuous time LDQBDs and discrete time LDQBDs. We refer to Alfa [4] who modelled a discrete MAP/PH/1 retrial vacation queue as a discrete time LDQBD. Alfa used our algorithms as a basis for his analysis.

We have constructed computer code for the algorithms presented in Chapters 3 and 4 and used this to obtain the numerical results presented in this chapter. In order to test the accuracy of our computer code we employed several methods. In Section 5.2 we present an analytical expression obtained by Diamond and Alfa [20] for the matrices $\boldsymbol{R}_k$ for the M/PH/1 retrial queue. This expression allows us to compute the matrices $\boldsymbol{R}_k$ by evaluating only one matrix inverse. We compared the values of $\boldsymbol{R}_k$ computed from this expression with the values obtained via Algorithm 3.4.4. As a more general test, we used our code to compute $\boldsymbol{R}_k$ and $\boldsymbol{R}_{k+1}$ for some value of k and then computed the matrix $\boldsymbol{Q}_0^{(k)} + \boldsymbol{R}_k\boldsymbol{Q}_1^{(k+1)} + \boldsymbol{R}_k\boldsymbol{R}_{k+1}\boldsymbol{Q}_2^{(k+2)}$ which should equal the zero matrix. We found that if we took $\epsilon = 10^{-10}$ in Algorithm 3.4.4 then the absolute value of the maximum element of $\boldsymbol{Q}_0^{(k)} + \boldsymbol{R}_k\boldsymbol{Q}_1^{(k+1)} + \boldsymbol{R}_k\boldsymbol{R}_{k+1}\boldsymbol{Q}_2^{(k+2)}$ was less than 10^{-14}. In order to test the accuracy of the vector $\boldsymbol{x}(K^*)$, we calculated $\boldsymbol{x}(K^*)\boldsymbol{Q}(K^*)$ where $\boldsymbol{Q}(K^*)$ is given by (4.3.1). This should be equal to the zero vector.

We begin this chapter by considering the M/PH/1 retrial queue which has been analysed by Diamond and Alfa [20]. We use the analytical expressions obtained by Diamond and Alfa for the $\boldsymbol{R}_k$ matrices for this queue to verify our computer code. We also investigate the change in the value of L in Algorithm 3.4.4 when this algorithm is used to compute the matrices $\boldsymbol{R}_k$ for different values of k. Following this we consider a M/M/c retrial queue and we use our algorithms to replicate some performance measures that have been obtained by other authors. We also investigate the values of L required by Algorithm 3.4.4 to compute various $\boldsymbol{R}_k$ matrices for this queue. We then show how some other retrial queues can be modelled using the LDQBD structure but we do not provide numerical results. Next we consider the M/M/∞ queue in a random environment and the PH/M/∞ queue. We compute the marginal stationary probabilities over the levels for these two queues. At the end of the chapter we consider some bivariate queueing systems that can be modelled as LDQBDs on an irregular state space. We reproduce some numerical results obtained by other authors but we see that our algorithms are not suitable for heavy traffic situations.

5.2 Retrial models

In this section we model a number of retrial queues as LDQBDs and analyse them using the algorithms presented in Chapter 4. Firstly we give a general definition of a retrial queue.

Suppose customers arrive to a system according to some arrival process. All customers in the system are either receiving service or are not receiving service, in which case they are said to be in *orbit*. On arrival to the system a customer either moves directly into orbit or tries to obtain service. If a customer tries to obtain service and fails, then they move into orbit. When a customer goes into orbit it waits a period of time and then tries to obtain service. If the customer is unsuccessful, the customer may keep re-attempting to obtain service until it is successful, in which case the customer is said to be *persistent* or alternatively, the customer may make a certain number of re-attempts and then leave the system.

Retrial queues have received a great deal of attention in the literature where they have been used to model a variety of communications systems. (See for example Yang and Templeton [96], Falin [24], de Kok [51] and Greenberg [32]). Various techniques such as generating function approaches and approximation methods have been used to analyse retrial queues. However, only recently has the matrix-analytic approach been employed. We note that Alfa [4] has used the algorithms presented here as a basis for analysing a MAP/PH/1 vacation queue and Diamond and Alfa [20] modelled a M/PH/1 retrial queue as a LDQBD.

5.2.1 M/PH/1 retrial queue

Diamond and Alfa [20] considered a retrial queue in which customers arrive to a single server system in a Poisson stream with rate λ. If the server is free the customer moves straight into service and if the server is busy the customer moves into the orbiting pool. Customers in orbit retry according to a Poisson process with rate θ and the service time distribution for each customer is of phase type with representation $(\boldsymbol{S}, \boldsymbol{\alpha})$. For more details on phase type distributions, see Neuts [72, Chapter 2].

Diamond and Alfa modelled this retrial queue as a continuous time LDQBD on the state space $\mathcal{S} = \{-1\} \cup \{(0,j) : 1 \leq j \leq M\} \cup \{(k,j) : k \geq 1, 0 \leq j \leq M\}$. The queue is in the state $\{-1\}$ when there are no customers in orbit and the server is idle. The index k represents the number of customers in orbit and the index j represents the phase of service where $j = 0$ means the server is idle. M is the dimension of the underlying process of service time distribution. The q-matrix for this process is given by

$$Q = \begin{pmatrix} Q_1^{(-1)} & Q_0^{(-1)} & 0 & 0 & \cdots \\ Q_2^{(0)} & Q_1^{(0)} & Q_0^{(0)} & 0 & \cdots \\ 0 & Q_2^{(1)} & Q_1^{(1)} & Q_0^{(1)} & \cdots \\ 0 & 0 & Q_2^{(2)} & Q_1^{(2)} & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix} \tag{5.2.1}$$

where

$$Q_1^{(-1)} = -\lambda, \tag{5.2.2}$$

$$Q_0^{(-1)} = \lambda\boldsymbol{\alpha}', \tag{5.2.3}$$

$$Q_2^{(0)} = \boldsymbol{S}^0, \tag{5.2.4}$$

$$Q_1^{(0)} = \boldsymbol{S} - \lambda\boldsymbol{I}, \tag{5.2.5}$$

$$Q_0^{(0)} = \begin{pmatrix} 0 & \lambda\boldsymbol{I} \end{pmatrix}, \tag{5.2.6}$$

$$Q_2^{(1)} = \begin{pmatrix} \theta\boldsymbol{\alpha}' \\ \boldsymbol{0} \end{pmatrix}, \tag{5.2.7}$$

$$Q_1^{(k)} = \begin{pmatrix} -(k\theta + \lambda) & \lambda\boldsymbol{\alpha}' \\ \boldsymbol{S}^0 & \boldsymbol{S} - \lambda\boldsymbol{I} \end{pmatrix} \quad k \geq 1, \tag{5.2.8}$$

$$Q_0^{(k)} = \lambda \begin{pmatrix} 0 & 0 \\ 0 & \boldsymbol{I} \end{pmatrix} \quad k \geq 1, \tag{5.2.9}$$

$$Q_2^{(k)} = \begin{pmatrix} 0 & k\theta\boldsymbol{\alpha}' \\ \boldsymbol{0} & \boldsymbol{0} \end{pmatrix} \quad k \geq 2, \tag{5.2.10}$$

where $\boldsymbol{S}^0$ is a $M \times 1$ vector such that $(\boldsymbol{S} + \boldsymbol{S}^0)\boldsymbol{e} = \boldsymbol{0}$ and $\boldsymbol{\alpha}'$ is the transpose of the vector $\boldsymbol{\alpha}$. It is shown in Falin [24] that this queue is positive recurrent provided $\lambda(-\boldsymbol{\alpha}\boldsymbol{S}^{-1}\boldsymbol{e}) < 1$.

Diamond and Alfa [20] showed that the matrices $\boldsymbol{R}_k, k \geq -1$ can be written explicitly as

$$\begin{aligned}
\boldsymbol{R}_{-1} &= \boldsymbol{\alpha}'\boldsymbol{R}, && (5.2.11)\\
\boldsymbol{R}_0 &= \left(\frac{\lambda}{\theta}\boldsymbol{e}, \boldsymbol{R} + \frac{\lambda}{\theta}\boldsymbol{e}\boldsymbol{\alpha}'\boldsymbol{R}\right), && (5.2.12)\\
\boldsymbol{R}_k &= \begin{pmatrix} \boldsymbol{0} & \boldsymbol{0} \\ \dfrac{\lambda}{(k+1)\theta}\boldsymbol{e} & \boldsymbol{R} + \dfrac{\lambda}{(k+1)\theta}\boldsymbol{e}\boldsymbol{\alpha}'\boldsymbol{R} \end{pmatrix}, && (5.2.13)
\end{aligned}$$

where $\boldsymbol{R} = -\lambda(\boldsymbol{S} - \lambda\boldsymbol{I} + \lambda\boldsymbol{e}\boldsymbol{\alpha}')^{-1}$.

Diamond and Alfa actually showed the more general result that provided the matrices $\boldsymbol{Q}_2^{(k)}, k \geq 0$ have rank one, then explicit expressions can always be obtained for the matrices $\boldsymbol{R}_k$. Furthermore, the explicit expressions only involve one matrix inverse. We refer to [20] for more details. We note that it is a simple matter to show that when $\boldsymbol{Q}_2^{(k)}, k \geq 0$ have rank one, explicit expressions can also be obtained for $\boldsymbol{G}_k$. Liu and Zhao [59] also obtained explicit expressions for the matrices $\boldsymbol{G}$ and $\boldsymbol{R}$ for particular cases of level independent M/G/1–type and GI/M/1–type processes.

Validation of computer programs

Since simple expressions exist for $\boldsymbol{G}_k$ and $\boldsymbol{R}_k$ for this model, it is not necessary to use any of the algorithms presented in Chapter 4 to evaluate $\boldsymbol{G}_k$ and $\boldsymbol{R}_k$. However, the algorithms in Chapter 4 do of course still apply and the explicit expressions for $\boldsymbol{G}_k$ and $\boldsymbol{R}_k$ provide us with a means of validating the computer programs that we have written for the algorithms. We have done this for a range of parameter values and found that our programs provide results that agree with the explicit expressions.

Even though the matrices $\boldsymbol{R}_k$ can be found explicitly, in order to compute the stationary distribution we must still decide where to truncate the process. Diamond and Alfa [20] do not discuss methods of truncation and it seems that a general trial and error approach should be adopted.

Rather than generate the stationary distribution or any performance measures for this model we used Algorithm 3.4.4 to compute $\boldsymbol{R}_k$ for various values of k and we studied the value of the truncation parameter L. Recall that L is the value of ℓ at which we truncate the infinite sum in equation (3.3.18).

We assumed that the service time distribution has the following phase type distribution;

$$0.05841196E_{10}(1.168239125) + 0.94158804E_{10}(18.83176088) \tag{5.2.14}$$

which is the mixture of two ten stage Erlang distributions considered by Ramaswami and Latouche [81]. Since we have a mixture of two ten stage Erlang distributions, the resulting phase type distribution has 20 phases. The matrix $\boldsymbol{S}$ is given by

$$\boldsymbol{S} = \begin{pmatrix} \boldsymbol{S}_1 & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{S}_2 \end{pmatrix} \tag{5.2.15}$$

where

$$\boldsymbol{S}_1 = \begin{pmatrix} -\lambda_1 & \lambda_1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -\lambda_1 & \lambda_1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -\lambda_1 & \lambda_1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -\lambda_1 & \lambda_1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -\lambda_1 & \lambda_1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -\lambda_1 & \lambda_1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -\lambda_1 & \lambda_1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & -\lambda_1 & \lambda_1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -\lambda_1 & \lambda_1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -\lambda_1 \end{pmatrix}, \tag{5.2.16}$$

and

$$\boldsymbol{S}_2 = \begin{pmatrix} -\lambda_2 & \lambda_2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -\lambda_2 & \lambda_2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -\lambda_2 & \lambda_2 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -\lambda_2 & \lambda_2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -\lambda_2 & \lambda_2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -\lambda_2 & \lambda_2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -\lambda_2 & \lambda_2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & -\lambda_2 & \lambda_2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -\lambda_2 & \lambda_2 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -\lambda_2 \end{pmatrix}, \tag{5.2.17}$$

with

$$\lambda_1 = 1.168239125, \qquad \lambda_2 = 18.83176088. \tag{5.2.18}$$

$\boldsymbol{S}_0$ and $\boldsymbol{\alpha}$ are given by

$$\boldsymbol{S}^0 = (0,0,0,0,0,0,0,0,0,\lambda_1,0,0,0,0,0,0,0,0,0,\lambda_2)', \tag{5.2.19}$$

and

$$\boldsymbol{\alpha} = (p_1,0,0,0,0,0,0,0,0,0,p_2,0,0,0,0,0,0,0,0,0)' \tag{5.2.20}$$

where

$$p_1 = 0.05841196, \qquad p_2 = 0.94158804. \tag{5.2.21}$$

We note that with this choice of phase distribution, $-\boldsymbol{\alpha}\boldsymbol{S}^{-1}\boldsymbol{e} = 1$ and hence the queue is stable provided $\lambda < 1$. We present some results in Table 5.2.1 for $\theta = 0.25$. We see from Table 5.2.1 that for a fixed value of λ, the value of L remains constant as the value of k increases. We will see later that in other models the value of L can decrease as the value of k increases.

λ	k	L	λ	k	L	λ	k	L
0.5	1	7	0.7	1	8	0.9	1	9
0.5	10	7	0.7	10	8	0.9	10	9
0.5	100	7	0.7	100	8	0.9	100	9

Table 5.2.1: Value of L when Algorithm 3.4.4 is used to compute $\boldsymbol{R}_k$ for the M/PH/1 retrial queue with $\theta = 0.25$ and phase type distribution given by equation (5.2.14).

5.2.2 M/M/c retrial queue

Neuts and Rao [74] considered a queueing system in which customers arrive according to a Poisson stream, to a system where there are c servers. On arrival, each customer goes directly into orbit and begins generating requests for service according to a Poisson process with rate θ. The customers keep generating requests for service until they are successful. Let the arrival stream have rate λ and suppose the service time for each customer is negative exponentially distributed with mean $1/\mu$. This system can be modelled as a continuous time LDQBD on the state space $\mathcal{S} = \{(k, j) : k \geq 0, 0 \leq j \leq c\}$ where k is the number of customers in orbit and j is the number of busy servers. The matrices $\boldsymbol{Q}_0^{(k)}$, $\boldsymbol{Q}_1^{(k)}$, and $\boldsymbol{Q}_2^{(k)}$, are given by

$$\boldsymbol{Q}_0^{(k)} = \lambda \boldsymbol{I}, \tag{5.2.22}$$

$$\boldsymbol{Q}_1^{(k)} = \begin{pmatrix} -k\theta - \lambda & & & & \\ \mu & -k\theta - \mu - \lambda & & & \\ & 2\mu & -k\theta - 2\mu - \lambda & & \\ & & \ddots & \ddots & \\ & & & c\mu & -c\mu - \lambda \end{pmatrix}, \tag{5.2.23}$$

$$\boldsymbol{Q}_2^{(k)} = \begin{pmatrix} 0 & k\theta & & & \\ & 0 & k\theta & & \\ & & \ddots & \ddots & \\ & & & 0 & k\theta \\ & & & & 0 \end{pmatrix}. \tag{5.2.24}$$

Neuts and Rao analysed this queue by setting $\boldsymbol{Q}_1^{(k)} = \boldsymbol{Q}_1^{(M)}$, and $\boldsymbol{Q}_2^{(k)} = \boldsymbol{Q}_2^{(M)}\ \forall\ k \geq M$ where M is essentially chosen by a trial and error approach. Neuts and Rao did not give an analytical form for the stationary distribution for this system but given the theory in chapter 4 this is now possible. We note that Falin [23] showed that the queue is stable if and only if $\rho = \lambda/c\mu < 1$.

Firstly, note that Conditions 1 and 2 presented in Section 4.4.4 do not hold. We therefore choose to analyse the process using Algorithm 4.4.1. We made an initial guess for K^* and then we increased K^* until we found a value that was sufficiently large. In order to determine how much we should increase K^* by at each iteration, we considered the size of $\boldsymbol{x}_{K^*}(K^*)$. For larger values of $\boldsymbol{x}_{K^*}(K^*)$ we made larger increases in the value of K^*. We used our algorithm to replicate some of the results presented by Neuts and Rao. The performance measures we computed were L_i, V_i and P_b which are the mean and variance of the number of customers in orbit and the probability that all servers are busy respectively. We set $\lambda = 1$ and computed the performance measures for various values of c, θ and ρ. The results are presented in Table 5.2.2 together with the value of L required in Algorithm 3.4.4.

c	θ	ρ	K^*	L_i	V_i	P_b	L
7	0.05	0.5	50	21.084	22.436	0.049	5
7	0.01	0.5	170	105.034	111.037	0.047	5
4	0.5	0.9	140	13.777	133.190	0.727	6
4	0.05	0.9	50	71.455	554.233	0.693	6
4	0.01	0.9	600	326.978	2418.641	0.688	7
3	0.1	0.95	400	90.828	1653.921	0.866	8
3	0.05	0.95	500	163.922	2929.623	0.864	8

Table 5.2.2: Some performance measures for the M/M/c retrial queue.

The results in Table 5.2.2 agree with the results in Tables 3 and 4 of Neuts and Rao [74]. We note that the value of L increases as ρ increases which is what we expect given the physical interpretation of Algorithm 3.4.4.

We also investigated the value of L required in Algorithm 3.4.4 to compute $\boldsymbol{R}_k$ for different values of k but with all parameters kept constant. The results are presented in Table 5.2.3 for the parameter values $c = 7$, $\theta = 0.05$ and $\lambda = 1$. From Table 5.2.3 we see that if we keep all parameter values constant, the value of L decreases as k increases. These results should be compared with the results in Table 5.2.1 which show that for the M/PH/1 retrial queue, the value of L remains constant when k increases.

ρ	k	L	ρ	k	L	ρ	k	L
0.5	1	6	0.7	1	7	0.9	1	9
0.5	10	6	0.7	10	6	0.9	10	8
0.5	100	5	0.7	100	6	0.9	100	7

Table 5.2.3: Value of L when Algorithm 3.4.4 is used to compute $\boldsymbol{R}_k$ for the M/M/c retrial queue with $c = 7$, $\theta = 0.05$ and $\lambda = 1$.

Variations of the M/M/c retrial queue

Neuts and Rao considered three variations of the $M/M/c$ retrial model presented above. In the first variation customers arriving to the system move immediately into service if a customer is available and they move into the pool only if all servers are busy. This system can be modelled by a continuous time LDQBD on the same state space as we considered before with

$$\boldsymbol{Q}_0^{(k)} = \begin{pmatrix} 0 & & & & \\ & 0 & & & \\ & & \ddots & & \\ & & & 0 & \\ & & & & \lambda \end{pmatrix}, \tag{5.2.25}$$

$$\boldsymbol{Q}_1^{(k)} = \begin{pmatrix} -k\theta-\lambda & \lambda & & & \\ \mu & -k\theta-\mu-\lambda & \lambda & & \\ & 2\mu & -k\theta-2\mu-\lambda & \ddots & \\ & & \ddots & \ddots & \lambda \\ & & & c\mu & -c\mu-\lambda \end{pmatrix} \tag{5.2.26}$$

and $\boldsymbol{Q}_2^{(k)}$ is given by (5.2.24).

In the second variation after each unsuccessful retrial customers leave the orbiting pool with probability α. For this LDQBD we have

$$\boldsymbol{Q}_1^{(k)} = \begin{pmatrix} -k\theta-\lambda & & & & \\ \mu & -k\theta-\mu-\lambda & & & \\ & 2\mu & -k\theta-2\mu-\lambda & & \\ & & \ddots & \ddots & \\ & & & c\mu & -c\mu-k\alpha\theta-\lambda \end{pmatrix}, \tag{5.2.27}$$

$$\boldsymbol{Q}_2^{(k)} = \begin{pmatrix} 0 & k\theta & & & \\ & 0 & k\theta & & \\ & & \ddots & \ddots & \\ & & & 0 & k\theta \\ & & & & k\alpha\theta \end{pmatrix}. \tag{5.2.28}$$

and $\boldsymbol{Q}_0^{(k)}$ is given by (5.2.22).

In the third variation, as well as retrying at rate θ, customers in orbit renege at rate

δ. For this process we have

$$Q_1^{(k)} = \begin{pmatrix} -k(\theta+\delta)-\lambda & & & \\ \mu & -k(\theta+\delta)-\mu-\lambda & & \\ & \ddots & \ddots & \\ & & c\mu & -c\mu-k\delta-\lambda \end{pmatrix}, \tag{5.2.29}$$

$$Q_2^{(k)} = \begin{pmatrix} k\delta & k\theta & & & \\ & k\delta & k\theta & & \\ & & \ddots & \ddots & \\ & & & k\delta & k\theta \\ & & & & k\delta \end{pmatrix} \tag{5.2.30}$$

and $Q_0^{(k)}$ is given by (5.2.22).

All the variations on the original model can be analysed in a similar way to how we analysed the original model, although we do not present any numerical results here.

5.2.3 M/PH/1 retrial queue with Markov modulated Poisson process retrials

In the following two sections we show how to model two more retrial queues, although again we do not present any numerical results. Diamond and Alpha [20] proposed that the M/PH/1 retrial queue with a Markov modulated Poisson process retrial process could be analysed via matrix-analytic techniques. We show in this section that such a system can be modelled as a LDQBD and analysed via the algorithms in Chapter 4.

Consider a queue in which customers arrive to the system according to a Poisson stream with rate λ and the service time for each customer is of phase type with representation $(\boldsymbol{S}, \boldsymbol{\alpha})$. Arriving customers move into service if the server is free and if it is busy they go into orbit. We assume that the rate at which orbiting customers retry depends on the state of a separate Markov process $Y(t)$ on a state space $\{1, \ldots, N\}$ with generator $\hat{\boldsymbol{Q}}$. When $Y(t)$ is in state ℓ, orbiting customers are assumed to retry at a rate θ_ℓ. This queue can be modelled as a LDQBD on the state space $\{-1\} \cup \{(0,j) : 1 \leq j \leq M\} \cup \{(k,j,\ell) : k \geq 1, 0 \leq j \leq M, 1 \leq \ell \leq N\}$ where k represents the number of customers in orbit, j represents the phase of service and ℓ represents the phase of the Markov modulated Poisson process. The state $\{-1\}$ represents the state where there are no customers in orbit and the server is free. The states $(0, j)$ represent those states of the system where there are no customers in orbit and the server is busy. The q-matrix for the LDQBD has the same form as that given in equation (5.2.1) and the matrices $\boldsymbol{Q}_i^{(k)}$ are given by

$$\boldsymbol{Q}_1^{(-1)} = -\lambda, \tag{5.2.31}$$

$$\boldsymbol{Q}_0^{(-1)} = \lambda\boldsymbol{\alpha}', \tag{5.2.32}$$

$$\boldsymbol{Q}_2^{(0)} = \boldsymbol{S}^0, \tag{5.2.33}$$

$$\boldsymbol{Q}_1^{(0)} = \boldsymbol{S} - \lambda\boldsymbol{I}_m, \tag{5.2.34}$$

$$\boldsymbol{Q}_0^{(0)} = \begin{pmatrix} \boldsymbol{0} & \lambda\boldsymbol{I}_m \end{pmatrix}, \tag{5.2.35}$$

$$\boldsymbol{Q}_2^{(k)} = \begin{pmatrix} \boldsymbol{0} & k\Delta(\boldsymbol{\theta}) \otimes \boldsymbol{\alpha} \\ \boldsymbol{0} & \boldsymbol{0} \end{pmatrix} \qquad k \geq 1, \tag{5.2.36}$$

$$\boldsymbol{Q}_1^{(k)} = \begin{pmatrix} \hat{\boldsymbol{Q}} - \lambda\boldsymbol{I}_n - k\Delta(\boldsymbol{\theta}) & \lambda\boldsymbol{\alpha}' \otimes \boldsymbol{I}_n \\ \boldsymbol{S}^0 \otimes \boldsymbol{I}_n & (\hat{\boldsymbol{Q}} - \lambda\boldsymbol{I}_n) \otimes \boldsymbol{I}_m + \boldsymbol{S} \otimes \boldsymbol{I}_n \end{pmatrix} \qquad k \geq 1, \tag{5.2.37}$$

$$\boldsymbol{Q}_0^{(k)} = \begin{pmatrix} \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{0} & \lambda\boldsymbol{I}_n \end{pmatrix} \qquad k \geq 1, \tag{5.2.38}$$

$$\Delta(\boldsymbol{\theta}) = diag(\theta_1, \ldots, \theta_N) \tag{5.2.39}$$

where $\otimes$ is the Kronecker product. (See Neuts [72, page 53] for a definition.)

It can be shown that this queue is positive recurrent if $\lambda(-\boldsymbol{\alpha}\boldsymbol{S}^{-1}\boldsymbol{e}) < 1$ which is the same condition as required for the M/PH/1 retrial queue. Again we adopted a trial and error approach to find a value of K^* for this process. The qualitative behaviour of this queue is similar to that of the M/PH/1 retrial queue which is to be expected.

5.2.4 PH/PH/1 retrial queue

The $PH/PH/1$ retrial queue can also be modelled as a LDQBD. Here we give a brief description of this queue.

Suppose customers arrive to a single server queue according to a phase type distribution with representation $(\boldsymbol{T}, \boldsymbol{\alpha})$. Customers move directly into service if the server is free and they go into orbit if the server is busy. The service time of each customer is of phase type with representation $(\boldsymbol{S}, \boldsymbol{\beta})$ and customers in orbit retry according to a Poisson process with rate θ. This queue can be modelled as a LDQBD on the state space $\{(-1,\ell) : 1 \leq \ell \leq N\} \cup \{(0,j,\ell) : 1 \leq j \leq M, 1 \leq \ell \leq N\} \cup \{(k,j,\ell) : k \geq 1, 0 \leq j \leq M, 1 \leq \ell \leq N\}$ where ℓ represents the phase of the arrival process, j represents the phase of the service process, and k represents the number in orbit. The states $(-1,\ell)$ represent those states of the system when there are no customers in orbit and the server is free and the states $(0,j,\ell)$ represent those states of the system where there are no customers in orbit and the server is busy. The q-matrix for the LDQBD has the same form as that given in equation (5.2.1) and the matrices $\boldsymbol{Q}_i^{(k)}$ are given by

$$\boldsymbol{Q}_1^{(-1)} = \boldsymbol{T}, \tag{5.2.40}$$

$$\boldsymbol{Q}_0^{(-1)} = \boldsymbol{T}^0 \cdot \boldsymbol{\alpha} \otimes \boldsymbol{\beta}', \tag{5.2.41}$$

$$\boldsymbol{Q}_2^{(0)} = \boldsymbol{S}^0 \otimes \boldsymbol{I}_n, \tag{5.2.42}$$

$$\boldsymbol{Q}_1^{(0)} = \boldsymbol{I}_m \otimes \boldsymbol{T} + \boldsymbol{S} \otimes \boldsymbol{I}_n, \tag{5.2.43}$$

$$\boldsymbol{Q}_0^{(0)} = \begin{pmatrix} \boldsymbol{0} & \boldsymbol{I}_m \otimes \boldsymbol{T}^0 \cdot \boldsymbol{\alpha} \end{pmatrix}, \tag{5.2.44}$$

$$\boldsymbol{Q}_2^{(1)} = \begin{pmatrix} \theta\boldsymbol{\beta}' \otimes \boldsymbol{I}_n \\ 0 \end{pmatrix}, \tag{5.2.45}$$

$$Q_1^{(k)} = \begin{pmatrix} \boldsymbol{T} - k\theta \boldsymbol{I}_n & \boldsymbol{T}^0 \cdot \boldsymbol{\alpha} \otimes \boldsymbol{\beta}' \\ \boldsymbol{S}^0 \otimes \boldsymbol{I}_n & \boldsymbol{S} \otimes \boldsymbol{I}_n + \boldsymbol{I}_m \otimes \boldsymbol{T} \end{pmatrix} \qquad k \geq 1, \tag{5.2.46}$$

$$Q_0^{(k)} = \begin{pmatrix} 0 & 0 \\ 0 & \boldsymbol{I}_m \otimes \boldsymbol{T}^0 \cdot \boldsymbol{\alpha} \end{pmatrix} \qquad k \geq 1, \tag{5.2.47}$$

$$Q_2^{(k)} = \begin{pmatrix} 0 & k\theta\boldsymbol{\beta} \otimes \boldsymbol{I}_n \\ 0 & 0 \end{pmatrix} \qquad k \geq 2. \tag{5.2.48}$$

We note that there are NM phases at each level. This clearly places a restriction on the values of N and M that we can consider. This dimensionality problem frequently arises when adopting the matrix-analytic approach. Recently Latouche and Ramaswami [57] showed how a PH/PH/1 queue can be modelled as a QBD with $N + M$ phases at each level where N and M are the dimensions of the phase type processes. This of course means that a much wider range of values for N and M can be considered. A possible area for further research is to attempt to extend the modelling techniques of Latouche and Ramaswami to the case of a PH/PH/1 retrial queue.

5.3 M/M/∞ queue in a random environment

In this section we consider an M/M/∞ queue in a random environment. This is an infinite server queue whose arrival rate and service rate are determined by the state of a M-state irreducible Markov process $E(t)$. If $E(t)$ is in state j and there are k customers in the queue then the arrival and service rates for the queue are λ_j and $k\mu_j$ respectively. If $E(t)$ changes from state j to state ℓ, the arrival and service rates change to λ_ℓ and $k\mu_\ell$ instantaneously.

Neuts [72, page 274] gave a brief discussion of the M/M/∞ queue in a random environment. He suggested that methods similar to those employed in Neuts and Ramaswami [82] could be used to find the moments of the queue lengths for the various states of $E(t)$. Neuts did not give an analytical expression for the stationary distribution but he observed that a numerical approximation for the stationary distribution could be obtained by solving a truncated version of the global balance equations.

The M/M/∞ queue in a random environment can be modelled as a LDQBD on the state space $\{(k,j) : k \geq 0, 1 \leq j \leq M\}$ where the state (k,j) represents k customers in the queue and $E(t)$ being in state j. M is the number of possible states for $E(t)$. Hence the stationary distribution has the simple analytic form given by equation (4.2.1). If $E(t)$ has q-matrix $\hat{\boldsymbol{Q}}$ and we set $\boldsymbol{\lambda} = (\lambda_1, \ldots, \lambda_M)$, and $\boldsymbol{\mu} = (\mu_1, \ldots, \mu_M)$ then the matrices $\boldsymbol{Q}_0^{(k)}$, $\boldsymbol{Q}_1^{(k)}$ and $\boldsymbol{Q}_2^{(k)}$ are given by

$$\boldsymbol{Q}_0^{(k)} = \boldsymbol{D}(\boldsymbol{\lambda}) \qquad k \geq 0, \tag{5.3.1}$$

$$\boldsymbol{Q}_1^{(k)} = \hat{\boldsymbol{Q}} - \boldsymbol{D}(\boldsymbol{\lambda} + k\boldsymbol{\mu}) \qquad k \geq 0, \tag{5.3.2}$$

$$\boldsymbol{Q}_2^{(k)} = \boldsymbol{D}(k\boldsymbol{\mu}) \qquad k \geq 1, \tag{5.3.3}$$

where $\boldsymbol{D}(\boldsymbol{v})$ is a diagonal matrix with diagonal elements equal to the elements of the vector $\boldsymbol{v}$.

As mentioned in Chapter 4, in order to compute the stationary distribution we must determine a value for K^*. Condition 1 holds for this LDQBD provided $\mu_i > 0 \ \forall i$, in which case we can apply Algorithm 4.4.4. As suggested in Algorithm 4.4.4 we consider the birth-death process $\bar{Y}(t)$ on the non negative integers with transition rates

$$\bar{q}(k, k+1) = [\boldsymbol{\lambda}]_{max} \qquad k \geq 0, \tag{5.3.4}$$

$$\bar{q}(k, k-1) = k[\boldsymbol{\mu}]_{min} \qquad k \geq 1. \tag{5.3.5}$$

It is clear that $\bar{Y}(t)$ is positive recurrent. If we set $\lambda^* = [\boldsymbol{\lambda}]_{max}$, $\mu^* = [\boldsymbol{\mu}]_{min}$ and $a^* = \lambda^*/\mu^*$ then $\bar{Y}(t)$ has stationary distribution

$$\bar{\ell}_k = e^{-a^*} \frac{(a^*)^k}{k!} \qquad k \geq 0. \tag{5.3.6}$$

We choose K^* to be the minimum value of K such that

$$\sum_{k=K+1}^{\infty} \bar{\ell}_k < \epsilon \tag{5.3.7}$$

for some tolerance ϵ. Readers experienced in computational probability will be familiar with this type of computation. However, for completeness we give a brief description of how this can be done.

Firstly observe that equation (5.3.7) is equivalent to

$$\sum_{k=0}^{K} \bar{\ell}_k > 1 - \epsilon \tag{5.3.8}$$

so we want the smallest value of K such that (5.3.8) holds. We cannot use (5.3.6) to compute $\sum_{k=0}^{K} \bar{\ell}_k$ directly since numerical difficulties arise if we attempt to compute e^{-a^*} for large values of a^*. To overcome these difficulties, we observe that provided $a > 1$, the distribution $\{\bar{\ell}_k, k \geq 0\}$ has its maximum value at $m = \lfloor a^* \rfloor$ and decreases monotonically either side of m. Hence K^* can be found as follows. Firstly consider sets of the form $\{m - L \leq k \leq m + L\}$ for $0 \leq L \leq m$ and calculate a normalised invariant measure over the set. Since we are dealing with a finite state space, this is easy to do and in particular it does not require us to calculate e^{-a^*}. Furthermore, if we have the invariant measure for the set of states $\{m - L \leq k \leq m + L\}$, then it does not require much additional effort to compute the invariant measure for the set of states $\{m - L - 1 \leq k \leq m + L + 1\}$.

The normalised invariant measures for the set $\{m - L \leq k \leq m + L\}$ will be greater than the stationary probabilities $\bar{\ell}_k$. Hence if we can find an $0 \leq L \leq m$ such that the normalised invariant measure is sufficiently small for the state $m + L$ then we can take $K^* = m + L$. The justification for doing this is that if the normalised invariant measure is small then, $\bar{\ell}_{m+L}$ is small. Since $\bar{\ell}_k$ is decreasing for all $k > m + L$, we can also say that the probability of being in the set $\{k : k \geq m + L + 1\}$ is small. That is, the condition (5.3.7) is likely to hold.

If we cannot find an L such that the above conditions hold, then we consider sets of the form $\{0, 1, \ldots, 2m + L\}$ for $L \geq 0$. Again we calculate a normalised invariant measure over these sets for increasing values of L. We set $K^* = 2m + L$ for the smallest value of L such that the normalised invariant measure is sufficiently small for the state $2m + L$.

In Chapter 4, Section 4.4.4 we showed that provided Condition 1 holds, it is possible to construct a LDQBD that stochastically dominates the original LDQBD. If we put

$$[\bar{Q}_0^{(0)}]_{ij} = [Q^{(0)}]_{ij}, \tag{5.3.9}$$

$$[\bar{Q}_0^{(k)}]_{ij} = [\lambda]_{max}, \tag{5.3.10}$$

$$[\bar{Q}_2^{(1)}]_{ij} = 0, \tag{5.3.11}$$

$$[\bar{Q}_2^{(k)}]_{ij} = \frac{(k-1)[\mu]_{min}}{M}, \tag{5.3.12}$$

$$[\bar{Q}_1^{(k)}]_{ij} = [Q_1^{(k)}]_{ij}, \tag{5.3.13}$$

then the resulting LDQBD $\bar{X}(t)$ on the state space $\{(k,j) : k \geq 1, 1 \leq j \leq M\}$ stochastically dominates the original LDQBD. Now the marginal stationary probabilities ℓ_k^d over the levels for $\bar{X}(t)$ are given by

$$\ell_k^d = D\frac{(M[\lambda]_{max}/[\mu]_{min})^{k-1}}{(k-1)!} \qquad k \geq 1 \tag{5.3.14}$$

where D is a normalising constant.

Hence a possible way to choose K^* is to set it equal to the smallest value of K such that

$$\sum_{k=K+1}^{\infty} \ell_k^d < \epsilon. \tag{5.3.15}$$

However it is clear that this will provide a much larger value for K^* than the previous method. We will see that this is indeed the case when we consider a numerical example.

The first two numerical examples we present are taken from Bright and Taylor [12]. Consider the case where $\hat{Q}$ is given by

$$\hat{Q} = \begin{pmatrix} -8 & 8 \\ 5 & -5 \end{pmatrix} \tag{5.3.16}$$

and μ and λ are given by $\mu = (2,1)$ and $\lambda = (40,5)$. Since $\mu_i > 0, i = 1,2$ Condition 1 holds and we can use Algorithm 4.4.4. In order to compute K^* we chose the smallest value of K such that (5.3.8) is satisfied with $\epsilon = 10^{-10}$. We found a value of 88 for K^*.

It is clear that this LDQBD is positive recurrent so we used Truncation rule 3.4.4 with $\epsilon = 10^{-10}$ to compute $\boldsymbol{R}_{K^*-1}$. In Table 5.3.1 we give the marginal probabilities ℓ_k of being in level k. In order to see if the value $K^* = 88$ was sufficiently large, we computed the stationary distribution with $K^* = 138$. We found that the absolute value of the differences of the two results was less than 10^{-14} in all cases.

k	ℓ_k	k	ℓ_k
0	2.43033×10^{-5}	18	4.36024×10^{-2}
1	2.12737×10^{-4}	19	3.34413×10^{-2}
2	9.61911×10^{-4}	20	2.47627×10^{-2}
3	2.99837×10^{-3}	21	1.77270×10^{-2}
4	7.25313×10^{-3}	22	1.22830×10^{-2}
5	1.45280×10^{-2}	23	8.24654×10^{-3}
6	2.50976×10^{-2}	24	5.36972×10^{-3}
7	3.84463×10^{-2}	25	3.39415×10^{-3}
8	5.32737×10^{-2}	26	2.08432×10^{-3}
9	6.77654×10^{-2}	27	1.24447×10^{-3}
10	8.00201×10^{-2}	28	7.22940×10^{-4}
11	8.84792×10^{-2}	29	4.08889×10^{-4}
12	9.22315×10^{-2}	30	2.25306×10^{-4}
13	9.11276×10^{-2}	31	1.21022×10^{-4}
14	8.57099×10^{-2}	32	6.34068×10^{-5}
15	7.70106×10^{-2}	33	3.24203×10^{-5}
16	6.62927×10^{-2}	34	1.61860×10^{-5}
17	5.48054×10^{-2}	>34	$< 10^{-5}$

Table 5.3.1: The marginal probabilities ℓ_k for the M/M/∞ queue in a random environment with $\boldsymbol{Q}$ given by (5.3.16), $\boldsymbol{\mu} = (2, 1)$ and $\boldsymbol{\lambda} = (40, 5)$.

As we mentioned previously, the LDQBD with transition rates given by (5.3.9)-(5.3.13) stochastically dominates the original LDQBD. We calculated a value of K^* by considering the marginal stationary probabilities over the levels for this dominating process and we found that $K^* = 180$. As expected, this is a much larger value than the value of 88 that we obtained previously. As we mentioned in Chapter 4, this method is not an efficient way to calculate K^*.

We now alter the above example by changing $\boldsymbol{\mu}$ to $\boldsymbol{\mu} = (2, 0)$. This corresponds to a situation where the server goes on a vacation or shuts down whenever $E(t)$ is in state two. In this situation Condition 1 fails to hold so we need to use another method to determine a value of K^*. In order to make an initial guess for K^* we used the results for the previous example. Since we have set $\mu_2 = 0$, we expect that we will require a value of K^* at least as high as in the previous example. For this reason we guessed an initial value of $K^* = 100$ which is slightly higher than the value of K^* for the previous example. We also set $\epsilon = 10^{-10}$ and we found that our initial guess proved to be a sufficiently large choice. The marginal probabilities ℓ_k of being in level k are presented in Table 5.3.2. For this example we haven't proved that the process is actually positive recurrent although the infinite server capacity in phase one certainly suggests that it will be. The numerical results in Table 5.3.2 also indicate positive recurrence.

The following example is similar to an example presented in Neuts [72]. Neuts modelled the M/M/1 queue in a random environment as a QBD. In a numerical example Neuts considered the random environment specified by

$$\boldsymbol{Q} = \begin{pmatrix} -1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -1 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & -1 \end{pmatrix}, \tag{5.3.17}$$

with

$$\boldsymbol{\mu} = (6.25, 6.25, 6.25, 6.25, 6.25, 6.25, 6.25, 6.25), \tag{5.3.18}$$

and

$$\boldsymbol{\lambda} = (21, 5, 5, 5, 1, 1, 1, 1). \tag{5.3.19}$$

Here the process $E(t)$ moves sequentially through each of the environments and the mean sojourn time in each environment is one. A salient feature of the vectors $\boldsymbol{\mu}$ and $\boldsymbol{\lambda}$ is that

k	ℓ_k	k	ℓ_k
0 to 5	$< 10^{-5}$	30	4.07063×10^{-2}
6	1.54870×10^{-5}	31	3.34230×10^{-2}
7	4.99438×10^{-5}	32	2.67954×10^{-2}
8	1.41353×10^{-4}	33	2.10015×10^{-2}
9	3.56725×10^{-4}	34	1.61113×10^{-2}
10	8.12863×10^{-4}	35	1.21114×10^{-2}
11	1.68958×10^{-3}	36	8.93141×10^{-3}
12	3.23061×10^{-3}	37	6.46767×10^{-3}
13	5.72303×10^{-3}	38	4.60375×10^{-3}
14	9.45034×10^{-3}	39	3.22418×10^{-3}
15	1.46231×10^{-2}	40	2.22364×10^{-3}
16	2.13018×10^{-2}	41	1.51153×10^{-3}
17	2.93325×10^{-2}	42	1.01352×10^{-3}
18	3.83202×10^{-2}	43	6.70883×10^{-4}
19	4.76513×10^{-2}	44	4.38713×10^{-4}
20	5.65699×10^{-2}	45	2.83620×10^{-4}
21	6.42890×10^{-2}	46	1.81386×10^{-4}
22	7.01145×10^{-2}	47	1.14829×10^{-4}
23	7.35521×10^{-2}	48	7.20016×10^{-5}
24	7.43740×10^{-2}	49	4.47413×10^{-5}
25	7.26348×10^{-2}	50	2.75666×10^{-5}
26	6.86384×10^{-2}	51	1.68492×10^{-5}
27	6.28697×10^{-2}	52	1.02210×10^{-5}
28	5.59082×10^{-2}	> 53	$< 10^{-5}$
29	4.83435×10^{-2}		

Table 5.3.2: The marginal probabilities ℓ_k for the M/M/∞ queue in a random environment with $\boldsymbol{Q}$ given by (5.3.16), $\boldsymbol{\mu} = (2, 0)$ and $\boldsymbol{\lambda} = (40, 5)$.

for phase one the arrival rate greatly exceeds the service rate and hence there will be a buildup of customers in that phase. Of course this buildup will be less significant in the infinite server case.

We consider the M/M/∞ queue in a random environment with $\boldsymbol{Q}$ and $\boldsymbol{\mu}$ given by (5.3.17) and (5.3.18) respectively and we set

$$\boldsymbol{\lambda} = (210, 50, 50, 50, 10, 10, 10, 10). \tag{5.3.20}$$

Since every element of $\boldsymbol{\mu}$ is positive, Condition 1 holds for this process and we can therefore use the method described above to determine a value for K^*. We found a value of 77 for K^*. Using this value of K^*, we computed the marginal probabilities ℓ_k of being in level k. The results are presented in Table 5.3.3.

k	ℓ_k	k	ℓ_k
0	5.83838×10^{-1}	13	3.19390×10^{-6}
1	2.27254×10^{-1}	14	7.54001×10^{-7}
2	9.23815×10^{-2}	15	1.66385×10^{-7}
3	4.50569×10^{-2}	16	3.44647×10^{-8}
4	2.50998×10^{-2}	17	6.72620×10^{-9}
5	1.40063×10^{-2}	18	1.24090×10^{-9}
6	7.13977×10^{-3}	19	2.17055×10^{-10}
7	3.23404×10^{-3}	20	3.60933×10^{-11}
8	1.30266×10^{-3}	21	5.71951×10^{-12}
9	4.70512×10^{-4}	22	8.65613×10^{-13}
10	1.53806×10^{-4}	23	1.25370×10^{-13}
11	4.58865×10^{-5}	24	1.74089×10^{-14}
12	1.25860×10^{-5}	> 24	$< 10^{-14}$

Table 5.3.3: The marginal probabilities ℓ_k for the M/M/∞ queue in a random environment with $\boldsymbol{Q}$ given by (5.3.17), $\boldsymbol{\mu}$ given by (5.3.18) and $\boldsymbol{\lambda}$ given by (5.3.20).

We considered two performance measures, both of which are conditional on the phase. Table 5.3.4 gives the mean conditional queue lengths and the probability of the queue being empty.

Phase	Mean	P[empty]
1	3.64832×10^{-1}	1.04784×10^{-2}
2	1.36528×10^{-1}	4.63677×10^{-2}
3	1.05038×10^{-1}	5.44649×10^{-2}
4	1.00695×10^{-1}	5.58998×10^{-2}
5	3.11303×10^{-2}	9.84669×10^{-2}
6	2.15352×10^{-2}	1.05321×10^{-1}
7	2.02118×10^{-2}	1.06346×10^{-1}
8	2.00292×10^{-2}	1.06494×10^{-1}

Table 5.3.4: The mean conditional queue lengths and the conditional probability of the queue being empty for the M/M/∞ queue in a random environment with $\boldsymbol{Q}$ given by (5.3.17), $\boldsymbol{\mu}$ given by (5.3.18) and $\boldsymbol{\lambda}$ given by (5.3.20).

5.4 PH/M/∞ queue

The PH/M/∞ queue is an infinite server queue with a Markovian service time distribution and a phase type interarrival time distribution. This queue can be modelled as a LDQBD on the state space $\{(k, j) : k \geq 0, 1 \leq j \leq M\}$ where the state (k, j) represents k customers being present and the interarrival time process being in phase j. M is the dimension of the phase type distribution.

We considered the phase type distribution specified by equation (5.2.14). If we let each customer be served at rate μ then the matrices $\boldsymbol{Q}_0^{(k)}$, $\boldsymbol{Q}_1^{(k)}$ and $\boldsymbol{Q}_2^{(k)}$ are given by

$$\boldsymbol{Q}_0^{(k)} = \boldsymbol{S}^0 \cdot \boldsymbol{\alpha}', \tag{5.4.1}$$

$$\boldsymbol{Q}_1^{(k)} = \begin{pmatrix} \boldsymbol{S}_1 & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{S}_2 \end{pmatrix} - k\mu \boldsymbol{I}, \tag{5.4.2}$$

$$\boldsymbol{Q}_2^{(k)} = k\mu \boldsymbol{I} \tag{5.4.3}$$

where $\boldsymbol{S}_0$, $\boldsymbol{\alpha}$, $\boldsymbol{S}_1$ and $\boldsymbol{S}_2$ are given by equations (5.2.19), (5.2.20), (5.2.16) and (5.2.17) respectively.

Condition 1 holds for this process so we used Algorithm 4.4.4 to compute the stationary distribution. With $\mu = 0.5$ we found $K^* = 84$. The marginal probabilities ℓ_k are presented in Table 5.4.1.

k	ℓ_k	k	ℓ_k
0	2.94726×10^{-1}	10	4.00557×10^{-5}
1	1.68517×10^{-1}	11	4.09573×10^{-6}
2	1.57537×10^{-1}	12	3.30731×10^{-7}
3	1.55108×10^{-1}	13	2.12891×10^{-8}
4	1.19684×10^{-1}	14	1.10142×10^{-9}
5	6.70135×10^{-2}	15	4.61353×10^{-11}
6	2.71629×10^{-2}	16	1.57479×10^{-12}
7	8.09367×10^{-3}	17	4.40641×10^{-14}
8	1.80460×10^{-3}	> 17	$< 10^{-14}$
9	3.06062×10^{-4}		

Table 5.4.1: The marginal probabilities ℓ_k for the PH/M/∞ queue with phase type distribution given by equation (5.2.14) and $\mu = 0.5$.

We also investigated the change in the value of L in Algorithm 3.4.4 when computing $\boldsymbol{R}_k$ for different values of k. The results are presented in Table 5.4.2 for $\mu = 0.5$ and in Table 5.4.3 for $\mu = 0.1$. For $\mu = 0.1$ we obtained a value of 280 for K^*.

k	L
0-10	4
11-36	3
37-84($= K^*$)	2

Table 5.4.2: Value of L when Algorithm 3.4.4 is used to compute $\boldsymbol{R}_k$ for the PH/M/∞ queue with phase type distribution given by equation (5.2.14) and $\mu = 0.5$.

From the results in Table 5.4.2 we see that the value of L decreases as k increases. We saw in Chapter 4 that to calculate $\boldsymbol{R}_k$ requires the calculation of $4.2^L - L - 3$ inverses. So for example, to compute $\boldsymbol{R}_0$ requires the calculation of approximately $4.2^4 = 64$ inverses and to compute $\boldsymbol{R}_{K^*} = \boldsymbol{R}_{84}$ requires approximately $4.2^2 = 16$ inverses. We observe similar result in Table 5.4.3 for $\mu = 0.1$. This observation suggests that if we wish to calculate say $\boldsymbol{R}_{k1}$, it may be more efficient to use Algorithm 3.4.4 to compute $\boldsymbol{R}_{k2}$ for some $k2 > k1$ and then to use the recursive relationship (4.3.15) to obtain $\boldsymbol{R}_{k2}$. For example, suppose we required $L = 7$ to calculate $\boldsymbol{R}_{k1}$, for some value k_1. This

k	L
0-9	7
10-29	6
30-66	5
67-189	4
190-280($=K^*$)	3

Table 5.4.3: Value of L when Algorithm 3.4.4 is used to compute $\boldsymbol{R}_k$ for the PH/M/∞ queue with phase type distribution given by equation (5.2.14) and $\mu = 0.1$.

requires approximately $4.2^7 = 512$ inverses. Now suppose in order to calculate $\boldsymbol{R}_{k1+40}$, we required only $L = 4$. This requires approximately $4.2^4 = 64$ inverses. If we use the recursion (4.3.15) to calculate $\boldsymbol{R}_{k1}$, we need to calculate a further 40 inverses giving a total of 104 inverses. The second method is clearly more efficient.

We saw in Table 5.2.1 for the M/PH/1 retrial queue that the value of L remains constant for increasing k. For the M/M/c retrial queue the value of L decreased as k increased, see Table 5.2.3, but the decrease was not as pronounced as for the PH/M/∞ queue. These results can be explained by considering the forms of the matrices $\boldsymbol{Q}_i^{(k)}$. For the PH/M/∞ queue, $\boldsymbol{Q}_2^{(k)} = k\mu\boldsymbol{I}$ and the matrices $\boldsymbol{Q}_1^{(k)}$ and $\boldsymbol{Q}_0^{(k)}$ are essentially independent of k. Hence, at any phase at level k, the process moves down to level $k-1$ at a rate $k\mu$. Therefore as k increases, the down rates at every phase increase and hence if the process starts in level k, sample paths that reach levels much higher than level k are less likely. For the M/M/c retrial queue, the process moves down to level $k-1$ at rate $k\theta$ from every phase at level k except phase c. The fact that the transition rate for phase c in level k to level $k-1$ is zero results in a less pronounced decrease in L as k increases than for the PH/M/∞ queue. For the M/PH/1 retrial queue, $\boldsymbol{Q}_2^{(k)}$ is given by (5.2.10) which depends upon k. Again the matrices $\boldsymbol{Q}_1^{(k)}$ and $\boldsymbol{Q}_0^{(k)}$ are essentially independent of k. From the form of (5.2.10) we see that the process moves down from level k to level $k-1$ according to a positive rate only if the process is in phase zero. Hence even if k increases, if the process starts in level k, long sample paths are still likely since although the down rate for phase zero increases, the down rates are still zero over the remaining phases.

It is clear that if we increase the traffic intensity then a larger value of K^* will be required. An interesting question to consider is will the number of iterations required to compute $\boldsymbol{R}_{K^*}$ also increase? That is, will the value of L increase? We present some results in Table 5.4.4.

μ	K^*	L
0.5	84	2
0.3	119	2
0.1	280	2
0.01	2153	2

Table 5.4.4: Values of K^* and the number of iterations required to calculate $\boldsymbol{R}_{K^*}$ for the PH/M/∞ queue with phase type distribution given by equation (5.2.14).

From Table 5.4.4 we see that as μ decreases (which corresponds to an increase in the traffic intensity), although there is a steady increase in K^*, the value of L remains constant at a value of two. Hence the same effort is required to compute $\boldsymbol{R}_{K^*}$ and the only increased computational effort results from the fact that the state space $\{(k,j) : 0 \leq k \leq K^*, 1 \leq j \leq M\}$ increases in size.

The phase type arrival process is a special case of a more general class of arrival process, namely the Markovian arrival process (MAP). Lucantoni, Meier-Hellstein and Neuts [61] introduced the MAP as a generalisation of the phase type renewal process and the Markov modulated Poisson process. An attractive feature of the MAP is that it can be used to model arrival processes where there is correlation between interarrival times. It is a simple matter to model the MAP/M/∞ queue as a LDQBD. Before we show how this can be done, we give a brief description of the MAP.

A MAP can be viewed as a Markov process $(N(t), J(t))$ on the state space $\{(n,j) : n \geq 0, 1 \leq j \leq M\}$ with q-matrix given by

$$Q = \begin{pmatrix} \boldsymbol{D}_0 & \boldsymbol{D}_1 & 0 & 0 & \cdots \\ 0 & \boldsymbol{D}_0 & \boldsymbol{D}_1 & 0 & \cdots \\ 0 & 0 & \boldsymbol{D}_0 & \boldsymbol{D}_1 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix} \qquad (5.4.4)$$

where $\boldsymbol{D}_0$ and $\boldsymbol{D}_1$ are $M \times M$ matrices. $\boldsymbol{D}_1$ has positive elements, $\boldsymbol{D}_0$ has positive off-diagonal elements and negative diagonal elements and $(\boldsymbol{D}_0 + \boldsymbol{D}_1)e = \boldsymbol{0}$. $N(t)$ represents the number of arrivals in $(0, t)$ and $J(t)$ represents the auxiliary state or phase of the MAP.

An alternative way to think of the MAP is as follows. The matrix $\boldsymbol{D} = \boldsymbol{D}_0 + \boldsymbol{D}_1$ is the q-matrix for a M-state Markov process. In this Markov process, some transitions correspond to arrival epochs for the MAP and some do not. The rate $[\boldsymbol{D}_1]_{ij}$ is the rate at which transitions occur from state i to state j with an arrival. The rate $[\boldsymbol{D}_0]_{ij}$ is the rate at which transitions occur from state i to state j without an arrival.

Lucantoni [60] generalised the MAP to the batch Markovian arrival process (BMAP) by allowing batch arrivals. Lucantoni observed that the BMAP is equivalent to Neuts' versatile arrival process (N-process) which was developed in [70]. Neuts developed his process using a matrix structure and showed that it could be viewed as a natural generalisation of the Poisson process. Since the N-process appeared before the BMAP, many results were obtained using it. However, the BMAP has a simpler notation than the N-process and hence we prefer to use the BMAP.

The MAP/M/∞ queue can be modelled as a LDQBD on the state space $\{(k, j) : k \geq 0, 1 \leq j \leq M\}$ where k is the number of customers in the queue and j is the phase of the MAP. The matrices $\boldsymbol{Q}_0^{(k)}$, $\boldsymbol{Q}_1^{(k)}$ and $\boldsymbol{Q}_2^{(k)}$ are given by

$$\boldsymbol{Q}_0^{(k)} = \boldsymbol{D}_1, \tag{5.4.5}$$

$$\boldsymbol{Q}_1^{(k)} = \boldsymbol{D}_0 - k\mu\boldsymbol{I}, \tag{5.4.6}$$

$$\boldsymbol{Q}_2^{(k)} = k\mu\boldsymbol{I}. \tag{5.4.7}$$

We have analysed some MAP/M/∞ queues using the algorithms presented in Chapter 4 and we observed similar behaviour to that for the PH/M/∞ queue. The detailed results are not presented here.

5.5 Some bivariate queueing models

In this section we consider some two dimensional queueing models that can be modelled either as QBDs or LDQBDs. By modelling these systems as LDQBDs we obtain an exact analytical expression for the stationary distribution. In certain cases, the algorithms presented in Chapter 4 can then be used to obtain numerical values for the stationary distribution.

5.5.1 The symmetric shortest queue problem

Consider a system consisting of two single server queues where the service time at both queues is the same and is exponentially distributed. Further, suppose that customers arrive to this system according to a Poisson stream and they join the shorter of the two queues. If both queues have the same length then a customer joins each queue with probability 0.5. This system is called the symmetric shortest queue system.

The symmetric shortest queue problem was first considered by Haight [33] and has since received a great deal of attention in the literature. Kingman [47] showed that if the arrival rate is λ and each server serves at rate μ then the stationary distribution exists if $\lambda/\mu < 2$. Kingman also derived an expression for the generating function of the stationary probabilities using complex function methods. Flatto and McKean [27] analysed this problem using a generating function approach and Cohen and Boxma [15] have analysed the problem in terms of Reimann-Hilbert boundary value problems. Recently, Adan, Wessels and Zijm [1] showed that the stationary distribution can be expressed as an infinite linear combination of product forms. They also gave relationships for the product forms and for the co-efficients in the linear combination. Their method did not involve generating function arguments and furthermore, their results lead to an efficient numerical algorithm.

As well as the approaches described above, there has also been a number of numerical approaches to modelling the symmetric shortest queue system. Rao and Posner [86] and Gertsbakh [30] showed that by truncating one of the state variables, it is possible to model this problem using a matrix-geometric approach. Halfin [35] used linear programming techniques to derive upper and lower bounds on the stationary probabilities and Knessl, Matkowsky, Schuss and Tier [50] derived asymptotic expressions for the stationary probabilities.

In the next section we show how the symmetric shortest queue problem can be approximately modelled as a QBD. The results we present are taken from Rao and Posner [86]. We then show how to model this system as a LDQBD.

5.5.2 The shortest queue problem modelled as a QBD

The shortest queue problem can be modelled as a continuous time Markov process on the state space $\{(k,j) : k \geq 0, j \geq 0\}$ where k represents the number of customers in queue one and j represents the number of customers in queue two. Now suppose we modify the shortest queue system by assuming that queue two can have at most M customers in it. If customers arrive when queue two has M customers present then they move into queue one regardless of how many customers are in queue one. If we choose M large enough then it is clear that this modified system will be a close approximation to the original shortest queue system. Rao and Posner [86] showed that the modified shortest queue system can be modelled as a LDQBD that is eventually level independent. The state space of this LDQBD is given by $\mathcal{S} = \{(k,j) : k \geq 0, 1 \leq j \leq M\}$ and the q-matrix has the form

$$\boldsymbol{Q} = \begin{pmatrix} \boldsymbol{A}_1^{(0)} & \boldsymbol{A}_0^{(0)} & & & & & \\ \boldsymbol{A}_2 & \boldsymbol{A}_1^{(1)} & \boldsymbol{A}_0^{(1)} & & & & \\ & \ddots & \ddots & \ddots & & & \\ & & \boldsymbol{A}_2 & \boldsymbol{A}_1^{(N)} & \boldsymbol{A}_0 & & \\ & & & \boldsymbol{A}_2 & \boldsymbol{A}_1 & \boldsymbol{A}_0 & \\ & & & & \ddots & \ddots & \ddots \end{pmatrix} \tag{5.5.1}$$

where all the matrices have dimension $M+1 \times M+1$. We refer to [86] for expressions for these matrices.

The value of M can be chosen by considering a M/M/1 queue with arrival rate λ and service rate μ. If the stationary distribution for the single server queue is given by $\pi(k)$ then we can choose M to be the smallest value of m such that $\sum_{k=m+1}^{\infty} \pi(k) < \epsilon$ for some tolerance ϵ. Once a value is chosen for M, the system can be analysed using Algorithm 3.6.2.

5.5.3 The symmetric shortest queue problem modelled as a LDQBD

In this section we present an alternative way of modelling the symmetric shortest queue system. As mentioned in the previous section, the shortest queue problem can be modelled as a continuous time Markov process on the state space $\bar{\mathcal{S}} = \{(n_1, n_2) : n_1 \geq 0, n_2 \geq$

$0\}$ where n_1 represents the number of customers in queue one and n_2 represents the number of customers in queue two. Now partition $\bar{\mathcal{S}}$ into levels $k \geq 0$ such that level k is the set of states in $\bar{\mathcal{S}}$ such that $n_1 + n_2 = k$. It is clear that $\bar{\mathcal{S}}$ is equivalent to the set $\mathcal{S} = \{(k, j) : k \geq 0, 0 \leq j \leq k\}$ where k is the total number of customers in the system and j is the total number in queue one. By partitioning the state space in this fashion, the only possible transitions out of state (k, j) are to state $(k+1, j)$ (representing an arrival to queue two), to state $(k+1, j+1)$ (representing an arrival to queue one), to state $(k-1, j)$ (representing a departure from queue two), or to $(k-1, j-1)$ (representing a departure from queue two). Given this structure, it is clear that the shortest queue problem can be modelled as a LDQBD on the state space $\mathcal{S}$ where the matrices $\boldsymbol{Q}_0^{(k)}$, $\boldsymbol{Q}_1^{(k)}$ and $\boldsymbol{Q}_2^{(k)}$ have dimensions $k+1 \times k+2$, $k+1 \times k+1$ and $k+1 \times k$ respectively. $\boldsymbol{Q}_0^{(k)}$ and $\boldsymbol{Q}_2^{(k)}$ and are given by

$$\boldsymbol{Q}_0^{(k)} = \begin{pmatrix} 0 & \lambda & & & & & & \\ & & \lambda & & & & & \\ & & & \ddots & & & & \\ & & & & \lambda & & & \\ & & & & \lambda & & & \\ & & & & & \lambda & & \\ & & & & & & \ddots & \\ & & & & & & \lambda & 0 \end{pmatrix} \qquad k \text{ odd}, \tag{5.5.2}$$

$$\boldsymbol{Q}_0^{(k)} = \begin{pmatrix} 0 & \lambda & & & & & & \\ & & \ddots & & & & & \\ & & & \lambda & & & & \\ & & & \lambda/2 & \lambda/2 & & & \\ & & & & \lambda & & & \\ & & & & & \ddots & & \\ & & & & & & \lambda & 0 \end{pmatrix} \qquad k \text{ even}, \tag{5.5.3}$$

$$Q_2^{(k)} = \begin{pmatrix} \mu & & & & \\ \mu & \mu & & & \\ & \ddots & \ddots & & \\ & & \ddots & \ddots & \\ & & & \mu & \mu \\ & & & & \mu \end{pmatrix} \tag{5.5.4}$$

and $Q_1^{(k)}$ has zeros on the off diagonal and its diagonal elements are negative and such that $(Q_0^{(k)} + Q_1^{(k)} + Q_2^{(k)})e = 0$.

We note that for all the LDQBDs we have considered so far, there has been a constant number of phases at each level. However, the present process has $k+1$ phases at level k. Another notable feature of this LDQBD is that the k dependency of the matrices $Q_0^{(k)}$, $Q_1^{(k)}$ and $Q_2^{(k)}$ arises solely through their dimensions. That is, the individual elements of the matrices are independent of k.

As mentioned earlier, Kingman [47] derived an expression for the generating function of the stationary distribution for this system and Adan, Wessels and Zijm [1] expressed the stationary distribution as an infinite linear combination of product forms. Since we have modelled the symmetric shortest queue system as a LDQBD, the stationary distribution has the analytical form given by equation (4.2.1). This is a much simpler expression than that given in [1] and it also has a clear probabilistic interpretation.

To compute the stationary distribution for this LDQBD we note that Condition 1 holds. We therefore consider the birth-death process $\bar{Y}(t)$ on the non negative integers with transition rates

$$\bar{q}(k, k+1) = \lambda \qquad k \geq 0, \tag{5.5.5}$$

$$\bar{q}(k, k-1) = \mu \qquad k \geq 1. \tag{5.5.6}$$

The stationary distribution $\bar{\ell}_k$ for $\bar{Y}(t)$ only exists if $\lambda < \mu$. However, the symmetric shortest queue system has a stationary distribution provided $\lambda < 2\mu$. Therefore considering $\bar{Y}(t)$ will only provide us with a method for selecting K^* when $\lambda < \mu$, that is when the traffic intensity is low. We adopt a trial and error approach for finding K^* for the case $\lambda \geq \mu$.

When analysing this LDQBD, a potential problem arises due to the increasing number of phases at each level. Since there are $k+1$ phases at each level, the matrix $\widetilde{U}_k^{\ell}$ has dimensions $k+1 \times k+2^{\ell}+1$ and hence increases in size as both k and ℓ increase. This contrasts with all previous LDQBDs we have analysed, where $\widetilde{\boldsymbol{U}}^{\ell}_{k}$ had the same dimension for all k and ℓ. Since $\widetilde{\boldsymbol{U}}^{\ell}_{k}$ increases in size with k and ℓ, the storage required and the computational effort required to compute $\widetilde{\boldsymbol{U}}^{\ell}_{k}$ can become excessive. For example, suppose $K^{*}=200$ and a value of $L=7$ is needed in Algorithm 3.4.4 to compute $\boldsymbol{R}_{200}$. We hence need to compute the $\widetilde{UD}$-pair $\widetilde{UD}(7,200)$. From equations (3.3.21)-(3.3.22) we see that to do this we need to compute the inverse of $\left[\boldsymbol{I}-\widetilde{\boldsymbol{U}}^{6}_{328}\widetilde{\boldsymbol{D}}^{6}_{392}-\widetilde{\boldsymbol{D}}^{6}_{328}\widetilde{\boldsymbol{U}}^{6}_{264}\right]^{-1}$ which has size 329×329. As mentioned in Section 3.7.3, when computing $\widetilde{UD}(7,200)$, it is necessary to have in storage certain $\widetilde{UD}$-pairs $\widetilde{UD}(\ell,k)$ for $\ell<7$ and various values of k. Hence a large amount of storage space is required. In practice this means that if K^{*} and the value of L required to compute $\boldsymbol{R}_{K^{*}}$ are too large, our algorithms will not be suitable for analysing this LDQBD. Hence for low values of λ/μ, our algorithms are suitable, but as λ/μ approaches two, our algorithms are essentially limited by the space restrictions of the computer and the time constraints of the analyst.

We used our algorithms to replicate the results in the first two rows of Table 1 of Adan, Wessels and Zijm [1]. These two rows correspond to values of λ/μ equal to 0.6 and 1. The values in rows three and four of Table 1 of [1] correspond to values of 1 and 1.8 for λ/μ and we found that the space required for these cases was excessive. We note that we have not attempted to optimise our computer code to minimise the storage space required. We discussed briefly in Section 3.7.3 how to reduce the storage required but it may be possible to improve upon the arguments presented there.

Before concluding our discussion of the symmetric shortest queue system, we briefly consider the asymmetric shortest queue system. In the asymmetric system server one serves at rate μ_1 and server two serves at rate μ_2. It is clear that we can model the asymmetric shortest queue system as a LDQBD. Furthermore, the same method of analysis as that described above can be employed. Hence the analysis of the asymmetric system is no more complex than the analysis of the symmetric system. We note that this is in contrast to the results of Adan, Wessels and Zijm [2]. There the authors extended their results in [1] for the symmetric case to the asymmetric case but there is a significant increase in the complexity of their analysis.

Adan, Wessels and Zijm [3] also considered the case of an asymmetric shortest queue system with threshold jockeying. That is, if the difference between the two queue lengths is greater than some value T, then one customer moves from the longest to the shortest queue. In this situation the analysis is simpler than in the case where there is no jockeying. They also show that the asymmetric system with jockeying can be modelled as a LDQBD that is eventually level independent.

5.5.4 Two parallel queues

Here we consider a system consisting of two M/M/1 queues that behave independently of one another except when one of them is empty. In this case, the server at the empty queue goes to help the server at the non empty queue and the service rate changes. A more detailed description is as follows.

The system consists of two single server queues, queue one and queue two. Customers arrive to queue one according to a Poisson process with rate λ_1 and customers arrive to queue two according to a Poisson process with rate λ_2. The two arrival streams are assumed to be independent of each other. If a customer arrives at either queue and finds the server is free then it moves straight into service, otherwise it queues up. The server at queue one serves customers at rate μ_1 and the server at queue two serves customers at rate μ_2. If queue one becomes empty, the server at queue one goes to help the server at queue two and the combined service rate at queue two becomes γ_2. Similarly, if queue two becomes empty, the server at queue two goes to help the server at queue one and the combined service rate at queue one becomes γ_1.

Rao and Posner [85] approximated this system by assuming that queue two can have at most M customers present. With this assumption, the new system can then be modelled as a QBD on the state space $\{(n_1, n_2) : n_1 \geq 0, 0 \leq n_2 \leq M\}$ where n_1 represents the number in queue one and n_2 represents the number in queue two.

An alternative approach to that of Rao and Posner is to model this system as a LDQBD. This leads to an exact expression for the stationary probabilities. The system can be modelled as a LDQBD on the state space $\{(k, j) : k \geq 0, 0 \leq j \leq k\}$ where k is the total number of customers in the system and j is the number of customers in queue one. The matrices $Q_0^{(k)}$ and $Q_2^{(k)}$ are given by

$$Q_0^{(k)} = \begin{pmatrix} \lambda_2 & \lambda_1 & & \\ & \ddots & \ddots & \\ & & \lambda_2 & \lambda_1 \end{pmatrix}, \tag{5.5.7}$$

$$Q_2^{(k)} = \begin{pmatrix} \gamma_2 & & & \\ \mu_1 & \mu_2 & & \\ & \ddots & \ddots & \\ & & \mu_1 & \mu_2 \\ & & & \gamma_1 \end{pmatrix}, \tag{5.5.8}$$

and $Q_1^{(k)}$ has zeros on the off diagonals and its diagonal elements are negative and such that $(Q_0^{(k)} + Q_1^{(k)} + Q_2^{(k)})e = \mathbf{0}$.

Again we note that the stationary distribution has a simple closed form expression given by equation (4.2.1). Since Condition 1 holds for this LDQBD we consider the birth-death process $\bar{Y}(t)$ with transition rates

$$\bar{q}(k, k+1) = \lambda_1 + \lambda_2 \qquad k \geq 0, \tag{5.5.9}$$

$$\bar{q}(k, k-1) = \min(\gamma_1, \gamma_2, \mu_1 + \mu_2) \qquad k \geq 1. \tag{5.5.10}$$

Provided the stationary distribution $\bar{\ell}_k$ exists, we can take K^* to be the smallest value of K such that $\sum_{k=K+1}^{\infty} \bar{\ell}_k < \epsilon$ holds. A problem with this approach is that the stationary distribution only exists if $\lambda_1 + \lambda_2 < \min(\gamma_1, \gamma_2, \mu_1 + \mu_2)$. An alternative approach to choosing K^* is as follows. Consider two M/M/1 queues, one with arrival rate λ_1 and service rate μ_1 and the other with arrival rate λ_2 and service rate μ_2. If the stationary distributions for each of these queues are given by $\pi_i(k)$, then we can take $K_i^*, i = 1, 2$ to be the smallest values of K_i such that $\sum_{k=K_i+1}^{\infty} \pi_i(k) < \epsilon$ for some tolerance ϵ. Now provided $\gamma_i > \mu_i, i = 1, 2$, each of these M/M/1 queues will perform worse than the corresponding queues in the parallel system. Hence we can take $K^* = K_1^* + K_2^*$.

In Table 5.5.1 we present the values of x_0, the stationary probability of being in state $(0,0)$ and $\sum_{i=0}^{K^*}[\boldsymbol{x}_k]_0$, the stationary probability of queue one being empty. The results in Table 5.5.1 agree with the results given in the first four rows of Table 1 in Rao and Posner [85].

The last two rows of Table 1 in Rao and Posner correspond to the parameter values $\lambda_1 = 0.7$, $\mu_1 = 1$, $\gamma_1 = 7/5$, $\lambda_2 = 0.7$, $\mu_2 = 1$, $\gamma_2 = 4/3$ and $\lambda_1 = 0.9$, $\mu_1 = 1$, $\gamma_1 = 9/7$,

λ_1	μ_1	γ_1	λ_2	μ_2	γ_2	x_0	K^*	$\sum_{i=0}^{K^*}[\boldsymbol{x}_k]_0$
0.3	1.0	1.5	0.2	1.0	2.0	0.706	16	0.787
0.5	1.0	5/3	0.3	1.0	1.5	0.527	29	0.659
0.6	1.0	1.5	0.5	1.0	5/3	0.359	59	0.526
0.6	1.0	1.5	0.6	1.0	1.5	0.293	82	0.507

Table 5.5.1: Some performance measures for the system of two parallel queues

$\lambda_2 = 0.7$, $\mu_2 = 1$, $\gamma_2 = 7/5$. For these parameter values we found values of 123 and 225 for K^*. These parameter values required a large amount of storage space and we were unable to obtain values for x_0 and $\sum_{i=0}^{K^*}[\boldsymbol{x}_k]_0$ in these cases.

Before concluding this section we note that there are many more systems that have a similar structure to the two systems we have considered here. We refer to van Houtum [37] and Cohen and Boxma [15] for further examples. Although these systems can be modelled using the LDQBD structure, the algorithms we have presented appear to be limited to the case of low traffic intensities. We note that for the models we have considered, the matrices $\boldsymbol{Q}_i^{(k)}, i = 0, 1, 2$ have a large number of elements that are zero. Given this structure, it may be possible to develop algorithms that exploit this fact. We feel that our algorithms may be useful as a starting point for developing alternative algorithms.

Chapter 6

Duality results for M/G/1–type and GI/M/1–type processes

6.1 Introduction

In this chapter we consider M/G/1–type and GI/M/1–type Markov processes. Both these processes are extensions of LDQBDs. Central to their analysis are the matrices $\{\boldsymbol{G}_k,\ k \geq 1\}$ and $\{\boldsymbol{R}_k,\ k \geq 0\}$ for M/G/1–type and GI/M/1–type processes respectively. Our main motivation for considering such processes is the work of Ramaswami [79] and Asmussen and Ramaswami [7]. Ramaswami [79] showed that given a level independent M/G/1–type process, one can construct a level independent GI/M/1–type process such that the matrix $\boldsymbol{G}$ for the M/G/1–type process and the matrix $\boldsymbol{R}$ for the GI/M/1–type process satisfy a so called duality relationship. The GI/M/1–type process constructed from the M/G/1–type process is called the dual process. Asmussen and Ramaswami [7] gave an alternative proof of Ramaswami's duality results by appealing to time reversal results.

Ramaswami presented his results in the more general context of Markov renewal processes but his results can of course, be specialised to the corresponding continuous time Markov processes. We will consider only continuous time Markov processes in this chapter but analogous results exist for discrete time processes.

We show that Ramaswami's dual process can be interpreted as the time reverse of a process similar in structure to the M/G/1–type process. In order to do this we give a detailed discussion of the time reverse of a Markov process with respect to an invariant measure $\boldsymbol{m}$. When considering invariant measures for continuous time Markov processes one needs to consider the issue of regularity. Hence we give a discussion of regularity and provide conditions for regularity of M/G/1–type and GI/M/1–type processes.

By interpreting Ramaswami's dual process in the manner described above we construct an alternative dual process to Ramaswami's dual for level independent processes.

At the end of the chapter we construct via time reversal methods the most sensible dual process for level dependent M/G/1–type and GI/M/1–type processes.

The rest of this chapter is set out as follows. Firstly we present some background theory for continuous time Markov processes. We give definitions of regularity, invariant measures and the time reverse of a Markov process with respect to an invariant measure. We also provide probabilistic interpretations of invariant measures and the time reverse process. Following this we define Markov processes of M/G/1–type and GI/M/1–type and give some background results for such processes. We then define the time reverse of a M/G/1–type and GI/M/1–type process. We use the time reverse processes to consider situations under which the level and phase variables for M/G/1–type and GI/M/1–type processes are independent. In Section 6.8 we give a presentation of Ramaswami's dual process and we then consider the probabilistic interpretation of the dual process. Finally we present an alternative to Ramaswami's dual process and we define a dual process for level dependent M/G/1–type and GI/M/1–type processes.

6.2 Some theory for continuous time Markov processes

The results we present in this section are taken from a variety of sources. In particular, we refer to results from Kelly [40] and Pollett and Taylor [78].

In the following we consider a continuous time Markov process $X(t)$ on a state space $\mathcal{S}$ with q-matrix given by $\boldsymbol{Q} = (q(i,j),\ i,j \in \mathcal{S})$. We assume that $\boldsymbol{Q}$ is stable, conservative and irreducible. Kelly [40] presented results for processes that may fail to be regular but we will restrict our attention to regular processes. Reuter [87] showed that if $\boldsymbol{Q}$ is stable and conservative then $\boldsymbol{Q}$ is regular if and only if the equation

$$\boldsymbol{Q}\boldsymbol{x} = \alpha \boldsymbol{x} \tag{6.2.1}$$

has no bounded, non-trivial solution for some (and then for all) $\alpha > 0$.

If $X(t)$ fails to be regular then, with probability greater then zero, the sample paths for $X(t)$ stop after a finite time T. This time is referred to as the explosion time or the first exit time. Furthermore, in the interval of time $[0, T]$, the process $X(t)$ will have made infinitely many transitions. Hence an irregular Markov process can be interpreted intuitively as able to start in an initial state, evolve through an infinite number of states and then stop, all within a finite time. Kelly [39] suggested that when the process stops, we can think of the process as having "run out of instructions."

Establishing that a process is regular via Reuter's condition is not usually easy to do. For various types of Markov processes with special structures there exist other necessary and sufficient conditions or sufficient conditions for regularity. Feller [25] showed that a pure birth process on the positive integers with transition rates $q(i, i+1), i \geq 1$ is regular if and only if

$$\sum_{i=0}^{\infty} \frac{1}{q(i, i+1)} = \infty. \tag{6.2.2}$$

Dobrušin [21] and Reuter [87] gave straightforward necessary and sufficient conditions for determining if birth-death processes are regular. For more general Markov processes, other methods are needed to prove regularity. A sufficient condition for a general Markov process to be regular is that $-q(i, i)$ be bounded above for all i. This condition is easy to apply but for many regular processes this condition fails to hold. For example, this condition cannot be used to show that an M/M/∞ queue is regular. Another sufficient condition for regularity is that the jump chain be recurrent. This condition can be used to show that an M/M/∞ queue is regular but in general it is not particularly easy to apply in practice.

Pollett and Taylor [78] used results of Yan and Chen [95] to derive a sufficient condition for the regularity of processes whose state space can be divided into levels and whose transition rates between these levels have a particular structure. Corollary 2 in [78] can be stated as follows.

Lemma 6.2.1 *Consider a Markov process on a state space $\mathcal{S}$ and let $\mathcal{S}_0, \mathcal{S}_1, \ldots$ be a partition of $\mathcal{S}$ with $\mathcal{S}_k$ non-empty. If the transition rates between states $x, y \in \mathcal{S}$ are such that*

$$q(x, y) > 0, \quad x \in \mathcal{S}_k \Rightarrow y \in \bigcup_{i=0}^{k+1} \mathcal{S}_i \tag{6.2.3}$$

and

$$\sup_{x \in \mathcal{S}_k} -q(x,x) < \infty, \quad k \geq 0 \tag{6.2.4}$$

then the process is regular if the pure-birth process with rates

$$\bar{q}(i,i+1) = \sup_{x \in \mathcal{S}_k} \sum_{y \in \mathcal{S}_{k+1}} q(x,y) \quad k \geq 1 \tag{6.2.5}$$

is regular.

If we refer to the subsets $\mathcal{S}_k$ as levels, the condition (6.2.3) implies that the process can move to any lower levels but only to the level immediately above the present level. In fact it is skip-free to the right in the sense we used in Section 2.2.1. Condition (6.2.4) states that the total transition rate out of states must be bounded for each level. We will use this lemma in the sequel to derive a sufficient condition for regularity of a GI/M/1–type process.

We now move to the definition of an invariant measure. Let $\boldsymbol{P}(t)$ be the transition probability matrix for $X(t)$. That is, $\boldsymbol{P}(t) = (P_{ij}(t), i, j \in \mathcal{S})$ where $P_{ij}(t) = P(X(t) = j|X(0) = i)$, $\forall t \geq 0$. We now define the concept of an invariant measure for $\boldsymbol{P}(t)$ and also an invariant measure for a conservative generator matrix $\boldsymbol{Q}$.

Definition 6.2.1 *A vector $\boldsymbol{m} = (m_k, k \in \mathcal{S})$ of positive numbers satisfying*

$$\boldsymbol{m}\boldsymbol{P}(t) = \boldsymbol{m} \qquad \forall t > 0 \tag{6.2.6}$$

is an invariant measure for $\boldsymbol{P}(t)$.

Definition 6.2.2 *A vector $\boldsymbol{m} = (m_k, k \in \mathcal{S})$ of positive numbers satisfying*

$$\boldsymbol{m}\boldsymbol{Q} = 0 \tag{6.2.7}$$

is an invariant measure for $\boldsymbol{Q}$.

To find an invariant measure for $\boldsymbol{Q}$ we simply need to find a positive vector $\boldsymbol{m}$ that satisfies equation (6.2.7). The question that we are interested in is when is an invariant measure for $\boldsymbol{Q}$ also an invariant measure for $\boldsymbol{P}(t)$?

Kendall and Reuter [42] showed that if $X(t)$ is recurrent, then (6.2.7) has a solution that is unique up to constant multiples. Furthermore, this solution is known to be invariant for $\boldsymbol{P}(t)$. If the solution can be chosen such that $\boldsymbol{me} = 1$ then $X(t)$ is positive recurrent. If the solution is such that $\boldsymbol{me} = \infty$ then $X(t)$ is null recurrent. In the positive recurrent case, $\boldsymbol{m}$ can be interpreted as the stationary distribution and the ratio $\frac{m_i}{m_j}$ can be interpreted as the limit as t tends to infinity of the ratio of the time spent in state i to time spent in state j in an interval of time $[0, t]$. The ratio $\frac{m_i}{m_j}$ has the same interpretation in the null recurrent case.

If $X(t)$ is transient then a non-trivial solution to (6.2.7) may not exist. If such a solution does exist, and $\boldsymbol{Q}$ is regular, then $\boldsymbol{m}$ will be such that $\boldsymbol{me} = \infty$. It can be shown that when $X(t)$ is transient, (6.2.7) has a non-trivial solution if and only if the jump chain has an invariant measure for its transition probability matrix $\boldsymbol{P}^J(n)$. Note that when $X(t)$ is transient, the jump chain will also be transient. The following lemma which appears as a corollary to Theorem 2 in Harris [36] gives a sufficient condition for a discrete time transient Markov chain to have an invariant measure for $\boldsymbol{P}^J(n)$.

Lemma 6.2.2 *If $P^J(i, j)$ are the single step transition probabilities for a transient, irreducible Markov chain and $\boldsymbol{P}^J(n)$ is the transition probability matrix, then the chain has an invariant measure for $\boldsymbol{P}^J(n)$ provided that for all i, $P^J(k, i) = 0$ except for a finite set of values of k.*

Hence when $X(t)$ is transient, by considering the jump chain and appealing to Lemma 6.2.2, we can determine if there exists an invariant measure for $\boldsymbol{Q}$. To determine if this vector is also invariant for $\boldsymbol{P}(t)$, we can use the following theorem which is adapted from the theorem in Section 3 of Kelly [40].

Lemma 6.2.3 *Suppose $\boldsymbol{m}$ is an invariant measure for $\boldsymbol{Q}$ and define $\boldsymbol{Q}^R = (Q^R(i, j),\ i, j \in \mathcal{S})$ by*

$$q^R(i, j) = \frac{m_j}{m_i} q(j, i). \tag{6.2.8}$$

Then $\boldsymbol{m}$ is an invariant measure for $\boldsymbol{P}(t)$ if and only if the Markov process with q-matrix $\boldsymbol{Q}^R$ is regular.

Derman [19] gave an interpretation of an invariant measure $\boldsymbol{m}$ for $\boldsymbol{P}(t)$ for a transient process. The interpretation also holds for recurrent processes and is as follows. Suppose at time $t = 0$ we place N_k particles in each state $k \in \mathcal{S}$ where N_k are independent random variables following a Poisson distribution with mean m_k. From time $t = 0$ onwards, let the particles move independently from state to state, each moving according to a Markov process with q-matrix $\boldsymbol{Q}$. If we denote the number of particles in state $k \in \mathcal{S}$ at time $t > 0$ by $N_k(t)$ then $N_k(t)$ has a Poisson distribution with mean m_k and $N_k(t)$, $k \in \mathcal{S}$ are independent random variables.

We now define the time-reverse of $X(t)$ with respect to an invariant measure $\boldsymbol{m}$.

Definition 6.2.3 *If $\boldsymbol{m}$ is an invariant measure for $\boldsymbol{P}(t)$ then the time-reverse of $X(t)$ with respect to $\boldsymbol{m}$ is the the Markov process $X^R(t)$ on the state space $\mathcal{S}$ with q-matrix given by $\boldsymbol{Q}^R = (q^R(i,j),\ i,j \in \mathcal{S})$ where $q^R(i,j)$ is given by equation (6.2.8).*

Note that, by Lemma 6.2.3, since $\boldsymbol{m}$ is an invariant measure for $\boldsymbol{P}(t)$, $\boldsymbol{Q}^R$ must be regular and hence by Reuter [87] there is a unique Markov process with q-matrix $\boldsymbol{Q}^R$. The process $X^R(t)$ is therefore uniquely defined by Definition 6.2.3.

As mentioned previously, when $X(t)$ is positive recurrent and $\boldsymbol{m}$ is chosen such that $\boldsymbol{me} = 1$, $\boldsymbol{m}$ gives the stationary distribution for $X(t)$. In this case $\boldsymbol{m}$ is also the stationary distribution for $X^R(t)$ and $X^R(t)$ is called the *reversed process* of $X(t)$. For more details on the reversed process we refer to Kelly [39]. Although the reversed process only exists when $X(t)$ is positive recurrent, the time-reverse of $X(t)$ can exist when $X(t)$ is not positive recurrent.

We state the following result that was observed by Kelly [40].

Lemma 6.2.4 *If $\boldsymbol{m}$ is an invariant measure for $\boldsymbol{P}(t)$ then it is also an invariant measure for $\boldsymbol{P}^R(t)$.*

Lemma 6.2.4 can be interpreted in terms of the particle system as follows. Suppose $X(t)$, $t \geq 0$ starts at time $t = 0$ with N_k particles in state $k \in \mathcal{S}$ where N_k are independent Poisson random variables with mean m_k. Further, suppose we let the process run until a time τ with particles moving independently, each according to a Markov process with q-matrix $\boldsymbol{Q}$. Similarly, suppose that we start the process $X^R(t)$ at time $t = 0$ with N_k particles in state $k \in \mathcal{S}$ where N_k are independent Poisson random variables with mean m_k. Lemma 6.2.4 states that the distribution of particles at any time $t > 0$ in the process $X(\tau - t)$, $t \geq 0$ is the same as the distribution of particles at time $t > 0$ in the process $X^R(t)$.

6.3 M/G/1–type processes

A continuous time Markov process $X_M(t)$ on a state space $\mathcal{S} = \{(k,j) : k \geq 0, 1 \leq j \leq M(k)\}$ is called a level dependent M/G/1–type process if its q-matrix has the block partitioned form

$$Q_M = \begin{pmatrix} A_1^{(0)} & A_2^{(0)} & A_3^{(0)} & A_4^{(0)} & \cdots \\ A_0^{(1)} & A_1^{(1)} & A_2^{(1)} & A_3^{(1)} & \cdots \\ 0 & A_0^{(2)} & A_1^{(2)} & A_2^{(2)} & \cdots \\ 0 & 0 & A_0^{(3)} & A_1^{(3)} & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix} \qquad (6.3.1)$$

where $A_0^{(k)}$, $k \geq 1$ are $M(k) \times M(k-1)$ matrices and $A_\ell^{(k)}$, $\ell \geq 1$, $k \geq 0$ are $M(k) \times M(k+\ell-1)$ matrices. In the case where $M(k) = M$ $\forall k \geq 0$ and Q_M has the form

$$Q_M = \begin{pmatrix} A_1^{(0)} & A_2^{(0)} & A_3^{(0)} & A_4^{(0)} & \cdots \\ A_0 & A_1 & A_2 & A_3 & \cdots \\ 0 & A_0 & A_1 & A_2 & \cdots \\ 0 & 0 & A_0 & A_1 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}, \qquad (6.3.2)$$

where A_ℓ, $\ell \geq 0$ and $A_\ell^{(0)}$, $\ell \geq 1$ are $M \times M$ matrices, $X_M(t)$ is called a level independent M/G/1–type process. For the M/G/1–type processes we consider, we assume that Q_M is stable and conservative.

For a general level dependent M/G/1–type process it does not seem possible to obtain a simple test for regularity. Hence the method used to prove regularity will depend upon the particular structure of the process under consideration. However, in the case of a level independent M/G/1–type process it is clear that due to the homogeneity of the process, $-q((k,j),(k,j)), 1 \leq j \leq M$ is independent of k for all $k \geq 1$. Hence there are only $2M$ possible values for $-q((k,j),(k,j))$, M possible values for the states in levels one and above and M for the states in level zero. Therefore, since we have assumed that the process is stable, $-q((k,j),(k,j))$ must be bounded above and hence all level independent M/G/1–type processes are regular. For simplicity we will only consider regular level dependent M/G/1–type processes for the rest of this chapter.

We now present some results for M/G/1–type processes. Neuts [73] dealt extensively with level independent M/G/1–type processes. More recently, Ramaswami [80] presented results for level dependent M/G/1–type processes. Note that a level dependent M/G/1–type process can be thought of as an extension of the LDQBD. We therefore adopt the notation we used for LDQBDs. In the state (k,j), k is referred to as the level of the state and j is referred to as the phase of the state. From the structure of $\boldsymbol{Q}_M$, we see that from level k, the process can either move down to level $k-1$, stay at level k or move up to any level $k+\ell$, $\ell \geq 1$. Due to this property we say that $X_M(t)$ is skip free to the left.

Recall that for a LDQBD, the (i,j)th element of the matrix $\boldsymbol{G}_k, k \geq 1$ is defined to be the probability that, starting in the state (k,i), the process eventually reaches level $k-1$ and does so through the state $(k-1,j)$. In the case of a M/G/1–type process we make the same definition for $\boldsymbol{G}_k$. Ramaswami [80] has shown using methods similar to those used in Chapter 2 that the matrices $\{\boldsymbol{G}_k, k \geq 1\}$ are the minimal non negative solutions to the equations

$$\boldsymbol{0} = \boldsymbol{A}_0^{(k)} + \sum_{\ell=0}^{\infty} \boldsymbol{A}_{\ell+1}^{(k)} \boldsymbol{G}_{k+\ell} \cdots \boldsymbol{G}_k \qquad k \geq 1. \tag{6.3.3}$$

In [80] it is also shown that if $X_M(t)$ is positive recurrent then the stationary distribution $\boldsymbol{x} = (\boldsymbol{x}_0, \boldsymbol{x}_1, \dots)$ satisfies the recursion

$$\boldsymbol{x}_k = \left(\sum_{\ell=0}^{k-1} \boldsymbol{x}_\ell \boldsymbol{B}(\ell,k)\right)\left(-\boldsymbol{B}(k,k)\right)^{-1} \tag{6.3.4}$$

where

$$\boldsymbol{B}(\ell,k) = \boldsymbol{A}_{k-\ell+1}^{(\ell)} + \sum_{j=0}^{\infty} \boldsymbol{A}_{k-\ell+2+j}^{(\ell)} \boldsymbol{G}_{k+1+j} \dots \boldsymbol{G}_{k+1}. \tag{6.3.5}$$

As was the case for LDQBDs, it seems that there does not exist any simple test for positive recurrence of a level dependent M/G/1–type process. However, Neuts [73] has derived conditions for a level independent M/G/1–type process to be positive recurrent. Neuts actually considered discrete time M/G/1–type chains but the following results for continuous time processes can be proved using methods similar to those used in the proof of Theorem 2.8.3.

Theorem 6.3.1 *Consider a continuous time level independent M/G/1–type process where* $\boldsymbol{A} = \sum_{\ell=0}^{\infty} \boldsymbol{A}_\ell$ *is an irreducible, aperiodic q-matrix. Let* $\boldsymbol{\pi}$ *be such that* $\boldsymbol{\pi A} = \boldsymbol{0}$ *and* $\boldsymbol{\pi e} = 1$ *and define* $\boldsymbol{\beta} = \sum_{\ell=0}^{\infty} \ell A_\ell \boldsymbol{e}$. *Then*

1. *if* $\boldsymbol{\pi\beta} < 0$ *then the M/G/1–type process is positive recurrent,*
2. *if* $\boldsymbol{\pi\beta} = 0$ *then the M/G/1–type process is null recurrent,*
3. *if* $\boldsymbol{\pi\beta} > 0$ *then the M/G/1–type process is transient.*

It is also possible to characterise the ergodic nature of a level independent M/G/1–type process by considering the spectral radius of the matrix $\boldsymbol{G}$. We state the following theorem which can be derived from results in Neuts [73].

Theorem 6.3.2 *Consider a continuous time level independent M/G/1–type process where* $\boldsymbol{A} = \sum_{\ell=0}^{\infty} \boldsymbol{A}_\ell$ *is an irreducible, aperiodic q-matrix. If* $\boldsymbol{G}$ *is the minimal non negative solution to equation (6.3.3) then*

1. *the M/G/1–type process is recurrent if and only if the spectral radius of* $\boldsymbol{G}$ *is equal to 1,*
2. *the M/G/1–type process is transient if and only if the spectral radius of* $\boldsymbol{G}$ *is less than 1.*

6.4 GI/M/1–type processes

A continuous time Markov process $X_{GI}(t)$ on a state space $\mathcal{S} = \{(k,j) : k \geq 0, 1 \leq j \leq M(k)\}$ is called a GI/M/1–type process if its q-matrix has the block partitioned form

$$Q_{GI} = \begin{pmatrix} D_1^{(0)} & D_0^{(0)} & 0 & 0 & \cdots \\ D_2^{(1)} & D_1^{(1)} & D_0^{(1)} & 0 & \cdots \\ D_3^{(2)} & D_2^{(2)} & D_1^{(2)} & D_0^{(2)} & \cdots \\ D_4^{(3)} & D_3^{(3)} & D_2^{(3)} & D_1^{(3)} & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix} \tag{6.4.1}$$

where $D_0^{(k)}$, $k \geq 0$ are $M(k) \times M(k+1)$ matrices and $D_\ell^{(k)}$, $\ell \geq 1$, $k \geq \ell - 1$ are $M(k) \times M(k - \ell + 1)$ matrices. In the case where $M(k) = M \; \forall k \geq 0$ and Q_{GI} has the form

$$Q_{GI} = \begin{pmatrix} D_1^{(0)} & D_0 & 0 & 0 & \cdots \\ D_2^{(1)} & D_1 & D_0 & 0 & \cdots \\ D_3^{(2)} & D_2 & D_1 & D_0 & \cdots \\ D_4^{(3)} & D_3 & D_2 & D_1 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix} \tag{6.4.2}$$

where D_ℓ, $\ell \geq 0$ and $D_\ell^{(\ell-1)}$, $\ell \geq 1$ are $M \times M$ matrices, $X_{GI}(t)$ is called a level independent GI/M/1–type process. For the GI/M/1–type processes we consider, we assume that Q_{GI} is stable and conservative. From the structure of Q_{GI}, we see that from level k, the process can either move up to level $k+1$, stay at level k or move down to any level $k - \ell$, $1 \leq \ell \leq k$. Due to this property we say that $X_{GI}(t)$ is skip free to the right.

We now apply Lemma 6.2.1 to obtain a sufficient condition for regularity of a level dependent GI/M/1–type process. If we let $\mathcal{S}_k$, $k \geq 0$ represent all states in level k then since $\boldsymbol{Q}_{GI}$ is skip free to the right, condition (6.2.3) is satisfied. It is clear that condition (6.2.4) holds so we have that a level independent GI/M/1–type process is regular if the pure birth process with rates given by

$$q(k,k+1) = \sup_i \sum_{j=0}^{M(k+1)} [\boldsymbol{D}_0^{(k)}]_{ij} \tag{6.4.3}$$

is regular. That is, using the result of Feller [25] and equation (6.2.2), if

$$\sum_{k=0}^{\infty} \left[\sup_i \sum_{j=0}^{M(k+1)} [\boldsymbol{D}_0^{(k)}]_{ij} \right]^{-1} = \infty. \tag{6.4.4}$$

Note that for many practical processes condition (6.4.4) will be easy to verify. In particular, in the case of a level independent GI/M/1–type process, equation (6.4.4) will always hold due to the homogeneity of the process and hence all level independent GI/M/1–type processes will be regular. For simplicity we will only consider regular level dependent GI/M/1–type processes for the rest of this chapter.

GI/M/1–type processes can also be thought of as an extension of LDQBDs. The seminal work on level independent GI/M/1–type processes was done by Neuts [72]. Ramaswami [80] extended the theory to the case of level dependent GI/M/1–type processes. We now state some of these results but omit the proofs since they follow exactly the same lines as the proofs of the corresponding results for LDQBDs presented in Chapter 2.

If we define the matrix $\boldsymbol{R}_k$, $k \geq 0$ in the same way that it was defined in Chapter 2 then it is shown in Ramaswami [80] that $\{\boldsymbol{R}_k, k \geq 0\}$ are the minimal non negative solutions to the equations

$$0 = \boldsymbol{D}_0^{(k)} + \sum_{\ell=0}^{\infty} \boldsymbol{R}_k \cdots \boldsymbol{R}_{k+\ell} \boldsymbol{D}_{\ell+1}^{(k+\ell+1)} \qquad k \geq 0. \tag{6.4.5}$$

It is also shown in [80] that the stationary distribution $\boldsymbol{x} = (\boldsymbol{x}_0, \boldsymbol{x}_1, \dots)$ of $X_{GI}(t)$ satisfies the recursion

$$\boldsymbol{x}_{k+1} = \boldsymbol{x}_k \boldsymbol{R}_k \qquad k \geq 0. \tag{6.4.6}$$

As was the case for level dependent M/G/1–type processes, it does not seem possible to obtain easily verifiable conditions for positive recurrence of level dependent GI/M/1–type processes. Neuts [72] derived conditions for positive recurrence of a level independent GI/M/1–type process. We state these conditions in the following theorem.

Theorem 6.4.1 *Consider a continuous time level independent GI/M//1–type process where* $\boldsymbol{D} = \sum_{\ell=0}^{\infty} \boldsymbol{D}_\ell$ *is an irreducible, aperiodic q-matrix. Let* $\boldsymbol{\pi}$ *be such that* $\boldsymbol{\pi D} = \boldsymbol{0}$ *and* $\boldsymbol{\pi e} = 1$ *and define* $\boldsymbol{\beta} = \sum_{\ell=0}^{\infty} \ell D_\ell \boldsymbol{e}$*. Then*

1. *if* $\boldsymbol{\pi\beta} > 0$ *then the GI/M/1–type process is positive recurrent,*
2. *if* $\boldsymbol{\pi\beta} = 0$ *then the GI/M/1–type process is null recurrent,*
3. *if* $\boldsymbol{\pi\beta} < 0$ *then the GI/M/1–type process is transient.*

The following theorem characterises the ergodic nature of a level independent GI/M/1-type process in terms of the spectral radius of the matrix $\boldsymbol{R}$. We refer to Neuts [72] for the proof.

Theorem 6.4.2 *Consider a continuous time level independent GI/M//1–type process where* $\boldsymbol{D} = \sum_{\ell=0}^{\infty} \boldsymbol{D}_\ell$ *is an irreducible, aperiodic q-matrix. If* $\boldsymbol{R}$ *is the minimal non negative solution to equation (6.4.5) then*

1. *the GI/M/1–type process is positive recurrent if and only if the spectral radius of* $\boldsymbol{R}$ *is less than one,*
2. *the GI/M/1–type process is null recurrent or transient if and only if the spectral radius of* $\boldsymbol{R}$ *is equal to one.*

6.5 The time-reverse of M/G/1–type and GI/M/1–type processes

In this section we consider the time-reverse of M/G/1–type and GI/M/1–type processes.

6.5.1 Level dependent M/G/1–type and GI/M/1–type processes

Consider an irreducible, continuous time level dependent M/G/1–type process $X_M(t)$ with q-matrix $\boldsymbol{Q}_M$ given by (6.3.1) and transition probability matrix $\boldsymbol{P}(t)$. If $X_M(t)$ is recurrent then we know from Section 6.2 that an invariant measure always exists. To determine if an invariant measure exists when $X_M(t)$ is transient we consider the jump chain. The jump chain of $X_M(t)$ is the discrete time Markov chain with transition probability matrix

$$P_M = \begin{pmatrix} \bar{A}_1^{(0)} & \bar{A}_2^{(0)} & \bar{A}_3^{(0)} & \bar{A}_4^{(0)} & \cdots \\ \bar{A}_0^{(1)} & \bar{A}_1^{(1)} & \bar{A}_2^{(1)} & \bar{A}_3^{(1)} & \cdots \\ 0 & \bar{A}_0^{(2)} & \bar{A}_1^{(2)} & \bar{A}_2^{(2)} & \cdots \\ 0 & 0 & \bar{A}_0^{(3)} & \bar{A}_1^{(3)} & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix} \qquad (6.5.1)$$

where $\bar{A}_\ell^{(k)} = \Phi_k^{-1} A_\ell^{(k)} + \delta_{\ell 1} I$, $\Phi_k = -\text{diag}([A_1^{(k)}]_{ii})$ and $\delta_{\ell 1}$ is the Kronecker delta. For any state (k, j), the only states (ℓ, i) for which $P_{(\ell,i),(k,j)}$ is potentially greater then zero, are the states in the levels $0, 1, \ldots, k+1$. These states make up a finite set of states and hence by Lemma 6.2.2, the jump chain has an invariant measure and therefore the equation $mQ_M = 0$ has a non negative solution m. Lemma 6.2.3 must be applied to determine if m is an invariant measure for $P_M(t)$.

If such an invariant measure m exists, then the q-matrix for the time-reverse of $X_M(t)$, $X_M^R(t)$ is given by

$$Q_M^R = \begin{pmatrix} D_1^{(0)} & D_0^{(0)} & 0 & 0 & \cdots \\ D_2^{(1)} & D_1^{(1)} & D_0^{(1)} & 0 & \cdots \\ D_3^{(2)} & D_2^{(2)} & D_1^{(2)} & D_0^{(2)} & \cdots \\ D_4^{(3)} & D_3^{(3)} & D_2^{(3)} & D_1^{(3)} & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix} \qquad (6.5.2)$$

where

$$\boldsymbol{D}_0^{(k)} = \boldsymbol{\Delta}_k^{-1}(\boldsymbol{A}_0^{(k+1)})'\boldsymbol{\Delta}_{k+1} \qquad k \geq 0, \tag{6.5.3}$$

$$\boldsymbol{D}_1^{(k)} = \boldsymbol{\Delta}_k^{-1}(\boldsymbol{A}_1^{(k)})'\boldsymbol{\Delta}_k \qquad k \geq 0, \tag{6.5.4}$$

$$\boldsymbol{D}_{\ell+1}^{(k)} = \boldsymbol{\Delta}_k^{-1}(\boldsymbol{A}_{\ell+1}^{(k-\ell)})'\boldsymbol{\Delta}_{k-\ell} \qquad k \geq 1,\ 1 \leq \ell \leq k, \tag{6.5.5}$$

$$\boldsymbol{\Delta}_k = \mathrm{diag}(\boldsymbol{m}_k) \qquad k \geq 0. \tag{6.5.6}$$

So we see that $X_M^R(t)$ is a level dependent GI/M/1–type process.

Now consider an irreducible, continuous time level dependent GI/M/1–type process $X_{GI}(t)$ with q-matrix $\boldsymbol{Q}_{GI}$ given by (6.4.1) and transition probability matrix $\boldsymbol{P}_{GI}(t)$. As was the case for the level dependent M/G/1-type process, if $X_{GI}(t)$ is recurrent then we know from Section 6.2 that an invariant measure always exists. When $X_{GI}(t)$ is transient, unlike in the case of a level dependent M/G/1–type process, we are unable to show that $\boldsymbol{m}\boldsymbol{Q}_{GI} = \boldsymbol{0}$ has a positive solution. If a positive solution $\boldsymbol{m}$ to this equation does exist, then Lemma 6.2.3 can be applied to determine if $\boldsymbol{m}$ is an invariant measure for $\boldsymbol{P}_{GI}(t)$.

If an invariant measure $\boldsymbol{m}$ exists for a level dependent GI/M/1–type process $X_{GI}(t)$, then the q-matrix for the time-reverse of $X_{GI}(t)$, $X_{GI}^R(t)$, is given by

$$\boldsymbol{Q}_{GI}^R = \begin{pmatrix} \boldsymbol{A}_1^{(0)} & \boldsymbol{A}_2^{(0)} & \boldsymbol{A}_3^{(0)} & \boldsymbol{A}_4^{(0)} & \cdots \\ \boldsymbol{A}_0^{(1)} & \boldsymbol{A}_1^{(1)} & \boldsymbol{A}_2^{(1)} & \boldsymbol{A}_3^{(1)} & \cdots \\ \boldsymbol{0} & \boldsymbol{A}_0^{(2)} & \boldsymbol{A}_1^{(2)} & \boldsymbol{A}_2^{(2)} & \cdots \\ \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{A}_0^{(3)} & \boldsymbol{A}_1^{(3)} & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix} \tag{6.5.7}$$

where

$$A_0^{(k)} = \Delta_k^{-1}(D_0^{(k-1)})'\Delta_{k-1} \qquad k \geq 1, \tag{6.5.8}$$

$$A_{\ell+1}^{(k)} = \Delta_k^{-1}(D_{\ell+1}^{(k)})'\Delta_{k+\ell} \qquad k \geq 0,\ \ell \geq 0, \tag{6.5.9}$$

$$\Delta_k = \mathrm{diag}(\boldsymbol{m}_k) \qquad k \geq 0 \tag{6.5.10}$$

and we see that this is the q-matrix for a level dependent M/G/1–type process.

6.5.2 Level independent M/G/1–type and GI/M/1–type processes

We now consider the time-reverse of level independent M/G/1–type and GI/M/1–type processes.

Consider an irreducible, continuous time level independent M/G/1–type process $X_M(t)$ with q-matrix $\boldsymbol{Q}_M$ given by (6.3.2). As mentioned in the previous section, if $X_M(t)$ is recurrent then an invariant measure always exists. We showed in the previous section that when $X_M(t)$ is transient, the equation $\boldsymbol{m}\boldsymbol{Q}_M = \boldsymbol{0}$ has a positive solution $\boldsymbol{m}$. Lemma 6.2.3 states that $\boldsymbol{m}$ will be an invariant measure for $\boldsymbol{P}_M(t)$ if the process with q-matrix given by

$$q^R((k,j),(\ell,m)) = \frac{m(\ell,m)}{m(k,j)}q((\ell,m),(k,j)) \tag{6.5.11}$$

is regular. It is easy to show that $-q^R((k,j),(\ell,m)) = -q((k,j),(\ell,m))$. Since we showed in Section 6.3 that $-q((k,j),(\ell,m))$ is bounded above, $-q^R((k,j),(\ell,m))$ is also bounded above and hence from results in Section 6.2 the process with q-matrix given by (6.5.11) is regular. Hence when $X_M(t)$ is transient, an invariant measure for $\boldsymbol{P}_M(t)$ always exists and furthermore, any positive solution to $\boldsymbol{m}\boldsymbol{Q}_M = \boldsymbol{0}$ is an invariant measure for $\boldsymbol{P}_M(t)$.

We can now state that the time-reverse of a level independent M/G/1–type process with invariant measure equal to $\boldsymbol{m}$ is, in general, a level dependent GI/M/1–type process with q-matrix given by

$$Q_M^R = \begin{pmatrix} D_1^{(0)} & D_0^{(0)} & 0 & 0 & \cdots \\ D_2^{(1)} & D_1^{(1)} & D_0^{(1)} & 0 & \cdots \\ D_3^{(2)} & D_2^{(2)} & D_1^{(2)} & D_0^{(2)} & \cdots \\ D_4^{(3)} & D_3^{(3)} & D_2^{(3)} & D_1^{(3)} & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix} \tag{6.5.12}$$

where

$$\boldsymbol{D}_0^{(k)} = \boldsymbol{\Delta}_k^{-1}(\boldsymbol{A}_0)'\boldsymbol{\Delta}_{k+1} \qquad k \geq 0, \tag{6.5.13}$$

$$\boldsymbol{D}_\ell^{(k)} = \boldsymbol{\Delta}_k^{-1}(\boldsymbol{A}_\ell)'\boldsymbol{\Delta}_{k-\ell+1} \qquad k \geq 1,\ 1 \leq \ell \leq k, \tag{6.5.14}$$

$$\boldsymbol{D}_{k+1}^{(k)} = \boldsymbol{\Delta}_k^{-1}(\boldsymbol{A}_{k+1}^{(0)})'\boldsymbol{\Delta}_0 \qquad k \geq 0, \tag{6.5.15}$$

$$\boldsymbol{\Delta}_k = \text{diag}(\boldsymbol{m}_k) \qquad k \geq 0. \tag{6.5.16}$$

Now consider an irreducible, continuous time level independent GI/M/1–type process $X_{GI}(t)$ with q-matrix $\boldsymbol{Q}_{GI}$ given by (6.4.2). We stated in the previous section that if $X_{GI}(t)$ is recurrent then an invariant measure for $\boldsymbol{P}_{GI}(t)$ always exists. If $X_{GI}(t)$ is transient and $\boldsymbol{m}\boldsymbol{Q}_{GI} = \boldsymbol{0}$ has a positive solution then Lemma 6.2.3 can be applied to determine if $\boldsymbol{m}$ is an invariant measure for $\boldsymbol{P}_{GI}(t)$. Using similar methods to those used above for level independent M/G/1–type processes, it can be shown via Lemma 6.2.3 that any positive solution to $\boldsymbol{m}\boldsymbol{Q}_{GI} = \boldsymbol{0}$ is always an invariant measure for $\boldsymbol{P}_{GI}(t)$.

The time-reverse of a level independent GI/M/1–type process with q-matrix given by (6.4.2) and invariant measure equal to $\boldsymbol{m}$ is in general a level dependent M/G/1–type process with q-matrix given by

$$Q_{GI}^{R} = \begin{pmatrix} A_1^{(0)} & A_2^{(0)} & A_3^{(0)} & A_4^{(0)} & \cdots \\ A_0^{(1)} & A_1^{(1)} & A_2^{(1)} & A_3^{(1)} & \cdots \\ 0 & A_0^{(2)} & A_1^{(2)} & A_2^{(2)} & \cdots \\ 0 & 0 & A_0^{(3)} & A_1^{(3)} & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix} \tag{6.5.17}$$

where

$$A_0^{(k)} = \Delta_k^{-1}(D_0)'\Delta_{k-1} \qquad k \geq 1, \tag{6.5.18}$$

$$A_\ell^{(0)} = \Delta_k^{-1}(D_\ell^{(\ell-1)})'\Delta_{\ell-1} \qquad \ell \geq 1, \tag{6.5.19}$$

$$A_\ell^{(k)} = \Delta_k^{-1}(D_\ell)'\Delta_{k+\ell-1} \qquad k \geq 1,\ \ell \geq 1, \tag{6.5.20}$$

$$\Delta_k = \mathrm{diag}(\boldsymbol{m}_k) \qquad k \geq 0. \tag{6.5.21}$$

An interesting question to pose is when will the time-reverse of a level independent M/G/1–type or GI/M/1–type process be level independent? In order to derive conditions under which this holds, we need to consider processes for which the phase and level variables are independent. We devote the next section to such processes.

6.6 Independence of level and phase

In this section we consider M/G/1–type and GI/M/1–type processes on state spaces where the number of phases at each level is a constant. That is, $\mathcal{S}$ is such that $M(k) = M\ \forall k \geq 0$. This enables us to make the following definitions.

Definition 6.6.1 *An invariant measure* $\boldsymbol{m} = (\boldsymbol{m}_0, \boldsymbol{m}_1, \dots)$ *for a level dependent M/G/1–type or GI/M/1–type process on the state space* $\mathcal{S} = \{(k,j) : k \geq 0, 1 \leq j \leq M\}$ *is said to have the phase independent of level property if* $\boldsymbol{m}$ *has the form*

$$\boldsymbol{m}_k = \boldsymbol{\theta} \prod_{\ell=0}^{k-1} \eta_\ell \qquad k \geq 0 \tag{6.6.1}$$

where $\boldsymbol{\theta}$ *is a non negative vector independent of* k, *and* $\eta_\ell, \ell \geq 0$ *are positive scalars.*

Definition 6.6.2 *An invariant measure* $\boldsymbol{m} = (\boldsymbol{m}_0, \boldsymbol{m}_1, \dots)$ *for a level independent M/G/1–type or GI/M/1–type process on the state space* $\mathcal{S} = \{(k,j) : k \geq 0, 1 \leq j \leq M\}$ *is said to have the phase independent of level property if* $\boldsymbol{m}$ *has the form*

$$\boldsymbol{m}_k = \boldsymbol{\theta} \eta^k \qquad k \geq 0 \tag{6.6.2}$$

where $\boldsymbol{\theta}$ *is a non negative vector independent of* k, *and* η *is a positive scalar.*

It is of interest to determine under what conditions an invariant measure for a M/G/1–type or GI/M/1–type process has the phase independent of level property. In general this is difficult to do. The following theorem gives necessary and sufficient conditions for a positive recurrent level dependent GI/M/1–type process to have the phase independent of level property. The theorem is a generalisation of Theorem 2.1 in Ramaswami and Taylor [83]. Ramaswami and Taylor modelled a variety of queueing networks as LDQBDs and derived various known results for queueing networks by applying their Theorem 2.1.

Theorem 6.6.1 *Consider a level dependent GI/M/1–type process on a state space* $\mathcal{S} = \{(k,j) : k \geq 0, 1 \leq j \leq M\}$ *with q-matrix given by (6.4.1). Then the process is positive recurrent and the stationary distribution* $\boldsymbol{x}$ *has the phase independent of level property if and only if there exists a non negative vector* $\boldsymbol{\xi}$ *with* $\boldsymbol{\xi e} = 1$ *and positive scalars* η_k *such that,*

$$\boldsymbol{\xi R}_k = \eta_k \boldsymbol{\xi} \quad k \geq 0, \tag{6.6.3}$$

$$\sum_{k=0}^{\infty} \prod_{\ell=0}^{k-1} \eta_\ell < \infty, \tag{6.6.4}$$

and

$$\boldsymbol{\xi} \left(\sum_{k=0}^{\infty} \left(\prod_{\ell=0}^{k-1} \eta_\ell \right) \boldsymbol{D}_{k+1}^{(k)} \right) = 0, \tag{6.6.5}$$

in which case

$$\boldsymbol{x}_k = K\boldsymbol{\xi}\prod_{\ell=0}^{k-1}\eta_\ell \qquad k \geq 0 \tag{6.6.6}$$

where K is a normalising constant.

Proof: Firstly assume that the process is positive recurrent and $\boldsymbol{x}$ has the phase independent of level property. Hence $\boldsymbol{x}_k = (\boldsymbol{me})^{-1}\boldsymbol{m}_k$ where $\boldsymbol{m}_k$ is given by equation (6.6.1). We know from the first block of the global balance equations that $\boldsymbol{x}$ satisfies

$$\sum_{k=0}^{\infty}\boldsymbol{x}_k\boldsymbol{D}_{k+1}^{(k)} = \boldsymbol{0} \tag{6.6.7}$$

which upon substituting for $\boldsymbol{x}_k$ becomes

$$\boldsymbol{\theta}\left(\sum_{k=0}^{\infty}\left(\prod_{\ell=0}^{k-1}\eta_\ell\right)\boldsymbol{D}_{k+1}^{(k)}\right) = \boldsymbol{0}. \tag{6.6.8}$$

This is equation (6.6.5) with $\boldsymbol{\xi} = \boldsymbol{\theta}$. Since $\sum_{k=0}^{\infty}\boldsymbol{x}_k\boldsymbol{e} = 1$ it is easy to show that (6.6.4) holds and hence provided we can show $\boldsymbol{\theta}\boldsymbol{R}_k = \eta_k\boldsymbol{\theta}$ $\forall k \geq 0$, we have the required result. From equation (6.4.6) we have $\forall k \geq 0$,

$$\begin{aligned}
\boldsymbol{x}_{k+1} &= \boldsymbol{x}_k\boldsymbol{R}_k & (6.6.9)\\
\Rightarrow \boldsymbol{\theta}\left(\prod_{\ell=0}^{k}\eta_\ell\right) &= \boldsymbol{\theta}\left(\prod_{\ell=0}^{k-1}\eta_\ell\right)\boldsymbol{R}_k & (6.6.10)\\
\Rightarrow \boldsymbol{\theta}\eta_k &= \boldsymbol{\theta}\boldsymbol{R}_k & (6.6.11)
\end{aligned}$$

and the desired result holds.

Now assume that there exists a non negative vector $\boldsymbol{\xi}$ with $\boldsymbol{\xi e} = 1$ and positive scalars η_k such that (6.6.3), (6.6.4) and (6.6.5) hold. We know from Markov chain theory that the process is positive recurrent if we can find a positive solution $\boldsymbol{m}$ such that $\boldsymbol{me} < \infty$ to the global balance equations

$$\begin{aligned}
\sum_{k=0}^{\infty}\boldsymbol{m}_k\boldsymbol{D}_{k+1}^{(k)} &= \boldsymbol{0} & & (6.6.12)\\
\sum_{k=0}^{\infty}\boldsymbol{m}_{k+\ell}\boldsymbol{D}_{k}^{(k+\ell)} &= \boldsymbol{0} & \ell \geq 0. & (6.6.13)
\end{aligned}$$

By repeatedly applying equation (6.6.3), equation (6.6.5) reduces to

$$\left(\sum_{k=0}^{\infty}\left(\boldsymbol{\xi}\prod_{\ell=0}^{k-1}\boldsymbol{R}_\ell\right)\boldsymbol{D}_{k+1}^{(k)}\right) = \boldsymbol{0}. \tag{6.6.14}$$

Hence $\boldsymbol{m}_k = \boldsymbol{\xi} \prod_{\ell=0}^{k-1} \boldsymbol{R}_\ell$ satisfies equation (6.6.12) which is the first block of the global balance equations. By using equations (6.6.3) and (6.4.5) it can be shown that $\boldsymbol{m}_k = \boldsymbol{\xi} \prod_{\ell=0}^{k-1} \boldsymbol{R}_\ell$ also satisfies the equations (6.6.13). Now we have

$$\begin{aligned} \boldsymbol{m}\boldsymbol{e} &= \sum_{k=0}^{\infty} \boldsymbol{\xi} \prod_{\ell=0}^{k-1} \boldsymbol{R}_\ell \boldsymbol{e} \\ &= \sum_{k=0}^{\infty} \prod_{\ell=0}^{k-1} \eta_\ell \boldsymbol{\xi} \boldsymbol{e} \\ &= \sum_{k=0}^{\infty} \prod_{\ell=0}^{k-1} \eta_\ell \\ &< \infty \end{aligned}$$

where the last inequality results from equation (6.6.4). Hence the process is positive recurrent and $\boldsymbol{x}_k$ is given by $\boldsymbol{x}_k = K\boldsymbol{m}_k$ which by applying (6.6.3) is equivalent to (6.6.6). □

In the case of a level independent GI/M/1–type process, we have the following theorem. We omit the proof as it is similar to that of theorem 6.6.1.

Theorem 6.6.2 *Consider a level independent GI/M/1–type process with q-matrix given by (6.4.2). Then the process is positive recurrent and the stationary distribution $\boldsymbol{x}$ has the phase independent of level property if and only if there exists a non negative vector $\boldsymbol{\xi}$ with $\boldsymbol{\xi}\boldsymbol{e}^t = 1$ and a positive scalar $\eta < 1$, such that*

$$\boldsymbol{\xi}\boldsymbol{R} = \eta\boldsymbol{\xi}, \tag{6.6.15}$$

and

$$\boldsymbol{\xi} \sum_{k=0}^{\infty} \eta^k \boldsymbol{D}_{k+1}^{(k)} = \boldsymbol{0}, \tag{6.6.16}$$

in which case

$$\boldsymbol{x}_k = K\eta^\ell\boldsymbol{\xi} \qquad k \geq 0 \tag{6.6.17}$$

where K is a normalising constant.

6.7 Level independence of the time reverse process

We now return to the question we posed at the end of Section 6.5.2, namely, when will the time reverse of a level independent M/G/1–type or GI/M/1–type process be level independent? This question is answered for the case of a M/G/1–type process in the following theorem.

Theorem 6.7.1 *If $X_M(t)$ is a level independent M/G/1–type process with an invariant measure $\boldsymbol{m}$, then $X_M^R(t)$, the time-reverse of $X_M(t)$ with respect to $\boldsymbol{m}$, is a level independent GI/M/1–type process if and only if the invariant measure for $X_M(t)$ (and $X_M^R(t)$) has the phase independent of level property.*

Proof: Firstly assume that the invariant measure has the phase independent of level property. This implies that $\boldsymbol{m}_k = \boldsymbol{\theta}\eta^k$. We need to show that the (i,j)th element of each of the matrices $\boldsymbol{D}_0^{(k)}, k \geq 0$, $\boldsymbol{D}_1^{(k)}, k \geq 0$ and $\boldsymbol{D}_{\ell+1}^{(k)}, k \geq 1,\ 1 \leq \ell < k$ is independent of k. We have the following

$$\begin{aligned} [\boldsymbol{D}_0^{(k)}]_{ij} &= \frac{[\boldsymbol{m}_{k+1}]_j}{[\boldsymbol{m}_k]_i}[\boldsymbol{A}_0]_{ji} \qquad\qquad k \geq 0 \\ &= \frac{\eta^{k+1}[\boldsymbol{\theta}]_j}{\eta^k[\boldsymbol{\theta}]_i}[\boldsymbol{A}_0]_{ji} \\ &= \frac{\eta[\boldsymbol{\theta}]_j}{[\boldsymbol{\theta}]_i}[\boldsymbol{A}_0]_{ji}, \end{aligned}$$

$$\begin{aligned} [\boldsymbol{D}_1^{(k)}]_{ij} &= \frac{[\boldsymbol{m}_k]_j}{[\boldsymbol{m}_k]_i}[\boldsymbol{A}_1]_{ji} \qquad\qquad k \geq 0 \\ &= \frac{\eta^{k}[\boldsymbol{\theta}]_j}{\eta^k[\boldsymbol{\theta}]_i}[\boldsymbol{A}_1]_{ji} \\ &= \frac{[\boldsymbol{\theta}]_j}{[\boldsymbol{\theta}]_i}[\boldsymbol{A}_1]_{ji}, \end{aligned}$$

$$\begin{aligned} [\boldsymbol{D}_{\ell+1}^{(k)}]_{ij} &= \frac{[\boldsymbol{m}_{k-\ell}]_j}{[\boldsymbol{m}_k]_i}[\boldsymbol{A}_{\ell+1}]_{ji} \qquad\qquad k \geq 1,\ 1 \leq \ell < k \\ &= \frac{\eta^{k-\ell}[\boldsymbol{\theta}]_j}{\eta^k[\boldsymbol{\theta}]_i}[\boldsymbol{A}_{\ell+1}]_{ji} \\ &= \frac{[\boldsymbol{\theta}]_j}{\eta^\ell[\boldsymbol{\theta}]_i}[\boldsymbol{A}_{\ell+1}]_{ji} \end{aligned}$$

which are all independent of k.

Now assume that the time-reverse is level independent. We must show for some constant $\eta > 0$, that $\boldsymbol{m}_{k+1} = \boldsymbol{m}_k\eta$ holds for all $k \geq 0$. From Definition 6.6.2, since $\boldsymbol{m}_0$ is a non negative vector, this is sufficient for showing that $\boldsymbol{m}$ has the phase independent of level property.

We have the following

$$[D_0^{(k+1)}]_{ij} = [D_0^{(k)}]_{ij} \qquad \forall k \geq 1 \tag{6.7.1}$$

$$\Longrightarrow \frac{[\boldsymbol{m}_{k+2}]_j}{[\boldsymbol{m}_{k+1}]_i}[A_0]_{ji} = \frac{[\boldsymbol{m}_{k+1}]_j}{[\boldsymbol{m}_k]_i}[A_0]_{ji} \tag{6.7.2}$$

$$\Longrightarrow \frac{[\boldsymbol{m}_{k+2}]_j}{[\boldsymbol{m}_{k+1}]_j} = \frac{[\boldsymbol{m}_{k+1}]_i}{[\boldsymbol{m}_k]_i}. \tag{6.7.3}$$

Now equation (6.7.3) holds for all i and j and all $k \geq 1$ so applying the equation recursively gives

$$\frac{[\boldsymbol{m}_{k+1}]_j}{[\boldsymbol{m}_k]_j} = \frac{[\boldsymbol{m}_1]_i}{[\boldsymbol{m}_0]_i} \qquad \forall i, j,\ \forall k \geq 0. \tag{6.7.4}$$

Now since equation (6.7.4) holds for all i, j, by fixing j and varying i, it is clear that $\frac{[\boldsymbol{m}_1]_i}{[\boldsymbol{m}_0]_i}$ is independent of i. Hence we have that $[\boldsymbol{m}_{k+1}]_j = K[\boldsymbol{m}_k]_j\ \forall k \geq 0$ where K is a constant. This completes the proof. ☐

The following theorem is proved in a similar way to the previous theorem.

Theorem 6.7.2 *If $X_{GI}(t)$ is a level independent GI/M/1–type process with an invariant measure $\boldsymbol{m}$, then $X_{GI}^R(t)$, the time-reverse of $X_{GI}(t)$ with respect to $\boldsymbol{m}$, is a level independent M/G/1–type process if and only if the invariant measure for $X_{GI}(t)$ (and $X_{GI}^R(t)$) has the phase independent of level property.*

6.8 Ramaswami's duality result

In this section we present a special case of a duality result obtained by Ramaswami [79] for level independent M/G/1–type and GI/M/1–type processes. Ramaswami actually considered Markov renewal processes of M/G/1–type and GI/M/1–type. Such processes have the same skip free structure as the Markov processes we are considering here but the q-matrices are replaced by transition kernels consisting of matrices of distribution functions. For such processes it is necessary to consider the matrices $\boldsymbol{G}(z, s)$ and $\boldsymbol{R}(z, s)$ respectively which are matrices of double transforms. Ramaswami's duality result gives a relationship between $\boldsymbol{G}(z, s)$ and $\boldsymbol{R}(z, s)$ for two suitably defined Markov renewal processes of M/G/1–type and GI/M/1–type. We state the result only for the

Markov processes we are considering. Ramaswami proved his result by using the fact that $G(z,s)$ is the unique solution in the class of double transforms to a particular matrix equation. We give a direct algebraic proof rather than use this result. The same method of proof will be used to prove results in later sections. The special case of Ramaswami's result that we present is equivalent to Corollary 5.1 in Asmussen [6]. Asmussen derived his result using a probabilistic version of Wiener-Hopf factorisation methods.

The following theorem is based on Theorem 2.1 in Ramaswami [79].

Theorem 6.8.1 *Consider a level independent M/G/1–type process with q-matrix given by (6.3.2). Let* $\boldsymbol{A} = \sum_{\ell=0}^{\infty} \boldsymbol{A}_\ell$ *and assume that* $\boldsymbol{A}$ *is irreducible. Further, let* $\boldsymbol{a}$ *be the invariant probability vector of* $\boldsymbol{A}$, $\boldsymbol{\Delta} = diag(\boldsymbol{a})$ *and* $\boldsymbol{G}$ *be the minimal non negative solution to equation (6.3.3).*

If we define $\boldsymbol{D}_\ell$ *by*

$$\boldsymbol{D}_\ell = \boldsymbol{\Delta}^{-1}\boldsymbol{A}_\ell'\boldsymbol{\Delta} \qquad \ell \geq 0, \tag{6.8.1}$$

then the matrix $\boldsymbol{D} = \sum_{\ell=0}^{\infty} \boldsymbol{D}_\ell$ *is an irreducible q-matrix and the matrix* $\boldsymbol{R}$ *for the GI/M/1–type process with* $\boldsymbol{D}_\ell$ *given by (6.8.1) is given by*

$$\boldsymbol{R} = \boldsymbol{\Delta}^{-1}\boldsymbol{G}'\boldsymbol{\Delta}. \tag{6.8.2}$$

Proof: We give an algebraic proof similar to that given in Ramaswami [79]. The matrix $\boldsymbol{D}$ is given by

$$\begin{aligned} \boldsymbol{D} &= \sum_{\ell=0}^{\infty} \boldsymbol{D}_\ell && (6.8.3)\\ &= \boldsymbol{\Delta}^{-1}\sum_{\ell=0}^{\infty} \boldsymbol{A}_\ell'\boldsymbol{\Delta} && (6.8.4)\\ &= \boldsymbol{\Delta}^{-1}\boldsymbol{A}'\boldsymbol{\Delta}. && (6.8.5) \end{aligned}$$

Hence $\boldsymbol{D}$ must be irreducible since $\boldsymbol{A}$ is irreducible. The off diagonal elements of $\boldsymbol{D}$ are non negative and since

$$\begin{aligned} \boldsymbol{D}\boldsymbol{e} &= \boldsymbol{\Delta}^{-1}\boldsymbol{A}'\boldsymbol{\Delta}\boldsymbol{e} && (6.8.6)\\ &= \boldsymbol{\Delta}^{-1}\boldsymbol{A}'\boldsymbol{a}' && (6.8.7)\\ &= \boldsymbol{0}, && (6.8.8) \end{aligned}$$

we have that $\boldsymbol{D}$ is a q-matrix.

From results in Sections 6.3 and 6.4 we have that the matrix $\boldsymbol{R}$ for the GI/M/1–type process with $\boldsymbol{D}_\ell$ given by (6.8.1) is the minimal non negative solution to

$$\boldsymbol{0} = \sum_{\ell=0}^{\infty} \boldsymbol{R}^\ell \boldsymbol{\Delta}^{-1} \boldsymbol{A}'_\ell \boldsymbol{\Delta}. \tag{6.8.9}$$

The matrix $\boldsymbol{G}$ for the M/G/1–type process is the minimal non negative solution to

$$\boldsymbol{0} = \sum_{\ell=0}^{\infty} \boldsymbol{A}_\ell \boldsymbol{G}^\ell. \tag{6.8.10}$$

Pre-multiplying (6.8.9) by $\boldsymbol{\Delta}$, post-multiplying by $\boldsymbol{\Delta}^{-1}$ and taking transposes of both sides gives

$$\boldsymbol{0} = \sum_{\ell=0}^{\infty} \boldsymbol{A}_\ell (\boldsymbol{\Delta}^{-1} \boldsymbol{R}' \boldsymbol{\Delta})^\ell. \tag{6.8.11}$$

Comparing equation (6.8.11) to equation (6.8.10), since $\boldsymbol{G}$ is the minimal non negative solution to (6.8.10), we have that

$$\boldsymbol{\Delta}^{-1} \boldsymbol{R}' \boldsymbol{\Delta} \geq \boldsymbol{G} \tag{6.8.12}$$

which implies

$$\boldsymbol{R} \geq \boldsymbol{\Delta}^{-1} \boldsymbol{G}' \boldsymbol{\Delta} \tag{6.8.13}$$

since $\boldsymbol{a}$ is a positive vector.

Pre-multiplying (6.8.10) by $\boldsymbol{\Delta}$, post-multiplying by $\boldsymbol{\Delta}^{-1}$ and taking transposes of both sides gives

$$\boldsymbol{0} = \sum_{\ell=0}^{\infty} (\boldsymbol{\Delta}^{-1} \boldsymbol{G}' \boldsymbol{\Delta})^\ell \boldsymbol{\Delta}^{-1} \boldsymbol{A}'_\ell \boldsymbol{\Delta}. \tag{6.8.14}$$

Comparing equation (6.8.14) to equation (6.8.9), since $\boldsymbol{R}$ is the minimal non negative solution to (6.8.9), we have that

$$\boldsymbol{\Delta}^{-1} \boldsymbol{G}' \boldsymbol{\Delta} \geq \boldsymbol{R}. \tag{6.8.15}$$

which together with equation (6.8.13) implies that $\boldsymbol{R} = \boldsymbol{\Delta}^{-1} \boldsymbol{G}' \boldsymbol{\Delta}$. □

Ramaswami called the GI/M/1–type process constructed in Theorem 6.8.1 the dual of the M/G/1–type process. We will refer to the relationship (6.8.2) as Ramaswami's duality relationship.

Although Ramaswami only defined the dual of a M/G/1–type process, it is clear that it is also possible to define the dual of a GI/M/1 type process. In order to do this we state the following theorem which can be proved in the same way as Theorem 6.8.1.

Theorem 6.8.2 *Consider a level independent GI/M/1–type process with q-matrix given by (6.4.2). Let* $\boldsymbol{D} = \sum_{\ell=0}^{\infty} \boldsymbol{D}_\ell$ *and assume that* $\boldsymbol{D}$ *is irreducible. Further, let* $\boldsymbol{d}$ *be the invariant probability vector of* $\boldsymbol{D}$, $\boldsymbol{\Delta} = diag(\boldsymbol{d})$ *and* $\boldsymbol{R}$ *be the minimal non negative solution to equation (6.4.5).*

If we define $\boldsymbol{A}_\ell$ *by*

$$\boldsymbol{A}_\ell = \boldsymbol{\Delta}^{-1}\boldsymbol{D}_\ell'\boldsymbol{\Delta} \qquad \ell \geq 0, \tag{6.8.16}$$

then the matrix $\boldsymbol{A} = \sum_{\ell=0}^{\infty} \boldsymbol{A}_\ell$ *is an irreducible q-matrix and the matrix* $\boldsymbol{G}$ *for the M/G/1–type process with* $\boldsymbol{A}_\ell$ *given by (6.8.16) is given by*

$$\boldsymbol{G} = \boldsymbol{\Delta}^{-1}\boldsymbol{R}'\boldsymbol{\Delta}. \tag{6.8.17}$$

We will refer to the M/G/1–type process constructed in Theorem 6.8.2 as Ramaswami's dual of a GI/M/1–type process.

When considering dual processes it is usual to consider the dual of the dual process. The dual of a M/G/1 type process is a GI/M/1–type process with $\boldsymbol{D}_\ell$ given by

$$\boldsymbol{D}_\ell = \boldsymbol{\Delta}^{-1}\boldsymbol{A}_\ell'\boldsymbol{\Delta} \tag{6.8.18}$$

with $\boldsymbol{\Delta} = \text{diag}(\boldsymbol{a})$ where $\boldsymbol{a}$ is the unique solution to $\boldsymbol{aA} = \boldsymbol{0}$ such that $\boldsymbol{ae} = 1$. Now the dual of this GI/M/1-type process is a M/G/1–type process with matrices $\bar{\boldsymbol{A}}_\ell$ given by

$$\begin{aligned} \bar{\boldsymbol{A}}_\ell &= \bar{\boldsymbol{\Delta}}^{-1}(\boldsymbol{D}_\ell)'\bar{\boldsymbol{\Delta}} && (6.8.19)\\ &= \bar{\boldsymbol{\Delta}}^{-1}(\boldsymbol{\Delta}^{-1}(\boldsymbol{A}_\ell)'\boldsymbol{\Delta})'\bar{\boldsymbol{\Delta}} && (6.8.20) \end{aligned}$$

with $\bar{\boldsymbol{\Delta}} = \text{diag}(\boldsymbol{d})$ where $\boldsymbol{d}$ is the unique solution to $\boldsymbol{dD} = \boldsymbol{0}$ subject to $\boldsymbol{de} = 1$. Now if we set $\boldsymbol{d} = \boldsymbol{a}$ then $\bar{\boldsymbol{\Delta}} = \boldsymbol{\Delta}$ and then

$$\begin{aligned} &\boldsymbol{dD} && (6.8.21)\\ &= \boldsymbol{a}\boldsymbol{\Delta}^{-1}\boldsymbol{A}'\boldsymbol{\Delta} && (6.8.22)\\ &= \boldsymbol{e}'\boldsymbol{A}'\boldsymbol{\Delta} && (6.8.23)\\ &= \boldsymbol{0}. && (6.8.24) \end{aligned}$$

So because $\boldsymbol{dD} = \boldsymbol{0}$ has a unique solution such that $\boldsymbol{de} = 1$, $\boldsymbol{d}$ must be equal to $\boldsymbol{a}$. This implies that the dual process of the dual of a M/G/1–type process is the original

M/G/1–type process. It can be shown in a similar fashion that the dual process of the dual of a GI/M/1–type process is the original GI/M/1–type process. We now state these two results as a theorem.

Theorem 6.8.3 *The dual process of the dual of a M/G/1-type (GI/M/1–type) process is the original M/G/1–type (GI/M/1–type) process.*

The relationships (6.8.2) and (6.8.17) have a number of applications. For example, if we derive a result for M/G/1–type (or GI/M/1–type) processes then the corresponding result for GI/M/1–type (or M/G/1–type) processes can be obtained directly via the duality relationships. Latouche and Ramaswami [56] derived a closed form expression for the matrix $\boldsymbol{G}$ for a QBD and they used the relationship (6.8.2) to derive a closed form expression for the matrix $\boldsymbol{R}$. The following theorems follow from equations (6.8.2) and (6.8.17).

Theorem 6.8.4

1. *The matrix $\boldsymbol{R}$ for the dual of a level independent M/G/1–type process and the matrix $\boldsymbol{G}$ for the original M/G/1–type process have the same eigenvalues.*
2. *The matrix $\boldsymbol{G}$ for the dual of a level independent GI/M/1–type process and the matrix $\boldsymbol{R}$ for the original GI/M/1–type process have the same eigenvalues.*

Proof: Part *1* follows directly from equation (6.8.2) and *2* follows from equation (6.8.17). □

The following two theorems describe the relationship between the ergodic characteristics of M/G/1–type and GI/M/1–type processes and their duals.

Theorem 6.8.5

1. *The dual process of a recurrent level independent M/G/1–type process is either transient or null recurrent.*
2. *The dual process of a transient level independent M/G/1–type process is positive recurrent.*

Proof: First we prove *1*. From Theorem 6.3.2, since the M/G/1–type process is recurrent, the spectral radius of $\boldsymbol{G}$ is equal to one and hence from Theorem 6.8.4 the spectral radius of $\boldsymbol{R}$ for the dual process is equal to one. From Theorem 6.4.2 the dual process must therefore be transient or null recurrent.

Now consider part *2*. If the M/G/1–type process is transient, the spectral radius of $\boldsymbol{G}$ is less than one and hence the spectral radius of $\boldsymbol{R}$ is less than one which implies that the dual is positive recurrent. □

The following theorem is proved in the same way as Theorem 6.8.5.

Theorem 6.8.6 *1. The dual process of a positive recurrent level independent GI/M/1–type process is transient.*

2. The dual process of a transient or null recurrent level independent GI/M/1–type process is recurrent.

6.9 An alternative proof of Ramaswami's duality result

Asmussen and Ramaswami [7] gave an alternative proof of Ramaswami's duality result. They proved the duality result in its more general Markov renewal process form. The method of proof was in terms of time reversal of sample paths. In this section we give a brief outline of their arguments. To avoid unnecessary complications, we present their results for the case of Markov processes. We will only consider Ramaswami's dual of a M/G/1–type process but the discussion can be easily extended to the dual of a GI/M/1–type process.

Consider a Markov process $X_M(t) = (L_M(t), J_M(t))$ on the state space $\{(\ell, j) : \ell = \pm 1, \pm 2, \dots, 1 \le j \le M\}$ with q-matrix of the form

$$\boldsymbol{Q}_M = \begin{pmatrix} \ddots & \vdots & \vdots & \vdots & \vdots & \\ \cdots & \boldsymbol{A}_1 & \boldsymbol{A}_2 & \boldsymbol{A}_3 & \boldsymbol{A}_4 & \cdots \\ \cdots & \boldsymbol{A}_0 & \boldsymbol{A}_1 & \boldsymbol{A}_2 & \boldsymbol{A}_3 & \cdots \\ \cdots & \boldsymbol{0} & \boldsymbol{A}_0 & \boldsymbol{A}_1 & \boldsymbol{A}_2 & \cdots \\ \cdots & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{A}_0 & \boldsymbol{A}_1 & \cdots \\ & \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix} \qquad (6.9.1)$$

where the matrix $\boldsymbol{A} = \sum_{\ell=0}^{\infty} \boldsymbol{A}_\ell$ is an irreducible, aperiodic q-matrix. We assume that the matrix $\boldsymbol{Q}_M$ is stable and conservative. Because of the structure of the state space, the process is referred to as a "doubly infinite" process. Note that the process has the skip-free to the left property. Hence we refer to this process as a doubly infinite M/G/1–type process. Due to the homogeneity of the matrix $\boldsymbol{Q}_M$, there are only M possible values of $-q_M((k,j),(k,j))$ and hence since $\boldsymbol{Q}_M$ is stable, $-q_M((k,j),(k,j))$ must be bounded above. Therefore the process $X_M(t)$ is regular.

Asmussen and Ramaswami defined a second doubly infinite process, $X_{GI}(t) = (L_{GI}(t), J_{GI}(t))$ on the same state space as that above, with q-matrix given by

$$\boldsymbol{Q}_{GI} = \begin{pmatrix} \ddots & \vdots & \vdots & \vdots & \vdots & \\ \cdots & \boldsymbol{D}_1 & \boldsymbol{D}_0 & 0 & 0 & \cdots \\ \cdots & \boldsymbol{D}_2 & \boldsymbol{D}_1 & \boldsymbol{D}_0 & 0 & \cdots \\ \cdots & \boldsymbol{D}_3 & \boldsymbol{D}_2 & \boldsymbol{D}_1 & \boldsymbol{D}_0 & \cdots \\ \cdots & \boldsymbol{D}_4 & \boldsymbol{D}_3 & \boldsymbol{D}_2 & \boldsymbol{D}_1 & \cdots \\ & \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix} \tag{6.9.2}$$

with

$$\boldsymbol{D}_\ell = \boldsymbol{\Delta}^{-1}\boldsymbol{A}_\ell'\boldsymbol{\Delta} \qquad \ell \geq 0 \tag{6.9.3}$$

and

$$\boldsymbol{\Delta} = \mathrm{diag}(\boldsymbol{a}) \tag{6.9.4}$$

where $\boldsymbol{a}$ is the solution to $\boldsymbol{aA} = \boldsymbol{0}$ such that $\boldsymbol{ae} = 1$. Note that this process has the skip-free upward property and hence we refer to it as a doubly infinite GI/M/1–type process. Asmussen and Ramaswami refer to this doubly infinite process as the "dual" of the original doubly infinite process since it is constructed in a similar fashion to Ramaswami's dual of an M/G/1–type process. It is easy to show that this doubly infinite process is regular.

It is clear that the matrix $\boldsymbol{A}$ is the q-matrix for the partially observed process $J_M(t)$ and the matrix $\boldsymbol{D}$ is the q-matrix for the partially observed process $J_{GI}(t)$. Asmussen and Ramaswami observed that the relationship (6.9.3) implies $\boldsymbol{D} = \boldsymbol{\Delta}^{-1}\boldsymbol{A}\boldsymbol{\Delta}$ which implies that $J_{GI}(t)$ is the time reverse of $J_M(t)$ with respect to the invariant measure $\boldsymbol{a}$. In fact, since $J_M(t)$ is positive recurrent, $J_{GI}(t)$ is the reversed process of $J_M(t)$. Therefore from results for reversed processes, we can say that

$$[\boldsymbol{a}]_{j_0} P^M_{0,j_0}(j_0 \to j_1 \to \ldots \to j_n) = [\boldsymbol{a}]_{j_n} P^{GI}_{0,j_n}(j_n \to j_{n-1} \to \ldots \to j_0) \qquad (6.9.5)$$

where $P^M_{0,j_0}(j_0 \to j_1 \to \ldots \to j_n)$ is the probability that $X_M(t)$ moves sequentially through the phases $j_0, j_1, \ldots, j_n$, given it started in the state $(0, j_0)$ and $P^{GI}_{0,j_n}(j_n \to j_{n-1} \to \ldots \to j_0)$ has a similar interpretation. We note that equation (6.9.5) corresponds to equation (3.2) in Asmussen and Ramaswami [7].

Although $J_{GI}(t)$ is the time reverse of $J_M(t)$, the matrices $\boldsymbol{G}$ and $\boldsymbol{R}$ relate to the bivariate processes $X_M(t)$ and $X_{GI}(t)$. Hence it would be useful if we could extend the relationship (6.9.5) to the doubly infinite processes.

Asmussen and Ramaswami realised that it is difficult to find a direct relationship between $X_M(t)$ and $X_{GI}(t)$ since they are both processes on doubly infinite state spaces. In order to obtain a relationship that could be used to derive Ramaswami's duality result, Asmussen and Ramaswami considered the discrete time processes $K^M_n, n \geq 0$ and $K^{GI}_n, n \geq 0$ which are defined to be the processes of changes in levels at each transition point for $X_M(t)$ and $X_{GI}(t)$ respectively. The processes K^M_n and K^{GI}_n are such that $K^M_n \geq -1, n \geq 0$ and $K^{GI}_n \leq 1, n \geq 0$. That is, these processes reside in singly infinite state spaces. It is this property that enabled Asmussen and Ramaswami to obtain their relationship (3.6) which they then use to derive Ramaswami's duality result. It is straightforward to use equation (3.6) in [7] to derive the following result for the Markov processes $X_M(t)$ and $X_{GI}(t)$;

$$\begin{aligned}
&[\boldsymbol{a}]_{j_0} P^M_{0,j_0}((j_1, K_1) \to (j_2, K_2) \to \ldots \to (j_n, K_n)) \\
&\qquad = [\boldsymbol{a}]_{j_n} P^{GI}_{0,j_n}((j_{n-1}, -K_n) \to \ldots \to (j_1, -K_2) \to (j_0, -K_1)).
\end{aligned} \qquad (6.9.6)$$

So by introducing the processes K^M_n and K^{GI}_n, Asmussen and Ramaswami derived the relationship (6.9.6) which is essentially a time reversal result. On the left hand side of (6.9.6) we consider $X^M(t)$ moving sequentially through the phases $j_0, j_1, \ldots, j_n$ and making jumps between levels of sizes $K_1, K_2, \ldots, K_n$. On the right hand side of (6.9.6) we consider $X_{GI}(t)$ moving through the phases $j_n, j_{n-1}, \ldots, j_0$ and making jumps between levels of sizes $K_n, K_{n-1}, \ldots, K_1$. That is, in moving from $X_M(t)$ to $X_{GI}(t)$, we reverse the original order of the phases and we reverse the order and the direction of the jump sizes between levels.

6.10 A probabilistic interpretation of Ramaswami's duality result

The need to consider the processes K_n^M and K_n^{GI} arises essentially because the doubly infinite processes $X_M(t)$ and $X_{GI}(t)$ are transient. However, in Section 6.2 we have seen how to construct and interpret time reverses for transient processes. In this section we show that $X_{GI}(t)$ is the time reverse of $X_M(t)$ with respect to a particular invariant measure. This enables us to give a probabilistic interpretation of Ramaswami's duality result. In the next section we will use the probabilistic interpretations presented here to define an alternative dual process. In the rest of the chapter, when we refer to an invariant measure for a process $X(t)$ with q-matrix $\boldsymbol{Q}$ and transition probability matrix $\boldsymbol{P}(t)$, we will mean an invariant measure for both $\boldsymbol{Q}$ and $\boldsymbol{P}(t)$.

Suppose $\boldsymbol{m} = (\ldots \boldsymbol{m}_{-1}, \boldsymbol{m}_0, \boldsymbol{m}_1, \ldots)$ is a positive vector satisfying

$$\sum_{\ell=0}^{\infty} \boldsymbol{m}_{\nu-\ell} \boldsymbol{A}_\ell = \boldsymbol{0} \qquad \nu \in \mathbb{Z}. \tag{6.10.1}$$

Then from Lemma 6.2.3 $\boldsymbol{m}$ is an invariant measure for $X_M(t)$ if the Markov process with q-matrix given by

$$q_M^R((k,j),(\ell,m)) = \frac{[\boldsymbol{m}_\ell]_m}{[\boldsymbol{m}_k]_j} q_M((\ell,m),(k,j)) \tag{6.10.2}$$

is regular. Using methods to those we have used earlier, it can be shown that

$$q_M^R((k,j),(\ell,m)) = q_M((k,j),(\ell,m)) \tag{6.10.3}$$

and hence $q_M^R((k,j),(\ell,m))$ is bounded above since $q_M((k,j),(\ell,m))$ is bounded above. Therefore the process with q-matrix (6.10.2) is regular.

Putting $\boldsymbol{m} = (\ldots \boldsymbol{a}, \boldsymbol{a}, \boldsymbol{a}, \ldots)$ we have

$$\sum_{\ell=0}^{\infty} \boldsymbol{m}_{\nu-\ell} \boldsymbol{A}_\ell \tag{6.10.4}$$

$$= \sum_{\ell=0}^{\infty} \boldsymbol{a} \boldsymbol{A}_\ell \tag{6.10.5}$$

$$= \boldsymbol{a} \boldsymbol{A} \tag{6.10.6}$$

$$= \boldsymbol{0} \tag{6.10.7}$$

and hence $\boldsymbol{m} = (\ldots \boldsymbol{a}, \boldsymbol{a}, \boldsymbol{a}, \ldots)$ is an invariant measure for the doubly infinite M/G/1–type process. It is easy to see that the time reverse with respect to this invariant measure, of the doubly infinite M/G/1–type process is equal to $X_{GI}(t)$. Hence we have that Ramaswami's dual process of the original doubly infinite process is the time reverse of the original process with respect to the invariant measure $\boldsymbol{m} = (\ldots \boldsymbol{a}, \boldsymbol{a}, \boldsymbol{a}, \ldots)$. This is a much stronger relationship than the relationship (6.9.6) derived by Asmussen and Ramaswami.

Now note that for the doubly infinite M/G/1–type process we can still define the matrix $\boldsymbol{G}$ in the same way as we did for the singly infinite process. $\boldsymbol{G}$ will then be the minimum non negative solution to equation (6.3.3). Similarly, for the doubly infinite GI/M/1–type process we can define the matrix $\boldsymbol{R}$ in the same way as we did for the singly infinite process and $\boldsymbol{R}$ will then be the minimum non negative solution to equation (6.4.5). Having defined $\boldsymbol{G}$ and $\boldsymbol{R}$, the same algebraic arguments as those used in the proof of Theorem 6.8.1 give us that

$$\boldsymbol{R} = \boldsymbol{\Delta}^{-1}\boldsymbol{G}'\boldsymbol{\Delta} \tag{6.10.8}$$

or equivalently

$$a_i[\boldsymbol{R}]_{ij} = a_j[\boldsymbol{G}]_{ji} \qquad \forall i, j. \tag{6.10.9}$$

By appealing to the particle system interpretation of an invariant measure given in Section 6.2 it is possible to give a probabilistic interpretation of equation (6.10.9).

Suppose $X_M(t)$ starts at time $t = 0$ with N_j particles in phase j (at each level), where N_j are independent Poisson random variables with mean a_j, and we let the particles move between the states according to a Markov process with q-matrix given by (6.9.1). Similarly, suppose $X_{GI}(t)$ starts at time $t = 0$ with N_j particles in phase j (at each level), where N_j are independent Poisson random variables with mean a_j, and we let the particles move between the states according to a Markov process with q-matrix given by (6.9.2). From Lemma 6.2.4 we have that for any $\tau > 0$, the distribution of particles at time $t > 0$ in the process $X_M(\tau - t)$ is the same as the distribution of particles at time $t > 0$ in the process $X_{GI}(t)$. That is, we can say that the distribution of particles in the forward time process is the same as the distribution of particles in the time reverse process.

To assist in obtaining a probabilistic interpretation of equation (6.10.9) we consider the jump chains $X_M^J(n)$ and $X_{GI}^J(n)$ of the processes $X_M(t)$ and $X_{GI}(t)$ respectively.

$X_M^J(n)$ is a discrete time Markov chain with transition probability matrix given by

$$P_M^J = \begin{pmatrix} \ddots & \vdots & \vdots & \vdots & \vdots & \\ \cdots & \bar{A}_1 & \bar{A}_2 & \bar{A}_3 & \bar{A}_4 & \cdots \\ \cdots & \bar{A}_0 & \bar{A}_1 & \bar{A}_2 & \bar{A}_3 & \cdots \\ \cdots & 0 & \bar{A}_0 & \bar{A}_1 & \bar{A}_2 & \cdots \\ \cdots & 0 & 0 & \bar{A}_0 & \bar{A}_1 & \cdots \\ & \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix} \tag{6.10.10}$$

where $\bar{A}_i = \Phi^{-1}A_i + \delta_{i1}I, i \geq 0$ where $\Phi = -\text{diag}([A_1]_{ii})$. $X_{GI}^J(n)$ is a discrete time Markov chain with transition probability matrix given by

$$P_{GI}^J = \begin{pmatrix} \ddots & \vdots & \vdots & \vdots & \vdots & \\ \cdots & \bar{D}_1 & \bar{D}_0 & 0 & 0 & \cdots \\ \cdots & \bar{D}_2 & \bar{D}_1 & \bar{D}_0 & 0 & \cdots \\ \cdots & \bar{D}_3 & \bar{D}_2 & \bar{D}_1 & \bar{D}_0 & \cdots \\ \cdots & \bar{D}_4 & \bar{D}_3 & \bar{D}_2 & \bar{D}_1 & \cdots \\ & \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix} \tag{6.10.11}$$

where $\bar{D}_i = \Phi^{-1}D_i + \delta_{i1}I, i \geq 0$ where $\Phi = -\text{diag}([D_1]_{ii})$. It is clear that the particle interpretation given above also holds for the jump chains. That is, for any $\tau > 0$, the distribution of particles at time $n > 0$ in the process $X_M^J(\tau - n)$ is the same as the distribution of particles at time $t > 0$ in the process $X_{GI}^J(n)$.

We define the matrices G^J and R^J for the processes $X_M^J(n)$ and $X_{GI}^J(n)$ in a similar way to how we defined the matrices G and R for the processes $X_M(t)$ and $X_{GI}(t)$. From the probabilistic interpretations of the matrices G^J and G, it is clear that $G^J = G$. By applying the same arguments used to prove equation (2.4.20) in Chapter 2 , we can show that

$$[R]_{ij} = \frac{[\Phi]_{ii}[R^J]_{ij}}{[\Phi]_{jj}} \tag{6.10.12}$$

Consider the set of all sample paths for $X_M^J(n)$ that start in state (k, j) at time zero and first visit level $k - 1$ at time $m \geq 1$ and do so through state $(k - 1, i)$. Call this set of sample paths $\mathcal{A}_m$. If we define the set $\mathcal{A}$ by $\mathcal{A} = \bigcup_{m=1}^{\infty} \mathcal{A}_m$ then from the definition of G^J and since $G^J = G$, we have that the probability mass of the sample paths in the set $\mathcal{A}$ is equal to $[G]_{ji}$. Now consider the set of all sample paths for $X_{GI}^J(n)$ that start in state

$(k-1,i)$, are in state (k,j) at time $m \geq 1$ and have not visited level $k-1$ in between. Denote this set of sample paths by $\mathcal{B}_m$. If we define the set $\mathcal{B}$ by $\mathcal{B} = \bigcup_{m=1}^{\infty} \mathcal{B}_m$ then the probability mass of the sample paths in the set $\mathcal{B}$ is equal to $[\boldsymbol{R}^J]_{ij}$. Hence from equation (6.10.12), we see that to obtain $[\boldsymbol{R}]_{ij}$, we need to multiply the probability mass of the sample paths in $\mathcal{B}$ by $[\boldsymbol{\Phi}]_{ii}/[\boldsymbol{\Phi}]_{jj}$.

Observe that each sample path in the set $\mathcal{B}$ is the time reverse of a sample path in the set $\mathcal{A}$. That is, if a sample path in $\mathcal{B}$ consists of the sequence of states $(k-1,i),(k_1,j_1),\ldots,(k_N,j_N),(k,j)$ then the sample path consisting of the sequence of states $(k,j),(k_N,j_N),\ldots,(k_1,j_1),(k-1,i)$ is in set $\mathcal{A}$.

Since the distribution of particles in $X_M^J(\tau-n)$ is the same as the distribution of particles at time $n>0$ in the process $X_{GI}^J(n)$, it is reasonable to expect that there is some relationship between $[\boldsymbol{G}^J]_{ji}$ and $[\boldsymbol{R}^J]_{ij}$ and hence some relationship between $[\boldsymbol{G}]_{ji}$ and $[\boldsymbol{R}]_{ij}$. Now although the distribution of particles in the process $X_M^J(\tau-n)$ is the same as the distribution of particles in $X_{GI}^J(n)$, the probability mass of the paths in $\mathcal{A}$ will not in general be equal to the probability mass of paths in $\mathcal{B}$. That is, $[\boldsymbol{G}^J]_{ji}$ is not equal to $[\boldsymbol{R}^J]_{ij}$. Similarly, for the continuous time processes $X_M(t)$ and $X_{GI}(t)$, $[\boldsymbol{G}]_{ji}$ is not equal to $[\boldsymbol{R}]_{ij}$. However, $[\boldsymbol{G}]_{ji}$ and $[\boldsymbol{R}]_{ij}$ are related by equation (6.10.9) which can be interpreted as a set of "flux" equations.

Since a_j is the distribution of particles in state (k,j) we can define $a_j[\boldsymbol{G}]_{ji}$ as the "downward flux" of particles under the invariant measure in $X_M(t)$, from state (k,j) to state $(k-1,i)$ avoiding levels $k-1$ and below. Similarly, define $a_i[\boldsymbol{R}]_{ij}$ to be the "upward flux" of particles under the invariant measure in $X_{GI}(t)$, from state $(k-1,i)$ to state (k,j) avoiding levels $k-1$ and below. Hence equation (6.10.9) can be interpreted as stating that the upward flux in the time-reverse process is equal to the downward flux in the forward time process. That is, the downward flux in the original process is balanced by the upward flux in the dual process.

6.11 An alternative dual process

In the previous section we saw that Ramaswami's dual of a doubly infinite M/G/1–type process is the time reverse with respect to a particular invariant measure. This observation suggests that if we take the time reverse with respect to a different invariant measure then we may obtain a different dual process. In this section we derive an alternative dual process in this fashion.

Consider a doubly infinite level independent M/G/1–type process with q-matrix given by (6.9.1). We showed in the previous section that a positive vector $\boldsymbol{m} = (\ldots \boldsymbol{m}_{-1}, \boldsymbol{m}_0, \boldsymbol{m}_1, \ldots)$ is an invariant measure for this process if it is a solution to

$$\sum_{\ell=0}^{\infty} \boldsymbol{m}_{\nu-\ell} \boldsymbol{A}_\ell = \boldsymbol{0} \qquad \nu \in \mathbb{Z}. \tag{6.11.1}$$

Let ϕ be the spectral radius of $\boldsymbol{G}$. From results in Neuts [73] it can be shown that the matrix $\sum_{\ell=0}^{\infty} \boldsymbol{A}_\ell \phi^\ell$ has spectral radius equal to zero. Hence there exists a positive vector $\boldsymbol{\alpha}$ such that

$$\boldsymbol{\alpha} \left(\sum_{\ell=0}^{\infty} \boldsymbol{A}_\ell \phi^\ell \right) = \boldsymbol{0}. \tag{6.11.2}$$

If we put $\boldsymbol{m}_k = \boldsymbol{\alpha} \left(\frac{1}{\phi} \right)^k$, $k \in \mathbb{Z}$ where $\boldsymbol{\alpha}$ is the solution to (6.11.2) such that $\boldsymbol{\alpha e} = 1$ then

$$\sum_{\ell=0}^{\infty} \boldsymbol{m}_{\nu-\ell} \boldsymbol{A}_\ell \tag{6.11.3}$$

$$= \sum_{\ell=0}^{\infty} \boldsymbol{\alpha} \left(\frac{1}{\phi} \right)^{\nu-\ell} \boldsymbol{A}_\ell \tag{6.11.4}$$

$$= \left(\frac{1}{\phi} \right)^{\nu} \boldsymbol{\alpha} \left(\sum_{\ell=0}^{\infty} \boldsymbol{A}_\ell \phi^\ell \right) \tag{6.11.5}$$

$$= \boldsymbol{0} \tag{6.11.6}$$

where the last equality results from equation (6.11.2). Hence $\boldsymbol{m}_k = \boldsymbol{\alpha} \left(\frac{1}{\phi} \right)^k$, $k \in \mathbb{Z}$ is an invariant measure for the doubly infinite M/G/1–type process. Note that when $\phi = 1$, that is, when the process is recurrent, the above invariant measure is the same as the invariant measure used in the construction of Ramaswami's dual.

The time-reverse with respect to the invariant measure $\boldsymbol{m}_k = \boldsymbol{\alpha} \left(\frac{1}{\phi} \right)^k$ of the doubly infinite M/G/1–type process is a doubly infinite GI/M/1–type process with q-matrix of the same form as (6.9.2) where

$$\boldsymbol{D}_\ell = \phi^{\ell-1} \boldsymbol{\Delta}^{-1} \boldsymbol{A}_\ell' \boldsymbol{\Delta} \qquad \ell \geq 0 \tag{6.11.7}$$

with $\boldsymbol{\Delta} = \mathrm{diag}(\boldsymbol{\alpha})$. The discussion in Section 6.10 suggests that the $\boldsymbol{R}$ matrix for the level independent GI/M/1–type process with matrices $\boldsymbol{D}_\ell$ given by (6.11.7) and the $\boldsymbol{G}$

matrix for the original M/G/1–type process may satisfy a duality relationship similar to equation (6.8.2). In the following theorem we see that such a relationship does hold.

Theorem 6.11.1 *Consider a level independent M/G/1–type process with q-matrix given by (6.3.2). Let* $\boldsymbol{A} = \sum_{\ell=0}^{\infty} \boldsymbol{A}_\ell$ *and assume that* $\boldsymbol{A}$ *is irreducible. Further, let* ϕ *be the spectral radius of* $\boldsymbol{G}$*, which is the minimal non negative solution to equation (6.3.3) and let* $\boldsymbol{\alpha}$ *be the solution to*

$$\boldsymbol{\alpha}\left(\sum_{\ell=0}^{\infty} \boldsymbol{A}_\ell \phi^\ell\right) = \mathbf{0} \tag{6.11.8}$$

subject to $\boldsymbol{\alpha e} = 1$.

If we define $\boldsymbol{D}_\ell$ *by*

$$\boldsymbol{D}_\ell = \phi^{\ell-1} \boldsymbol{\Delta}^{-1} \boldsymbol{A}_\ell' \boldsymbol{\Delta} \qquad \ell \geq 0, \tag{6.11.9}$$

where $\boldsymbol{\Delta} = diag(\boldsymbol{\alpha})$ *then the matrix* $\boldsymbol{D} = \sum_{\ell=0}^{\infty} \boldsymbol{D}_\ell$ *is an irreducible q-matrix and the matrix* $\boldsymbol{R}$ *for the GI/M/1–type process with* $\boldsymbol{D}_\ell$ *given by (6.11.9) is given by*

$$\boldsymbol{R} = \frac{1}{\phi} \boldsymbol{\Delta}^{-1} \boldsymbol{G}' \boldsymbol{\Delta}. \tag{6.11.10}$$

Proof: The matrix $\boldsymbol{D}$ is given by

$$\begin{aligned} \boldsymbol{D} &= \boldsymbol{\Delta}^{-1} \sum_{\ell=0}^{\infty} \phi^{\ell-1} \boldsymbol{A}_\ell' \boldsymbol{\Delta} \\ &= \boldsymbol{\Delta}^{-1} \left(\sum_{\ell=0}^{\infty} \phi^{\ell-1} \boldsymbol{A}_\ell\right)' \boldsymbol{\Delta}. \end{aligned}$$

Hence $\boldsymbol{D}$ is irreducible since $\boldsymbol{A}$ is irreducible and $\boldsymbol{D}$ has non negative elements. We have that

$$\begin{aligned} \boldsymbol{De} &= \boldsymbol{\Delta}^{-1} \sum_{\ell=0}^{\infty} \phi^{\ell-1} \boldsymbol{A}_\ell' \boldsymbol{\Delta e} \\ &= \boldsymbol{\Delta}^{-1} \sum_{\ell=0}^{\infty} \phi^{\ell-1} \boldsymbol{A}_\ell' \boldsymbol{\alpha}' \\ &= \boldsymbol{\Delta}^{-1} \frac{1}{\phi} \left[\boldsymbol{\alpha} \left(\sum_{\ell=0}^{\infty} \boldsymbol{A}_\ell \phi^\ell\right)\right]' \\ &= \mathbf{0} \end{aligned}$$

where the last equality follows from equation (6.11.8). Hence $\boldsymbol{D}$ is a q-matrix.

From results in Sections 6.3 and 6.4 we have that the $\boldsymbol{R}$ matrix for the GI/M/1–type process with $\boldsymbol{D}_\ell$ given by (6.11.9) is the minimal non negative solution to

$$0 = \sum_{\ell=0}^{\infty} \boldsymbol{R}^\ell \phi^{\ell-1} \boldsymbol{\Delta}^{-1} \boldsymbol{A}'_\ell \boldsymbol{\Delta}. \tag{6.11.11}$$

The matrix $\boldsymbol{G}$ for the M/G/1–type chain is the minimal non negative solution to

$$0 = \sum_{\ell=0}^{\infty} \boldsymbol{A}_\ell \boldsymbol{G}^\ell. \tag{6.11.12}$$

Pre-multiplying (6.11.11) by $\phi\boldsymbol{\Delta}$, post-multiplying by $\boldsymbol{\Delta}^{-1}$ and taking transposes of both sides gives

$$0 = \sum_{\ell=0}^{\infty} \boldsymbol{A}_\ell (\phi \boldsymbol{\Delta}^{-1} \boldsymbol{R}' \boldsymbol{\Delta})^\ell. \tag{6.11.13}$$

Hence we have that $\phi\boldsymbol{\Delta}^{-1}\boldsymbol{R}'\boldsymbol{\Delta}$ is a non negative solution to (6.11.12) and since $\boldsymbol{G}$ is the minimal non negative solution to (6.11.12), we have that

$$\phi \boldsymbol{\Delta}^{-1} \boldsymbol{R}' \boldsymbol{\Delta} \geq \boldsymbol{G} \tag{6.11.14}$$

which, since $\boldsymbol{\alpha}$ is positive implies that

$$\boldsymbol{R} \geq \frac{1}{\phi} \boldsymbol{\Delta}^{-1} \boldsymbol{G}' \boldsymbol{\Delta}. \tag{6.11.15}$$

Now observe that

$$\sum_{\ell=0}^{\infty} (\frac{1}{\phi} \boldsymbol{\Delta}^{-1} \boldsymbol{G}' \boldsymbol{\Delta})^\ell \phi^{\ell-1} \boldsymbol{\Delta}^{-1} \boldsymbol{A}'_\ell \boldsymbol{\Delta} = \frac{1}{\phi} \boldsymbol{\Delta}^{-1} \left(\sum_{\ell=0}^{\infty} \boldsymbol{A}_\ell \boldsymbol{G}^\ell \right)' \boldsymbol{\Delta} \tag{6.11.16}$$

$$= 0 \tag{6.11.17}$$

where the last equality follows from equation (6.11.12). Hence we have that $\frac{1}{\phi}\boldsymbol{\Delta}^{-1}\boldsymbol{G}'\boldsymbol{\Delta}$ is a non negative solution to (6.11.11) and hence since $\boldsymbol{R}$ is the minimal non negative solution to (6.11.11), we have that

$$\frac{1}{\phi} \boldsymbol{\Delta}^{-1} \boldsymbol{G}' \boldsymbol{\Delta} \geq \boldsymbol{R} \tag{6.11.18}$$

which together with equation (6.11.15) implies that equation (6.11.10) holds. □

We propose the GI/M/1–type process defined in Theorem 6.11.1 as an alternative dual process to Ramaswami's dual. However, note that our alternative dual is only different to Ramaswami's dual when the process is not positive recurrent.

Theorem 6.8.5 states that the ergodic character of Ramaswami's dual process of a M/G/1–type process depends upon the ergodic character of the original process. The following theorem shows that the ergodic character of our dual process is independent of the ergodic character of the original process.

Theorem 6.11.2 *Our dual process of a level independent M/G/1–type process is either transient or null recurrent.*

Proof: It is clear from equation (6.11.10) that the spectral radius of $\boldsymbol{R}$ is alway equal to one and hence the result follows from Theorem 6.4.2. □

Now consider a doubly-infinite GI/M/1–type process with q-matrix of the form given in equation (6.9.2). It is easy to show that a positive vector $\boldsymbol{m} = (\ldots \boldsymbol{m}_{-1}, \boldsymbol{m}_0, \boldsymbol{m}_1, \ldots)$ is an invariant measure for this process if it is a solution to

$$\sum_{\ell=0}^{\infty} \boldsymbol{m}_{\ell-\nu} \boldsymbol{D}_\ell = \boldsymbol{0} \qquad \nu \in \mathbb{Z}. \tag{6.11.19}$$

Suppose θ is the spectral radius of $\boldsymbol{R}$ and $\boldsymbol{\beta}$ is the corresponding left eigenvector. By pre-multiplying equation (6.4.5) by $\boldsymbol{\beta}$ we obtain

$$\boldsymbol{\beta}\left(\sum_{\ell=0}^{\infty} \boldsymbol{D}_\ell \theta^\ell\right) = \boldsymbol{0}. \tag{6.11.20}$$

Using equation (6.11.20) it follows that $\boldsymbol{m}_k = \boldsymbol{\beta}\theta^k$, $k \in \mathbb{Z}$ is an invariant measure for the doubly infinite GI/M/1–type process. Note that when $\theta = 1$, that is, when the process is transient or null recurrent, the above invariant measure is the same as the invariant measure used in the construction of Ramaswami's dual. The time-reverse of the doubly infinite GI/M/1–type process with respect to the invariant measure $\boldsymbol{m}_k = \boldsymbol{\beta}\theta^k$, $k \in \mathbb{Z}$ is a doubly infinite M/G/1–type process with q-matrix of the same form as (6.9.1) where

$$\boldsymbol{A}_\ell = \theta^{\ell-1} \boldsymbol{\Delta}^{-1} \boldsymbol{D}'_\ell \boldsymbol{\Delta} \qquad \ell \geq 0 \tag{6.11.21}$$

where $\boldsymbol{\Delta} = \mathrm{diag}(\boldsymbol{\beta})$. The following theorem gives a duality relationship between the $\boldsymbol{G}$ matrix for the M/G/1–type process with matrices $\boldsymbol{A}_\ell$ given by (6.11.21) and the $\boldsymbol{R}$ matrix for the original level independent GI/M/1–type process. The proof is similar to that of Theorem 6.11.1.

Theorem 6.11.3 *Consider a level independent GI/M/1–type process with q-matrix given by (6.4.2). Let* $\boldsymbol{D} = \sum_{\ell=0}^{\infty} \boldsymbol{D}_\ell$ *and assume that* $\boldsymbol{D}$ *is irreducible. Further, let* $\boldsymbol{R}$ *be the* *minimal non negative solution to equation (6.4.2) and denote the spectral radius of* $\boldsymbol{R}$ *by* θ *and the corresponding left eigenvector by* $\boldsymbol{\beta}$.

If we define $\boldsymbol{A}_\ell$ *by*

$$\boldsymbol{A}_\ell = \theta^{\ell-1}\boldsymbol{\Delta}^{-1}\boldsymbol{D}'_\ell\boldsymbol{\Delta} \qquad \ell \geq 0, \tag{6.11.22}$$

where $\boldsymbol{\Delta} = diag(\boldsymbol{\beta})$ *then the matrix* $\boldsymbol{A} = \sum_{\ell=0}^{\infty}\boldsymbol{D}_\ell$ *is an irreducible q-matrix and the matrix* $\boldsymbol{G}$ *for the M/G/1–type process with* $\boldsymbol{A}_\ell$ *given by (6.11.22) is given by*

$$\boldsymbol{G} = \frac{1}{\theta}\boldsymbol{\Delta}^{-1}\boldsymbol{R}'\boldsymbol{\Delta}. \tag{6.11.23}$$

We propose the M/G/1–type process defined in Theorem 6.11.3 as an alternative dual process to Ramaswami's dual of a level independent GI/M/1–type process.

Theorem 6.8.6 states that the ergodic nature of Ramaswami's dual process of a GI/M/1–type process depends upon the ergodic nature of the original process. We show in the following theorem that our dual process is always recurrent.

Theorem 6.11.4 *Our dual process of a level independent GI/M/1–type process is always recurrent.*

Proof: It is clear from equation (6.11.23) that the spectral radius of $\boldsymbol{G}$ is always equal to one and hence the result follows from Theorem 6.3.2. ☐

We note that our dual of a M/G/1–type process and our dual of a GI/M/1–type process are both constructed in essentially the same way. In constructing our dual of a M/G/1–type process we find $\boldsymbol{\alpha}$ which is the left eigenvector of $\sum_{\ell=0}^{\infty}\boldsymbol{A}_\ell\phi^\ell$ corresponding to the eigenvalue zero. When constructing our dual of a GI/M/1–type process we find $\boldsymbol{\beta}$ which is the left eigenvector of $\sum_{\ell=0}^{\infty}\boldsymbol{D}_\ell\theta^\ell$ corresponding to the eigenvalue zero. Equation (6.11.22) is the same as equation (6.11.9) with $\boldsymbol{D}_\ell$ replaced by $\boldsymbol{A}_\ell$, ϕ replaced by θ and $\boldsymbol{\alpha}$ replaced by $\boldsymbol{\beta}$.

6.12 A dual process for level dependent M/G/1–type and GI/M/1–type processes

In this section we construct a dual process for level dependent M/G/1-type and GI/M/1–type processes. We do this by considering the time reverse interpretations of the dual processes for level independent M/G/1–type and GI/M/1–type processes.

It seems that Ramaswami's dual process for level independent M/G/1–type and GI/M/1–type processes cannot easily be extended to level dependent processes. The same statement holds for the alternative dual processes presented in Section 6.11. We have seen in the preceding two sections that both Ramaswami's dual process and the alternative dual processes constructed in Section 6.11 can be constructed in the following manner. Firstly the (singly infinite) process is extended to a doubly infinite process and then the time reverse of the doubly infinite process is taken with respect to a certain invariant measure. For Ramaswami's dual the time reverse is taken with respect to the invariant measure $\boldsymbol{m} = (\ldots \boldsymbol{a}, \boldsymbol{a}, \boldsymbol{a}, \ldots)$ and for the dual processes in Section 6.11 the invariant measures are given by $\boldsymbol{m}_k = \boldsymbol{\alpha}(1/\phi)^k, k \in \mathbb{Z}$ and $\boldsymbol{m}_k = \boldsymbol{\beta}\theta^k, k \in \mathbb{Z}$ respectively. The dual process is then defined to be the M/G/1-type (GI/M/1-type) process with the matrices $\boldsymbol{A}_\ell$ ($\boldsymbol{D}_\ell$) equal to the corresponding matrices for the time reverse process.

The discussion above suggests that an alternative dual process could simply be the time reverse of the singly infinite process. Of course, Theorems 6.7.1 and 6.7.2 tell us that the time reverse will in general be a level dependent process, and level independent if and only if the invariant measure has the phase independent of level property. We note here that when we consider doubly infinite processes, the phase and level variables are always independent and hence the time reverse is always level independent. Hence, when constructing dual processes for level independent processes, the reason that we consider the doubly infinite process is so that when we take the time reverse we will obtain a level independent dual process. When constructing a dual process for level dependent processes, in general we don't expect the dual to be level independent and hence we can simply take the dual to be the time reverse with respect to an invariant measure.

The following theorem gives a duality relationship relating the $\boldsymbol{G}_k$ matrices for a level dependent M/G/1–type process to the $\boldsymbol{R}_k$ matrices for its time reverse.

Theorem 6.12.1 *If $X_M(t)$ is a level dependent M/G/1–type process with q-matrix given by (6.3.1) and invariant measure $\boldsymbol{m} = (\boldsymbol{m}_0, \boldsymbol{m}_1, \ldots)$, then the matrices $\{\boldsymbol{R}_k, k \geq 0\}$ for the time-reverse process $X_M^R(t)$ are given by*

$$\boldsymbol{R}_k = \boldsymbol{\Delta}_k^{-1}\boldsymbol{G}'_{k+1}\boldsymbol{\Delta}_{k+1} \tag{6.12.1}$$

where the matrices $\{\boldsymbol{G}_k, k \geq 1\}$ are the minimal non negative solution to the equations (6.3.3) and $\boldsymbol{\Delta}_k = diag(\boldsymbol{m}_k)$.

Proof:

Recall that the matrices $\{\boldsymbol{G}_k, k \geq 1\}$ are the minimal non negative solutions to the equations

$$\mathbf{0} = \boldsymbol{A}_0^{(k)} + \sum_{\ell=0}^{\infty} \boldsymbol{A}_{\ell+1}^{(k)} \boldsymbol{G}_{k+\ell} \dots \boldsymbol{G}_k \qquad k \geq 1 \tag{6.12.2}$$

and the matrices $\{\boldsymbol{R}_k,\ k \geq 0\}$ are the minimal non negative solutions to the equations

$$\begin{aligned}\mathbf{0} &= \boldsymbol{D}_0^{(k)} + \sum_{\ell=0}^{\infty} \boldsymbol{R}_k \dots \boldsymbol{R}_{k+\ell} \boldsymbol{D}_{\ell+1}^{(k+\ell+1)} \\ &= \boldsymbol{\Delta}_k^{-1}(\boldsymbol{A}_0^{(k+1)})' \boldsymbol{\Delta}_{k+1} + \sum_{\ell=0}^{\infty} \boldsymbol{R}_k \dots \boldsymbol{R}_{k+\ell} \boldsymbol{\Delta}_{k+\ell+1}^{-1}(\boldsymbol{A}_{\ell+1}^{(k+1)})' \boldsymbol{\Delta}_{k+1} \qquad k \geq 0\end{aligned} \tag{6.12.3}$$

since the matrices $\boldsymbol{D}_\ell^{(k)}$ are given by equations (6.5.3)-(6.5.5).

Now if we substitute $\boldsymbol{G}_k = \boldsymbol{\Delta}_k^{-1} \boldsymbol{R}_{k-1}' \boldsymbol{\Delta}_{k-1},\ k \geq 1$ into the right hand side of (6.12.2) we obtain

$$\boldsymbol{A}_0^{(k)} + \sum_{\ell=0}^{\infty} \boldsymbol{A}_{\ell+1}^{(k)} \boldsymbol{\Delta}_{k+\ell}^{-1} \boldsymbol{R}_{k+\ell-1}' \dots \boldsymbol{R}_{k-1}' \boldsymbol{\Delta}_{k-1} \tag{6.12.4}$$

$$= \boldsymbol{A}_0^{(k)} + \boldsymbol{\Delta}_k^{-1} \left[\sum_{\ell=0}^{\infty} \boldsymbol{R}_{k-1} \dots \boldsymbol{R}_{k+\ell-1} \boldsymbol{\Delta}_{k+\ell}^{-1} (\boldsymbol{A}_{\ell+1}^{(k)})' \boldsymbol{\Delta}_k\right]' \boldsymbol{\Delta}_{k-1} \tag{6.12.5}$$

$$= \boldsymbol{\Delta}_k^{-1} \left[\boldsymbol{\Delta}_{k-1}^{-1} (\boldsymbol{A}_0^{(k)})' \boldsymbol{\Delta}_k + \sum_{\ell=0}^{\infty} \boldsymbol{R}_{k-1} \dots \boldsymbol{R}_{k+\ell-1} \boldsymbol{\Delta}_{k+\ell}^{-1} \boldsymbol{A}_{\ell+1}^{(k)} \boldsymbol{\Delta}_k\right]' \boldsymbol{\Delta}_{k-1} \tag{6.12.6}$$

$$= \mathbf{0} \tag{6.12.7}$$

where we have used equation (6.12.3) to obtain the final equality. Hence $\{\boldsymbol{\Delta}_k^{-1} \boldsymbol{R}_{k-1}' \boldsymbol{\Delta}_{k-1}, k \geq 1\}$ is a solution to the equations (6.12.2). Now since $\{\boldsymbol{G}_k, k \geq 1\}$ are the minimal non negative solutions to (6.12.2), we must have

$$\boldsymbol{\Delta}_k^{-1} \boldsymbol{R}_{k-1}' \boldsymbol{\Delta}_{k-1} \geq \boldsymbol{G}_k \qquad k \geq 1 \tag{6.12.8}$$

which, since $\boldsymbol{m}_k$ is positive, implies that

$$\boldsymbol{R}_{k-1} \geq \boldsymbol{\Delta}_{k-1}^{-1} \boldsymbol{G}_k' \boldsymbol{\Delta}_k \qquad k \geq 1. \tag{6.12.9}$$

Now if we substitute $\boldsymbol{R}_k = \boldsymbol{\Delta}_k^{-1} \boldsymbol{G}_{k+1}' \boldsymbol{\Delta}_{k+1},\ k \geq 0$ into the right hand side of (6.12.3) then we can show this equals the zero matrix. Hence $\{\boldsymbol{\Delta}_k^{-1} \boldsymbol{G}_{k+1}' \boldsymbol{\Delta}_{k+1}, k \geq 0\}$ is a solution to the equations (6.12.3) and since $\{\boldsymbol{R}_k, k \geq 0\}$ are the minimal non negative solutions to (6.12.3), we have that

$$\boldsymbol{\Delta}_k^{-1} \boldsymbol{G}'_{k+1} \boldsymbol{\Delta}_{k+1} \geq \boldsymbol{R}_k \qquad k \geq 0. \tag{6.12.10}$$

Equations (6.12.9) and (6.12.10) imply that

$$\boldsymbol{R}_k = \boldsymbol{\Delta}_k^{-1} \boldsymbol{G}'_{k+1} \boldsymbol{\Delta}_{k+1}. \tag{6.12.11}$$

☐

It is possible to interpret the duality relationship (6.12.1) as a set of flux equations in the same way that we interpreted equation (6.10.8) in Section 6.10.

We note that the discrete time analogue of equation (6.12.1) was stated in Section 3.3 and used to obtain an explicit expression for the matrix $\boldsymbol{R}_k$.

We can prove the following theorem in the same way as we proved Theorem 6.12.1.

Theorem 6.12.2 *If $X_{GI}(t)$ is a level dependent GI/M/1–type process with q-matrix given by (6.4.1) and invariant measure $\boldsymbol{m} = (\boldsymbol{m}_0, \boldsymbol{m}_1, \dots)$, then the matrices $\{\boldsymbol{G}_k, k \geq 1\}$ for the time-reverse process $X_{GI}^R(t)$ are given by*

$$\boldsymbol{G}_k = \boldsymbol{\Delta}_k^{-1} \boldsymbol{R}'_{k-1} \boldsymbol{\Delta}_{k-1} \tag{6.12.12}$$

where the matrices $\{\boldsymbol{R}_k, k \geq 0\}$ are the minimal non negative solution to the equations (6.4.5) and $\boldsymbol{\Delta}_k = diag(\boldsymbol{m}_k)$.

Chapter 7

Quasi-stationary and limiting-conditional distributions for continuous time QBDs

7.1 Introduction

In this Chapter we consider quasi-stationary distributions for continuous time QBDs. Bean, Bright, Latouche, Pearce, Pollett and Taylor [8] considered quasi-stationary distributions for discrete time QBDs and the results we present here are continuous time analogues of their results. Before proceeding we give an informal discussion of quasi-stationary distributions.

Quasi-stationary distributions arise when we consider stochastic processes that consist of a single transient communicating class and one or more absorbing states. When the time to absorption is long in such processes, it is possible for the process to exhibit stationary-like behaviour. When this is the case we say that a quasi-stationary distribution exists. Moreover, it may be the case that the probabilities of the process being in the transient states, conditional on absorption not having taken place approach a limit, known as a limiting-conditional distribution. We give more formal definitions of and discuss the relationship between quasi-stationary and limiting-conditional distributions in the next section.

Quasi-stationary distributions have been used in the modelling of a wide variety of physical systems. Dambrine and Moreau [16, 17] and Parsons and Pollett [75] used quasi-stationary distributions to model the concentration of a catalyst in chemical reactions where the catalyst can become exhausted. Quasi-stationary distributions have also been used to model a number of biological phenomena, see for example Klein [49], Pollett [76] and Day and Possingham [18]. Problems in queueing theory have also been analysed via quasi-stationary distributions, see Kyprianou [52, 53], and Makimoto [62].

The initial work on quasi-stationary distributions considered processes on a finite state space. For such processes a quasi-stationary distribution always exists. However, in the case of an infinite state space, a quasi-stationary distribution may not exist. Early work on infinite state space processes was done by Vere-Jones [92, 93] and Kingman [48]. Recently Kesten [43] has made very significant advances in this area.

To determine the quasi-stationary distribution for a process one needs to compute the decay parameter for the process. In general this is difficult to do. Some processes for which it is possible are the Galton-Watson branching process, simple birth-and-death processes and GI/M/1 queues (see Kyprianou [53]). Recently Kijima [45] derived an algebraic equation for the decay parameter of PH/PH/1 queues. In fact, Kijima's results also hold for the more general level independent M/G/1-type and GI/M/1-type processes. In [45] Kijima showed that for PH/PH/1 queues the algebraic equations can be solved by using the Laplace Stieltjes transforms of the inter arrival and service time distributions. Kijima also gave the form of the quasi-stationary distribution for the special cases of the M/PH/1 and PH/M/1 queues. Makimoto [62] extended the work of Kijima by giving an explicit expression for the quasi-stationary distribution of a PH/PH/c queue. Makimoto's expression was in terms of the solution to a matrix equation but Makimoto did not discuss methods for solving this equation.

Bean, Bright, Latouche, Pearce, Pollett and Taylor [8] extended the work of Kijima [45] and Makimoto [62] by considering limiting-conditional distributions for general discrete time QBDs. The class of PH/PH/c queues is included in this class. Bean, Bright, Latouche, Pearce, Pollett and Taylor used results in Kesten [43] to show under a mild assumption that the limiting-conditional distribution always exists. They gave an explicit expression for this distribution and developed an efficient computational procedure for computing the distribution.

Kijima [44] showed that the limiting-conditional distribution of a uniform, continuous time Markov process is equal to the limiting-conditional distribution of its uniformised process. Continuous time QBDs are uniform due to their homogeneity so to obtain the limiting-conditional distribution we can apply the results of Bean, Bright, Latouche, Pearce, Pollett and Taylor [8] to the uniformised process. In this chapter we consider the uniformised process but we state all our results directly in terms of the continuous time QBD. We also propose physical interpretations in terms of the continuous time QBD. We mention that it is possible to derive the results via a direct approach, that is not via the uniformised process. This would involve a similar analysis to that employed in [8].

We begin the chapter by presenting some definitions and results for quasi-stationary and limiting-conditional distributions for continuous time Markov processes. We then consider continuous time absorbing QBDs and derive the form of the limiting-conditional distribution. Finally we give a computational procedure for computing the limiting-conditional distribution.

7.2 Quasi-stationary and limiting-conditional distributions for continuous time Markov processes

In this section we give some background theory for quasi-stationary and limiting-conditional distributions for continuous time Markov processes. For more details we cite [93], [26], [77], [64] and [5].

Consider a continuous time Markov process $\{X(t), t \geq 0\}$ on a state space $\mathcal{S} = \{0, 1, \ldots\}$ with q-matrix given by $\boldsymbol{Q} = (q_{ij}, i, j \in \mathcal{S})$. We assume that $\boldsymbol{Q}$ is stable and conservative, that is, $q_i = -q_{ii} < +\infty \; \forall i$ and $\sum_{j \in \mathcal{S}} q_{ij} = 0 \; \forall i$.

Assume $\boldsymbol{Q}$ has the form

$$\boldsymbol{Q} = \begin{pmatrix} 0 & \mathbf{0}' \\ q_0 & \hat{\boldsymbol{Q}} \end{pmatrix} \tag{7.2.1}$$

where $\mathbf{0}'$ is a row vector of zeros and q_0 is a column vector. It is clear that state 0 is an absorbing state. We assume that the set of states $\mathcal{C} = \{1, 2, \ldots\}$ is an irreducible and aperiodic communicating class. Furthermore, we assume that $q_{i0} > 0$ for at least one $i \in \mathcal{C}$ which implies that there is a positive probability of reaching state 0. Let $\boldsymbol{P}(t)$ be the transition probability matrix for $X(t)$. That is, $\boldsymbol{P}(t) = (P_{ij}(t), i, j \in \mathcal{S})$ where $P_{ij}(t) = P(X(t) = j | X(0) = i)$, $\forall t \geq 0$. We define a quasi-stationary distribution as follows.

Definition 7.2.1 *Let* $\boldsymbol{\pi} = (\pi_1, \pi_2, \ldots)$ *be a positive vector such that* $\boldsymbol{\pi e} = 1$ *and define* $P_j(t), j \in \mathcal{S}, t \geq 0$ *by*

$$P_j(t) = \sum_{i \in \mathcal{C}} \pi_i P_{ij}(t). \tag{7.2.2}$$

$\boldsymbol{\pi}$ *is a quasi-stationary distribution on* $\mathcal{C}$ *if* $\forall t > 0$,

$$\frac{P_j(t)}{\sum_{i \in \mathcal{C}} P_i(t)} = \pi_j. \tag{7.2.3}$$

Equation (7.2.3) states that if $X(t)$ starts in $\mathcal{C}$ according to the distribution $\boldsymbol{\pi}$, then the probabilities of being in states in $\mathcal{C}$ at time t, conditional on absorption not having occurred, do not vary with time.

Closely related to the study of quasi-stationary distributions is the study of β–invariant measures. A β–invariant measure is defined as follows.

Definition 7.2.2 *A positive vector* $\boldsymbol{m} = (m_1, m_2, \ldots)$ *is a* β*–invariant measure on* $\mathcal{C}$ *for* $\boldsymbol{Q}$ *if*

$$\boldsymbol{mQ} = -\beta \boldsymbol{m}. \tag{7.2.4}$$

The relationship between quasi-stationary distributions and β–invariant measures is described by the following two lemmas. These lemmas are taken from Pollett [77] where they appear as Proposition 3 and Corollary 1 respectively.

Lemma 7.2.1 *If* $\boldsymbol{\pi} = (\pi_1, \pi_2, \ldots)$ *is a quasi-stationary distribution on* $\mathcal{C}$ *then* $\boldsymbol{\pi}$ *is a* β*–invariant measure on* $\mathcal{C}$ *for* $\boldsymbol{Q}$ *with*

$$\beta = \sum_{i \in \mathcal{C}} \pi_i q_{i0}. \tag{7.2.5}$$

Lemma 7.2.2 *If* $\boldsymbol{\pi} = (\pi_1, \pi_2, \ldots)$ *such that* $\boldsymbol{\pi e} = 1$ *is a* β*–invariant measure on* $\mathcal{C}$ *for* $\boldsymbol{Q}$ *then provided* $\sum_{j \in \mathcal{C}} \pi_j q_j < \infty$, $\boldsymbol{\pi}$ *is a quasi-stationary distribution on* $\mathcal{C}$ *and* β *is given by* (7.2.5).

We now define a limiting-conditional distribution.

Definition 7.2.3 *A positive vector* $\boldsymbol{\pi} = (\pi_1, \pi_2, \ldots)$ *such that* $\boldsymbol{\pi e} = 1$ *is said to be a limiting-conditional distribution if* $\forall i \in \mathcal{C}$

$$\lim_{t \to \infty} P(X(t) = j \mid X(0) = i \text{ and } X(t) \in \mathcal{C}) = \pi_j \qquad j \in \mathcal{C}, \tag{7.2.6}$$

independent of the initial state i.

In equation (7.2.6) we condition on $X(t)$ starting in state i. This is equivalent to $X(t)$ starting in $\mathcal{C}$ according to the degenerate distribution $\boldsymbol{e}_i$ where $\boldsymbol{e}_i$ is a vector with one in the ith position and zeros elsewhere. Thus a limiting-conditional distribution exists if the limit $\lim_{t\to\infty} P(X(t) = j \mid X(t) \in \mathcal{C}, \text{ and } X(0))$ is the same for all degenerate initial distributions. However, this does not necessarily say that $\lim_{t\to\infty} P(X(t) = j \mid X(t) \in \mathcal{C}, \text{ and } X(0))$ is invariant over the set of all possible initial distributions. In general, even when a limiting-conditional distribution exists, the left hand side of (7.2.6) may depend on the initial distribution. To see this recall that if $X(t)$ starts in $\mathcal{C}$ according to a quasi-stationary distribution then the probabilities of being in $\mathcal{C}$ will not vary with time. Hence if the quasi-stationary distribution is not a limiting-conditional distribution then the process cannot evolve to a limiting-conditional distribution. It is certainly possible for a process to have many quasi-stationary distributions. In fact Bean, Pollett and Taylor [9] have shown that a QBD has an uncountably infinite number of quasi-stationary distributions. However, only one of these quasi-stationary distributions can also be a limiting-conditional distribution.

Note that from an applications point of view, it is sensible to condition on $X(t)$ starting in a particular state when we define the limiting-conditional distribution. This is clear since if we are modelling a physical system, that system must have some fixed initial state. For example, a biological population has an initial population size and a queueing system has an initial number of customers in the buffer.

Using Lemma 7.2.2 we can find quasi-stationary distributions by firstly finding β-invariant measures $\boldsymbol{\pi}$ and then testing if $\sum_{j\in\mathcal{C}} \pi_j q_j < \infty$. Kesten [43, Theorem 2′] gives conditions under which a quasi-stationary distribution is also the limiting-conditional distribution. An important parameter in the determination of the limiting-conditional distribution is the decay parameter of $\mathcal{C}$. Kingman [48] proved the following result which will be used to define the decay parameter.

Lemma 7.2.3 *There exists a number $\lambda \geq 0$, independent of i and j such that*

$$\lim_{t\to\infty} \frac{\log P_{ij}(t)}{t} = -\lambda \qquad \forall i, j \in \mathcal{C}. \tag{7.2.7}$$

Definition 7.2.4 *The decay parameter of $\mathcal{C}$ is the number λ defined in Lemma 7.2.3.*

Theorem 2′ of Kesten [43] states that if a quasi-stationary distribution is also a limiting-conditional distribution then β defined by (7.2.5) is equal to the decay parameter of $\mathcal{C}$.

Define $N_{ij}(z)$ for $z \in \mathbb{R}$ and $i, j \in \mathcal{C}$ by

$$N_{ij}(z) = \int_0^\infty e^{zt} P_{ij}(t) dt. \tag{7.2.8}$$

It can be shown (see for example Seneta [89]) that for $z < \lambda$, $N_{ij}(z)$ is finite for all $i, j \in \mathcal{C}$ and for $z > \lambda$, $N_{ij}(z)$ is infinite for all $i, j \in \mathcal{C}$. Hence we have that

$$\lambda = \sup\{z : N_{ij}(z) \text{ is finite}\}. \tag{7.2.9}$$

The behaviour of $N_{ij}(z)$ at $z = \lambda$ can be used to classify the set $\mathcal{C}$. If $N_{ij}(z)$ diverges at $z = \lambda$ then $\mathcal{C}$ is said to be λ-recurrent (positive or null) and if $N_{ij}(z)$ converges at $z = \lambda$ then $\mathcal{C}$ is said to be λ-transient.

The above results suggest the following method for finding a limiting-conditional distribution. Firstly we determine the decay parameter λ, and then we find a λ-invariant measure $\boldsymbol{\pi}$ on $\mathcal{C}$ for $\boldsymbol{Q}$. If $\sum_{j \in \mathcal{C}} \pi_j q_j < \infty$ then $\boldsymbol{\pi}$ is a quasi-stationary distribution. We can then apply results from Kesten [43] to determine if $\boldsymbol{\pi}$ is a limiting-conditional distribution.

As mentioned in the previous section, Kijima [44] has shown that the limiting-conditional distribution for a uniform continuous time Markov process is the same as the limiting conditional distribution for its uniformised process. A continuous time Markov process is uniform if $q = \sup_i q_i < \infty$ and then its uniformised process is defined to be the discrete time Markov process with transition probability matrix equal to $\frac{1}{r}\boldsymbol{Q} + \boldsymbol{I}$ where $r \geq q$. As we mentioned in the introduction, we will obtain our results for continuous time QBDs by considering the uniformised process.

7.3 Continuous time absorbing QBDs

In this section we present some definitions and results for continuous time absorbing QBDs. These results will be used when we derive the limiting-conditional distribution for a continuous time absorbing QBD. In what follows we will refer to a matrix as finite if all of its entries are finite.

Let $X(t)$ be a continuous time QBD on the state space $\mathcal{S} = \{0\} \cup \mathcal{C}$ where $\mathcal{C} = \{(k, j) :$

$k \geq 1, 1 \leq j \leq M\}$ and let the q-matrix have the block partitioned form

$$Q = \begin{pmatrix} 0 & 0' & 0' & 0' & 0' & \cdots \\ Q_2 e & Q_1 & Q_0 & 0 & 0 & \cdots \\ 0 & Q_2 & Q_1 & Q_0 & 0 & \cdots \\ 0 & 0 & Q_2 & Q_1 & Q_0 & \cdots \\ 0 & 0 & 0 & Q_2 & Q_1 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix} \tag{7.3.1}$$

where $\boldsymbol{Q}_0$, $\boldsymbol{Q}_1$ and $\boldsymbol{Q}_2$ are $M \times M$ matrices and $\boldsymbol{e}$ is a $M \times 1$ vector of ones. We assume that $\boldsymbol{Q}$ is stable and conservative and that C is an irreducible and aperiodic communicating class. If we let $\hat{\boldsymbol{Q}} = \boldsymbol{Q}_0 + \boldsymbol{Q}_1 + \boldsymbol{Q}_2$ then the irreducibility of C and the fact that $\boldsymbol{Q}$ is conservative imply that $\hat{\boldsymbol{Q}}$ is an irreducible and conservative q-matrix.

Let $\boldsymbol{\nu}$ be a solution to $\boldsymbol{\nu}\hat{\boldsymbol{Q}} = \boldsymbol{0}$. We assume that $X(t)$ is such that $\boldsymbol{\nu}\boldsymbol{Q}_0\boldsymbol{e} < \boldsymbol{\nu}\boldsymbol{Q}_2\boldsymbol{e}$. From Theorem 2.8.3 this is equivalent to assuming that $X(t)$ would be positive recurrent if there was a positive rate from state 0 into level 1.

Let $\boldsymbol{N}_{11}(\beta)$ denote the $M \times M$ matrix whose (i,j)th entry is $\boldsymbol{N}_{(1,i)(1,j)}(\beta)$ as defined in (7.2.8) and define

$$\boldsymbol{R}(\beta) = \boldsymbol{Q}_0\boldsymbol{N}_{11}(\beta). \tag{7.3.2}$$

We now give a probabilistic interpretation for the matrix $\boldsymbol{R}(\beta)$. The interpretation we give is equivalent to that given in Ramaswami [80]. In the case where $\beta = 0$, we saw in Chapter 2 that for all $k \geq 1$, the (i,j)th element of $\boldsymbol{R}(0)$ is the expected time spent in the state (k,j), in units of the mean sojourn time in state $(k-1,i)$, before the process first returns to level $k-1$, given the process started in the state $(k-1,i)$. To interpret $\boldsymbol{R}(\beta)$ for $\beta \neq 0$, observe that $\boldsymbol{R}(\beta)$ can be written as

$$\boldsymbol{R}(\beta) = \boldsymbol{Q}_0 \int_0^\infty e^{\beta t} {}_0\boldsymbol{P}(\boldsymbol{1},\boldsymbol{1};t)dt \tag{7.3.3}$$

where we adopt the notation we introduced in Chapter 2. Because of the homogeneity of $\boldsymbol{Q}$, we have that ${}_0P(1,1;t) = {}_{\boldsymbol{k}-1}P(\boldsymbol{k},\boldsymbol{k};t)\ \forall k \geq 1$ and hence we can write

$$\boldsymbol{R}(\beta) = \boldsymbol{Q}_0 \int_0^{\infty} e^{\beta t}{}_{\boldsymbol{k}-1}P(\boldsymbol{k},\boldsymbol{k};t)dt. \tag{7.3.4}$$

Using this expression for $\boldsymbol{R}(\beta)$, the (i,j)th element of $\boldsymbol{R}(\beta)$ is given by

$$\begin{aligned}
[\boldsymbol{R}(\beta)]_{ij} &= \sum_{\ell=1}^{M}[\boldsymbol{Q}_0]_{i\ell}\int_{t=0}^{\infty} e^{\beta t}[{}_{\boldsymbol{k}-1}\boldsymbol{P}(\boldsymbol{k},\boldsymbol{k};t)]_{\ell j}dt & (7.3.5)\\
&= \int_{t=0}^{\infty} e^{\beta t}\left\{\sum_{\ell=1}^{M}\frac{[\boldsymbol{Q}_0]_{i\ell}}{-[\boldsymbol{Q}_1]_{ii}}(-[\boldsymbol{Q}_1]_{ii})[{}_{\boldsymbol{k}-1}\boldsymbol{P}(\boldsymbol{k},\boldsymbol{k};t)]_{\ell j}\right\}dt & (7.3.6)\\
&= \int_{t=0}^{\infty} e^{\beta t}d\left\{\sum_{\ell=1}^{M}\frac{[\boldsymbol{Q}_0]_{i\ell}}{-[\boldsymbol{Q}_1]_{ii}}(-[\boldsymbol{Q}_1]_{ii})\int_{\tau=0}^{t}[{}_{\boldsymbol{k}-1}\boldsymbol{P}(\boldsymbol{k},\boldsymbol{k};\tau)]_{\ell j}d\tau\right\}. & (7.3.7)
\end{aligned}$$

Now $\dfrac{[\boldsymbol{Q}_0]_{i\ell}}{-[\boldsymbol{Q}_1]_{ii}}$ is the probability that the process first moves into state (k,ℓ) given it started in state $(k-1,i)$ and $(-[\boldsymbol{Q}_1]_{ii})\int_0^t[{}_{\boldsymbol{k}-1}\boldsymbol{P}(\boldsymbol{k},\boldsymbol{k};\tau)]_{\ell j}d\tau$ is the expected sojourn time in the state (k,j), in an interval of time $[0,t]$, in units of the mean sojourn time in state $(k-1,i)$, given the process starts in the state (k,ℓ) and the process does not visit level $k-1$ in the interval $[0,t]$. Hence we can interpret the (i,j)th element of $\boldsymbol{R}(\beta)$ as the Laplace-Stieltjes transform of the expected sojourn time in the state (k,j), in an interval of time $[0,t]$, expressed in units of the mean sojourn time in state $(k-1,i)$, given the process starts in the state $(k-1,i)$, the process does not visit level $k-1$ in the interval $[0,t]$, and we set the time origin as the time at which the process first visits level k.

The following lemma describes when $\boldsymbol{R}(\beta)$ is finite.

Lemma 7.3.1 *The matrix $\boldsymbol{R}(\beta)$ is finite for all $\beta < \lambda$ and finite for $\beta = \lambda$ if and only if the process is λ-transient.*

Proof: If $\beta < \lambda$ then by the definition of λ, $\boldsymbol{N}_{11}(\beta)$ is finite and hence $\boldsymbol{R}(\beta)$ is finite. By the definition of λ-transience, $\boldsymbol{N}_{11}(\lambda)$ is finite if and only if the process is λ-transient. Hence $\boldsymbol{R}(\lambda)$ is finite if and only if the process is λ-transient. □

The following lemma can be shown in a similar way to Theorem 3.6 in Ramaswami [80]. Theorem 3.6 in [80] requires that $-\beta > 0$. Using Lemma 7.3.1 it can be seen that we only require that $-\beta > -\lambda$.

Lemma 7.3.2 *If the matrix $\boldsymbol{R}(\beta)$ is finite then it is the minimal non negative solution to the matrix-quadratic equation*

$$-\beta \boldsymbol{S} = \boldsymbol{Q}_0 + \boldsymbol{S}\boldsymbol{Q}_1 + \boldsymbol{S}^2\boldsymbol{Q}_2. \tag{7.3.8}$$

We now define some quantities which will be used in subsequent sections. Define $\boldsymbol{Q}^*(z)$ by

$$\boldsymbol{Q}^*(z) = \boldsymbol{Q}_0 + z\boldsymbol{Q}_1 + z^2\boldsymbol{Q}_2. \tag{7.3.9}$$

From Theorem 2.6 in Seneta [89] there exists an eigenvalue $\chi(z)$ of $\boldsymbol{Q}^*(z)$ such that $\chi(z)$ is real, it has strictly positive left and right eigenvectors associated with it and it is larger than the real part of any other eigenvalue of $\boldsymbol{Q}^*(z)$. Let $\boldsymbol{u}(z)$ and $\boldsymbol{v}(z)$ be the left and right eigenvectors of $\boldsymbol{Q}^*(z)$ corresponding to $\chi(z)$ and let them be normalized so that

$$\boldsymbol{u}(z)\boldsymbol{e} = 1 = \boldsymbol{u}(z)\boldsymbol{v}(z). \tag{7.3.10}$$

It can be shown using methods similar to those used in the proof of Lemma 1.3.3 in Neuts [72] that $\chi'(z)$ is given by

$$\chi'(z) = \boldsymbol{u}(z)(\boldsymbol{Q}_1 + 2z\boldsymbol{Q}_2)\boldsymbol{v}(z). \tag{7.3.11}$$

7.4 The form of the limiting-conditional distribution

In this section we derive an expression for the limiting-conditional distribution of a continuous time QBD. We do this by applying results in Bean, Bright, Latouche, Pearce, Pollett and Taylor [8] to the uniformised process. Due to their homogeneity, all continuous time QBDs are uniformisable. $\boldsymbol{A}_0$, $\boldsymbol{A}_1$ and $\boldsymbol{A}_2$ for the uniformised QBD are given by

$$\boldsymbol{A}_0 = \frac{1}{r}\boldsymbol{Q}_0, \tag{7.4.1}$$

$$\boldsymbol{A}_1 = \frac{1}{r}\boldsymbol{Q}_1 + \boldsymbol{I}, \tag{7.4.2}$$

$$\boldsymbol{A}_2 = \frac{1}{r}\boldsymbol{Q}_2 \tag{7.4.3}$$

where r is chosen such that $r \geq \sup_i \sum_j q_{ij}$.

For $0 \leq z \leq 1$ let $\bar{\chi}(z)$ be the maximal eigenvalue of the matrix

$$\boldsymbol{A}^*(z) = \boldsymbol{A}_0 + z\boldsymbol{A}_1 + z^2\boldsymbol{A}_2 \tag{7.4.4}$$

and $\bar{\boldsymbol{u}}(z)$ and $\bar{\boldsymbol{v}}(z)$ be the corresponding left and right eigenvectors, normalized so that

$$\bar{\boldsymbol{u}}(z)\boldsymbol{e} = 1 = \bar{\boldsymbol{u}}(z)\bar{\boldsymbol{v}}(z). \tag{7.4.5}$$

Equation (1.3.10) in Neuts [72] states that

$$\bar{\chi}'(z) = \bar{\boldsymbol{u}}(z)(\boldsymbol{A}_1 + 2z\boldsymbol{A}_2)\bar{\boldsymbol{v}}(z). \tag{7.4.6}$$

By substituting equations (7.4.1)-(7.4.3) into (7.4.4) it is easy to show that

$$\boldsymbol{A}^*(z) = \frac{1}{r}\boldsymbol{Q}^*(z) + z\boldsymbol{I}. \tag{7.4.7}$$

Hence we have that

$$\bar{\chi}(z) = \frac{1}{r}\chi(z) + z \tag{7.4.8}$$

which gives

$$\bar{\chi}'(z) = \frac{1}{r}\chi'(z) + 1 \tag{7.4.9}$$

Furthermore, it can be shown that

$$\boldsymbol{u}(z) = \bar{\boldsymbol{u}}(z) \quad \text{and} \quad \boldsymbol{v}(z) = \bar{\boldsymbol{v}}(z). \tag{7.4.10}$$

These results will be useful in proving the following theorem which gives the limiting-conditional distribution for a continuous time QBD.

Theorem 7.4.1 *The limiting-conditional distribution for $X(t)$ is given by*

$$\boldsymbol{\pi}_j = \frac{1}{c}\left(z_0^j\boldsymbol{b} + jz_0^{j-1}\boldsymbol{a} - \boldsymbol{b}\boldsymbol{R}(-\alpha)^j\right) \qquad j \geq 1 \tag{7.4.11}$$

where z_0 is the unique solution in $(0,1)$ to

$$\chi(z) = \chi'(z)z, \tag{7.4.12}$$

$$\boldsymbol{a} = \boldsymbol{u}(z_0), \tag{7.4.13}$$

$$\alpha = \boldsymbol{u}(z_0)(\boldsymbol{Q}_1 + 2z_0\boldsymbol{Q}_2)\boldsymbol{v}(z_0), \tag{7.4.14}$$

$\boldsymbol{b}$ is the unique solution to

$$\boldsymbol{b}(z_0^2\boldsymbol{Q}_2 + z_0(\boldsymbol{Q}_1 - \alpha\boldsymbol{I}) + \boldsymbol{Q}_0) = -\boldsymbol{a}(2z_0\boldsymbol{Q}_2 + \boldsymbol{Q}_1 - \alpha\boldsymbol{I}) \tag{7.4.15}$$

such that $\boldsymbol{b}\boldsymbol{e} = 0$, and

$$c = \frac{1}{(1-z_0)^2} - \boldsymbol{b}\boldsymbol{R}(-\alpha)(\boldsymbol{I} - \boldsymbol{R}(-\alpha))^{-1}\boldsymbol{e}. \tag{7.4.16}$$

Proof: Denote the limiting-conditional distribution for the uniformised process of $X(t)$ by $\bar{\boldsymbol{\pi}}_j$. From Theorem 10 in [8] we have that

$$\bar{\boldsymbol{\pi}}_j = \frac{1}{\bar{c}}\left(\bar{z}_0^j\bar{\boldsymbol{b}} + j\bar{z}_0^{j-1}\bar{\boldsymbol{a}} - \bar{\boldsymbol{b}}\bar{\boldsymbol{R}}(\bar{\alpha})^j\right) \qquad j \geq 1 \tag{7.4.17}$$

where $\bar{z}_0$ is the unique solution in $(0, 1)$ to

$$\bar{\chi}(z) = \bar{\boldsymbol{u}}(z)\frac{1}{r}(\boldsymbol{Q}_1 + r\boldsymbol{I} + 2z\boldsymbol{Q}_2)\bar{\boldsymbol{v}}(z)z, \tag{7.4.18}$$

$$\bar{\boldsymbol{a}} = \bar{\boldsymbol{u}}(\bar{z}_0), \tag{7.4.19}$$

$$\bar{\alpha} = (\bar{\boldsymbol{u}}(\bar{z}_0)\frac{1}{r}(\boldsymbol{Q}_1 + r\boldsymbol{I} + 2\bar{z}_0\boldsymbol{Q}_2)\bar{\boldsymbol{v}}(\bar{z}_0))^{-1}, \tag{7.4.20}$$

$\bar{\boldsymbol{b}}$ is the unique solution to

$$\bar{\boldsymbol{b}}(\bar{z}_0^2\boldsymbol{Q}_2 + \bar{z}_0(\boldsymbol{Q}_1 + (r - \frac{r}{\bar{\alpha}})\boldsymbol{I}) + \boldsymbol{Q}_0) = -\bar{\boldsymbol{a}}(2\bar{z}_0\boldsymbol{Q}_2 + \boldsymbol{Q}_1 + (r - \frac{r}{\bar{\alpha}})\boldsymbol{I}) \tag{7.4.21}$$

such that $\bar{\boldsymbol{b}}\boldsymbol{e} = 0$,

$\bar{\boldsymbol{R}}(\bar{\alpha})$ is the minimal non negative solution to

$$\boldsymbol{S} = \frac{\bar{\alpha}}{r}(\boldsymbol{Q}_0 + \boldsymbol{S}(\boldsymbol{Q}_1 + r\boldsymbol{I}) + \boldsymbol{S}^2\boldsymbol{Q}_2) \tag{7.4.22}$$

and

$$\bar{c} = \frac{1}{(1 - \bar{z}_0)^2} - \bar{\boldsymbol{b}}\bar{\boldsymbol{R}}(\bar{\alpha})(\boldsymbol{I} - \bar{\boldsymbol{R}}(\bar{\alpha}))^{-1}\boldsymbol{e}. \tag{7.4.23}$$

We know from Kijima [45] that $\boldsymbol{\pi}_j = \bar{\boldsymbol{\pi}}_j$ so we need to show that right hand side of equation (7.4.17) can be written in the same form as the right hand side of equation (7.4.11). We do this by showing that $\bar{z}_0 = z_0$, $\bar{\boldsymbol{a}} = \boldsymbol{a}$, $\bar{\boldsymbol{b}} = \boldsymbol{b}$, $\bar{\boldsymbol{R}}(\bar{\alpha}) = \boldsymbol{R}(-\alpha)$ and $\bar{c} = c$.

$\bar{z}_0$ satisfies (7.4.18) so we have

$$\bar{\chi}(\bar{z}_0) = \bar{\boldsymbol{u}}(\bar{z}_0)\frac{1}{r}(\boldsymbol{Q}_1 + r\boldsymbol{I} + 2\bar{z}_0\boldsymbol{Q}_2)\bar{\boldsymbol{v}}(\bar{z}_0)\bar{z}_0 \tag{7.4.24}$$

which upon substituting (7.4.8) and (7.4.10) becomes

$$\frac{1}{r}\chi(\bar{z}_0) + \bar{z}_0 = \boldsymbol{u}(\bar{z}_0)\frac{1}{r}(\boldsymbol{Q}_1 + r\boldsymbol{I} + 2\bar{z}_0\boldsymbol{Q}_2)\boldsymbol{v}(\bar{z}_0)\bar{z}_0 \tag{7.4.25}$$

$$\Rightarrow \chi(\bar{z}_0) = \boldsymbol{u}(\bar{z}_0)(\boldsymbol{Q}_1 + 2\bar{z}_0\boldsymbol{Q}_2)\boldsymbol{v}(\bar{z}_0)\bar{z}_0 \tag{7.4.26}$$

$$\Rightarrow \chi(\bar{z}_0) = \chi'(\bar{z}_0)\bar{z}_0 \tag{7.4.27}$$

where the last line follows from equation (7.3.11).

Hence $\bar{z}_0$ is a solution in (0,1) to (7.4.12). Now suppose there exists a $z^* \neq \bar{z}_0$ such that z^* is in $(0,1)$ and z^* satisfies equation (7.4.12). We have that

$$\chi(z^*) = \chi'(z^*)z^*, \tag{7.4.28}$$

which substituting in (7.4.8) and (7.4.9) gives

$$\begin{aligned} r\bar{\chi}(z^*) - rz^* &= (r\bar{\chi}'(z^*) - r)z^* && (7.4.29)\\ \Rightarrow \bar{\chi}(z^*) &= \bar{\chi}'(z^*)z^* && (7.4.30) \end{aligned}$$

which is equivalent to

$$\bar{\chi}(z^*) = \bar{\boldsymbol{u}}(z^*)\frac{1}{r}(\boldsymbol{Q}_1 + r\boldsymbol{I} + 2z^*\boldsymbol{Q}_2)\bar{\boldsymbol{v}}(z^*)z^*. \tag{7.4.31}$$

But this contradicts the fact that (7.4.18) has only the solution $\bar{z}_0$ in $(0,1)$ so the assumption that there exists a $z^* \neq \bar{z}_0$ such that z^* is in $(0,1)$ and z^* satisfies equation (7.4.12) must be false. Hence equation (7.4.12) has the unique solution in $(0,1)$ given by $z_0 = \bar{z}_0$.

It follows from equation (7.4.10) that $\bar{\boldsymbol{a}} = \boldsymbol{a}$. Equation (7.4.20) can be rearranged to give

$$\bar{\alpha} = \frac{r}{\alpha + r}. \tag{7.4.32}$$

Substituting this expression for $\bar{\alpha}$ along with $\bar{z}_0 = z_0$ and $\bar{\boldsymbol{a}} = \boldsymbol{a}$ into (7.4.21), gives us that $\bar{\boldsymbol{b}}$ is the unique solution to

$$\bar{\boldsymbol{b}}[z_0^2\boldsymbol{Q}_2 + z_0(\boldsymbol{Q}_1 - \alpha\boldsymbol{I}) + \boldsymbol{Q}_0] = -\boldsymbol{a}[2z_0\boldsymbol{Q}_2 + \boldsymbol{Q}_1 - \alpha\boldsymbol{I}], \tag{7.4.33}$$

which implies that $\boldsymbol{b} = \bar{\boldsymbol{b}}$. Substituting for $\bar{\alpha}$ into (7.4.22) and rearranging gives us that $\bar{\boldsymbol{R}}(\bar{\alpha})$ is the minimal non negative solution to

$$\alpha\boldsymbol{S} = \boldsymbol{Q}_0 + \boldsymbol{S}\boldsymbol{Q}_1 + \boldsymbol{S}^2\boldsymbol{Q}_2, \tag{7.4.34}$$

which comparing with (7.3.8) implies that $\bar{\boldsymbol{R}}(\bar{\alpha}) = \boldsymbol{R}(-\alpha)$. It is clear from the above that $\bar{c} = c$ and hence we have the result. □

Theorem 7.4.2 *The limiting-conditional distribution for $X(t)$ satisfies the relationship*

$$\boldsymbol{\pi}_j = \boldsymbol{\pi}_{j-1}\boldsymbol{R}(-\alpha) + z_0^{j-1}\boldsymbol{\pi}_1 \qquad j \geq 2. \tag{7.4.35}$$

Proof: Equation (5.10) in Bean, Bright, Latouche, Pearce, Pollett and Taylor [8] states that

$$\bar{\pi}_j = \bar{\pi}_{j-1}\bar{R}(\bar{\alpha}) + \bar{z}_0^{j-1}\bar{\pi}_1 \qquad j \geq 2 \tag{7.4.36}$$

where $\bar{\pi}$ is the limiting-conditional distribution for the uniformised process and $\bar{z}_0$, $\bar{\alpha}$ and $\bar{R}(\bar{\alpha})$ are as defined in the proof of Theorem 7.4.1. The result follows since we showed in the proof of Theorem 7.4.1 that $\bar{\pi}_j = \pi_j$, $\bar{z}_0 = z_0$ and $\bar{R}(\bar{\alpha}) = R(-\alpha)$. □

Theorem 7.4.3 *The decay parameter of $X(t)$ is given by $\lambda = -\alpha$ where α is given by equation (7.4.14).*

Proof: From the definitions of $R(\beta)$ and λ, we have that $\lambda = \sup\{z : R(z) \text{ is finite}\}$. Consider the uniformised process of $X(t)$ and for $y > 0$ define $\bar{R}(y)$ as the minimal non negative solution to

$$S = y\left(\frac{1}{r}Q_0 + S\left(\frac{1}{r}Q_1 + I\right) + S^2\frac{1}{r}Q_2\right). \tag{7.4.37}$$

It is shown in Bean, Bright, Latouche, Pearce, Pollett and Taylor [8] that $\bar{R}(y)$ is finite if and only if $y \leq \bar{\alpha}$ where $\bar{\alpha}$ is given by (7.4.32). Rearranging (7.4.37) gives

$$\left(\frac{r - ry}{y}\right)S = Q_0 + SQ_1 + S^2Q_2 \tag{7.4.38}$$

which implies that $\bar{R}(y) = R((ry - r)/y)$. Setting $x = (ry - r)/y$ gives $R(x) = \bar{R}(r/(r - x))$. Hence we have that $R(x)$ is finite whenever $\bar{R}(r/(r - x))$ is finite. That is, if and only if $r/(r - x) \leq \bar{\alpha}$ which from (7.4.32) is equivalent to $r/(r - x) \leq r/(\alpha + r)$. Now since $y > 0$, we have $r - x = r/y > 0$, so by choosing r such that $\alpha + r > 0$, we have that $R(x)$ is finite whenever

$$\begin{aligned} \frac{r}{r-x} &\leq \frac{r}{\alpha + r} \\ \Rightarrow r(\alpha + r) &\leq r(r - x) \\ \Rightarrow x &\leq -\alpha. \end{aligned}$$

Hence the decay parameter of $X(t)$ is given by $-\alpha$. □

Theorem 7.4.4 *$X(t)$ is λ-transient.*

Proof: The result follows since $\lambda = -\alpha$ and since we showed in the proof of Theorem 7.4.3 that $R(-\alpha)$ is finite. □

7.5 Calculating the limiting-conditional distribution

In this section we discuss how to compute the limiting-conditional distribution. The results we present here are continuous time versions of the discrete time results in Section 6 of Bean, Bright, Latouche, Pearce, Pollett and Taylor [8]. We note that some numerical examples were presented in [8] but we do not present any here.

Equation (7.4.11) can be used to compute the limiting-conditional distribution provided we know z_0, $\boldsymbol{a}$, α, $\boldsymbol{b}$, $\boldsymbol{R}(-\alpha)$ and c. We will show here how each of these quantities can be computed.

To find z_0, we need to find the unique solution in $(0, 1)$ to

$$\chi(z) - \chi'(z)z = 0. \tag{7.5.1}$$

Since $\boldsymbol{Q}_0$ is a non negative matrix and not equal to the zero matrix, $\chi(0) > 0$ and hence $\chi(0) - \chi'(0).0 > 0$. It is easy to show that $\chi(1) = 0$. From equation (7.4.9) we have that $\chi'(1) = r(\bar{\chi}'(1) - 1)$. Using results from Lemma 1.3.3 and Theorem 1.3.2 in Neuts [72], we can show that $\bar{\chi}'(1) > 1$ which implies $\chi'(1) > 0$. Hence we have that $\chi(1) - \chi'(1).1 < 0$. Given this and the fact that $\chi(0) - \chi'(0).0 > 0$, since z_0 is the unique solution in $(0, 1)$ to (7.5.1), we can find z_0 by using a bisection search. For a given value of z, if $\chi(z) - \chi'(z)z < 0$ then $z_0 < z$, and if $\chi(z) - \chi'(z)z > 0$ then $z_0 > z$. Neuts [72, page 40] mentions a similar technique for solving a related equation.

Finding the left and right eigenvectors of $\boldsymbol{Q}^*(z_0)$ corresponding to $\chi(z_0)$ enables us to calculate $\boldsymbol{a}$ and α.

Note that when we find z_0, we can only compute it to within some arbitrary accuracy and hence we can estimate α only to within some arbitrary accuracy. Since $\boldsymbol{R}(\beta)$ is infinite for $\beta > -\alpha$ it is essential that our estimate $\hat{\alpha}$ of α be such that $\hat{\alpha} \geq \alpha$ otherwise $\boldsymbol{R}(-\hat{\alpha})$ will be infinite. Some interesting implications follow from the fact that $\hat{\alpha}$ is not exactly equal to α and these are currently being investigated.

The vector $\boldsymbol{b}$ is the unique solution to equation (7.4.15) subject to $\boldsymbol{be} = 0$. To find $\boldsymbol{b}$ we simply replace the last column of the matrix $z_0^2\boldsymbol{Q}_2 + z_0(\boldsymbol{Q}_1 - \alpha\boldsymbol{I}) + \boldsymbol{Q}_0$ by the column vector $\boldsymbol{e}$ and replace the last element of the vector on the right hand side with a zero. The resulting system of equations has a unique solution that will be equal to the required vector $\boldsymbol{b}$.

In order to evaluate $\boldsymbol{R}(-\alpha)$ we present the following theorem which gives an expression for $\boldsymbol{R}(\beta)$. The theorem is a generalisation of Theorem 3.3.2 in Chapter 3.

Theorem 7.5.1 *The matrix $\boldsymbol{R}(\beta)$ for $\beta \leq \lambda$ is given by*

$$\boldsymbol{R}(\beta) = \sum_{\ell=0}^{\infty} \boldsymbol{C}_0^\ell \prod_{i=0}^{\ell-1} \boldsymbol{C}_2^{\ell-1-i} \tag{7.5.2}$$

where $\boldsymbol{C}_i^\ell$ is defined for $i = 0, 2$ as

$$\boldsymbol{C}_i^0 = \boldsymbol{Q}_i(-\boldsymbol{Q}_1 - \beta \boldsymbol{I})^{-1}, \tag{7.5.3}$$

$$\boldsymbol{C}_i^{\ell+1} = (\boldsymbol{C}_i^\ell)^2(\boldsymbol{I} - \boldsymbol{C}_0^\ell \boldsymbol{C}_2^\ell - \boldsymbol{C}_2^\ell \boldsymbol{C}_0^\ell)^{-1} \quad k \geq 0. \tag{7.5.4}$$

Proof: Again we prove the result by considering the uniformised process. We saw in the proof of Theorem 7.4.3 that $\boldsymbol{R}(\beta) = \bar{\boldsymbol{R}}(r/(r-\beta))$ for $r \geq 0$. From Theorem 11 in Bean, Bright, Latouche, Pearce, Pollett and Taylor [8], since $\beta \leq \lambda$, we have

$$\bar{\boldsymbol{R}}\left(\frac{r}{r-\beta}\right) = \sum_{\ell=0}^{\infty} \bar{\boldsymbol{C}}_0^\ell \prod_{i=0}^{\ell-1} \bar{\boldsymbol{C}}_2^{\ell-1-i} \tag{7.5.5}$$

where $\bar{\boldsymbol{C}}_i^\ell$ is defined for $i = 0, 2$ as

$$\bar{\boldsymbol{C}}_i^0 = \frac{r}{r-\beta}\boldsymbol{A}_i\left(-\frac{r}{r-\beta}\boldsymbol{A}_1 - \boldsymbol{I}\right)^{-1}, \tag{7.5.6}$$

$$\bar{\boldsymbol{C}}_i^{\ell+1} = (\bar{\boldsymbol{C}}_i^\ell)^2(\boldsymbol{I} - \bar{\boldsymbol{C}}_0^\ell \bar{\boldsymbol{C}}_2^\ell - \bar{\boldsymbol{C}}_2^\ell \bar{\boldsymbol{C}}_0^\ell)^{-1} \quad \ell \geq 0. \tag{7.5.7}$$

Upon substituting equations (7.4.1)-(7.4.3) and simplifying, we see that $\bar{\boldsymbol{C}}_i^\ell = \boldsymbol{C}_i^\ell$ for $i = 0, 2$ and hence the theorem is proved. ☐

In order to use equation (7.5.2) to compute $\boldsymbol{R}(-\alpha)$, we need to truncate the infinite sum at some point. The following lemma will be useful in developing a truncation rule.

Lemma 7.5.1 *The spectral radius of $\boldsymbol{R}(-\alpha)$ is equal to z_0 and $\boldsymbol{a}$ is the corresponding left eigenvector.*

Proof: Firstly we state that since we have assumed that the matrix $\boldsymbol{Q}$ given by equation (7.3.1) is irreducible, it can be shown using methods similar to those used in the proof of Lemma 1.2.4 in Neuts [72] that the spectral radius of $\boldsymbol{R}(-\alpha)$ is positive. The remainder of the proof is similar to the proof of Lemma 1.3.2 in Neuts [72]. If we let η be the spectral radius of $\boldsymbol{R}(-\alpha)$, then since η is positive, the left eigenvector $\boldsymbol{s}$ of $\boldsymbol{R}(-\alpha)$ corresponding to η is non negative and not equal to the zero vector. Using equation (7.3.8) it is easy to show that $\boldsymbol{s}$ is also a left eigenvector of $\boldsymbol{Q}^*(\eta)$ corresponding to the eigenvalue $\alpha\eta$. Since $\boldsymbol{s}$

is non negative and $\boldsymbol{u}(\eta)$ is positive, by using methods similar to those in Gantmacher [28, pages 63-64] it can be shown that $\boldsymbol{s} = \boldsymbol{u}(\eta)$ and $\chi(\eta) = \alpha\eta$. Now since $\alpha = \chi'(z_0)$, we have that η satisfies $\chi(\eta) = \chi'(z_0)\eta$ and we know from Theorem 7.4.1 that the unique solution to this equation is $z_0 = \eta$. Hence the spectral radius of $\boldsymbol{R}(-\alpha)$ is z_0 and the corresponding left eigenvector is given by $\boldsymbol{s} = \boldsymbol{u}(\eta) = \boldsymbol{u}(z_0) = \boldsymbol{a}$. □

If we define $\boldsymbol{R}^{(\ell)}(-\alpha)$ to be the sum of the first $\ell + 1$ terms in the sum in equation (7.5.2), then we can evaluate $\boldsymbol{R}(-\alpha)$ by truncating the infinite sum according to the following truncation rule.

Truncation rule 7.5.1 *Take* $\boldsymbol{R}(-\alpha) = \boldsymbol{R}^{(L)}(-\alpha)$ *where* L *is the smallest value of* ℓ *such that* $\|\boldsymbol{a}\boldsymbol{R}^{(\ell)}(-\alpha) - z_0\boldsymbol{a}\|_\infty < \epsilon$.

Note that Truncation rule 7.5.1 is a generalisation of Truncation rule 3.4.5 in Chapter 3.

From numerical experience in calculating $\boldsymbol{R}(-\alpha)$ via equation (7.5.2), we have observed that for large values of ℓ, the matrix $\boldsymbol{C}_0^\ell$ becomes zero to machine accuracy and the matrix $\boldsymbol{C}_2^\ell$ becomes infinite to machine accuracy. Hence in order to calculate the product $\boldsymbol{C}_0^\ell \prod_{i=0}^{\ell-1} \boldsymbol{C}_2^{\ell-1-i}$, for large values of ℓ, it is necessary to use some form of exponential scaling technique. This can be done by writing $\boldsymbol{C}_i^\ell, i = 0, 2$ in the form $\hat{\boldsymbol{C}}_i^\ell \times 10^{c_i}$ where the elements of $\hat{\boldsymbol{C}}_i^\ell$ are between one and ten and then the product $\boldsymbol{C}_i^\ell \boldsymbol{C}_j^\ell$ is given by $\hat{\boldsymbol{C}}_i^\ell \hat{\boldsymbol{C}}_j^\ell \times 10^{(c_i + c_j)}$.

In the case where $\beta = 0$, we have that $\forall \ell \geq 0$, $\boldsymbol{C}_0^\ell = \widetilde{\boldsymbol{U}}^\ell$ and $\boldsymbol{C}_2^\ell = \widetilde{\boldsymbol{D}}^\ell$ where $\widetilde{\boldsymbol{U}}^\ell$ and $\widetilde{\boldsymbol{D}}^\ell$ are given by equations (3.5.17)-(3.5.18). Since $\widetilde{\boldsymbol{U}}^\ell$ can be interpreted in the same way as $\widetilde{\boldsymbol{U}}_k^\ell$, we have from equation (3.7.11) in Chapter 3, for the case where $\beta = 0$,

$$\boldsymbol{C}_0^\ell = \boldsymbol{Q}_0 E\left[\int_{\tau=\delta(k)\wedge\delta(k+2^\ell)}^{\delta(k)\wedge\delta(k+2^{\ell+1})} \mathbb{I}(X(\tau) \in \boldsymbol{k} + \boldsymbol{2}^\ell)dt \mid X(0) \in \boldsymbol{k} + \boldsymbol{1}\right] \tag{7.5.8}$$

where $\delta(k) = \inf\{t > 0 : X(t) \in \boldsymbol{k}\}$. Similarly, when $\beta = 0$ we have from equation (3.7.12)

$$\boldsymbol{C}_2^\ell = \boldsymbol{Q}_2 E\left[\int_{\tau=\delta(k)\wedge\delta(k-2^\ell)}^{\delta(k)\wedge\delta(k-2^{\ell+1})} \mathbb{I}(X(\tau) \in \boldsymbol{k} - \boldsymbol{2}^\ell)dt \mid X(0) \in \boldsymbol{k} - \boldsymbol{1}\right]. \tag{7.5.9}$$

Given the interpretations (7.5.8) and (7.5.9) it is reasonable to expect that for the case $\beta \neq 0$, $\boldsymbol{C}_0^\ell$ and $\boldsymbol{C}_2^\ell$ have the interpretations

$$\boldsymbol{C}_0^\ell = \boldsymbol{Q}_0 E\left[\int_{\tau=\delta(k)\wedge\delta(k+2^\ell)}^{\delta(k)\wedge\delta(k+2^{\ell+1})} e^{\beta t}\mathbb{I}(X(\tau) \in \boldsymbol{k} + \boldsymbol{2}^\ell)dt \mid X(0) \in \boldsymbol{k} + \boldsymbol{1}\right] \tag{7.5.10}$$

and

$$\boldsymbol{C}_2^{\ell} = \boldsymbol{Q}_2 E\left[\int_{\tau=\delta(k)\wedge\delta(k-2^{\ell})}^{\delta(k)\wedge\delta(k-2^{\ell+1})} e^{\beta t}\mathbb{I}(X(\tau) \in \boldsymbol{k} - \boldsymbol{2^{\ell}})dt \mid X(0) \in \boldsymbol{k} - \boldsymbol{1}\right]. \tag{7.5.11}$$

We have attempted to prove the interpretations (7.5.10) and (7.5.11) hold for $\beta \neq 0$ but it seems that this is difficult to do. However, given the interpretations hold for $\beta = 0$, it is certainly sensible to conjecture that (7.5.10) and (7.5.11) hold for $\beta \neq 0$. We also conjecture that the ℓth element of the infinite sum in equation (7.5.2) has the interpretation

$$\boldsymbol{Q}_0^{(k)} E\left[\int_{\tau=\delta(k+2^{\ell})}^{\delta(k)\wedge\delta(k+2^{\ell+1})} e^{\beta t}\mathbb{I}(X(\tau) \in \boldsymbol{k} + \boldsymbol{1})dt \mid X(0) \in \boldsymbol{k} + \boldsymbol{1}\right]. \tag{7.5.12}$$

If we assume that the interpretations (7.5.10)-(7.5.12) hold then we can give the following interpretation of the expression (7.5.2).

Equation (7.5.12) indicates that the ℓth element of the sum in equation (7.5.2) considers only sample paths that have reached as high as level $k+2^{\ell}$ but have not reached as high as level $k+2^{\ell+1}$. Hence as ℓ tends to infinity, the lengths of the sample paths grow at an exponential rate. We mentioned in Chapter 3 that because of this structure, the probabilities of the various sample paths will tend to zero quickly as ℓ tends to infinity. A direct consequence of this is that when $\beta = 0$, a small value of L is usually sufficient.

We now consider the size of the truncation parameter L for the case $\beta > 0$. Equation (7.3.4) states that

$$\boldsymbol{R}(\beta) = \boldsymbol{Q}_0 \int_0^{\infty} e^{\beta t} {}_{\boldsymbol{k}-\boldsymbol{1}}P(\boldsymbol{k},\boldsymbol{k};t)dt \qquad k \geq 1. \tag{7.5.13}$$

Using methods similar to those used in Section 3.7.2 of Chapter 3 we can write

$$\boldsymbol{R}(\beta) = \boldsymbol{Q}_0 E\left[\int_{t=0}^{\delta(k-1)} e^{\beta t}\mathbb{I}(X(t) \in \boldsymbol{k})dt \mid X(0) \in \boldsymbol{k}\right]. \tag{7.5.14}$$

Hence to compute $\boldsymbol{R}(\beta)$ we need to consider all sample paths that start in level k and finish in level $k-1$ and don't visit level $k-1$ in between. Since for all $\beta > 0$, $e^{\beta t} \to \infty$ as $t \to \infty$, it is clear that for $\beta > 0$, we will need to consider longer sample paths than we do when $\beta = 0$. We have seen in previous chapters how Latouche and Ramaswami [56] observed that for $\beta = 0$, only a small value of L is usually needed. A value of L of the order of 6 to 8 often suffices. However, in the case where $\beta > 0$, the term $e^{\beta t}$ means that we need to consider much longer sample paths and often L is in the range $30 - 50$. Although we do not provide numerical results here, examples for discrete time processes are provided in Bean, Bright, Latouche, Pearce, Pollett and Taylor [8].

7.6 Quasi-stationary and limiting-conditional distributions for GI/M/1–type processes

In this section we briefly consider quasi-stationary and limiting-conditional distributions for GI/M/1–type processes. Kijima and Makimoto [46] obtained an explicit expression for a quasi-stationary distribution for a discrete time GI/M/1–type process and showed that under certain conditions this quasi-stationary distribution is also a limiting-conditional distribution. Continuous time analogues of the results presented by Kijima and Makimoto can be obtained by considering the uniformised chain. In this section we state these continuous time analogues without proof. Kijima and Makimoto did not discuss the computation of the limiting-conditional distribution. At the end of this section we propose a method for developing an algorithm for computing the limiting-conditional distribution.

In Chapter 6 we defined a continuous time (level independent) GI/M/1–type process. Here we consider an absorbing GI/M/1–type process. A continuous time Markov process on the state space $\mathcal{S} = \{0\} \cup \mathcal{C}$ where $\mathcal{C} = \{(k, j) : k \geq 1, 1 \leq j \leq M\}$ is an absorbing GI/M/1–type process if its q-matrix has the block partitioned form

$$Q = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & \cdots \\ Q_2^{(1)}e & Q_1 & Q_0 & 0 & 0 & \cdots \\ Q_3^{(2)}e & Q_2 & Q_1 & Q_0 & 0 & \cdots \\ Q_4^{(3)}e & Q_3 & Q_2 & Q_1 & Q_0 & \cdots \\ Q_5^{(4)}e & Q_4 & Q_3 & Q_2 & Q_1 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix} \qquad (7.6.1)$$

where $\boldsymbol{Q}_\ell, \ell \geq 0$, and $\boldsymbol{Q}_\ell^{(\ell-1)}, \ell \geq 2$ are $M \times M$ matrices. We assume that $\boldsymbol{Q}$ is stable and conservative and that C is an irreducible and aperiodic communicating class. If we let $\hat{\boldsymbol{Q}} = \sum_{\ell=0}^{\infty} \boldsymbol{Q}_\ell$ then the irreducibility of C and the fact that $\boldsymbol{Q}$ is conservative imply that $\hat{\boldsymbol{Q}}$ is an irreducible and conservative q-matrix. We note that most of the results we gave for absorbing QBDs can be extended to absorbing GI/M/1–type processes.

If we define $\boldsymbol{R}(\beta)$ according to equation (7.3.2) then Lemma 7.3.1 holds and Lemma 7.3.2 can be extended to the following.

Lemma 7.6.1 *If the matrix $\boldsymbol{R}(\beta)$ is finite then it is the minimal non negative solution to the matrix equation*

$$-\beta \boldsymbol{S} = \sum_{\ell=0}^{\infty} \boldsymbol{S}^\ell \boldsymbol{Q}_\ell. \tag{7.6.2}$$

For $0 \leq z \leq 1$ let $\chi(z)$ be the maximal eigenvalue of the matrix

$$\boldsymbol{Q}^*(z) = \sum_{\ell=0}^{\infty} z^\ell \boldsymbol{Q}_\ell \tag{7.6.3}$$

and $\boldsymbol{u}(z)$ and $\boldsymbol{v}(z)$ the corresponding left and right eigenvectors, normalized so that

$$\boldsymbol{u}(z)\boldsymbol{e} = 1 = \boldsymbol{u}(z)\boldsymbol{v}(z). \tag{7.6.4}$$

It can be shown using methods similar to those used in the proof of Lemma 1.3.3 in Neuts [72] that $\chi'(z)$ is given by

$$\chi'(z) = \boldsymbol{u}(z)\left(\frac{d}{dt}\boldsymbol{Q}^*(z)\bigg|_{z=z_0}\right)\boldsymbol{v}(z). \tag{7.6.5}$$

The following theorem gives the continuous time analogues of the results in Theorems 4.1 and 4.2 of Kijima and Makimoto [46].

Theorem 7.6.1 *If $\boldsymbol{\pi}_j, j \geq 1$ is given by*

$$\boldsymbol{\pi}_j = \frac{1}{c}\left(z_0^j \boldsymbol{b} + j z_0^{j-1} \boldsymbol{a} - \boldsymbol{b}\boldsymbol{R}(-\alpha)^j\right) \qquad j \geq 1 \tag{7.6.6}$$

where z_0 is the unique solution in $(0, 1)$ to

$$\begin{aligned}
\chi(z) &= \chi'(z)z, && (7.6.7)\\
\boldsymbol{a} &= \boldsymbol{u}(z_0), && (7.6.8)\\
\alpha &= \chi'(z_0), && (7.6.9)
\end{aligned}$$

$\boldsymbol{b}$ is the solution to

$$\boldsymbol{b}[\boldsymbol{Q}^*(z_0) - z_0\alpha\boldsymbol{I}] = -\boldsymbol{a}\left[\left(\frac{d}{dt}\boldsymbol{Q}^*(z)\Big|_{z=z_0}\right) - \alpha\boldsymbol{I}\right] \tag{7.6.10}$$

such that $\boldsymbol{be} = 0$, and

$$c = \frac{1}{(1-z_0)^2} - \boldsymbol{b}\boldsymbol{R}(-\alpha)(\boldsymbol{I} - \boldsymbol{R}(-\alpha))^{-1}\boldsymbol{e}, \tag{7.6.11}$$

then $\boldsymbol{\pi}_j, j \geq 1$ is a quasi-stationary distribution for $X(t)$.

Using results in Kesten [43], it can be shown that if for some K, $\boldsymbol{Q}_k = \boldsymbol{0}, \forall k \geq K$, then $\boldsymbol{\pi}_j$ given by (7.6.6) is a limiting-conditional distribution.

7.6.1 Calculating the limiting-conditional distribution

The quantities $\boldsymbol{a}$, $\boldsymbol{b}$, α and z_0 can be computed using similar methods to those we used when computing the limiting-conditional distribution for QBDs. However, at present there appears to be no reliable method for finding $\boldsymbol{R}(-\alpha)$. In this section we suggest a possible method for developing an algorithm for computing $\boldsymbol{R}(-\alpha)$.

In order to calculate $\boldsymbol{R}(-\alpha)$ for a QBD, we extended the logarithmic reduction algorithm of Latouche and Ramaswami [56]. This suggests that a possible approach for GI/M/1-type processes is to start with an algorithm that can compute the matrix $\boldsymbol{R}$ for a GI/M/1-type process and then attempt to extend this algorithm to obtain an algorithm that can compute the matrix $\boldsymbol{R}(-\alpha)$.

As mentioned in Chapter 4, Bini and Meini [10] have developed an algorithm for computing the matrix $\boldsymbol{G}$ for level independent M/G/1-type processes. Given the duality results we presented in Chapter 6, this algorithm can be used as a basis for computing the matrix $\boldsymbol{R}$ for GI/M/1-type processes. Hence we conjecture that the algorithm of Bini and Meini can be used as a basis for developing an algorithm for computing $\boldsymbol{R}(-\alpha)$. This is a topic for further research.

Chapter 8

Conclusions

In this thesis we have focused on several areas in the field of matrix-analytic methods. We began by considering LDQBDs, which are a generalisation of QBDs. We showed that although the LDQBD is a more general model than the QBD, an algorithmic analysis of these models is still possible. In Chapter 5 we illustrated the use of our algorithms by considering a variety of queueing models.

In Chapter 6 we considered a duality result that was obtained by Ramaswami [79]. This result gives a relationship between two suitably defined level independent M/G/1-type and GI/M/1-type processes. We observed that Ramaswami's dual process can be interpreted as the time reverse with respect to an invariant measure of a transient process and we used this interpretation to develop an alternative dual process. We also developed a dual process for level dependent M/G/1–type and GI/M/1–type processes.

In the last chapter of this thesis we considered quasi-stationary distributions for QBDs. Quasi-stationary distributions are often used in the analysis of processes that consist of a single transient communicating class and one or more absorbing states. We gave an explicit expression for the quasi-stationary distribution of a QBD and described how the expression can be used to compute the quasi-stationary probabilities numerically. The algorithmic methods we developed are extensions of the methods used to compute the stationary probabilities of an irreducible QBD.

Further work

Before concluding we suggest some possible extensions and applications of the results presented in this thesis.

The algorithms we provided for computing the stationary distribution of a LDQBD apply to a wide variety of LDQBDs. However, for certain processes it may be possible to develop more efficient algorithms by exploiting the particular structure of the process. For example, for the bivariate queueing models considered in Chapter 5, many of the elements of the matrices $\boldsymbol{Q}_i^{(k)}, i = 0, 1, 2$ are equal to zero yet our algorithms do not exploit this fact. In developing alternative algorithms for various processes, our algorithms can be used as a starting point for gaining insight into the behaviour of the processes under consideration.

In Chapter 6 we presented duality relationships for M/G/1-type and GI/M/1-type processes. In Chapter 2 we used our level dependent duality relationship to derive a closed form expression for $\boldsymbol{R}_k$ from a closed form expression for $\boldsymbol{G}_k$. In general, if we have a result for a M/G/1-type process, then the duality relationships enable us to determine the corresponding result for a GI/M/1-type process and vice versa. We feel that the alternative dual processes for level independent M/G/1-type and GI/M/1-type processes and the level dependent dual process should prove to be useful tools for theoretical investigations of M/G/1-type and GI/M/1-type processes.

Currently the results for quasi-stationary distributions for QBDs are being extended to the case of LDQBDs. It is suspected that the algorithms used to compute the stationary distribution for irreducible LDQBDs can be extended to compute the quasi-stationary distribution for absorbing LDQBDs. Finally we note that quasi-stationary distributions have great potential for modelling a wide variety of biological processes and efforts are currently being made to apply the results presented in Chapter 5 in these areas.

Bibliography

[1] I. J. B. F. Adan, J. Wessels, and W. H. M. Zijm. Analysis of the symmetric shortest queue problem. *Commun. Statist. - Stochastic Models*, 6(4):691–713, 1990.

[2] I. J. B. F. Adan, J. Wessels, and W. H. M. Zijm. Analysis of the asymmetric shortest queue problem. *Queueing Systems*, 8:1–58, 1991.

[3] I. J. B. F. Adan, J. Wessels, and W. H. M. Zijm. Analysis of the symmetric shortest queue problem with threshold jockeying. *Commun. Statist. - Stochastic Models*, 7(4):615–627, 1991.

[4] A. S. Alfa. A discrete MAP/PH/1 vacation queue with gated time-limited service. *Preprint, Department of mechanical and industrial engineering, University of Manitoba, Winnipeg, Manitoba, Canada*, 1994.

[5] W. J. Anderson. *Continuous-Time Markov Chains, An Applications-Oriented Approach.* Springer-Verlag, 1991.

[6] S. Asmussen. Aspects of matrix Wiener-Hopf factorisation in applied probability. *Math. Scientist*, 14:101–116, 1989.

[7] S. Asmussen and V. Ramaswami. Probabilistic interpretations of some duality results for the matrix paradigms in queueing theory. *Commun. Statist. - Stochastic Models*, 6(4):715–733, 1990.

[8] N. G. Bean, L. Bright, G. Latouche, C. E. M. Pearce, P. K. Pollett, and P. G. Taylor. The quasistationary behaviour of quasi-birth-and-death processes. *Submitted for publication*, 1994.

[9] N. G. Bean, P. K. Pollett, and P. G. Taylor. The quasistationary distributions of homogeneous quasi-birth-and-death processes. In *Proceedings of the 2nd Australia-Japan Workshop on Stochastic Models*, July 1996.

[10] D. Bini and B. Meini. On the solution of a nonlinear matrix equation arising in queueing problems. *Submitted for publication*, 1995.

[11] A. Brandt and G. Last. On the pathwise comparison of jump processes driven by stochastic intensities. *Preprint, Humbolt-Universität zu Berlin*, 1993.

[12] L. Bright and P. G. Taylor. Calculating the equilibrium distribution in level dependent quasi-birth-and-death processes. *Commun. Statist. - Stochastic Models*, 11(3):497–526, 1995.

[13] L. Bright and P.G. Taylor. Equilibrium distributions for level dependent quasi-birth-and-death processes. In *Proceedings of the first international conference on matrix-analytic methods in stochastic models*, Flint, Michigan, August 24-25 1995.

[14] K. L. Chung. *Markov chains with stationary transition probabilities*. Springer-Verlag, 1966.

[15] J. W. Cohen and O. J. Boxma. *Boundary value problems in queueing system analysis*. North-Holland, Amsterdam, 1983.

[16] S. Dambrine and M. Moreau. Note on the stochastic theory of a self-catalytic chemical reaction, I. *Physica*, 106A:559–573, 1981.

[17] S. Dambrine and M. Moreau. Note on the stochastic theory of a self-catalytic chemical reaction, II. *Physica*, 106A:574–588, 1981.

[18] J. R. Day and H. P. Possingham. A stochastic metapopulation model with variability in patch size and position. *Theoretical Population Biology*, 48:333–360, 1995.

[19] C. Derman. Some contributions to the theory of denumerable Markov chains. *Trans. Amer. Math. Soc.*, 79:541–555, 1955.

[20] J. E. Diamond and A. S. Alfa. Matrix analytical methods for M/PH/1 retrial queues. *Commun. Statist. - Stochastic Models*, 11(3):447–470, 1995.

[21] R. L. Dobrušin. Some classes of homogeneous denumerable Markov processes. *Teoriya Veroyatnostei i Ee Primeneniya*, 2:377–380, 1957.

[22] R.V. Evans. Geometric distribution in some two-dimensional queueing systems. *Operations Research*, 15:830–846, 1967.

[23] G. Falin. On sufficient conditions for ergodicity of multichannel queueing systems with repeated calls. *J. Appl. Prob.*, 16:447–448, 1984.

[24] G. Falin. A survey on retrial queues. *Queueing systems*, 7:127–168, 1990.

[25] W. Feller. *An introduction to probability theory and its applications.* John Wiley and Sons, 1966.

[26] D. C. Flaspohler. Quasi-stationary distributions for absorbing continuous-time denumerable Markov chains. *Ann. Inst. Statist. Math.*, 26:351–356, 1974.

[27] L. Flatto and H. P. McKean. Two queues in parallel. *Comm. Pure Appl. Math*, 30:255–263, 1977.

[28] F. R. Gantmacher. *The theory of matrices.* Chelsea, New York, 1959.

[29] D. P. Gaver, P. A. Jacobs, and G. Latouche. Finite birth-and-death models in randomly changing environments. *Adv. Appl. Prob.*, 16:715–731, 1984.

[30] I. Gertsbakh. The shorter queue problem: A numerical study using the matrix-geometric solution. *Eur. J. Oper. Res.*, 15:374–381, 1984.

[31] A. Graham. *Kronecker products and matrix calculus with applications.* Ellis Horwood Limited, 1981.

[32] B. S. Greenberg. M/G/1 queueing systems with returning customers. *J. Appl. Prob.*, 26:152–163, 1989.

[33] F. A. Haight. Two queues in parallel. *Biometrika*, 45:401–410, 1958.

[34] B. Hajek. Birth-death processes on the integers with phases and general boundaries. *J. Appl. Prob.*, 19:488–499, 1982.

[35] S. Halfin. The shortest queue problem. *J. Appl. Prob.*, 22:865–878, 1985.

[36] T. E. Harris. Transient Markov chains with stationary measures. *Proc. Amer. Math. Soc.*, 8:937–942, 1957.

[37] G. J. van Houtum. *New Approaches for Multi-Dimensional Queueing Systems.* PhD thesis, Eindhoven University of Technology, 1994.

[38] S. Karlin and M. Taylor. *A First Course in Stochastic Processes.* Academic press, 2nd edition, 1975.

[39] F. P. Kelly. *Reversibility and stochastic networks.* Wiley, London, 1979.

[40] F. P. Kelly. Invariant measures and the q-matrix. *In J.F.C. Kingman and G.E.H. Reuter (eds.), Probability, statistics and analysis. London Mathematical Society Lecture Notes Series*, 79:143–160, 1983.

[41] J. G. Kemeny and J. L. Snell. *Finite Markov Chains.* Springer-Verlag, 1976.

[42] D. G. Kendall and G. E. H. Reuter. The calculation of the ergodic projection for Markov chains and processes with a countable infinity of states. *Acta Math.*, 97:103–144, 1957.

[43] H. Kesten. A ratio limit theorem for (sub) Markov chains on $\{1, 2, \dots\}$ with bounded jumps. *Adv. Appl. Prob*, 17(3):652–691, 1995.

[44] M. Kijima. Some results for uniformisable semi-Markov processes. *Austral. J. Statist.*, 29:193–207, 1987.

[45] M. Kijima. Quasi-stationary distributions of single-server phase-type queues. *Mathematics of operations research*, 18(2):423–437, 1993.

[46] M. Kijima and N. Makimoto. Quasi-stationary distributions of Markov chains arising from queueing processes : A survey. Technical Report B-304, Tokyo Institute of Technology, August 1995.

[47] J. F. C. Kingman. Two similar queues in parallel. *Ann. Math. Statist.*, 32:1314–1323, 1961.

[48] J. F. C. Kingman. The exponential decay of Markov transition probabilities. *Proc. London Math. Soc.*, 13:337–358, 1963.

[49] D. R. Klein. The introduction, increase, and crash of reindeer on St. Matthew island. *J. Wildlife Man.*, 32:351–367, 1968.

[50] C. Knessl, B. J. Matkowsky, Z. Schuss, and C. Tier. Two parallel queues with dynamic routing. *IEEE Trans. Commun.*, 34:1170–1175, 1986.

[51] A. G. de Kok. Algorithmic methods for single server systems with repeated attempts. *Statistica Neerlandica*, 38:23–32, 1984.

[52] E. K. Kyprianou. On the quasi-stationary distribution of the virtual waiting time in queues with Poisson arrivals. *J. Appl. Prob.*, 8:495–507, 1971.

[53] E. K. Kyprianou. On the quasi-stationary distributions of the GI/M/1 queue. *J. Appl. Prob.*, 9:117–128, 1972.

[54] G. Latouche. Algorithmic analysis of a multiprogramming-multiprocessor computer system. *J. Assoc. Comp. Mach.*, 28:662–679, 1981.

[55] G. Latouche. A note on two matrices occurring in the solution of quasi-birth-and-death processes. *Commun. Statist. - Stochastic Models*, 3(2):251–257, 1987.

[56] G. Latouche and V. Ramaswami. A logarithmic reduction algorithm for quasi-birth-and-death processes. *J. Appl. Prob.*, 30:650–674, 1993.

[57] G. Latouche and V. Ramaswami. The PH/PH/1 queue at epochs of queue size changes. Technical Report 310, Université Libre de Bruxelles, 1995.

[58] G. Latouche and P. G. Taylor. Invariant measures for null-recurrent and transient quasi-birth-and-death processes. *Preprint, Department of Applied Mathematics, The university of Adelaide, South Australia, Australia*, 1996.

[59] D. Liu and Y. Q. Zhao. Determination of explicit solution for a general class of Markov processes. In *Proceedings of the first international conference on matrix-analytic methods in stochastic models*, Flint, Michigan, August 24-25 1996.

[60] D. M. Lucantoni. New results on the single server queue with a batch Markovian arrival process. *Commun. Statist. - Stochastic Models*, 7:1–46, 1991.

[61] D. M. Lucantoni, K. S. Meier-Hellstern, and M. F. Neuts. A single-server queue with server vacations and a class of non-renewal arrival processes. *Adv. Appl. Prob.*, 22:676–705, 1990.

[62] N. Makimoto. Quasi-stationary distributions in a PH/PH/c queue. *Commun. Statist. - Stochastic Models*, 9(2):195–212, 1993.

[63] W. A. Massey. Stochastic orderings for Markov processes on partially ordered spaces. *Mathematics of Operations Research*, 12:350–367, 1987.

[64] M. G. Nair and P. K. Pollett. On the relationship between μ–invariant measures and quasi-stationary distributions for continuous-time Markov chains. *Adv. Appl. Prob.*, 25:82–102, 1993.

[65] V. Naoumov. Matrix-multiplicative approach to quasi-birth-and-death processes analysis. In *Proceedings of the first international conference on matrix-analytic methods in stochastic models*, Flint, Michigan, August 24-25 1995.

[66] R. D. Nelson. A performance evaluation of a general parallel processing model. *Performance Evaluation Review*, 18:13–26, 1990.

[67] M. F. Neuts. Further results on the M/M/1 queue with varying rates. *Opsearch*, 15:158–168, 1978.

[68] M. F. Neuts. Markov chains with applications in queueing theory, which have a matrix-geometric invariant measure. *Adv. Appl. Prob.*, 10:185–212, 1978.

[69] M. F. Neuts. The M/M/1 queue with randomly varying arrival and service rates. *Opsearch*, 15:139–157, 1978.

[70] M. F. Neuts. A versatile Markovian point process. *Journal of Applied Probability*, 16:764–779, 1979.

[71] M. F. Neuts. The probabilistic significance of the rate matrix in matrix-geometric invariant vectors. *J. Appl. Prob.*, 17:291–296, 1980.

[72] M. F. Neuts. *Matrix Geometric Solutions in Stochastic Models*. Johns Hopkins, 1981.

[73] M. F. Neuts. *Structured Stochastic Matrices of M/G/1 Type and their Applications.* Marcel Dekker, New York, 1989.

[74] M. F. Neuts and B. M. Rao. Numerical investigation of a multiserver retrial model. *Queueing Systems*, 7:169–190, 1990.

[75] R. W. Parsons and P. K. Pollett. Quasistationary distributions for some auto catalytic reactions. *J. of Statist. Phys.*, 46:249–254, 1987.

[76] P. K. Pollett. On a model for interference between searching insect parasites. *J. Austral. Math. Soc. Ser. B*, 31:133–150, 1990.

[77] P. K. Pollett. Analytical and computational methods for modelling the long-term behaviour of evanescent random processes. In *Proceedings of the 12th national conference of the Australian Society for Operations Research*, pages 514–535, 1993.

[78] P. K. Pollett and P. G. Taylor. On the problem of establishing the existence of stationary distributions for continuous-time Markov chains. *Probability in the Engineering and Information Sciences*, 7:529–543, 1993.

[79] V. Ramaswami. A duality theorem for the matrix paradigms in queueing theory. *Commun. Statist. - Stochastic Models*, 6(1):151–161, 1990.

[80] V. Ramaswami. Matrix analytic methods: A tutorial overview with some extensions and new results. In *Proceedings of the first international conference on matrix-analytic methods in stochastic models*, Flint, Michigan, August 24-25 1995.

[81] V. Ramaswami and G. Latouche. An experimental evaluation of the matrix-geometric method for the GI/PH/1 queue. *Commun. Statist. - Stochastic Models*, 5(4):629–667, 1989.

[82] V. Ramaswami and M. F. Neuts. Some explicit formulas and computational methods for infinite-server queues with phase-type arrivals. *J. Appl. Prob.*, 17(2):498–514, 1980.

[83] V. Ramaswami and P. G. Taylor. An operator-analytic approach to product-form networks. *Commun. Statist. - Stochastic Models*, 12(1), 1996.

[84] V. Ramaswami and P. G. Taylor. Some spectral properties of the rate operators in level-dependent quasi-birth-and-death processes with a countable number of phases. *Commun. Statist - Stochastic Models*, 12(1), 1996.

[85] B. M. Rao and M. J. M. Posner. Parallel exponential queues with dependent service rates. *Comput. and Ops. Res.*, 13(6):681–692, 1986.

[86] B. M. Rao and M. J. M. Posner. Algorithmic and approximation analyses of the shorter queue model. *Naval Research Logistics*, 34:381–398, 1987.

[87] G. E. H. Reuter. Denumerable Markov processes and the associated contraction semi-groups on ℓ. *Acta Mathematica*, 97:1–46, 1957.

[88] K. A. Ross and C. R. B. Wright. *Discrete Mathematics.* Prentice Hall, 2nd edition, 1988.

[89] E. Seneta. *Non-negative matrices and Markov chains.* Springer-Verlag, 1981.

[90] E. Smeitink. *Stochastic Models for Repairable Systems.* PhD thesis, Vrije University, 1992.

[91] D. Stoyan. *Comparison Methods for Queues and Other Stochastic Models.* Wiley series in probability and mathematical statistics, 1983.

[92] D. Vere-Jones. Geometric ergodicity in denumerable Markov chains. *Quart. J. Math. Oxford, Ser. 2*, 13:7–28, 1962.

[93] D. Vere-Jones. Some limit theorems for evanescent processes. *Austral. J. Statist.*, 11:67–78, 1969.

[94] V. Wallace. *The solution of quasi birth and death processes arising from multiple access computer systems.* PhD thesis, University of Michigan, 1969.

[95] S. J. Yan and M. F. Chen. Multidemensional q-processes. *Chinese annals of mathematics*, 7:90–110, 1986.

[96] T. Yang and J. G. C. Templeton. A survey of retrial queues. *Queueing systems*, 2:201–233, 1987.

[97] J. Ye and S. Li. Folding algorithm: A computational method for finite QBD processes with level-dependent transitions. *IEEE Trans. Commun.*, 42:625–639, 1994.

[98] M. Zukerman. Applications of matrix-geometric solutions for queueing performance evaluation of a hybrid switching system. *J. Aust. Math. Soc. Ser. B*, 31:219–239, 1989.

[99] M. Zukerman and P. Kirton. Applications of matrix-geometric solutions to the analysis of the bursty data queue in a B-ISDN switching system. In *Proceedings of GLOBECOM '88*, pages 1635–1639, 1988.