TAIL BEHAVIOUR FOR STATIONARY DISTRIBUTIONS FOR TWO-DIMENSIONAL STOCHASTIC MODELS

by

Lani Haque, B.Sc., M.Sc.

A thesis submitted to
the Faculty of Graduate Studies and Research
in partial fulfilment of
the requirements for the degree of

Doctor of Philosophy

School of Mathematics and Statistics
Ottawa-Carleton Institute for Mathematics and Statistics
Carleton University
Ottawa, Ontario, Canada
November 2003

© Copyright
2003, Lani Haque

National Library
of Canada

Bibliothèque nationale
du Canada

Acquisitions and
Bibliographic Services

Acquisitions et
services bibliographiques

395 Wellington Street
Ottawa ON K1A 0N4
Canada

395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file Votre référence

ISBN: 0-612-89891-1

Our file Notre référence

ISBN: 0-612-89891-1

The author has granted a non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of this thesis in microform, paper or electronic formats.

L'auteur a accordé une licence non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de cette thèse sous la forme de microfiche/film, de reproduction sur papier ou sur format électronique.

The author retains ownership of the copyright in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

L'auteur conserve la propriété du droit d'auteur qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

In compliance with the Canadian Privacy Act some supporting forms may have been removed from this dissertation.

Conformément à la loi canadienne sur la protection de la vie privée, quelques formulaires secondaires ont été enlevés de ce manuscrit.

While these forms may be included in the document page count, their removal does not represent any loss of content from the dissertation.

Bien que ces formulaires aient inclus dans la pagination, il n'y aura aucun contenu manquant.

Canada

Abstract

It is well known that for a QBD process or Markov chain of $GI/M/1$ type with finitely many phases the spectrum of the rate matrix R characterizes the stationary probability measure. For a QBD process with both level and phase infinite it is of interest to know whether the spectral properties of the infinite rate matrix analogously characterizes the stationary probability measure. Under some conditions, which will be presented, it can be shown that the spectrum of R guarantees a stationary probability measure whose tail is geometric, i.e. $\pi_{mi} \sim c\eta^m x_i$ where η is the eigenvalue of R with corresponding left eigenvector $\mathbf{x} = (x_0, x_1, x_2, \dots)$. The method by which η and $\mathbf{x}$ may be found will be described using a QBD process with an infinite block repeating structure. The situation where the infinite blocks have a tridiagonal structure will be considered first. Then the method will be generalized to when the infinite blocks have an $M/G/1$ type structure. While we have only considered the method for a QBD process it also holds for any infinite block repeating structure such as $M/G/1$, $GI/M/1$ and $GI/G/1$ type structures. Our concern is not with this structure but rather with the internal structure of the infinite blocks. It is the internal structure that is important for determining η and related quantities, i.e. $\mathbf{x}$ and $\mathbf{y}^T$. What makes this method attractive is its simplicity. Finding η reduces to determining the zeros of a polynomial in the first situation and an infinite series in the latter.

Acknowledgements

I would like to express my thanks and gratitude to Professor Y. Zhao for introducing me to the area of Applied Probability as well as for his advice and direction during the course of this project. I would also like to express my appreciation to Jiashan Tang for helpful discussions related to Chapter 5. My sincere thanks to the Natural Science and Engineering Research Council (NSERC), Ontario Graduate Scholarships (OGS) and to the School of Mathematics and Statistics at Carleton University for their financial support during my academic program.

I would also like to thank my family and friends for their endless support throughout this period. Finally, I would like to acknowledge Patrick Boily. Without his tireless technical help and support with figures and formatting, it would have been impossible to produce this thesis.

Contents

Acceptance Form	ii
Abstract	iii
Acknowledgements	iv
1 Introduction	1
2 Preliminaries	9
2.1 Censoring	9
2.2 β -invariance	23
3 Determination of η and $\mathbf{x}$ for Irreducible, Scalar, Tridiagonal P	39
3.1 Tail Theorem	42
3.2 Examples	55
3.2.1 Two-Demand Model	55
3.2.2 Model Motivated by Wang's Model	64
3.2.3 Priority Queue	76
3.2.4 Symmetric Shortest Queue	84
4 Determination of η and $\mathbf{x}$ for Generalized P	90
4.1 Generalizations	90

4.1.1	Irreducible $M/G/1$ Type	91
4.1.2	Irreducible $GI/M/1$ Type	97
4.2	Examples	102
4.2.1	Tridiagonal structure	103
4.2.2	Geometric Jumps	104
5	Determination of η and $\mathbf{x}$ for Reducible, Scalar, Tridiagonal P	112
5.1	Reducible Case	113
5.2	Examples	123
5.2.1	Gated Random Order of Service	123
5.2.2	Priority Queue	130
6	Application: Two-Stage Inventory-Queue Model	134
6.1	Introduction	135
6.2	Mathematical model	138
6.2.1	The echelon base-stock policy	139
6.2.2	The Kanban base-stock policy	144
6.3	Tail asymptotics of the echelon base-stock model	146
6.4	Starvation probability	148
6.5	Numerical studies	153
6.6	Conclusions	155
7	Further Research	170
	Bibliography	172

List of Tables

3.1	η, α and $\bar{\alpha}$ values for Wang's modified Model discussed in Section 3.2.2 for various values of μ_0, μ_1, μ_2 and p	74
3.2	R and G measure values for Wang's modified model discussed in Section 3.2.2 for various values of μ_0, μ_1, μ_2 and p	75
6.1	Comparison of calculated p_s^E and p_s^K for various values of μ_1 and μ_2	157
6.2	Comparison of calculated p_s^E and p_s^K for various values of μ_1 and μ_2	158
6.3	Comparison of calculated p_s^E and p_s^K for various values of μ_1 and μ_2	159
6.4	Comparison of calculated p_s^E and p_s^K for various values of μ_1 and μ_2	160
6.5	Comparison of calculated p_s^E and p_s^K for various values of μ_1 and μ_2	161
6.6	Comparison of calculated p_s^E and p_s^K for various values of μ_1 and μ_2	162
6.7	Approximation of p_s^E when $\mu_1/\mu_2 \approx \lambda/\mu_2$	163
6.8	Approximation, p_t^E , of p_s^E using result from Theorem 18	164
6.9	Improved approximation for p_s^E using $a_k = 1/(k+1)$ and $b_k = k/(k+1)$ in (6.4.3).	165
6.10	Comparison of calculated p_s^E for when $\mu_1 \geq \mu_2$ and when $\mu_1 \leq \mu_2$	166
6.11	Comparison of calculated p_s^E for when $\mu_1 \geq \mu_2$ and when $\mu_1 \leq \mu_2$	167
6.12	Comparison of calculated p_s^E for when $\mu_1 \geq \mu_2$ and when $\mu_1 \leq \mu_2$	168
6.13	Comparison of calculated p_s^E for when $\mu_1 \geq \mu_2$ and when $\mu_1 \leq \mu_2$	169

List of Figures

3.1	Transition diagram for Two Demand Model.	57
3.2	Wang's model.	66
3.3	Transition diagram for the model motivated by Wang's Model.	67
3.4	Transition diagram for Priority queue.	76
3.5	Transition diagram for the Symmetric Shortest Queue.	85
4.1	Diagram for Geometric Jump model.	105
4.2	Transition rate diagram for geometric jump model where $p_i = (1-p)p^i$ for $i = 0, 1, 2, \dots$	106
5.1	Transition rate diagram for the Gated Random Order of Service Model.	124
6.1	Transition diagram for echelon controlled two-stage inventory queue. .	140

Chapter 1

Introduction

In the literature, a number of methods have been used to determine the stationary distribution for a two dimensional Markov process.

The compensation method, developed by Adan, [1], works for the class of two dimensional nearest neighbour random walks, those with unit steps, for which no one step transitions are permitted to the north, north east, and east of a state space. An illustration of the use of this method may be found in [4]. This method tries to solve the equilibrium equations by a linear combination of products. First, a basis of product form solutions satisfying the equilibrium equations inside the state space is characterized. This basis is used to make a linear combination that also satisfies the equilibrium equations for the boundary states. A procedure is given to find the initial pairs of elements. These pairs are called feasible pairs. Subsequent elements are chosen so as to compensate for the error created by the previous term on one of the two boundaries. As the error created on one boundary is compensated for, another error is created on the other boundary. Continuing in this way, an infinite series of products is created. Limitations of this method is that it does not work for two dimensional random walks with north, north east and east nearest neighbour

transitions. In higher dimensions the restrictions increase. The advantage of this method is that when the conditions are met an explicit solution could be obtained.

Another method is the boundary value method. This is a complex function method. A complex function method aims at solving the functional equation of the generating function of the equilibrium distribution. The equilibrium equations are solved by introducing the generating function for the equilibrium distribution and studying the functional equations it should satisfy. Usually, the functional equations present difficulties when trying to study them. For a class of two dimensional random walks and queueing problems, techniques have been developed to reduce those functional equations to standard Riemann Hilbert boundary value problems and to singular integral equations for complex valued functions. Cohen and Boxma go through the general theory of this, [16].

The uniformization method is another complex function method. The equilibrium equations for the two dimensional equilibrium distribution lead to a functional equation for the generating function. The generating function $f(x, y)$ of the two dimensional equilibrium distribution is introduced. The equilibrium equations for the distribution leads to a functional equation for $f(x, y)$. A simple form of the kernel $K(x, y)$ allows the following approach. $K(x, y) = 0$ defines an algebraic curve which is uniformized in a convenient way. An uniformizing variable p is introduced and the variables x and y are written in terms of p , $x = x(p)$, $y = y(p)$. Now $f(x(p), 0)$ and $f(0, y(p))$ are considered. These are shown to be analytic functions of p in certain regions of p . Also, for all p such that $|x(p)| \leq 1$ and $|y(p)| \leq 1$ for which $K(x(p), y(p)) = 0$ the boundedness of $f(x(p), y(p))$ results in a relation between $f(x(p), 0)$ and $f(0, y(p))$. This relation is used to continue $f(x(p), 0)$ and $f(0, y(p))$ meromorphically onto the plane $0 < |p| < \infty$. Starting with a certain pole another pole may be found that is linked to it by the fact that together, the kernel equals

zero. Flatto and McKean apply this method to the symmetric shortest queue, [20].

The precedence relation method is not restricted to just two queues, however produces no exact results, but only bounds. It is a systematic approach for constructing bounds for the average “cost” in Markov chains. For example, the shortest queue model can be represented by a continuous time Markov chain with state space (m, n) where m and n are the lengths of the shortest and longest queues respectively. The total “cost” would be the total number of customers, $m + n$, in the system. The average “cost” would be the mean number of customers in the system. The steps involved are, first the continuous time Markov chain is converted into a discrete time Markov chain by uniformizing the transition rates. This is done in such a way that the “cost” of the discrete Markov chain is the same as for the continuous Markov chain. The second step is to identify precedence relations. For example, precedence between states, e.g. $m = (m_1, m_2)$ has precedence over $n = (n_1, n_2)$, or is more attractive than $n = (n_1, n_2)$ if $V_t(m) \leq V_t(n)$ for $t \geq 0$, where $V_t(m)$ is the expected cost in the first t periods, (mean number of customers by time t), when starting at $m = (m_1, m_2)$. The third step is to construct the bounds. This is done by modifying the original Markov chain by redirecting at some states one or more transitions. For example, suppose at state (m_1, m_2) there are transitions to (n_1, n_2) and (s_1, s_2) . Redirecting a transition from (m_1, m_2) to (n_1, n_2) to (m_1, m_2) to (s_1, s_2) means the new transition probability from (m_1, m_2) to (n_1, n_2) is 0 and the new transition probability from (m_1, m_2) to (s_1, s_2) is $p_{ms} + p_{mn}$ = the old transition probability from (m_1, m_2) to (s_1, s_2) plus the old transition probability from (m_1, m_2) to (n_1, n_2) . We now have a modified Markov chain. The cost per period is not altered. The average “cost” of the new Markov chain is, say, $\tilde{g}$. The relation between the new average cost and the old average cost depends on the following. If the new chain was obtained by redirecting transitions to more or less attractive states, then it holds for all $t \geq 0$ that the expected t -period

cost vector in the new chain is less or greater than the one in the original chain, respectively. This implies $\tilde{g} \leq$ or $\geq g$, respectively. This result is crucial in constructing the bounds. Two modified chains can be created, one directing transitions to less attractive states and another to more attractive states. By the above fact this will give rise to an upper and lower bound. The precedence relation method is used in [61, 60] to construct bounds for the mean waiting time for the shortest queue model.

The power series algorithm is a numerical-analytical method for analyzing certain Markov processes. Some system parameter ρ is chosen and the steady state distribution is considered as a function of the system parameter, e.g. for the shortest queue model ρ could be the work load. The steady state queue length probabilities are functions of ρ . Assume these probabilities are analytic functions of ρ when traffic is light. They can be determined by the coefficients of the power series expansion in ρ around $\rho = 0$. These power series are substituted into the balance and normalization equations of the Markov process. Now we have equalities with analytic functions on either side of the equality and this does not hold unless all the coefficients of the power series expansion are equal. Equating the coefficients gives equations from which the calculation of the coefficients is possible. The power series algorithm is flexible. There is not much structure except for the assumption that it must be applied to a Markov process. Disadvantages are, it depends on the dimension of the system because it directly depends on the balance equations. There are no useful error bounds and it is sensitive to external parameter values. Blanc, [10], uses the power series algorithm to solve the shortest queue model.

A survey/review of a number of the methods stated above may be found in [2].

The paper by Bean and Nielsen, [8], is interesting in that it considers the factors that determine the decay rate of discrete phase-type distributions when there are countably infinite number of phases. This translates into a countably infinite number

of levels, i.e. the level variable takes on values, $0, 1, 2, \dots$. However, the phase variable takes on a finite number of values. So the transition matrix does not represent a *true* two-dimensional Markov chain. However, they are able to give a bound on the decay rate η . Namely, $1/\eta = \min(\beta, \gamma)$, [8, Theorem 1], where β is the convergence radius of the transition matrix and γ is the convergence radius of another related matrix. The authors consider the level dependent QBD case, level dependent $M/G/1$ case and the level independent QBD case. In the case, when the number of phases in each level is finite, various properties about the rate matrix R and stationary distribution π for the Markov chain are known. For example, $\pi_m \sim c\eta^m \mathbf{x}$ as $m \rightarrow \infty$, where η is the Perron-Frobenius eigenvalue of R , $\mathbf{x}$ is the corresponding left eigenvector of η and c is a multiplicative constant. When a matrix M is finite dimensional there is an inverse relationship between the radius of convergence for M and the dominant eigenvalue for M , i.e. $\alpha = 1/\eta$ where α is the radius of convergence for M and η is the Perron-Frobenius eigenvalue of M , [54, p.29].

When an irreducible Markov chain of $GI/G/1$ type has finitely many phases the spectrum of the rate matrix R characterizes the Markov chain, [63]. When there are a countably infinite number of phases this is not so straight forward. Takahashi *et al.*, [57], give conditions under which it is possible to show that for a QBD process with countably many levels and phases the stationary distribution has a geometric tail along the level direction, i.e. $\pi_{mn} \sim c\eta^m x_n$ as $m \rightarrow \infty$. Miyazawa, [47], improves upon the conditions given by Takahashi *et al.*, [57], for a process of $M/G/1$ type, rather than just a QBD process. In [48] Miyazawa and Zhao extend the main result of [47] which relaxes the conditions as well as the modelling assumptions presented by Takahashi *et al* in [57]. However, of these last three papers, [57, 47, 48], a means of determining η and $\mathbf{x} = (x_0, x_1, x_2, \dots)$ remains unaddressed. In this dissertation we show that η is in fact the reciprocal of the radius of convergence for R and η acts

as the eigenvalue for R . We also go on to show that under the conditions in [48] P is $1/\eta$ -positive. The main contribution of this dissertation is a method to determine η and $\mathbf{x}$ and its application to various modeling examples.

In this dissertation we look at the asymptotic behaviour of the stationary distribution for two-dimensional Markov chains. The case where the stationary distribution has a light tail and more specifically a geometric tail along one dimension is considered. Others who have considered the asymptotic behaviour for two-dimensional Markov chains are Foley and McDonald, [21], and McDonald, [45]. They consider a system of two servers each with a dedicated stream of customers and another stream of smart customers that join the shortest queue. They give an exact asymptotic description of the steady state distribution of the system, [21, Theorem 5]. They consider, when the system is strongly pooled, if the proportion of smart customers is larger then the two servers overload in tandem, weakly pooled, both queues overloaded but in different proportions. The limiting distribution in these situations is also considered. They also give exact asymptotics of the mean time until overload. In Flatto and Hahn, [19], the generating function approach is used to derive the asymptotic behaviour of the stationary distribution for a two demand model. Further details of this model may be found in Section 3.2.1 of this dissertation. Fujimoto *et al.*, [24], are concerned with the geometric decay property of the stationary distribution in a two stage tandem queueing system, $PH/PH/1 \rightarrow /PH/1$. They give a threshold value at which the asymptotic behaviour changes. Bean *et. al*, [7], give a detailed examination of the asymptotic properties of the tandem queue studied in [24] and provide an explanation for their results.

The method outlined in this dissertation has similarities to the one outlined in McDonald, [45], and Foley and McDonald, [21], where Markov additive processes, MAP, are used. The method in this dissertation is a special more simplified case

of the MAP structure. In our situation the phase is the Markov component of the MAP and the additive part is a the result of repeating structure along the level direction of the infinite blocks A , B and C . However, in McDonald, [45], and Foley and McDonald, [21], there is an additional condition, C7 in [21], which appears to be a drift condition. This condition is not present in the set of conditions by Takahashi et. al, [57], in Theorem 15. However, Takahashi et. al, [57], have the additional requirement, $\eta^{-1}\mathbf{x}A\mathbf{y}^T \neq \eta\mathbf{x}C\mathbf{y}^T$, given in Theorem 15 as 4. or in Condition 1 as (d). In Miyazawa and Zhao, [48], neither of these conditions are required.

In Chapter 2, background material and discussions relevant to remaining chapters are presented. Section 2.1 looks at the idea of censoring and Section 2.2 introduces the radius of convergence of a matrix and related topics. The contributions of this dissertation may be found in Chapters 3 - 6 of which the main ones will be highlighted below. In Chapter 3 the method used to determine η and $\mathbf{x}$ is presented in Section 3.1 for the case when P , given in equation 3.1.2, has a scalar, tridiagonal structure. The main contribution of this section may be found in Theorem 16. The procedure used to find $\mathbf{x}$ may also be found in this sections. Section 3.2 looks at examples of the method's use by applying it to various models. New results obtained by applying the method to Markovian models may be found in Sections 3.2.2 and 3.2.3. In Chapter 4 the method is generalized to consider the situation when P has an $M/G/1$ or $GI/M/1$ type scalar structure. In section 4.1 the equivalent of Theorem 16 may be found in Theorems 20 and 23 for the $M/G/1$ type and $GI/M/1$ type type structures of P , respectively. In Chapter 5 we consider the case when P is reducible and has a repeating, tridiagonal, scalar structure. Among the contributions of this chapter in section 5.1 Theorem 26 characterizes the states of P under these new conditions and gives a way of finding η . One of the main contributions is by means of an example which is used to illustrate that α -positivity is not sufficient to conclude that π_{mn} has

a geometric tail along one direction. This can be found in section 5.2.1. In Chapter 6 the method presented in Chapter 3 is applied to a two-stage inventory queueing model under the echelon base-stock control policy. The reason this example merits its own chapter is because not only is the method applied but it is also modified to give the best possible practical results for the starvation probability. The main contribution of this chapter may be found in sections 6.3-6.5. Section 6.3 considers the tail behaviour of the stationary distribution along the level direction. In section 6.4 and approximation of the starvation probability is found using results from section 6.3. In section 6.5 numerical studies of the starvation probability are done for different service rates for the two stages under the echelon base-stock control policy. Numerical results are also given to compare the starvation probability for the two-stage inventory queue under both the echelon and Kanban base-stock control policies. Chapter 7 is dedicated to a brief discussion on further research.

Chapter 2

Preliminaries

A number of concepts, definitions, ideas and conventions that will be used throughout this dissertation will be introduced and discussed in this chapter. Among the conventions used, we will consider a scalar to be a matrix with one row and one column. Thus, the ideas of irreducibility and stochasticity may be applied to scalars. Also, all bold, lower case Roman letters are row vectors, i.e. one row, finite or infinitely many columns, and “T” will denote the transpose of a matrix.

2.1 Censoring

Let $X = \{X_n; n = 1, 2, \dots\}$ be a discrete time irreducible Markov chain with state space S . Let $E \subseteq S$ be a non empty subset of S . Suppose there exists a sequence $\{n_t\}_{t=1}^{\infty}$ such that successive visits of X_n to E take place at times $0 < n_1 < n_2 < \dots$. Then the process $X^E = \{X_t^E = X_{n_t}; t = 1, 2, \dots\}$ is called a *censored process* with censoring set E , [22, 34, 38, 64]. Censored Markov chains may also be called *restricted*, *watched* or *embedded* Markov chains. The transition probability matrices for the original and censored Markov chains will be denoted by P and P^E respectively. The

sample paths of the censored process X^E are paths of the Markov chain X with transitions in the set E^c omitted. The process X^E may also be thought of as being obtained by watching the Markov chain X while only in E or by censoring it from E^c , [22, 38, 41]. Censoring may also be thought of as restricting the range of the Markov chain X . That is X^E may be thought of as being obtained from X by omitting the times n for which $X_n \notin E$.

Above we only considered the case when the Markov chain is irreducible. If we consider a Markov chain that is reducible then there is more than one class of states and each class is irreducible and either recurrent or transient. Consequently, the state space may be partitioned into irreducible classes that are either recurrent or transient. Without loss of generality, we will assume that the censored process starts in the censoring set E . This means that there is a non empty sequence $\{n_t\}_{t=1}^\infty$ such that successive visits of X_n to E take place at times $0 < n_1 < n_2 < \dots$. If the censoring set is contained within or intersects a recurrent class, then with probability one the Markov chain will return to that class and hence the censoring set. Thus, the sequence $\{n_t\}_{t=1}^\infty$ will be infinite and the censored process X^E will have infinitely many terms. If the censoring set is contained in a transient class then there is a positive probability that the Markov chain will never return to that class and hence the censoring set. Thus, there is a positive probability that the sequence $\{n_t\}_{t=1}^\infty$ is finite and hence the censored process has finitely many terms.

Another definition of a censored Markov chain is the one by Freedman [22]. Let $X = \{X(t); 0 \leq t < \infty\}$ be a Markov chain with state space S . Let $S_j(\omega) = \{t; X(t, \omega) = j, j \in S\}$ be a level set where a *level set* is defined as the set of times at which the sample path of the process X for some fixed ω , is at state $j \in S$. Let $E \subseteq S$ be finite and $S_E(\omega) = \{t; X(t, \omega) \in E\} = \bigcup_{j \in E} S_j(\omega)$ for some fixed ω . Let

$$\mu_E(t, \omega) : S_E(\omega) \rightarrow [0, \infty)$$

where ω is fixed and $\mu_E(t, \omega)$ = Lebesgue measure on the set $[0, t] \cap S_E(\omega)$, be a continuous, strictly increasing, onto map on $S_E(\omega)$. Also, $\mu_E(0, \omega) = 0$ and $\mu_E(\infty, \omega) = \lim_{t \rightarrow \infty} \mu_E(t, \omega) = \infty$. $\mu_E(t, \omega)$ may be thought of as the time on the time scale of the censored process X^E corresponding to time t on the time scale of the original Markov chain X . For fixed ω , let $\gamma_E(t, \omega)$ be the greatest value s such that $\mu_E(s, \omega) = t$ or $\gamma_E(t, \omega) = \max\{s; \mu_E(s, \omega) = t\}$. Then $\gamma_E(t, \omega)$ is a right continuous function, strictly increasing with range $S_E(\omega)$. When $t = 0$, $\gamma_E(0, \omega) = \max\{s; \mu_E(s, \omega) = 0\}$ is the largest value s such that $\mu_E(s, \omega) = 0$. This means we need to find all values s such that the measure of the set $[0, s] \cap \{t; X(t, \omega) \in E\}$ is zero. The set $[0, s] \cap \{t; X(t, \omega) \in E\}$ has measure zero when:

1. $s = 0$,
2. $\{t; X(t, \omega) \in E\}$ is empty or discrete, or
3. $[0, s] \cap \{t; X(t, \omega) \in E\}$ is empty or discrete.

The first two cases are trivial so we will consider the third case. Since the interval $[0, s]$ is continuous and $X(t, \omega)$ is a continuous function for fixed ω , $\{t; X(t, \omega) \in E\}$ will consist of the union of continuous intervals. The only way $[0, s] \cap \{t; X(t, \omega) \in E\}$ will have a measure of zero is when their intersection is a point. This occurs when $s = s_0$, the first or smallest value for which $X(s_0, \omega) \in E$. Or, $s_0 = \inf\{t; X(t, \omega) \in E\}$. Since there is only one such value s_0 ,

$$\gamma_E(0, \omega) = \max\{s; \mu_E(s, \omega) = 0\} = s_0 = \inf\{t; X(t, \omega) \in E\}.$$

Also, $\gamma_E(\infty, \omega) = \infty$. $\gamma_E(t, \omega)$ may be thought of as the largest time on the time scale for the Markov chain X corresponding to time t on the time scale for the restriction of X on E . Now the restriction X^E of X on E may be defined as the process,

$$X^E(t) = X(\gamma_E(t)),$$

namely,

$$X^E(t, \omega) = X(\gamma_E(t, \omega), \omega).$$

We also may write $X^E(\mu_E(t, \omega), \omega) = X(t, \omega)$. This restricted process X^E is called a *censored process* with censoring set E . This definition means that X^E is the process X watched only when in the set E . The times when X is not in E are erased from the time axis. The time axis is then collapsed to form a continuous line again [22].

The idea of a censored process for a Markov chain may be generalized to a censored process for an arbitrary stochastic process as is done in Chao and Miyazawa [15]. Let $X = \{X_n; n = 1, 2, \dots\}$ be a discrete time stochastic process with countable state space S and a unique limiting distribution π . The limiting distribution may be defined as

$$\lim_{m \rightarrow \infty} p_{ij}^{(m)} = \pi_j,$$

for all i , where $i, j \in S$ and $p_{ij}^{(m)}$ is the probability that given the stochastic process starts in state i it will be at state j on the m^{th} step. Let $E \subseteq S$ be such that the m^{th} time, of which X exits E , tends to infinity as the time n tends to infinity. So, the process visits E infinitely often. Let τ_n be the first point in time when the accumulated time, with which the process X is in E , is n . In other words,

$$\tau_n = \inf \left\{ k \geq n; \sum_{j=0}^k \chi_{\{X_j \in E\}} = n \right\},$$

where χ_A is the indicator function of the set A , i.e.

$$\chi_A(x) = \begin{cases} 1, & x \in A \\ 0, & x \notin A \end{cases}.$$

Therefore,

$$\sum_{j=0}^{\tau_n} \chi_{\{X_j \in E\}} = n.$$

Define the stochastic process $X^E = \{X_n^E; n = 1, 2, \dots\}$ with state space E as,

$$X_n^E = X_{\tau_n}, \quad n = 1, 2, 3, \dots$$

X^E is the truncated process from X whose state space is E . X^E is called the *censored process* of X on the truncated state space E .

Lemma 1 [15, Lemma 1.1] *The stationary distribution of the censored stochastic process X_n^E of X_n on E is the truncated distribution of π on E .*

As a special case, the above lemma for Markov chains also holds and may be found in [41, Lemma 2].

From the definition of a censored process we have, intuitively, as the finite subset E increases to the entire state space S that the censored process X^E approximates X more and more closely. The following theorem expresses this idea.

Theorem 1 [22, Theorem 48] *For each $t > 0$, as E increases to S , the censored process $X^E(t)$ converges to $X(t)$, where $X = \{X(t); t \geq 0\}$ is a Markov chain on the state space S with initial state $i \in S$, i.e. $X(0) = i$.*

This theorem considers the censored process itself. However, sometimes it is easier to consider the transition matrix of the stochastic process. The following lemma takes this into account while still expressing the same idea as in the Theorem 1.

Lemma 2 [64, Lemma 3] *Let P be a transition matrix with state space S and let E_n for $n = 1, 2, \dots$ be an increasing sequence of subsets of S such that $\lim_{n \rightarrow \infty} E_n = S = \bigcup_{n=1}^{\infty} E_n$. Then for any i, j fixed in E_n , $\lim_{n \rightarrow \infty} P_{ij}^{E_n} = P_{ij}$.*

This follows from Theorem 1 above. We will briefly describe how Lemma 2 follows from Theorem 1. $P_{ij}^{E_n}$ in Lemma 2 is given by

$$P_{ij}^{E_n} = P(X^{E_n}(t) = j | X^{E_n}(0) = i)$$

where $i, j \in E_n$. For fixed t , let

$$A_n = \{\omega; X^{E_n}(t, \omega) = j\}, \quad B_n = \{\omega; X^{E_n}(0, \omega) = i\}, \quad \text{and } C_n = A_n \cap B_n.$$

Now $P_{ij}^{E_n}$ may be rewritten as

$$P_{ij}^{E_n} = \frac{P(C_n)}{P(B_n)}. \quad (2.1.1)$$

Since $E_n \uparrow S$ by Theorem 1 $X^{E_n}(t)$ converges to $X(t)$. As a result, $A_n \uparrow A$, $B_n \uparrow B$ and $C_n \uparrow C$ where

$$A = \{\omega; X(t, \omega) = j\}, \quad B = \{\omega; X(0, \omega) = i\},$$

and

$$C = \{\omega; X(t, \omega) = j \text{ and } X(0, \omega) = i\}.$$

Due to the continuity from below property in [9, p.25]

$$\lim_{n \rightarrow \infty} P(A_n) = P(A), \quad \lim_{n \rightarrow \infty} P(B_n) = P(B)$$

and hence

$$\lim_{n \rightarrow \infty} P(C_n) = P(C).$$

Taking the limit as n tends to infinity of both sides of (2.1.1) we get,

$$\lim_{n \rightarrow \infty} P_{ij}^{E_n} = \lim_{n \rightarrow \infty} \frac{P(C_n)}{P(B_n)} = \frac{P(C)}{P(B)}$$

and notice that

$$\frac{P(C)}{P(B)} = P(A|B) = P(X(t) = j | X(0) = i) = P_{ij},$$

as required.

The following theorem is useful when dealing with the convergence of computational algorithms, however further discussions on computational algorithms will not be presented here.

Theorem 2 [64, Theorem 7] *Let $P = (P_{ij})_{i,j=0,1,2,\dots}$ be the transition matrix of a recurrent Markov chain on the non-negative integers. For an integer $k \geq 0$, let $P(k) = (p_{i,j}(k))_{i,j=0,1,2,\dots,k}$ be a matrix such that, $p_{ij}(k) = p_{ij}$, for $i, j \leq \omega$, and $P(k)$ is either stochastic or substochastic matrix. For any fixed $n \geq 0$ let $E_n = \{0, 1, 2, \dots, n\}$ be the censoring set. Then,*

$$\lim_{k \rightarrow \infty} P^{E_n}(k) = P^{E_n}.$$

The definitions of censored processes by Freedman [22], Zhao [64], and Chao and Miyazawa [15] are all equivalent. In the definition by Zhao, [64], the censored process may be thought of as a subsequence of the original process. The time when the original Markov chain is in the set E , n_i , may be thought of as a function of the time on the time scale for the censored process. For example the first random variable X_1^E in the censored process occurs at the n_1^{th} random variable in the original Markov chain, the i^{th} random variable in the censored process, X_i^E , occurs at the n_i^{th} random variable in the original Markov chain and so on. In the definition by Freedman, [22], the censored process' time scale is a function, $\gamma_E(t)$, of the original Markov chain's time scale, i.e. the i^{th} random variable of the censored process, $X_E(i)$, occurs at the $\gamma_E(i)^{th}$ random variable of the original Markov chain, $X(\gamma_E(i))$. These two definitions describe the same process.

The following is a list of properties of a censored process.

Lemma 3 [11, Lemma 2.2] *If we let P be the transition matrix of a Markov chain, which is possibly stochastic, and let E be a subset of the state spaces then,*

1. *P is irreducible if and only if P^E is irreducible for all E .*
2. *P is recurrent if and only if P^E is recurrent for all E .*
3. *P is transient if and only if P^E is transient for all E .*

4. if P is irreducible, then P is recurrent if and only if P^E is recurrent for some E .
5. if P is irreducible, the P is transient if and only if P^E is transient for some E .
6. if $E_1 \subseteq E_2$ then $P^{E_1} = (P^{E_2})^{E_1}$.

Proof:

1. If $E = S$, the whole state space, then $P^E = P^S = P$ is irreducible.

Conversely, if P is irreducible, any state $i \in S$ can communicate with any other state $j \in S$. So, any $i \in E \subset S$ communicates with any $j \in E \subset S$ which means P^E is irreducible since $0 < P_{ij}^E$ is the transition probability of $i \in E$ to $j \in E$.

2. Take E to be the entire sample space S . Then we have $P^E = P$ which is recurrent.

Conversely, let $E \subset S$ be any subset. Since $j \in E \subset S$ is recurrent, the process will eventually return to j . Omitting the transitions of the process that are not in E we have that the censored process P^E returns to j given that it started at j within a finite length of time. That is,

$$1 = P(T_j < \infty | X_0 = j) \leq P^E(T_j^E < \infty | X_0 = j) = 1$$

where $T_j = \min\{n; X_n = j\}$, $T_j^E = \min\{n; X_n^E = j\}$ and X_n^E is the censored process.

Going from the original process to the censored process reduces the time scale for traveling from one state to another because all the the transition in the censored set are omitted. That is,

$$T_j \geq T_j^E$$

or

$$\{\omega; T_j(\omega) < \infty\} \subseteq \{\omega; T_j^E(\omega) < \infty\},$$

or

$$1 = P(T_j < \infty | X_0 = j) \leq P(T_j^E < \infty | X_0 = j),$$

or

$$1 = P(T_j^E < \infty | X_0 = j).$$

3. Take E to be the entire state space S . Then we have $P^E = P^S = P$ is transient.

Conversely, let $E \subset S$ and $j \in E$ be transient. So, $1 > P(T_j < \infty | X_0 = j)$ which means there is a positive probability of the process not returning to state j , given that it started at state j . Also,

$$T_j \geq T_j^E$$

or

$$\{\omega; T_j(\omega) < \infty\} \subseteq \{\omega; T_j^E(\omega) < \infty\},$$

or

$$P(T_j < \infty | X_0 = j) \leq P(T_j^E < \infty | X_0 = j) \leq 1.$$

If $P(T_j^E < \infty | X_0 = j) = 1$ then $j \in E$ is a recurrent state in the censored process and so is in the original process, a contradiction. Therefore, $P(T_j^E < \infty | X_0 = j) < 1$ and P^E is also transient.

4. Suppose P is irreducible. By 1., P^E is irreducible and by 2., P^E is recurrent for all E .

Conversely, if for some $E \subset S$, P^E is recurrent. Since P is irreducible, so is P^E . Irreducible P means every state in S communicates with every other state in S .

For any $j \notin E$ and $i \in E$, i and j communicate, since P is irreducible. P^E is recurrent means each state in E is recurrent but not necessarily in E , i.e. there may be transitions in E^c but the original process eventually goes back to E and back to the starting state in E . By Corollary 4.2.4., [53], the entire state space is recurrent, thus P is recurrent.

Note: $E = \{j\}$, a single state is sufficient.

5. We are given that P is irreducible so it is either transient or recurrent. If for some $E \subset S$ let P^E be transient. By 1., P^E is irreducible. If P is recurrent then by 4., P^E is recurrent, a contradiction. So, P must be transient [32, pp 23].

The converse, follows from 3.

6. P^{E_2} is the process with all the transition in E_2^c of the original process P omitted.

For any $E_1 \subset E_2$ we may take another further censored process $(P^{E_2})^{E_1}$. This is the process P^{E_2} with all the transition in $E_1^c \cap E_2$ omitted. This is the same as omitting all the transitions in $E_1^c \cup E_2^c = E_1^c$ since $E_2^c \subset E_1^c$. This is just the censored process P^{E_1} . Therefore, $(P^{E_2})^{E_1} = P^{E_1}$. ■

So far Markov chains with transition matrices consisting of scalar entries have been considered. Markov chains with block-structured transition matrices are those Markov chains whose transition probability matrix have matrices as its entries. Block

repeating Markov chains are those with a block-structure and a repeating pattern among the block entries. For example, Markov chains of Toeplitz or $GI/G/1$ type have a transition matrix with the following form,

$$P = \begin{bmatrix} D_0 & D_1 & D_2 & D_3 & \cdots & \cdots \\ D_{-1} & C_0 & C_1 & C_2 & \cdots & \cdots \\ D_{-2} & C_{-1} & C_0 & C_1 & \cdots & \cdots \\ D_{-3} & C_{-2} & C_{-1} & C_0 & \cdots & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots & \\ \vdots & \vdots & \vdots & \vdots & & \ddots \end{bmatrix}, \quad (2.1.2)$$

where the D_i 's and C_i 's are finite matrices. For certain values of the blocks D_i and C_i we get special cases of the Toeplitz type Markov chains. For example, when $D_i = D_{-i} = C_i = C_{-i} = 0$ for $i \geq 2$ we get a quasi-birth-and-death process. When $C_i = D_i = 0$ for $i \geq 2$ we get Markov chains of $GI/M/1$ type. When, $D_{-i} = C_{-i} = 0$ for $i \geq 2$ we get Markov chains of $M/G/1$ type as well as others, [64].

There are four probabilistic measures R, G, A and B dominant in studying Markov chains of Toeplitz type. For example they may be used to express the stationary probability distribution and to characterize conditions of classification of the states. Numerically, these measures can be efficiently calculated. They are also invariant under censoring. So, they may be studied for the censored Markov chain, which is usually finite, rather than for the original Markov chain, which is usually infinite [64]. Let $X = \{X_n; n = 1, 2, \dots\}$ be a discrete time Markov chain with state space S . Partition S in the following way, $S = \bigcup_{i=1}^{\infty} L_i$ where L_i is defined as $L_i = \{(i, j); j = 1, 2, \dots, m_i\}$, for any state (i, j) , i is called a *level* and j a *phase*. Also let $L_{\leq i} = \bigcup_{k=0}^i L_k$ and $L_{\geq i} = \bigcup_{k=i}^{\infty} L_k = L_{\leq (i-1)}^c$. Let the transition matrix of the Markov chain be partitioned according to levels, giving, $P = (P_{ij})_{i \geq 0, 1 \leq j \leq m_i}$ where P_{ij} is an $m_i \times m_j$

matrix. Now the probabilistic matrices may be defined as follows,

$$R_{ij}(r, s) = E(\text{ number of visits to state } (j, s) \text{ before hitting } L_{\leq(j-1)} | X_0 = (i, r)),$$

$$G_{ij}(r, s) = P(\text{ hitting } (j, s) \text{ upon entering set } L_{\leq(i-1)} \text{ for the first time } | X_0 = (i, r)),$$

$$A_{ij}(r, s) = E(\text{ number of visits to } (j, s) \text{ before hitting } L_i | X_0 = (i, r)),$$

$$B_{ij}(r, s) = P(\text{ hitting } (j, s) \text{ upon entering } L_j \text{ for the first time } | X_0(i, r)).$$

Each of these matrices are $m_i \times m_j$ matrices and each description above represents its $(r, s)^{th}$ entry.

Note: These definitions for the R, G, A and B measures also hold for transition matrices with a more general, non-repeating structure.

What is interesting about these four probabilistic measures is that they are all invariant under censoring. Specifically, let P be as in (2.1.2). If we censor P with censoring set $L_{\geq n} = \cup_{i=0}^n L_i$ then the R and G measure for the censored Markov chain are $R_{ij}^{(n)}$ and $G_{ij}^{(n)}$, respectively. Then given, $i = 0 < j$ or $0 \leq i \leq j$, $R_{ij}^{(n)} = R_{ij}$ for all $n \geq j$ and given $0 \leq j < i$, $0 \leq j < i$, $G_{ij}^{(n)} = G_{ij}$ for all $n \geq i$. Similarly for the A and B measures. Let $A_{ij}^{(n)}$ and $B_{ij}^{(n)}$ be the A and B measures, respectively, for the censored Markov chain. Then

$$A_{0j}^{(n)} = A_{0j}, \text{ for all } n \geq j > 0$$

and

$$B_{i0}^{(n)} = B_{i0}, \text{ for all } n \geq i > 0$$

[63]. This means, the four probabilistic measures, R, G, B , and A for the censored Markov chain are the same as those for the original Markov chain. This invariance property under censoring allows the invariant probability measure for the censored process, which is usually finite, to be studied rather than for the original Markov

chain, which is infinite, giving results which are the same for the original infinite state Markov chain. An important property about censoring is that the censored process is also a Markov chain. This is shown in the following theorem.

Theorem 3 [64, Theorem 2] *Let P be the transition matrix of the Markov chain $\{X_n; n = 1, 2, \dots\}$ partitioned according to subsets E and E^c ,*

$$P = \begin{matrix} & \begin{matrix} E & E^c \end{matrix} \\ \begin{matrix} E \\ E^c \end{matrix} & \begin{pmatrix} T & U \\ D & Q \end{pmatrix} \end{matrix} \quad (2.1.3)$$

Then, the censored process is a Markov chain and its transition probability matrix is given by

$$P^E = T + U\hat{Q}D \quad (2.1.4)$$

with $\hat{Q} = \sum_{k=0}^{\infty} Q^k$. If P has a unique stationary distribution $\{\pi_k\}$ then the stationary distribution $\{\pi_k^E\}$ of the censored Markov chain is given by,

$$\pi_k^E = \frac{\pi_k}{\sum_{i \in E} \pi_i}, \quad k \in E.$$

Remark 1 *Notice that the last result in the above theorem, namely, $\pi_k^E = \frac{\pi_k}{\sum_{i \in E} \pi_i}$, $k \in E$, is an application of Lemma 1.*

$\hat{Q}$ is called the *fundamental matrix* of Q and represents the transitions of the original Markov chain that are outside the set E . The probabilistic interpretation for Q is as follows, $(\hat{Q})_{ij}$ represents the expected number of visits to state $j \in E^c$ before entering E given that the process started at state $i \in E^c$. The results for discrete time Markov chains also hold for regular continuous time Markov processes and may be obtained using various arguments. Uniformization may be used when the diagonal entries of the generator of the Markov chain are bounded, i.e. $\sup\{|Q_{ii}|; i \in S\} < \infty$. When

the diagonal elements of the generator are not bounded, the Markov chain embedded at transition epochs may be used. Of course, the process may be directly considered in continuous time.

In the discrete time case the transition matrix contains all the information about the Markov chain except the initial state of the process. In the continuous case, the generator is the matrix containing all the information. For any subset $E \subseteq S$ the generator Q may be written as,

$$Q = \begin{matrix} & \begin{matrix} E & E^c \end{matrix} \\ \begin{matrix} E \\ E^c \end{matrix} & \begin{pmatrix} T & U \\ D & Q_{E^c} \end{pmatrix} \end{matrix}.$$

In Theorem 3 the fundamental matrix made an appearance. The continuous time version of the fundamental matrix is also of importance. For any subset $E^c \subseteq S$ the matrix of the expected length of time spent in E^c before first entering E given the process starts in E^c is given by

$$N = \int_0^\infty \exp(Q_{E^c} t) dt.$$

The continuous time version of Theorem 3 is as follows.

Theorem 4 [38, Theorem 5.5.3] *Let $\{X(t); t \geq 0\}$ be an irreducible homogeneous positive recurrent Markov process on a discrete state space S with generator Q . Let E be an arbitrary subset of S . The censored process $\{X^E(t); t \geq 0\}$ is an irreducible, homogeneous, positive recurrent Markov process on E . Its generator Q^E is given by,*

$$Q^E = T + U N D.$$

Its stationary probability vector is proportional to π^E , i.e. we have that $\pi^E Q^E = 0$.

2.2 β -invariance

Let T be an infinite, non-negative, irreducible matrix. It is known for infinite series that there is a disk within which the series converges [44, Definition 5.10.1]. In order to find this disk the radius of convergence α for the infinite series must be determined. A similar notion exists for infinite matrices.

Theorem 5 [54, Theorem 6.1] *The power series*

$$T_{ij}(z) = \sum_{k=0}^{\infty} t_{ij}^{(k)} z^k, \quad i, j = 1, 2, \dots \quad (2.2.1)$$

all have a common convergence radius α , $0 \leq \alpha < \infty$ for each pair i, j .

Here $t_{ij}^{(k)}$ represents the $(i, j)^{th}$ entry of the matrix T^k , $k = 1, 2, 3, \dots$. This common convergence radius α , $0 \leq \alpha < \infty$, of the power series (2.2.1) is called the *convergence parameter* or *convergence radius* of the matrix T , [54, Definition 6.1].

Note: When T is finite $\alpha = 1/r$ where r is the Perron-Frobenius eigenvalue of T [54].

For an arbitrary matrix $M = [m_{ij}]_{i,j}$ the eigenvalue is any scalar λ for which $M\mathbf{y}^T = \lambda\mathbf{y}^T$ where $\mathbf{y}$ is a nonzero row vector $\mathbf{y} = (y_0, y_1, y_2, \dots)$. When M is finite, say $n \times n$, and lower triangular with scalar entries, the determinant of $\lambda I - M$ is given by

$$\det(\lambda I - M) = \prod_{i=1}^n (\lambda - m_{ii}).$$

Here the order of $\det(\lambda I - M)$ is n . From here we see that the eigenvalues of M are $\lambda_i = m_{ii}$ for $i = 1, 2, \dots, n$. Regardless of the order of $\det(\lambda I - M)$, when M is lower triangular, $\det(\lambda I - M)$ vanishes when $\lambda_i = m_{ii}$. For a lower triangular infinite matrix with scalar entries from the definition, the eigenvalues are found to be

$\lambda_i = m_{ii}$, for $i = 1, 2, 3, \dots$, which is consistent with the finite case, [17, p.13]. The result in Cooke, [17, p.13], is for a lower triangular infinite matrix. We will encounter upper triangular infinite matrices. To get a similar result we may apply the above to the transpose of an upper triangular matrix. If U is an infinite upper triangular matrix, U^T is lower triangular whose eigenvalues are the entries along the diagonal. Each eigenvalue has a corresponding right eigenvector, $\mathbf{x}^T$, say, i.e. $U^T \mathbf{x}^T = \lambda \mathbf{x}^T$. Taking the transpose we have $\mathbf{x}U = \lambda \mathbf{x}$, implying the entries along the diagonal of U are eigenvalues of U with a corresponding left eigenvector.

The indices may be defined as α -recurrent, α -null or α -transient as follows. Let $l_{ij}^{(1)} = t_{ij}$ and define $l_{ij}^{(k)}$ inductively as,

$$l_{ij}^{(k+1)} = \sum_{r < i; j \geq 0} l_{ir}^{(k)} t_{rj}, \quad i, j = 1, 2, \dots, k = 0, 1, 2, \dots$$

and

$$L_{ij}(z) = \sum_{k=0}^{\infty} l_{ij}^{(k)} z^k.$$

Note: $l = [l_{ij}]$ is the same as the B measure defined in the previous section.

Since $l_{ij}^{(k)} \leq t_{ij}^{(k)}$, $L_{ij}(z)$ is convergent for $|z| < \alpha$. An index i is called α -recurrent if

$$L_{ii}(\alpha-) = \lim_{z \rightarrow \alpha-} L_{ii}(z) = 1$$

and α -transient if

$$L_{ii}(\alpha-) < 1.$$

An α -recurrent index i is said to be α -positive if

$$\mu_i(\alpha) = \alpha L'_{ii}(\alpha-) = \sum_k k l_{ii}^{(k)} \alpha^k < \infty$$

and α -null if

$$\mu_i(\alpha) = \infty.$$

Note that for $0 \leq s < \alpha$

$$L_{ii}(s) = 1 - (T_{ii}(s))^{-1} < 1$$

so that

$$L_{ii}(\alpha-) < 1 \text{ if and only if } \sum_{k=0}^{\infty} t_{ii}^{(k)} s^k = T_{ii}(s) < \infty$$

[54, Definition 6.2]. When the matrix T represents the transition matrix of a Markov chain the indices represent states in the state space. The above definitions describing the character of an index are thus describing the character of a state. As a result, the terms index and state may be used interchangeably.

Another way of determining the character of a particular state i of a Markov chain with transition matrix T is by the following theorem in Karlin and Taylor, [33].

Theorem 6 [33, Theorem 5.1, p.66] *A state i of a Markov chain with transition matrix T is recurrent if and only if,*

$$\sum_{n=0}^{\infty} T_{ii}^n = \infty.$$

If state i is recurrent we can further classify the state as positive recurrent or null recurrent as follows:

Definition 1 [33, p.85]

1. If $0 < \pi_i = \lim_{n \rightarrow \infty} T_{ii}^n$ for state i then state i is called positive recurrent.
2. If $\pi_i = 0$ then state i is called null recurrent.

We will be primarily considering $T = \beta P$, where P is a transition matrix, for $0 \leq \beta < \infty$ where βP is referred to as the discounted transition matrix at rate β . The entry $(\beta P)_{ij}$ represents the total discounted reward with rate β that the Markov chain will be at state j given it started at state i . Consider

$$\sum_{n=0}^{\infty} (\beta P)_{ii}^n,$$

where $(\beta P)_{ii}^n$ represents the total discounted reward with rate β that given the Markov chain starts in state i , it returns to itself on the n^{th} step. P is β -recurrent or β -transient as follows:

Lemma 4

$$\begin{aligned} \sum_{n=0}^{\infty} \beta^n p_{ii}^n &= \infty \text{ if and only if } i \text{ is } \beta\text{-recurrent} \\ \sum_{n=0}^{\infty} \beta^n p_{ii}^n &< \infty \text{ if and only if } i \text{ is } \beta\text{-transient}, \end{aligned}$$

where $P = [p_{ij}]_{i,j=0}^{\infty}$.

Once a state i of P is found to be β -recurrent, the state i may be further classified as β -positive or β -null by considering

$$\pi_i = \lim_{n \rightarrow \infty} T_{ii}^n = \lim_{n \rightarrow \infty} (\beta P)_{ii}^n. \quad (2.2.2)$$

The classifications are:

Lemma 5 1. $\pi_i > 0$ then i is β -positive,

2. $\pi_i = 0$ then i is β -null,

for each i .

We may define an invariant measure for T . Any vector $\mathbf{x} \geq \mathbf{0}$ satisfying

$$\alpha \mathbf{x} T \leq \mathbf{x} \quad (2.2.3)$$

is called an α -subinvariant measure. When equality in (2.2.3) occurs then $\mathbf{x}$ is called an α -invariant measure. Similarly we may define the α -subinvariant vector $\mathbf{y}^T$ of T . That is, any $\mathbf{y}^T \geq \mathbf{0}^T$ satisfying

$$\alpha T \mathbf{y}^T \leq \mathbf{y}^T \quad (2.2.4)$$

is called an α -subinvariant vector of T . When equality in (2.2.4) occurs then $\mathbf{y}^T$ is called an α -invariant vector of T .

If an α -subinvariant measure $\mathbf{x}$ for T exists, a substochastic matrix P may be defined by,

$$p_{ij} = \frac{\alpha x_j t_{ji}}{x_i}, \quad i, j = 1, 2, \dots, \quad (2.2.5)$$

and the k^{th} step by

$$p_{ij}^{(k)} = \frac{\alpha^k x_j t_{ji}^{(k)}}{x_i}$$

which is the $(i, j)^{th}$ entry of P^k .

It is apparent that since T is irreducible so is P . Since results for a substochastic matrix P are more readily available, using the transformation in (2.2.5) these results may be translated into results for T . Since P is irreducible all its indices / states are either transient or recurrent. Thus, all the indices / states of T are either all α -transient or α -recurrent. The latter of which occurs only if P is stochastic. We have that T is called α -transient, α -positive or α -null, if one of its indices / states is α -transient, α -positive or α -null, respectively.

For any value $\beta > 0$ we may define a β -subinvariant and β -invariant measure and vector for T . These definitions are parallel to those for α -subinvariant and α -invariant measure and vector for T , respectively.

A vector $\mathbf{x} \geq \mathbf{0}$ is a β -subinvariant measure of T if $\beta \mathbf{x} T \leq \mathbf{x}$, elementwise and is a β -invariant vector if $\beta \mathbf{x} T = \mathbf{x}$.

A vector $\mathbf{y}^T \geq \mathbf{0}^T$ is a β -subinvariant vector of T if $\beta T \mathbf{y}^T \leq \mathbf{y}^T$, elementwise and is a β -invariant measure if $\beta T \mathbf{y}^T = \mathbf{y}^T$.

The following two theorems are of considerable value and importance in the following chapters. Theorem 7 gives a relationship between α -positivity and T . Theorem 8 is a limiting relation between the entries of $\alpha^k T^k$ and the entries of $\mathbf{x} \mathbf{y}^T$, where $\mathbf{x}$ and $\mathbf{y}^T$ are the α -invariant measure and vector, respectively, of T .

Theorem 7 [54, Theorem 6.4] Suppose $\mathbf{x}$ is a β -invariant measure and $\mathbf{y}^T$ is a β -invariant vector of T . Then T is α -positive if

$$\mathbf{x} \mathbf{y}^T = \sum_i x_i y_i < \infty,$$

in which case $\beta = \alpha$ where α is the radius of convergence of T , $\mathbf{x}$ is a multiple of the α -invariant measure of T and $\mathbf{y}$ is a multiple of the unique α -invariant vector of T .

Conversely, if T is α -positive and $\mathbf{x}$ and $\mathbf{y}$ are, respectively, an α -invariant measure and vector of T then $\mathbf{x} \mathbf{y}^T < \infty$.

Theorem 8 [54, Theorem 6.5] If T is an aperiodic α -positive matrix, then as $k \rightarrow \infty$

$$\alpha^k t_{ij}^k \rightarrow \frac{x_j y_i}{\sum_j x_j y_j} > 0,$$

where $\mathbf{x}, \mathbf{y}$ are an α -invariant measure and vector of T , respectively.

Let T be a nonnegative matrix satisfying the following conditions,

1. $T^k = \{t_{ij}^{(k)}\}$, for $k \geq 1$, are all elementwise finite.
2. T is irreducible

Also, let

$$\bar{x}_{ij} = (1 - \delta_{ij})L(1-) + \delta_{ij}$$

where δ is the Kronecker delta function,

$$\delta_{ij} = \begin{cases} 1, & i = j \\ 0, & i < j \end{cases}.$$

Theorem 9 [54, Theorem 6.3] *For an α -recurrent matrix T , an α -invariant measure always exists and is a constant multiple of $\{\bar{x}_{ij}\}$, for $j = 1, 2, \dots$. An α -subinvariant measure which is not invariant exists if and only if T is α -transient; one such is then given by $\{\bar{x}_{ij}\}$, for $j = 1, 2, \dots$.*

Remark 2 *Since there is no “asymmetric” assumption such as stochasticity present, which tends to endow T and T^t with somewhat divergent properties an analogous result as Theorem 9 holds for T^t . It is easily shown that the common convergence radius α is the same for T and T^t and the subinvariant vector properties of T may be derived from the subinvariant measure properties of T^t , [54, Chapter 6].*

The Markov chains considered in this dissertation will be primarily quasi-birth-death processes or of the $M/G/1$ or $GI/M/1$ type. The following results hold for either $M/G/1$ or $GI/M/1$ type Markov chains. Since a QBD process is a special case of both the $M/G/1$ and $GI/M/1$ types, the results for QBD processes follow. The

transition matrices for $M/G/1$ type and $GI/M/1$ type Markov chains are as follows

$$\begin{bmatrix} D_1 & D_2 & D_3 & \cdots & \cdots \\ D_0 & C_1 & C_2 & \cdots & \cdots \\ & C_0 & C_1 & \cdots & \cdots \\ & & C_0 & \cdots & \cdots \\ & & & \ddots & \\ & & & & \ddots \end{bmatrix} \quad (2.2.6)$$

and

$$\begin{bmatrix} D_1 & D_0 & & & \\ D_2 & C_1 & C_0 & & \\ D_3 & C_2 & C_1 & C_0 & \\ \vdots & \vdots & \vdots & & \ddots \end{bmatrix}, \quad (2.2.7)$$

respectively, where D_i and C_i for $i \geq 0$ are either finite or infinite matrices. The transition matrix for a QBD process is a special case of the above two by letting $D_i = 0 = C_i$ for all $i \geq 3$ in either case. Assumptions on the transition matrices will be made when necessary.

In the remainder of this chapter definitions and results will be for Markov chains of $M/G/1$ type, unless otherwise stated. Let

$$P = \begin{bmatrix} D_1 & D_2 & D_3 & \cdots \\ D_0 & C_1 & C_2 & \cdots \\ & C_0 & C_1 & \cdots \\ & & C_0 & \cdots \\ & & & \ddots \end{bmatrix}. \quad (2.2.8)$$

Before presenting any results some definitions and notation will be introduced. The matrix P may be partitioned as

$$P = \begin{bmatrix} T & H \\ L & Q \end{bmatrix}, \quad (2.2.9)$$

where Q is the matrix constructed by deleting boundary blocks of P and T, H and L are the boundary blocks of P partitioned accordingly. The *fundamental matrix* for any matrix X is the minimal nonnegative inverse of $I - X$, is equal to $\sum_{k=0}^{\infty} X^k$ and denoted by $\widehat{X}$, i.e.

$$\widehat{X} = \sum_{k=0}^{\infty} X^k = (I - X)_{min}^{-1}. \quad (2.2.10)$$

For a scalar $0 < \beta \leq \alpha$, where α is the radius of convergence for P , βP is called the *discounted transition matrix* at rate β . The following lemma may be used to establish a relationship between block entries of the fundamental matrix $\widehat{\beta P}$.

Lemma 6 (Lemma 3, [40]) *Let P be partitioned as in (2.2.9) and let βP be partitioned accordingly as*

$$\beta P = \begin{bmatrix} \beta T & \beta H \\ \beta L & \beta Q \end{bmatrix}, \quad 0 < \beta \leq \alpha. \quad (2.2.11)$$

Assume that $I - \beta P$ is invertible. Then the minimal nonnegative inverse, $\widehat{\beta P}$, of $I - \beta P$ is given by

$$\widehat{\beta P} = \begin{bmatrix} (I - \beta T - \beta H \widehat{\beta Q} \beta L)_{min}^{-1} & (I - \beta T - \beta H \widehat{\beta Q} \beta L)^{-1} \beta H \widehat{\beta Q} \\ \widehat{\beta Q} \beta L (I - \beta T - \beta H \widehat{\beta Q} \beta L)_{min}^{-1} & \widehat{\beta Q} + \widehat{\beta Q} \beta L (I - \beta T - \beta H \widehat{\beta Q} \beta L)_{min}^{-1} \end{bmatrix}$$

or equivalently

$$\widehat{\beta P} = \begin{bmatrix} \widehat{\beta T} + \widehat{\beta T} \beta H (I - \beta Q - \beta L \widehat{\beta T} \beta H)_{min}^{-1} & \widehat{\beta T} \beta H (I - \beta Q - \beta L \widehat{\beta T} \beta H)_{min}^{-1} \\ (I - \beta Q - \beta L \widehat{\beta T} \beta H)_{min}^{-1} \beta L \widehat{\beta T} & (I - \beta Q - \beta L \widehat{\beta T} \beta H)_{min}^{-1} \end{bmatrix}$$

where $(I - X)_{min}^{-1} = \sum_{i=0}^{\infty} X^i$ is the minimal nonnegative inverse of $I - X$.

Remark 3 [40, Remark 1] *By a sample path argument or Lemma 6, we can show that the fundamental matrix is invariant under censoring. Let E be a subset of the state space. Let βP be partitioned according to E and its complement E^c as in (2.2.11). And let the fundamental matrix $\widehat{\beta P}$ of βP be expressed as in (2.2.10). Then, the fundamental matrix of the censored matrix $(\beta P)^E$ is equal to the block-entry corresponding to the states in E in the fundamental matrix $\widehat{\beta P}$.*

Note: The definition of the fundamental matrix also holds when P has a more general non-repeating structure. Also, Lemma 6 is valid when P has a more general $GI/G/1$ type structure.

Consider the fundamental matrix, $\widehat{\beta Q}$, of βQ . Denote the $(1, 1)^{st}$ block-entry of $\widehat{\beta Q}$ by $N(\beta)$. In the previous section the R and G measures for any substochastic matrix were defined. They may also be similarly defined for βP . The G measure may also be expressed as

$$G(\beta) = N(\beta)\beta C_0 \quad (2.2.12)$$

or is the minimal nonnegative solution to the equations

$$G(\beta) = \sum_{k=0}^{\infty} \beta C_k G(\beta)^k \quad (2.2.13)$$

and

$$G_{10}(\beta) = N(\beta)\beta D_0 \quad (2.2.14)$$

[40, Remark 3]. The entries of the first column of the fundamental matrix $\widehat{\beta Q}$ may be expressed in terms of $G(\beta)$ and $N(\beta)$ as follows,

$$\widehat{Q}_{j1}(\beta) = G(\beta)^{j-1} N(\beta), \quad j \geq 1, \quad (2.2.15)$$

where $\widehat{\beta Q} = [\widehat{Q}_{ij}(\beta)]_{i,j \geq 1}$, [40, Lemma 4]. The R measure for βP consists of two sequences, $R_{0k}(\beta)$ and $R_k(\beta)$, for $k = 1, 2, 3, \dots$ defined by

$$R_{0k}(\beta) = \sum_{i=1}^{\infty} \beta D_{k+i} \widehat{Q}_{i1}(\beta) \quad (2.2.16)$$

and

$$R_k(\beta) = \sum_{i=1}^{\infty} \beta C_{k+i} \widehat{Q}_{i1}(\beta). \quad (2.2.17)$$

Using equation (2.2.15) equations (2.2.16) and (2.2.17) may be rewritten as,

$$R_{0k}(\beta) = \sum_{i=1}^{\infty} \beta D_{k+i} G(\beta)^{i-1} N(\beta) \quad (2.2.18)$$

and

$$R_k(\beta) = \sum_{i=1}^{\infty} \beta C_{k+i} G(\beta)^{i-1} N(\beta), \quad (2.2.19)$$

for $k = 1, 2, \dots$

Consider βP . Let $U_0(\beta)$ be the discounted transition matrix of the matrix βP censored with censoring set $\mathcal{L}_0$. Using the partition for βP in (2.2.11) and (2.1.4) $U_0(\beta)$ is given by

$$\begin{aligned} U_0(\beta) &= \beta T + \beta H \widehat{\beta Q} \beta L \\ &= \beta D_1 + \beta(D_2, D_3, \dots) \widehat{\beta Q} \beta(D_0, 0, 0, \dots)^t \\ &= \beta D_1 + \left(\sum_{i=1}^{\infty} \beta D_{i+1} \widehat{Q}_{ij}(\beta) \right)_{j=1}^{\infty} \beta(D_0, 0, 0, \dots)^t \\ &= \beta D_1 + \sum_{i=1}^{\infty} \beta D_{i+1} \widehat{Q}_{i1}(\beta) \beta D_0 \\ U_0(\beta) &= \beta D_1 + \sum_{i=1}^{\infty} \beta D_{i+1} G(\beta)^{i-1} N(\beta) \beta D_0. \end{aligned} \quad (2.2.20)$$

According to Lemma 6

$$N_0(\beta) = (I - U_0(\beta))^{-1}, \quad (2.2.21)$$

where $N_0(\beta)$ is the $(1, 1)^{st}$ block-entry of $\widehat{\beta P}$, [40, Theorem 1]. Let $u_0(\beta)$ and $n_0(\beta)$ be the maximal eigenvalues of $U_0(\beta)$ and $N_0(\beta)$, respectively. Due to the relationship

between $U_0(\beta)$ and $N_0(\beta)$, there is a relationship between $u_0(\beta)$ and $n_0(\beta)$. Specifically,

$$n_0(\beta) = \frac{1}{1 - u_0(\beta)}. \quad (2.2.22)$$

The radius of convergence for a nonnegative matrix has already been introduced. As a result the radius of convergence, α , for P and, $\bar{\alpha}$, for Q exist. Under the assumption that $C = \sum_{n=-\infty}^{\infty} C_n$ is irreducible, aperiodic and stochastic, Kijima, [35], provides a method for determining the radius of convergence for Q .

Lemma 7 [35, Kijima] *Let $\chi(z)$ be the Perron-Frobenius eigenvalue for*

$$C^*(z) = \sum_{k=0}^{\infty} C_k z^k, \quad 0 \leq z < z_0.$$

If $z_0 > 1$ then there always exists a unique ν such that $\chi(z) \geq \nu z$ for all $0 < z < z_0$, and there exists some θ with $0 < \theta \leq z_0$ such that $\chi(\theta) = \theta\nu$. If $\theta = z_0$, then $\nu = \chi(z_0)/z_0$. Otherwise, ν and θ can be determined by solving the simultaneous equations,

$$\chi(\theta) = \nu\theta, \quad \chi'(\theta) = \nu. \quad (2.2.23)$$

Remark 4 *Kijima only presented this result for transition matrices with an M/G/1 type structure. However, since Lemma 7 is concerned with the generating function, $C^*(z)$, for the transition matrix, rather than the transition matrix as a whole, it may be applied to transition matrices with other repeating structures. Such as, GI/M/1 type or GI/G/1 type by means of relabelling the subscripts of the repeating blocks or working with the transpose of the transition matrix. Further details as to how it may be applied to a GI/M/1 type structure will be given in Section 4.1.2.*

Kijima then goes on to show that ν is in fact the reciprocal of the radius of convergence, $\bar{\alpha}$, for Q and that Q is $\bar{\alpha}$ -transient.

Theorem 10 [35, Kijima] *Let ν be as in Lemma 7. Then $\bar{\alpha} = 1/\nu$ is the radius of convergence of Q and Q is $\bar{\alpha}$ -transient.*

Based on the relationship (2.2.20) Li and Zhao, [40], were able to characterize the states of P .

Theorem 11 [40, Theorem 3]

- (i) *If for all $0 < \beta \leq \bar{\alpha}$, $u_0(\beta) < 1$, then $N_0(\bar{\alpha}) < \infty$ and $\alpha = \bar{\alpha}$. Therefore, P is α -transient.*
- (ii) *If there exists a β^* with $0 < \beta^* \leq \bar{\alpha}$ such that $u_0(\beta^*) = 1$, then $\alpha = \beta^*$ and $N_0(\alpha) = \infty$. Therefore, P is α -recurrent.*

If (ii) holds and P is α -recurrent, in [40] Li and Zhao provide a means of further characterizing the states as α -positive or α -null.

Theorem 12 [40, Theorem 4] *If $\sum_{k=1}^{\infty} kD_kG(\alpha)^{k-1} < \infty$, $\sum_{k=1}^{\infty} kC_kG(\alpha)^{k-1} < \infty$ and $\alpha \leq \bar{\alpha}$ then the α -recurrent Markov chain is α -positive, otherwise, it is α -null.*

As a Remark they consider the special case when P in (2.2.11) is a QBD process and provide necessary and sufficient conditions for P to be α -positive. These conditions are summarized in the following Remark.

Remark 5 [40, Remark 8] *If $D_i = 0 = C_i$ for $i \geq 3$ then P is a level-independent QBD process with boundary. In this case Theorem 12 illustrates that the α -recurrent QBD process is α -positive if and only if $\alpha < \bar{\alpha}$.*

For the matrix P the RG-factorization can be stated by the following theorem.

Theorem 13 [40, Theorem 5] *For the matrix P of M/G/1 type in equation (2.2.6), $I - \beta P$ can be factorized as*

$$I - \beta P = (I - R_U(\beta))(I - U_D(\beta))(I - G_L(\beta)), \quad (2.2.24)$$

where

$$I - R_U(\beta) = \begin{bmatrix} I - R_{01}(\beta) & -R_{02}(\beta) & \cdots & \cdots \\ & I & -R_1(\beta) & \cdots \\ & & I & -R_1(\beta) \\ & & & \ddots \\ & & & & \ddots \end{bmatrix},$$

$U_D(\beta)$ is the block diagonal matrix $U_D(\beta) = \text{diag}(U_0(\beta), U(\beta), U(\beta), \dots)$ and

$$I - G_L(\beta) = \begin{bmatrix} I & & & \\ -G_{10}(\beta) & I & & \\ & -G(\beta) & I & \\ & & -G(\beta) & I \\ & & & \ddots & \ddots \end{bmatrix}.$$

While the factorization in (2.2.24) has only been presented for P of $M/G/1$ type and irreducible, it can be obtained for any arbitrary matrix P . Similar results to Theorem 11 and Theorem 12 hold for when P is irreducible and of $GI/M/1$ type. The corresponding results may be found in Theorems 6 and 7 in [39], respectively, and in a more general setting in [6]. By using the factorization in equation (2.2.24) the α -invariant measure for P may be found. This result is presented in the following theorem.

Theorem 14 [40, Theorem 6] *If P is α -recurrent, the unique, up to multiplication by a positive constant, α -invariant measure is given by*

$$\begin{aligned} \pi_0 &= x_0 \\ \pi_k &= \pi_0 R_{0,k}(\alpha) + \sum_{i=1}^{k-1} \pi_i R_{k-i}(\alpha) \end{aligned}$$

where x_0 is the unique, up to a multiple of positive constant, solution to $x_0(I - U_0(\alpha)) = 0$.

Let $\pi = (\pi_{i,j})$ be the stationary joint probability measure of P' , and also the infinitesimal generator Q , for a QBD process

$$P' = \begin{bmatrix} B_0 & A_0 & & & \\ C_1 & B & A & & \\ & C & B & A & \\ & & C & B & A \\ & & & \ddots & \ddots & \ddots \end{bmatrix}. \quad (2.2.25)$$

Partition π according to levels where for level i , $\pi_i = (\pi_{i,0}, \pi_{i,1}, \pi_{i,2}, \dots)$ for $i = 0, 1, \dots$. This partition allows us to characterize the tail behavior of π_i or the tail asymptotics along the level direction.

For convenience, we adopt the following often used convention. For two sequences of scalars a_i and b_i , we use $a_i \sim b_i$ to imply that

$$\lim_{i \rightarrow \infty} \frac{a_i}{b_i} = 1. \quad (2.2.26)$$

In this case, we say that the two sequences have the same tail. For two sequences of vectors $\mathbf{a}_i$ and $\mathbf{b}_i$, $\mathbf{a}_i \sim \mathbf{b}_i$ means that (2.2.26) is true for every pair of the corresponding entry-wise sequences.

We say that the joint distribution $\pi_{i,j}$ has a geometric tail along a direction, say the level direction, if

$$\pi_i \sim \eta^i \mathbf{x}$$

for some $0 < \eta < 1$ and a vector $\mathbf{x} \geq 0$.

As mentioned in the Introduction, Takahashi *et. al.*, [57], look at the tail behaviour of the stationary distribution of a QBD process with a countable number of phases.

Based on the assumptions in Theorem 15 below and on the assumption $\pi_1 \mathbf{z}^T < \infty$, where $\mathbf{z}$ is given in Theorem 15, they show, with the aid of Theorem 8, that

$$\pi_m \sim c\eta^m \mathbf{x} \text{ as } m \rightarrow \infty \quad (2.2.27)$$

where η and $\mathbf{x}$ are the scalar and row vector introduced in Theorem 15 below. The details for the assumptions to obtain (2.2.27) are given in Theorem 18 in chapter 3.

Theorem 15 [57, Theorem 1] *Assume that there exists a positive constant $\eta > 1$, a positive row vector $\mathbf{x}$ and a positive column vector $\mathbf{y}^T$ such that*

1. $\mathbf{x}(\eta^{-1}A + B + \eta C) = \mathbf{x}$
2. $(\eta^{-1}A + B + \eta C)\mathbf{y}^T = \mathbf{y}^T$
3. $\mathbf{x}\mathbf{e} < \infty, \mathbf{x}\mathbf{y}^T < \infty$
4. $\eta^{-1}\mathbf{x}A\mathbf{y}^T < \eta\mathbf{x}C\mathbf{y}^T$

Then

(a) $\mathbf{x}R = \eta\mathbf{x}$ and

(b) $\mathbf{z}^T = A(\eta^{-1}I - G)\mathbf{y}^T$ is a non-zero nonnegative vector satisfying $R\mathbf{z}^T = \eta\mathbf{z}^T$, where R and G are the R and G measure of P' in (2.2.25), respectively. The R and G measures have been defined in the previous section.

In what follows, we show that the joint probability distribution $\pi_{i,j}$ has a geometric tail along the direction of the level variable and at the same time explicitly determine the geometric parameter. We also explicitly determine the vector $\mathbf{x}$.

Chapter 3

Determination of η and $\mathbf{x}$ for Irreducible, Scalar, Tridiagonal P

Based on Theorem 8 by Seneta [54], Takahashi, Fujimoto and Makimoto [57], provided sufficient conditions under which the joint probability distribution $\pi_{i,j}$ of a QBD process is geometric along the level direction. The transition probability matrix of a level independent QBD process has the following block tridiagonal form,

$$P' = \begin{bmatrix} B_0 & A_0 & & & \\ C_1 & B & A & & \\ & C & B & A & \\ & & \ddots & \ddots & \ddots \end{bmatrix},$$

in which the blocks have no particular fixed structure. However, in this chapter we will only be interested in the situation where the blocks A_0, B_0, C_1, A, B and C have an infinite tridiagonal structure. The details of this condition are,

Condition 1 *There exists a positive constant η with $0 < \eta < 1$, a positive row vector $\mathbf{x}$ and a positive column vector $\mathbf{y}^T$ such that*

- a. $\mathbf{x}(\eta^{-1}A + B + \eta C) = \mathbf{x}$;
- b. $(\eta^{-1}A + B + \eta C)\mathbf{y}^T = \mathbf{y}^T$;
- c. $\mathbf{x}\mathbf{e} < \infty, \mathbf{x}\mathbf{y}^T < \infty$;
- d. $\eta^{-1}\mathbf{x}A\mathbf{y}^T \neq \eta\mathbf{x}C\mathbf{y}^T$,

which can be found in Theorem 15.

However, finding such an η and determining the corresponding vectors $\mathbf{x}$ and $\mathbf{y}^T$ are all very hard, which is the main cause of the limitation of Theorem 15. In the following condition we convert Condition 1 into a condition such that it can be easily verified.

The stationary probability vector $\pi = (\pi_0, \pi_1, \pi_2, \dots)$ for the Markov chain defined above can be expressed into matrix-geometric form $\pi_i = \pi_1 R^{i-1}$ for $i \geq 0$, where R is the rate matrix for P' . Therefore, if the asymptotic behavior of powers of the infinite rate matrix R can be characterized, then we expect the characterization of the tail asymptotics of π_i . Theorem 8 essentially says that if R is α -positive, π_i has a geometric tail along the level direction, i.e., $\pi_i \sim c\eta^i \mathbf{x}$, where $\alpha = 1/\eta$ is the radius of convergence of R and $\mathbf{x}$ is the α -invariant measure of R . Condition 1 provides a sufficient condition for R to be α -positive with $\alpha = 1/\eta$, but we do not know how to compute R in general. In the following condition Condition 1 is reduced to two. Namely,

Condition 2 Let $\alpha(\eta)$ be the radius of convergence of the matrix $P(\eta) = A + \eta B + \eta^2 C$.

- a. There exists an $0 < \eta < 1$ such that $P(\eta)$ is $\alpha(\eta)$ -positive and $\alpha(\eta) = 1/\eta$.

- b. $\eta^{-1}\mathbf{x}A\mathbf{y}^T \neq \eta\mathbf{x}C\mathbf{y}^T$, where $\mathbf{x}$ and $\mathbf{y}^T$ are an $\alpha(\eta)$ -invariant measure and vector of $P(\eta)$, respectively, and $\mathbf{x}\mathbf{e}^T < \infty$.

Condition 2a. is presented as Theorem 16. The proof of Condition 2a. or Theorem 16 may be reduced to showing two points, Conditions 3a. and 3b. below.

Condition 3 a. *There exists a solution η to the equation $u_0(1/\eta) = 1$;*

b. *The η given in a. satisfies $1/\eta < \bar{\alpha}$;*

- c. $\eta^{-1}\mathbf{x}A\mathbf{y}^T \neq \eta\mathbf{x}C\mathbf{y}^T$, where $\mathbf{x}$ and $\mathbf{y}^T$ are an $\alpha(\eta)$ -invariant measure and vector of $P(\eta)$, respectively, and $\mathbf{x}\mathbf{e}^T < \infty$.

It follows from Theorem 7 that Condition 2a. implies Conditions 1a., b. and the second part of c. Therefore, Condition 2 implies Condition 1.

Within the framework of Markov renewal processes Miyazawa and Zhao [48] improve the conditions upon which the result in Theorem 2 of Takahashi *et. al* [57] is based, by proving the theorem without the use of Condition 2b. This gives a new set of conditions namely,

Condition 4 *Let $\alpha(\eta)$ be the radius of convergence of the matrix $P(\eta)$. There exists an $0 < \eta < 1$ such that $P(\eta)$ is $\alpha(\eta)$ -positive and $\alpha(\eta) = 1/\eta$.*

The main aim for the above discussion and Condition 3 is to show that the joint distribution π_{mn} has a geometric tail along the level direction, [57, Theorem 2] i.e.

$$\pi_m \sim c\eta^m \mathbf{x}, \quad \text{as } m \rightarrow \infty$$

shown in Theorem 18.

3.1 Tail Theorem

In this section the main theorem, Theorem 16, upon which the remainder of this dissertation is based will be introduced and proved. The remainder of the chapter is devoted to applications of Theorem 16 to well known queueing models.

Consider a quasi-birth-death process with a countable number of phases in each level. We have an irreducible Markov chain $\{X(n); n \geq 0\}$ on a two-dimensional state space $S = \{(m, i); m, i = 0, 1, 2, \dots\}$ with transition probability matrix

$$P' = \begin{bmatrix} B_0 & A_0 & 0 & 0 & 0 & \dots \\ C_1 & B & A & 0 & 0 & \dots \\ 0 & C & B & A & 0 & \dots \\ 0 & 0 & C & B & A & \dots \\ \vdots & \vdots & & \ddots & \ddots & \ddots \end{bmatrix}, \quad (3.1.1)$$

after partitioning the state space into levels $\mathcal{L}_m = \{(m, i); i = 0, 1, 2, \dots\}$ for $m = 0, 1, 2, \dots$. Let the repeating blocks A, B and C each have a tridiagonal structure with scalar entries, i.e.

$$A = \begin{bmatrix} \overline{a_1} & \overline{a_0} & 0 & 0 & 0 & \dots \\ \overline{a_2} & a_1 & a_0 & 0 & 0 & \dots \\ 0 & a_2 & a_1 & a_0 & 0 & \dots \\ 0 & 0 & a_2 & a_1 & a_0 & \dots \\ \vdots & \vdots & & \ddots & \ddots & \ddots \end{bmatrix}, B = \begin{bmatrix} \overline{b_1} & \overline{b_0} & 0 & 0 & 0 & \dots \\ \overline{b_2} & b_1 & b_0 & 0 & 0 & \dots \\ 0 & b_2 & b_1 & b_0 & 0 & \dots \\ 0 & 0 & b_2 & b_1 & b_0 & \dots \\ \vdots & \vdots & & \ddots & \ddots & \ddots \end{bmatrix},$$

$$C = \begin{bmatrix} \overline{c_1} & \overline{c_0} & 0 & 0 & 0 & \dots \\ \overline{c_2} & c_1 & c_0 & 0 & 0 & \dots \\ 0 & c_2 & c_1 & c_0 & 0 & \dots \\ 0 & 0 & c_2 & c_1 & c_0 & \dots \\ \vdots & \vdots & & \ddots & \ddots & \ddots \end{bmatrix}.$$

Let $\pi = (\pi_0, \pi_1, \pi_2, \dots)$, where $\pi_m = (\pi_{m0}, \pi_{m1}, \pi_{m2}, \dots)$, be the invariant probability measure of P' .

It is quite challenging to determine the joint tail behaviour of π_{mn} and in some situations it may not be necessary. Our aim is to determine the rate of decay of π_{mn} as $m \rightarrow \infty$, for all n . We would like to reduce the dimension of the state space on which work, by one. This will then allow us to appeal to results on Markov chains whose state spaces have infinite levels and finite phases, more specifically in this chapter, equal to one dimension. For $0 < \eta < 1$ consider the matrix

$$P = P(\eta) = A + \eta B + \eta^2 C \quad (3.1.2)$$

which is substochastic due to our choice of η .

$$P = \begin{bmatrix} \gamma_0 & \lambda_0 & 0 & 0 & 0 & \dots \\ \mu_0 & \gamma & \lambda & 0 & 0 & \dots \\ 0 & \mu & \gamma & \lambda & 0 & \dots \\ 0 & 0 & \mu & \gamma & \lambda & \dots \\ \vdots & \vdots & & \ddots & \ddots & \end{bmatrix} \quad (3.1.3)$$

where,

$$\begin{aligned} \gamma_0 &= \overline{a_1} + \eta \overline{b_1} + \eta^2 \overline{c_1}, & \gamma &= a_1 + \eta b_1 + \eta^2 c_1, \\ \lambda_0 &= \overline{a_0} + \eta \overline{b_0} + \eta^2 \overline{c_0}, & \lambda &= a_0 + \eta b_0 + \eta^2 c_0, \\ \mu_0 &= \overline{a_2} + \eta \overline{b_2} + \eta^2 \overline{c_2}, & \mu &= a_2 + \eta b_2 + \eta^2 c_2. \end{aligned}$$

Further, assume P is irreducible. Also, denote by Q the submatrix of P without the

boundary values, i.e.

$$Q = Q(\eta) = \begin{bmatrix} \gamma & \lambda & 0 & 0 & 0 & \cdots \\ \mu & \gamma & \lambda & 0 & 0 & \cdots \\ 0 & \mu & \gamma & \lambda & 0 & \cdots \\ 0 & 0 & \mu & \gamma & \lambda & \cdots \\ \vdots & \vdots & & \ddots & \ddots & \ddots \end{bmatrix}, \quad (3.1.4)$$

which is also assumed to be irreducible. The radius of convergence $\bar{\alpha}$ for Q will be of interest to us shortly. Theorem 10 tells us that $\bar{\alpha}$ is the reciprocal of ν determined by Kijima in Lemma 7. Kijima makes the assumption that $\mu + \gamma + \lambda$ is irreducible, aperiodic and stochastic. Above $\mu + \gamma + \lambda$ is not stochastic because

$$\mu + \gamma + \lambda = (a_0 + a_1 + a_2) + \eta(b_0 + b_1 + b_2) + \eta^2(c_0 + c_1 + c_2) < 1$$

since $0 < \eta < 1$. We shall multiply Q by a scalar c so that $c(\mu + \gamma + \lambda) = 1$ is stochastic and then apply Lemma 7 and hence Theorem 10 to the matrix cQ . The question that arises now is, how is the radius of convergence, $\bar{\alpha}_c$, of cQ related to the radius of convergence, $\bar{\alpha}$, of Q ? This question is addressed in the following result.

Lemma 8 *Let P be a non-negative matrix with radius of convergence α . Let $c > 0$ be a scalar. The radius of convergence, α_c , for cP is given by $\alpha_c = \alpha/c$.*

Proof: By the definition of the radius of convergence for a matrix, the radius of convergence for cP is defined as

$$\begin{aligned} \alpha_c &= \sup\left\{z; \sum_{n=0}^{\infty} z^n (cP)^n < \infty\right\} = \sup\left\{z; \sum_{n=0}^{\infty} (cz)^n P^n < \infty\right\} \\ &= \sup\left\{y/c; \sum_{n=0}^{\infty} y^n P^n < \infty\right\} = \frac{1}{c} \sup\left\{y; \sum_{n=0}^{\infty} y^n P^n < \infty\right\} = \frac{\alpha}{c}. \quad \blacksquare \end{aligned}$$

In our particular situation we will take $c = (\mu + \gamma + \lambda)^{-1}$. Since a_i, b_i and c_i are all greater than zero for $i = 0, 1, 2$ and $0 < \eta < 1$, $0 < c$.

In the method outlined by Kijima in Lemma 7 to determine ν we need to determine the Perron-Frobenius eigenvalue, $\chi(z)$, for the generating function, $C^*(z) = c(\mu + \gamma z + \lambda z^2)$, of cQ . Since the generating function $C^*(z)$ for cQ is a scalar function the Perron-Frobenius eigenvalue for cQ is the generating function itself, i.e. $\chi(z) = C^*(z)$. Now we need to solve,

$$\chi(\theta) = \nu\theta \text{ and } \chi'(\theta) = \nu$$

for ν , or

$$\chi(\theta) = c(\mu + \gamma\theta + \lambda\theta^2) = \nu\theta \quad (3.1.5)$$

and

$$\chi'(\theta) = c(\gamma + 2\lambda\theta) = \nu. \quad (3.1.6)$$

Substituting ν from (3.1.6) into (3.1.5) and solving for θ we get,

$$\theta = \sqrt{\frac{\mu}{\lambda}}.$$

The positive square root is taken since $\theta > 0$. Plugging this value for θ into (3.1.6) we get,

$$\nu = c(\gamma + 2\sqrt{\lambda\mu}).$$

Taking the reciprocal of ν we have that the radius of convergence for cQ is,

$$\bar{\alpha}_c = \frac{1}{\nu} = \frac{1}{c(\gamma + 2\sqrt{\lambda\mu})}. \quad (3.1.7)$$

By Lemma 10 the radius of convergence for Q is given by

$$\bar{\alpha} = c\bar{\alpha}_c = \frac{1}{a_1 + \eta b_1 + \eta^2 c_1 + 2\sqrt{(a_0 + \eta b_0 + \eta^2 c_0)(a_2 + \eta b_2 + \eta^2 c_2)}}.$$

According to Theorem 11(ii), in order for P to be α -recurrent a value β^* with $0 < \beta^* \leq \bar{\alpha}$ such that $u_0(\beta^*) = 1$, is required, where $\alpha = \beta^*$ is the radius of convergence for P . Notice that conditions 1. and 2. in Theorem 15 may be rewritten as,

1. $\frac{1}{\eta} \mathbf{x}P = \mathbf{x}$ and
2. $\frac{1}{\eta} P\mathbf{y} = \mathbf{y}$,

respectively. From these two conditions we may hypothesize that $1/\eta$ is the radius of convergence for P . In order for P to be $1/\eta$ -recurrent, $1/\eta$ must satisfy the conditions of Theorem 11(ii). If we let $\beta^* = 1/\eta$ satisfy the inequality $0 < \beta^* \leq \bar{\alpha}$ then η satisfies the inequality,

$$(a_1 + \eta b_1 + \eta^2 c_1)^2 < 4((a_0 + \eta b_0 + \eta^2 c_0)(a_2 + \eta b_2 + \eta^2 c_2)) \leq (a_1 + \eta(b_1 - 1) + \eta^2 c_1)^2. \quad (3.1.8)$$

The converse of this will be shown in Theorem 16 to follow.

We consider βP , the discounted transition matrix at rate β , for $0 < \beta \leq \alpha$ where α is the radius of convergence for P . Let $U_0(\beta)$ be the censored matrix of P with censoring set equal to level zero and let $u_0(\beta)$ be the maximal eigenvalue of $U_0(\beta)$. Since,

$$U_0(\beta) = \beta\gamma_0 + \frac{\lambda_0\mu_0}{2\lambda\mu}(1 - \beta\gamma - \sqrt{(\beta\gamma - 1)^2 - 4\beta^2\lambda\mu}) \quad (3.1.9)$$

is a scalar function of β , $u_0(\beta) = U_0(\beta)$. Again, if we let $\beta^* = 1/\eta$ and substitute $\beta = \beta^*$ into $u_0(\beta) = 1$, we have that η is a zero of

$$\beta^2[(\lambda_0\mu_0)^2 - \lambda_0\mu_0\gamma_0\gamma + \lambda\mu\gamma_0^2] + \beta[-2\gamma_0\lambda\mu + \lambda_0\mu_0(\gamma_0 + \gamma)] + [\lambda\mu - \lambda_0\mu_0] = 0. \quad (3.1.10)$$

Notice that (3.1.10) is an equation of degree at most eight evaluated at η and will henceforth be denoted by,

$$q(x) = D_0 + D_1x + \cdots + D_8x^8 = 0 \quad (3.1.11)$$

where the coefficients D_i are functions of the entries of A , B and C . The polynomial in equation (3.1.11) will be referred to as the *characteristic polynomial* for P since the value we wish to find from it plays the role of the eigenvalue for P . $u_0(\beta) = 1$

or (3.1.11) is a key ingredient to the following result and its importance will soon be evident.

Theorem 16 *Suppose η is in the interval $(0, 1)$ and satisfies equation (3.1.11) and inequality (3.1.8). Then the radius of convergence of P given in (3.1.2) is $\alpha = \frac{1}{\eta}$ and P is $\frac{1}{\eta}$ -recurrent. Moreover, P is $\frac{1}{\eta}$ -positive, if $\alpha < \bar{\alpha}$. Also, the α -invariant measure $\mathbf{x}$ and α -invariant vector $\mathbf{y}$ of P exist and, $\mathbf{x}\mathbf{y}^T < \infty$ if $\alpha < \bar{\alpha}$.*

Proof: Let $\beta = \frac{1}{\eta}$. We would like to show that P is α -recurrent where $\alpha = \beta = \frac{1}{\eta}$. If we can show that $\beta = \frac{1}{\eta}$ is β^* in Theorem 11(ii) then P is $1/\eta$ -recurrent. That is, we need to show, $0 < \beta^* \leq \bar{\alpha}$ and $u_0(\beta^*) = 1$.

From the inequality in (3.1.8) we have,

$$\gamma^2 < 4\lambda\mu \leq (\gamma - \eta)^2,$$

which implies

$$\gamma^2 < 4\lambda\mu \tag{3.1.12}$$

and

$$4\lambda\mu \leq (\gamma - \eta)^2. \tag{3.1.13}$$

(3.1.12) implies $0 < (2\sqrt{\lambda\mu} - \gamma)(2\sqrt{\lambda\mu} + \gamma)$. Since $\lambda, \gamma, \mu \geq 0$, $\gamma < 2\sqrt{\lambda\mu}$. And, (3.1.13) implies

$$(2\sqrt{\lambda\mu} - \gamma + \eta)(2\sqrt{\lambda\mu} + \gamma - \eta) \leq 0$$

or

$$\eta \leq \gamma - 2\sqrt{\lambda\mu} \text{ or } \eta \geq \gamma + 2\sqrt{\lambda\mu}.$$

We know $\eta > 0$ so if $\eta \leq \gamma - 2\sqrt{\lambda\mu}$ then $2\sqrt{\lambda\mu} \leq \gamma$, which contradicts $\gamma < 2\sqrt{\lambda\mu}$ shown above. Therefore, $\eta \geq \gamma + 2\sqrt{\lambda\mu}$ or

$$\beta^* = \frac{1}{\eta} \leq \frac{1}{\gamma + 2\sqrt{\lambda\mu}} = \bar{\alpha},$$

as required.

If we let $x = \eta$ in (3.1.11) then we have, $q(\eta) = D_0 + D_1\eta + \cdots + D_8\eta^8 = 0$ which was shown above to be equivalent to $u_0(\frac{1}{\eta}) = 1$ or $u_0(\beta^*) = 1$ where $\beta^* = \frac{1}{\eta}$. By Theorem 11(ii), $\alpha = \frac{1}{\eta}$ and P is α -recurrent. If $\alpha = \frac{1}{\eta} < \bar{\alpha}$ and P is an α -recurrent QBD process, by Theorem 12 and Remark 5, P is α -positive.

Fortunately, P is only substochastic and irreducible and so satisfies the assumptions made on T in Chapter 2.2. As a result we may appeal to the Note after Theorem 9 to justify the existence of an α -invariant vector $\mathbf{y}^T$ for P when P is α -recurrent and an α -subinvariant vector $\mathbf{y}^T$ for P when P is α -transient.

Now that we have an α -invariant measure $\mathbf{x}$ for P and an α -invariant vector $\mathbf{y}$ for P and if P is α -positive, by Theorem 7, $\mathbf{x}\mathbf{y}^T < \infty$. ■

Corollary 17 *If $\alpha < \bar{\alpha}$ then P is α -positive. Moreover, $\mathbf{x}\mathbf{y}^T < \infty$.*

By Theorem 10 in [39], the α -invariant measure $\mathbf{x} = (x_0, x_1, x_2, \dots)$ of $P = P(\frac{1}{\alpha})$ is given by,

$$x_0 = z_0, \quad \text{and } x_k = z_0 R_{01}(\alpha) R^{k-1}(\alpha), \quad k \geq 1 \quad (3.1.14)$$

where $z_0 > 0$ is a nonnegative solution to $z_0(I - U_0(\beta)) = 0$. In our particular case $R(\alpha)$ is given by

$$R(\alpha) = \frac{1 - \alpha\gamma - \sqrt{(\alpha\gamma - 1)^2 - 4\alpha^2\lambda\mu}}{2\alpha\mu} \quad (3.1.15)$$

and $R_{01}(\alpha)$ by

$$R_{01}(\alpha) = \frac{\lambda_0}{\lambda} R(\alpha). \quad (3.1.16)$$

In order to find the α -invariant vector $\mathbf{y} = (y_0, y_1, y_2, \dots)$ of P the RG factorization given by equation (2.2.24) in Theorem 13 is used where $\beta = \alpha$. We want to solve

$$0 = (I - \alpha P)\mathbf{y} = (I - R_U(\alpha))(I - U_D(\alpha))(I - G_L(\alpha))\mathbf{y}.$$

Let $\mathbf{w}^T = (I - G_L(\alpha))\mathbf{y}^T$. Then we have

$$0 = (I - R_U(\alpha))(I - U_D(\alpha))\mathbf{w}^T. \quad (3.1.17)$$

The product $(I - R_U(\alpha))(I - U_D(\alpha))$ is given by,

$$\begin{bmatrix} I - U_0(\alpha) & -R_{01}(\beta)(I - U(\alpha)) & 0 & 0 & 0 & \dots \\ 0 & I - U(\alpha) & -R(\beta)(I - U(\alpha)) & 0 & 0 & \dots \\ 0 & 0 & I - U(\alpha) & -R(\beta)(I - U(\alpha)) & 0 & \dots \\ 0 & 0 & 0 & I - U(\alpha) & -R(\beta)(I - U(\alpha)) & \dots \\ \vdots & \vdots & & \ddots & \ddots & \ddots \end{bmatrix}$$

(3.1.17) reduces to

$$0 = (I - U_0(\alpha))w_0 - R_{01}(\alpha)(I - U(\alpha))w_1 \quad (3.1.18)$$

and

$$0 = (I - U(\alpha))w_n - R(\alpha)(I - U(\alpha))w_{n+1}, \quad n \geq 1. \quad (3.1.19)$$

Since $U_0(\alpha) = 1$, we may take $\mathbf{w} = (w_0, 0, 0, \dots)$ where $w_0 > 0$ satisfies $w_0(I - U_0(\alpha)) = 0$. Now we need to solve

$$\mathbf{w} = (I - G_L(\alpha))\mathbf{y}. \quad (3.1.20)$$

Simplifying (3.1.20) gives us,

$$\begin{aligned} 0 &= w_1 = -G_{10}(\alpha)y_0 + y_1 \Rightarrow y_1 = G_{10}(\alpha)y_0, \\ 0 &= w_2 = -G(\alpha)y_1 + y_2 \Rightarrow y_2 = G_{10}(\alpha)G(\alpha)y_0, \\ &\vdots \\ 0 &= w_n = -G(\alpha)y_{n-1} + y_n \Rightarrow y_n = G^{n-1}(\alpha)G_{10}(\alpha)y_0, \quad n \geq 2. \end{aligned}$$

Therefore, the α -invariant vector of P is $\mathbf{y} = (y_0, y_1, y_2, \dots)^T$ where

$$y_0 = w_0 > 0, \quad \text{and} \quad y_n = G_{10}(\alpha)G^{n-1}(\alpha)w_0, \quad n \geq 1, \quad (3.1.21)$$

where $G(\alpha)$ is given by

$$G(\alpha) = \frac{1 - \alpha\gamma - \sqrt{(\alpha\gamma - 1)^2 - 4\alpha^2\lambda\mu}}{2\alpha\lambda} = \frac{\mu}{\lambda}R(\alpha) \quad (3.1.22)$$

and $G_{10}(\alpha)$ by

$$G_{10}(\alpha) = \frac{\mu_0}{\mu}G(\alpha). \quad (3.1.23)$$

Using the above expressions for the α -invariant measure $\mathbf{x}$ and vector $\mathbf{y}$ the result $\mathbf{xy}^T < \infty$ reduces to,

$$\begin{aligned} \mathbf{xy}^T &= \sum_{i=0}^{\infty} x_i y_i = z_0 w_0 + \sum_{n=1}^{\infty} x_n y_n \\ &= z_0 w_0 \left(1 + R_{10}(\alpha) G_{10}(\alpha) \sum_{n=1}^{\infty} R^{n-1}(\alpha) G^{n-1}(\alpha) \right) \\ &= z_0 w_0 \left(1 + \frac{\lambda_0 \mu_0}{\lambda \mu} \sum_{n=1}^{\infty} \left(\frac{\mu}{\lambda} \right)^n R^{2n}(\alpha) \right) < \infty \end{aligned}$$

where $G(\alpha) = \frac{\mu}{\lambda}R(\alpha)$. So,

$$\mathbf{xy}^T < \infty \text{ if and only if } \sum_{n=1}^{\infty} \left(\frac{\mu}{\lambda} R^2(\alpha) \right)^n < \infty. \quad (3.1.24)$$

By Theorem 5.2 of [51] $\mathbf{x}$ is the $1/\eta$ -invariant measure of R . However, we require an α -invariant vector $\mathbf{z}$ for R as well, before Theorem 8 may be used to prove the asymptotic behaviour of π along the level direction. Condition 4. of Theorem 15

remains to be confirmed. It states $\eta^{-1}\mathbf{x}A\mathbf{y}^T < \eta\mathbf{x}C\mathbf{y}^T$ which reduces to,

$$\begin{aligned} & \frac{1}{\eta}\mathbf{x}A\mathbf{y}^T \\ &= \frac{1}{\eta} \left(y_0(x_0\bar{a}_1 + x_1\bar{a}_2) + y_1(\bar{a}_0 + x_1a_1 + x_2a_2) + \sum_{n=2}^{\infty} y_n(x_{n-1}a_0 + x_na_1 + x_{n+1}a_2) \right) \\ &< \eta \left(y_0(x_0\bar{c}_1 + x_1\bar{c}_2) + y_1(\bar{c}_0 + x_1c_1 + x_2c_2) + \sum_{n=2}^{\infty} y_n(x_{n-1}c_0 + x_nc_1 + x_{n+1}c_2) \right) \\ &= \eta\mathbf{x}C\mathbf{y}^T \end{aligned}$$

upon expansion. From the above we see that, if

$$\frac{1}{\eta}\bar{a}_i = \eta\bar{c}_i, \quad \text{and} \quad \frac{1}{\eta}a_i = \eta c_i, \quad \text{for } i = 0, 1, 2 \quad (3.1.25)$$

then

$$\frac{1}{\eta}\mathbf{x}A\mathbf{y}^T = \eta\mathbf{x}C\mathbf{y}^T.$$

and conversely. Otherwise, $\frac{1}{\eta}\mathbf{x}A\mathbf{y}^T < \eta\mathbf{x}C\mathbf{y}^T$.

Next we would like to show that,

$$\pi_m \sim c\eta^m \mathbf{x} \text{ as } m \longrightarrow \infty,$$

In order to do this we will make use of Theorem 8 and the matrix geometric representation of the solution, [49] i.e.

$$\pi_m = \pi_1 R^{m-1}, \quad m \geq 0$$

where R is the R measure of P' . In order to use Theorem 8 R must be irreducible and aperiodic. We will consider the submatrix $R(A)$ of R consisting of all non-zero rows of R . Another way of interpreting $R(A)$ is as follows. Let $\mathcal{L}_m(A) = \{(m, i) \in \mathcal{L}_m; (Ae)_i \geq 0\}$. According to the Lemma 4(a), [64], the R measure for P' may be

written as $R = A\hat{Q}'_{11}$ where $\hat{Q}'_{11}$ is the $(1, 1)^{st}$ block entry of the fundamental matrix of Q' . The i^{th} row of R is zero if $(m, i) \notin \mathcal{L}_m(A)$. We may rewrite R as,

$$R = \begin{matrix} & \mathcal{L}_m(A) & \overline{\mathcal{L}_m(A)} \\ \begin{matrix} \mathcal{L}_m(A) \\ \overline{\mathcal{L}_m(A)} \end{matrix} & \begin{pmatrix} R(A) & \overline{R(A)} \\ 0 & 0 \end{pmatrix} \end{matrix}$$

where $\overline{\mathcal{L}_m(A)} = \mathcal{L}_m^c(A)$. We denote $R(A)$ to be the submatrix of R on the index set $\mathcal{L}_m(A) \times \mathcal{L}_m(A)$. So, it is sufficient to consider just when $R(A)$ is irreducible and aperiodic. A sufficient condition for $R(A)$ to be irreducible is for Q' to be irreducible. If this is so, then $\hat{Q}_{11} \succeq 0$ thus $R(A) > 0$. Once $R(A)$ is irreducible and Theorem 16 holds then $R(A)$ is $1/\eta$ -positive. If it is further assumed that $R(A)$ is aperiodic then Theorem 8 can be applied to give,

$$R^m \sim \eta^m \frac{1}{\mathbf{x}\mathbf{z}^T} \mathbf{z}\mathbf{x}^T, \quad \text{as } m \rightarrow \infty$$

Theorem 18 [57, Theorem 2] *Assume Theorem 16 holds. If $R(A)$ is irreducible and aperiodic and if $\pi_1 \mathbf{z}^T < \infty$ then the stationary probability measure π of P' is geometric along the level variable, i.e.*

$$\pi_m \sim c\eta^m \mathbf{x} \quad \text{as } m \rightarrow \infty$$

where $c = \frac{\pi_1 \mathbf{z}}{\mathbf{x}\mathbf{z}}$ is a constant.

Proof: Since Theorem 16 holds thus so does Theorem 15. This result gives that there is a left and right eigenvector $\mathbf{x}$ and $\mathbf{z}$, respectively, of R for the eigenvalue η . More specifically, $\mathbf{z}^T = A(\eta^{-1}I - G)\mathbf{y}^T \leq \eta^{-1}A\mathbf{y}^T \leq \mathbf{y}^T$ which implies, $\mathbf{x}\mathbf{z}^T \leq \mathbf{x}\mathbf{y}^T < \infty$. Since R is not necessarily irreducible we cannot conclude that R is $1/\eta$ -positive.

However, $R(A)$ is irreducible and hence $1/\eta$ -positive. We may rewrite R as,

$$R = \begin{matrix} & \mathcal{L}_m(A) & \overline{\mathcal{L}_m(A)} \\ \mathcal{L}_m(A) & \begin{pmatrix} R(A) & \overline{R(A)} \\ 0 & 0 \end{pmatrix} \\ \overline{\mathcal{L}_m(A)} & \end{matrix}.$$

Since $\mathbf{x}$ is an $1/\eta$ -invariant measure of R ,

$$\eta \mathbf{x} = \mathbf{x} R = \begin{bmatrix} \mathbf{x}(A) & \overline{\mathbf{x}}(A) \end{bmatrix} \begin{bmatrix} R(A) & \overline{R(A)} \\ 0 & 0 \end{bmatrix}$$

holds. Upon expansion we get the following relations,

$$\begin{bmatrix} \eta \mathbf{x}(A) & \eta \overline{\mathbf{x}}(A) \end{bmatrix} = \begin{bmatrix} \mathbf{x}(A) R(A) & \mathbf{x}(A) \overline{R(A)} \end{bmatrix}$$

which implies

$$\eta \mathbf{x}(A) = \mathbf{x}(A) R(A)$$

$$\eta \overline{\mathbf{x}}(A) = \mathbf{x}(A) \overline{R(A)}.$$

By induction we can show that,

$$R^m = \begin{matrix} & \mathcal{L}_m(A) & \overline{\mathcal{L}_m(A)} \\ \mathcal{L}_m(A) & \begin{pmatrix} R^m(A) & R^{m-1}(A) \overline{R(A)} \\ 0 & 0 \end{pmatrix} \\ \overline{\mathcal{L}_m(A)} & \end{matrix}.$$

We want to show that,

$$\pi_m = \pi_1 R^{m-1} \sim \eta^m \pi_1 \frac{1}{\mathbf{x} \mathbf{z}^T} \mathbf{z}^T \mathbf{x}, \quad m \longrightarrow \infty.$$

We may write, $\pi_1 = \begin{bmatrix} \pi_1(A) & \overline{\pi}_1(A) \end{bmatrix}$, $\pi_m = \begin{bmatrix} \pi_m(A) & \overline{\pi}_m(A) \end{bmatrix}$ and $\mathbf{x} = \begin{bmatrix} \mathbf{x}(A) & \overline{\mathbf{x}}(A) \end{bmatrix}$

$$\begin{aligned} \pi_m &= \begin{bmatrix} \pi_m(A) & \overline{\pi}_m(A) \end{bmatrix} = \pi_1 R^{m-1} \\ &= \begin{bmatrix} \pi_1(A) & \overline{\pi}_1(A) \end{bmatrix} \begin{bmatrix} R^m(A) & R^{m-1}(A) \overline{R(A)} \\ 0 & 0 \end{bmatrix} \\ &= \begin{bmatrix} \pi_1(A) R^m(A) & \pi_1(A) R^{m-1}(A) \overline{R(A)} \end{bmatrix}. \end{aligned}$$

Since $R(A)$ is $1/\eta$ -positive and aperiodic, we may apply Theorem 8, to give,

$$\begin{aligned}\pi_1(A)R^m(A) &\sim \eta^m \frac{1}{\mathbf{x}\mathbf{z}^T} \pi_1(A)(\mathbf{z}^T(A)\mathbf{x}(A)) \quad m \longrightarrow \infty \\ &= \frac{\eta^m}{\mathbf{x}\mathbf{z}^T} (\pi_1(A)\mathbf{z}^T(A))\mathbf{x}(A)\end{aligned}$$

and

$$\begin{aligned}\pi_1(A)R^{m-1}(A)\bar{R}(A) &\sim \frac{\eta^{m-1}}{\mathbf{x}\mathbf{z}^T} \pi_1(A)(\mathbf{z}^T(A)\mathbf{x}(A))\bar{R}(A) \quad m \longrightarrow \infty \\ &= \frac{\eta^{m-1}}{\mathbf{x}\mathbf{z}^T} \pi_1(A)\mathbf{z}^T(A)(\mathbf{x}(A)\bar{R}(A)) \\ &= \frac{\eta^{m-1}}{\mathbf{x}\mathbf{z}^T} \pi_1(A)\mathbf{z}^T(A)\eta\bar{\mathbf{x}}(A) \\ &= \frac{\eta^m}{\mathbf{x}\mathbf{z}^T} \pi_1(A)\mathbf{z}^T(A)\bar{\mathbf{x}}(A)\end{aligned}$$

therefore,

$$\begin{aligned}\pi_m &= \begin{bmatrix} \pi_m(A) & \bar{\pi}_m(A) \end{bmatrix} \\ &\sim \eta^m \frac{\pi_1(A)\mathbf{z}^T(A)}{\mathbf{x}\mathbf{z}^T} \begin{bmatrix} \mathbf{x}(A) & \bar{\mathbf{x}}(A) \end{bmatrix} \quad m \longrightarrow \infty \\ &= \eta^m \frac{\pi_1\mathbf{z}^T}{\mathbf{x}\mathbf{z}^T} \mathbf{x} \\ &= c\eta^m \mathbf{x},\end{aligned}$$

where $c = \frac{\pi_1\mathbf{z}^T}{\mathbf{x}\mathbf{z}^T}$. ■

Note: The results, namely equation (3.1.10) and inequality (3.1.8), may be obtained by other means. For example, they may be obtained directly using the definition of α -positivity as in Miyazawa and Zhao, [48], or by using the method outlined by McDonald in [45] and Foley and McDonald in [21]. However, the reason we use the method presented here is so that it may be easily applied to more general models as mentioned in Chapter 7. All the steps outlined above will remain the same. The difficulties will arise when trying to find certain relevant quantities to see the method

through. For example, in Chapter 4 the R measure for P satisfies an infinite series rather than a polynomial. Chapter 7 presents some other generalizations and their potential challenges.

3.2 Examples

In this section we will look at some applications of Theorem 16 through examples. The first example is of the two-demand model studied by Flatto and Hahn, [19]. It is an illustration of when P is α -recurrent but not α -positive. The second example is a model motivated by the model studied by Wang, [62], and illustrates when Theorem 16 is satisfied. The numerical results given for this example are all new results. The third example is a priority queue example considered by Miller, [46]. In this example the level and phase variables considered by Miller are interchanged to give new results about the asymptotic character along the other direction. The final example is the classic symmetric shortest queue example which has been studied by many authors. It is an example of when Theorem 16 is satisfied. Further details of each example will be presented in the following sections.

3.2.1 Two-Demand Model

A double queue is analyzed in the paper by Flatto and Hahn, [19]. This double queue occurs when arriving customers to the system simultaneously place two demands handled independently by two different servers. Customer arrivals form a Poisson process with mean 1 and the servers have exponential service times with rate α_1 and β_1 . For convenience, the servers with rate α_1 and β_1 are referred to as α_1 - and β_1 -servers, respectively and their queues as α_1 - and β_1 -queues, respectively. The double queue is referred to as the (α_1, β_1) -queue. Let $X_1(t)$ and $X_2(t)$ represent the number

of customers waiting or in service at time t in the α_1 - and β_1 -queues, respectively. $X(t) = (X_1(t), X_2(t))$ is a Markov chain with state space $S = \{(i, j); i, j \geq 0, i, j = 0, 1, 2, \dots\}$. The chain is stable if and only if $\alpha_1, \beta_1 > 1$. Without loss of generality assume that $1 < \alpha_1 \leq \beta_1$.

An application of such a queueing system may be found in the inventory control of data base systems. A common inaccuracy which arises in such systems is the deviation in the data base from the number of items present in stock. This deviation is due to the fact that the updating in the system may be made at different times from when an item is actually added or removed from stock. If the deviation is due to removal of stock only then may the actual and recorded number of items be treated as two parallel queues created by the same arrival process, each having its own servicing time distribution.

The infinitesimal generator for such a system has a quasi-birth-death structure with infinite blocks each having an infinite tridiagonal structure. This may be found from the equilibrium equations given by Flatto and Hahn in [19] as follows,

$$(1 + \alpha_1 \chi_{\{i \geq 1\}} + \beta_1 \chi_{\{j \geq 1\}}) \pi_{ij} = \alpha_1 \pi_{i+1, j} + \beta_1 \pi_{i, j+1} + \chi_{\{i \geq 1, j \geq 1\}} \pi_{i-1, j-1}, \quad i, j \geq 0 \quad (3.2.1)$$

where χ_A is the indicator function,

$$\chi_A(x) = \begin{cases} 1, & \text{if } x \in A \\ 0, & \text{if } x \notin A \end{cases}.$$

The transition rate diagram may be found in figure 3.1. From here the equilibrium equations for the transition matrix are given by

$$\begin{aligned} (1 + \alpha_1 + \beta_1) \pi_{ij} &= \beta_1 \pi_{i, j+1} + \alpha_1 \pi_{i+1, j} + \pi_{i-1, j-1} \\ (1 + \alpha_1 + \beta_1) \pi_{i0} &= \beta_1 \pi_{i1} + \alpha_1 \pi_{i+1, 0} + \beta_1 \pi_{i0} \\ (1 + \alpha_1 + \beta_1) \pi_{00} &= (\alpha_1 + \beta_1) \pi_{00} + \beta_1 \pi_{01} + \alpha_1 \pi_{10}. \end{aligned}$$

Figure 3.1: Transition diagram for Two Demand Model.

The transition matrix also has a quasi-birth-death structure as in (3.1.1) where

$$\begin{aligned}
 B_0 &= \theta \begin{bmatrix} \alpha_1 + \beta_1 & 0 & & & \\ \beta_1 & \alpha_1 & 0 & & \\ & \beta_1 & \alpha_1 & 0 & \\ & & \ddots & \ddots & \ddots \\ & & & \ddots & \ddots \end{bmatrix}, A_0 = \theta \begin{bmatrix} 0 & 1 & & & \\ & 0 & 1 & & \\ & & 0 & 1 & \\ & & & 0 & 1 \\ & & & & \ddots & \ddots & \ddots \end{bmatrix}, \\
 C_1 &= \theta \begin{bmatrix} \alpha_1 & & & & \\ & \alpha_1 & & & \\ & & \alpha_1 & & \\ & & & \ddots & \\ & & & & \ddots \end{bmatrix}, B = \theta \begin{bmatrix} \beta_1 & & & & \\ \beta_1 & 0 & & & \\ & \beta_1 & 0 & & \\ & & \beta_1 & 0 & \\ & & & \ddots & \ddots & \ddots \end{bmatrix},
 \end{aligned}$$

$$A = \theta \begin{bmatrix} 0 & 1 & & & \\ & 0 & 1 & & \\ & & 0 & 1 & \\ & & & \ddots & \ddots \\ & & & & \ddots & \ddots \end{bmatrix}, \text{ and } C = \theta \begin{bmatrix} \alpha_1 & & & & \\ & \alpha_1 & & & \\ & & \alpha_1 & & \\ & & & \ddots & \\ & & & & \ddots \end{bmatrix},$$

where $\theta = (1 + \alpha_1 + \beta_1)^{-1}$.

In order to apply Theorem 16 we need the matrix,

$$\begin{aligned} P &= P(\eta) = A + \eta B + \eta^2 C \\ &= \begin{bmatrix} \gamma_0 & \lambda_0 & 0 & 0 & 0 & \dots \\ \mu_0 & \gamma & \lambda & 0 & 0 & \dots \\ 0 & \mu & \gamma & \lambda & 0 & \dots \\ 0 & 0 & \mu & \gamma & \lambda & \dots \\ \vdots & \vdots & & \ddots & \ddots & \ddots \end{bmatrix}, \end{aligned}$$

where

$$\begin{aligned} \gamma_0 &= \eta(1 - \theta(1 + \alpha_1)) + \eta^2 \theta \alpha_1, & \gamma &= \eta^2 \theta \alpha_1, \\ \lambda_0 &= \theta, & \lambda &= \theta = \lambda_0, \\ \mu_0 &= \eta \theta \beta_1, & \mu &= \eta \theta \beta_1 = \mu_0. \end{aligned}$$

The convergence radius $\bar{\alpha}$ of Q , constructed from P by deleting boundaries, is given by

$$\bar{\alpha} = \frac{1}{\gamma + 2\sqrt{\lambda\mu}} = \frac{1}{\eta^2 \theta \alpha_1 + 2\sqrt{\eta \theta^2 \beta_1}} = \frac{1 + \alpha_1 + \beta_1}{\eta^2 \alpha_1 + 2\sqrt{\eta \beta_1}}. \quad (3.2.2)$$

The convergence radius of P is given by $\alpha = 1/\eta$ where η is a root of the characteristic equation (3.1.10) satisfying (3.1.8). Letting $\beta = \frac{1}{\eta}$, substituting $\theta = \frac{1}{1 + \alpha_1 + \beta_1}$ and simplifying equation (3.1.10) we get

$$\frac{\beta_1^2}{(1 + \alpha_1 + \beta_1)^4} (1 - \alpha_1 \eta)(1 - \eta) = 0$$

which gives $\eta = \frac{1}{\alpha_1}, 1$ as the roots. Since $\alpha_1 > 1$, $0 < \frac{1}{\alpha_1} < 1$. So the radius of convergence of $P(\eta)$ or of the R measure of P' is $\frac{1}{\eta} = \alpha_1$.

Next we need to determine if $P(\eta)$ is $1/\eta$ -positive which is equivalent to determining if $\alpha \leq \bar{\alpha}$. The radius of convergence for Q is given by

$$\bar{\alpha} = \alpha_1 \frac{1 + \alpha_1 + \beta_1}{1 + 2\sqrt{\alpha_1\beta_1}},$$

where $\frac{1}{\eta} = \alpha_1$ has been substituted into the expression for $\bar{\alpha}$ given in (3.2.2). We need to determine if $\frac{1+\alpha_1+\beta_1}{1+2\sqrt{\alpha_1\beta_1}} \geq 1$. For $\alpha_1 < \beta_1$,

$$\begin{aligned} 0 &\leq (\sqrt{\alpha_1} - \sqrt{\beta_1})^2, \\ 0 &\leq \alpha_1 - 2\sqrt{\alpha_1\beta_1} + \beta_1, \\ 1 &\leq 1 + \alpha_1 + \beta_1 - 2\sqrt{\alpha_1\beta_1}, \\ 1 + 2\sqrt{\alpha_1\beta_1} &\leq 1 + \alpha_1 + \beta_1, \\ 1 &\leq \frac{1 + \alpha_1 + \beta_1}{1 + 2\sqrt{\alpha_1\beta_1}} \end{aligned}$$

as required. Therefore, $\alpha_1 = \alpha \leq \bar{\alpha}$. Since $\alpha \leq \bar{\alpha}$ by Theorem 16 the α -invariant measure $\mathbf{x}$ and vector $\mathbf{y}$ for $P(\eta)$ exist. We need to determine if $\frac{1}{\eta}\mathbf{x}A\mathbf{y}^T \neq \eta\mathbf{x}C\mathbf{y}^T$. This is equivalent to determining if,

$$\begin{aligned} \frac{1}{\eta}\mathbf{x}A\mathbf{y}^T &= \frac{1}{\eta}(x_0, x_1, \dots)A(y_0, y_1, \dots)^T = \frac{\theta}{\eta} \sum_{i=0}^{\infty} x_i y_{i+1} = \theta\alpha \sum_{i=0}^{\infty} x_i y_{i+1} \\ &\neq \theta \sum_{i=0}^{\infty} x_i y_i = \eta\theta\alpha \sum_{i=0}^{\infty} x_i y_i = \eta\theta\alpha\mathbf{x}\mathbf{y}^T = \eta\mathbf{x}C\mathbf{y}^T. \end{aligned}$$

We may use the expressions for the α -invariant measure $\mathbf{x}$ and vector $\mathbf{y}$ given in equations (3.1.14) and (3.1.21), respectively. The R measure of P is given by

$$R(\beta) = \frac{1 - \beta\eta^2\theta\alpha_1 - \sqrt{(1 - \beta\eta^2\theta\alpha_1)^2 - 4\beta^2\eta\theta^2\beta_1}}{2\eta\theta\beta\beta_1}.$$

When $\beta = \frac{1}{\eta}$ we have

$$R\left(\frac{1}{\eta}\right) = \frac{1 - \eta\theta\alpha_1 - \sqrt{(1 - \eta\theta\alpha_1)^2 - 4\theta^2\eta^{-1}\beta_1}}{2\theta\beta_1}.$$

When $\frac{1}{\eta} = \alpha_1$ or $\eta = \frac{1}{\alpha_1}$ then,

$$R(\alpha_1) = \frac{1 - \theta - \sqrt{(1 - \theta)^2 - 4\theta^2\alpha_1\beta_1}}{2\theta\beta_1}.$$

Substituting for $\theta = \frac{1}{1+\alpha_1+\beta_1}$ and simplifying we get

$$R(\alpha) = \frac{\alpha_1}{\beta_1}$$

since $1 < \alpha_1 < \beta_1$ and $R(\alpha)$ is non-negative. Using the following relations for $R_{01}(\beta)$, $G(\beta)$ and $G_{10}(\beta)$,

$$R_{01}(\beta) = \frac{\lambda_0}{\lambda}R(\beta), \quad G(\beta) = \frac{\mu}{\lambda}R(\beta), \quad G_{10}(\beta) = \frac{\mu_0}{\lambda}R(\beta),$$

and our above expression for $R(\beta)$ we get

$$R_{01}(\alpha) = \frac{\alpha_1}{\beta_1}, \quad G(\alpha) = 1, \quad G_{10}(\alpha) = 1.$$

We know for an α -recurrent P , its α -invariant measure and α -invariant vector are unique. By equation (3.1.14) the α -invariant measure of P is given by

$$x_0 = z_0 > 0, \tag{3.2.3}$$

$$x_n = z_0 R^n(\alpha) = z_0 \left(\frac{\alpha_1}{\beta_1}\right)^n, \quad n \geq 1, \tag{3.2.4}$$

where z_0 is the unique, up to multiplication by a positive constant, solution to $z_0(I - U_0(\alpha)) = 0$. In our situation, $U_0(\alpha)$ is a scalar equal to 1 so z_0 is any non-zero, non-negative value. So the α -invariant measure of P is

$$\mathbf{x} = \frac{\beta_1 - \alpha_1}{\beta_1} \left(1, \frac{\alpha_1}{\beta_1}, \left(\frac{\alpha_1}{\beta_1}\right)^2, \dots, \left(\frac{\alpha_1}{\beta_1}\right)^n, \dots \right).$$

By equation (3.1.21) the α -invariant vector for P is

$$\begin{aligned} y_0 &= w_0 \\ y_n &= G^{n-1}(\alpha)G_{10}(\alpha)y_0 = w_0, \quad n \geq 2, \end{aligned}$$

where w_0 satisfies $w_0(I - U_0(\alpha)) = 0$. Therefore, $\mathbf{y} = (1, 1, 1, \dots) = \mathbf{e}$. Now that we have the α -invariant measure and vector for P , we can also determine whether or not P is α -positive or not by considering whether or not $\mathbf{x}\mathbf{y}$ converges or not.

$$\begin{aligned} \mathbf{x}\mathbf{y}^T &= \frac{\beta_1 - \alpha_1}{\beta_1} \left(1, \frac{\alpha_1}{\beta_1}, \left(\frac{\alpha_1}{\beta_1} \right)^2, \dots, \left(\frac{\alpha_1}{\beta_1} \right)^n, \dots \right) (1, 1, 1, \dots)^T \\ &= \frac{\beta_1 - \alpha_1}{\beta_1} \sum_{i=0}^{\infty} \left(\frac{\alpha_1}{\beta_1} \right)^i = 1 \end{aligned}$$

since $\alpha_1 < \beta_1$. Therefore, P is α -positive, by Theorem 7.

Since $\pi \mathbf{e}^T = 1$, $\pi_1 \mathbf{z}^T \leq \pi_1 \mathbf{y}^T = \pi_1 \mathbf{e}^T \leq 1$. Therefore, $\pi_1 \mathbf{z}^T < \infty$. We may conclude that

$$\begin{aligned} \pi_m \sim c\eta^m \mathbf{x} &= \frac{c}{\alpha_1^m} \frac{\beta_1 - \alpha_1}{\beta_1} \left(1, \frac{\alpha_1}{\beta_1}, \dots, \left(\frac{\alpha_1}{\beta_1} \right)^n, \dots \right) \\ &= \frac{c_1}{\alpha_1^m} \left(1, \frac{\alpha_1}{\beta_1}, \dots, \left(\frac{\alpha_1}{\beta_1} \right)^n, \dots \right) \end{aligned}$$

which is consistent with the result obtained by Flatto and Hahn in Theorem 19, [19].

In the special case when $\alpha_1 = \beta_1$ the α -invariant measure and vector of P are

$$\mathbf{x} = z_0(1, 1, 1, \dots) \text{ and } \mathbf{y} = w_0(1, 1, 1, \dots).$$

Notice that $\mathbf{x}\mathbf{y}^T = z_0 w_0 \sum_{i=0}^{\infty} 1 = \infty$. The condition $\frac{1}{\eta} \mathbf{x} A \mathbf{y}^T < \eta \mathbf{x} C \mathbf{y}^T$ still holds when $\beta_1 = \alpha_1$ but it is the condition $\mathbf{x}\mathbf{y}^T < \infty$ that fails. When $\alpha_1 = \beta_1$, $\bar{\alpha}$ reduces to

$$\bar{\alpha} = \frac{1 + 2\alpha_1}{\eta^2 \alpha_1 + 2\sqrt{\eta \alpha_1}}.$$

When $\alpha_1 = \beta_1$ the polynomial in equation (3.1.10) reduces to

$$\frac{\alpha_1^2}{(1 + 2\alpha_1)^4} (1 - \alpha_1\eta)(1 - \eta) = 0.$$

which again implies $\eta = 1, \frac{1}{\alpha_1}$. Therefore $\eta = \frac{1}{\alpha_1} < 1$ since $1 < \alpha_1$. $\bar{\alpha}$ now becomes

$$\bar{\alpha} = \frac{1 + 2\alpha_1}{\eta^2\alpha_1 + 2\sqrt{\eta\alpha_1}} = \alpha_1 \frac{(1 + 2\alpha_1)}{(1 + 2\alpha_1)} = \alpha_1 = \frac{1}{\eta}.$$

Therefore by Theorem 11, P is $\alpha = \frac{1}{\eta}$ -recurrent. In order to determine if P is $\frac{1}{\eta}$ -positive Theorem 12 and Remark 5 may be used. Remark 5 states, *an α -recurrent QBD process is α -positive if and only if $\alpha \leq \bar{\alpha}$* . But from above we see that $\alpha = \frac{1}{\eta} = \bar{\alpha}$. So P is not $\frac{1}{\eta}$ -positive and thus, P is $1/\eta$ -null. This is consistent with Theorem 16. As a result π_m does not have an asymptotic tail as is shown by Flatto and Hahn in Theorem 19, [19]. We may also consider the convergence of $\mathbf{xy}^T$ to determine whether or not P is $1/\eta$ -positive. In this situation $\mathbf{xy}^T = \infty$ and thus P is not $1/\eta$ -positive, by Theorem 7.

In Theorem 19 below, see Theorems 7.1 and 7.2 in [19], results are provided for the asymptotic behaviour for the joint stationary probabilities $\pi_{i0}, \pi_{0j}, \pi_{ij}$ as $i \rightarrow \infty, j \rightarrow \infty$ for the cases $\alpha_1 < \beta_1$ and $\alpha_1 = \beta_1$.

Theorem 19 [19, Theorems 7.1 and 7.2]

1. For $\alpha_1 < \beta_1$ and a fixed j ,

$$\pi_{ij} \sim \frac{c_1}{\alpha_1^{i-j}\beta_1^j}, \text{ as } i \rightarrow \infty.$$

2. For $\alpha_1 < \beta_1$ and a fixed i ,

$$\pi_{ij} \sim \frac{c_2(1 + [1 - \left(\frac{\alpha_1}{\alpha'_3}\right)^{1/2}]i)}{j^{3/2}\alpha_1^{1/2}(\alpha'_3)^{j-i/2}}, \text{ as } j \rightarrow \infty$$

and when $i = 0$,

$$\pi_{0j} \sim \frac{c_2}{j^{2/3}(a'_3)^j} \text{ as } j \rightarrow \infty.$$

3. For $\alpha_1 = \beta_1$ and a fixed j ,

$$\pi_{ij} \sim \frac{c_3}{i^{1/2}\alpha_1^i}, \text{ as } i \rightarrow \infty$$

and when $i = 0$ is fixed,

$$\pi_{0i} \sim \frac{c_3}{i^{1/2}\alpha_1^i}, \text{ as } i \rightarrow \infty$$

where

$$\begin{aligned} a &= \sqrt{a_3 - \alpha_1}, & b &= \sqrt{a_3 - 1}, & c &= \sqrt{a_3 - \frac{\alpha_1}{\beta_1}}, \\ d &= \sqrt{a_3}, & c_1 &= \frac{(a-1)(\beta_1-\alpha_1)}{\alpha_1\beta_1}, \\ c_2 &= \frac{1}{4\sqrt{\pi}} \frac{(\alpha_1-1)(a+c)(b+c)d}{\alpha_1 ab} \left(\frac{1}{a} + \frac{1}{b} + \frac{1}{c} \right), & c_3 &= \sqrt{\frac{2}{\pi}} \left(\frac{\alpha_1-1}{\alpha_1} \right)^{3/2}, & a'_3 &= \left(\frac{\beta_1}{\alpha_1} \right) a_3. \end{aligned}$$

From this theorem we see that,

$$\pi_m = \begin{bmatrix} \pi_{m0} \\ \pi_{m1} \\ \vdots \\ \pi_{mn} \\ \vdots \end{bmatrix} \sim \begin{bmatrix} c_1/\alpha_1^m \\ c_1/\alpha_1^m \left(\frac{\alpha_1}{\beta_1} \right) \\ \vdots \\ c_1/\alpha_1^m \left(\frac{\alpha_1}{\beta_1} \right)^n \\ \vdots \end{bmatrix} = \frac{c_1}{\alpha_1^m} \begin{bmatrix} 1 \\ \alpha_1/\beta_1 \\ \vdots \\ (\alpha_1/\beta_1)^n \\ \vdots \end{bmatrix}, \text{ as } m \rightarrow \infty.$$

So, $\pi_m \sim c_1 \left(\frac{1}{\alpha_1}\right)^m \mathbf{x}$, as $m \rightarrow \infty$. But,

$$\begin{aligned} \bar{\pi}_n &= \begin{bmatrix} \pi_{0n} \\ \pi_{1n} \\ \vdots \\ \pi_{mn} \\ \vdots \end{bmatrix} = \begin{bmatrix} \frac{\frac{c_2}{n^{3/2}(a'_3)^n}}{1 + \left[1 - \left(\frac{\alpha_1}{a'_3}\right)^{1/2}\right] (a'_3)^{1/2}} \\ \frac{c_2}{n^{3/2}(a'_3)^n} \frac{1}{\alpha_1^{1/2}} \\ \vdots \\ \frac{\frac{c_2}{n^{3/2}(a'_3)^n}}{1 + \left[1 - \left(\frac{\alpha_1}{a'_3}\right)^{1/2}\right] m(a'_3)^{m/2}} \\ \frac{c_2}{n^{3/2}(a'_3)^n} \frac{1}{\alpha_1^{1/2}} \\ \vdots \end{bmatrix} \\ &\sim \frac{c_2}{n^{3/2}(a'_3)^n \alpha_1^{1/2}} \begin{bmatrix} \alpha_1^{1/2} \\ 1 + \left[1 - \left(\frac{\alpha_1}{a'_3}\right)^{1/2}\right] (a'_3)^{1/2} \\ \vdots \\ 1 + \left[1 - \left(\frac{\alpha_1}{a'_3}\right)^{1/2}\right] m(a'_3)^{m/2} \\ \vdots \end{bmatrix}. \end{aligned}$$

So, $\bar{\pi}_n \sim \frac{c_2}{n^{3/2}(a'_3)^n \alpha_1^{1/2}} \mathbf{y}$, as $n \rightarrow \infty$ which is not geometric.

3.2.2 Model Motivated by Wang's Model

The model considered by Wang [62] is a priority queueing system consisting of two queues sharing a processor and in which there is delayed feedback. New jobs arrive to the system at the initialization queue INIT. The jobs wait to be served by the central processor, CPU. The service time is an exponential random variable with mean service time $1/\mu_0$. Once a job has been served, or rather initialized, by the CPU the job then joins the process queue, PROC, to await service by the same central processor. The service time for a job waiting at the PROC queue is also an exponential random variable with a different mean service time of $1/\mu_1$. Once a job from the PROC queue has been served by the CPU it either leaves the system

with probability p or returns to the PROC queue after some delay with probability $1 - p$. This delay is modelled by awaiting service at an infinite server node called the INF-SERV queue. The INF-SERV queue represents a “thinking state” for networks in computer science. For more information on the interpretation refer to [62]. The service time at the INF-SERV queue is exponential with mean service time $1/\mu_2$. The rate of departure from this INF-SERV queue depends on the number of jobs in the queue. For example, if there are n jobs in the INF-SERV queue then the service rate is $n\mu_2$. We assume the system is always overloaded in the sense of, $P(\text{INIT queue length} \geq 1) = 1$ and the service time random variables for the INIT, PROC and INF-SERV queue are independent. A diagram of the model may be found in figure 3.2. Let $X_1(t)$ and $X_2(t)$ be the number of jobs in the PROC queue and the INF-SERV queue at time t , respectively. Then $X(t) = (X_1(t), X_2(t))$ is a Markov chain with state space $S = \{(i, j); i, j = 0, 1, 2, \dots\}$.

The model considered in this section is motivated by the model above considered by Wang, [62]. Wang’s model results in phase dependent blocks. To achieve independence, rather than have infinitely many servers in the INF-SERV queue only one server is present. The service times are still exponentially distributed with parameter μ_2 and independent. Now we denote the INF-SERV queue by ONE-SERV queue. In this situation the Markov chain $X(t) = (X_2(t), X_1(t))$ with state space $S = \{(i, j); i, j = 0, 1, 2, 3, \dots\}$ is considered. The stability conditions for this modified model are

$$\frac{\mu_0(1-p)}{\mu_1} < 1 \text{ and } \frac{\mu_0(1-p)}{\mu_2} < 1. \quad (3.2.5)$$

The infinitesimal generator for the modified model we are considering has a quasi-birth-death structure with infinite blocks each with a tridiagonal structure and may be found in Wang [62] by considering the situation when $n = 1$. The transition rate diagram in figure 3.3 may also be considered.

Figure 3.2: Wang's model.

From figure 3.3 the equilibrium equations are given by

$$(\mu_0 + \mu_1 + \mu_2)\pi_{i1} = \mu_0\pi_{i0} + \mu_2\pi_{i+1,0} + \mu_0\pi_{i1} + \mu_1p\pi_{i2} + \mu_1(1-p)\pi_{i-1,2}$$

$$(\mu_0 + \mu_1 + \mu_2)\pi_{ij} = \mu_0\pi_{ij} + \mu_1p\pi_{i,j+1} + \mu_2\pi_{i+1,j-1} + \mu_1(1-p)\pi_{i-1,j+1}$$

$$(\mu_0 + \mu_1 + \mu_2)\pi_{i0} = \mu_1\pi_{i0} + \mu_1(1-p)\pi_{i-1,1} + \mu_1p\pi_{i1}$$

$$(\mu_0 + \mu_1 + \mu_2)\pi_{0j} = \mu_1p\pi_{0,j+1} + \mu_2\pi_{1,j-1} + \mu_0\pi_{0j} + \mu_2\pi_{0j}.$$

The transition matrix also has a quasi-birth-and-death structure as in (3.1.1) where

$$A = \theta \begin{bmatrix} 0 & & & & \\ \mu_1(1-p) & 0 & & & \\ & \mu_1(1-p) & 0 & & \\ & & \mu_1(1-p) & 0 & \\ & & & \ddots & \ddots \end{bmatrix}, C = \theta \begin{bmatrix} 0 & \mu_2 & & & \\ & 0 & \mu_2 & & \\ & & 0 & \mu_2 & \\ & & & \ddots & \ddots \end{bmatrix},$$

Figure 3.3: Transition diagram for the model motivated by Wang's Model.

and

$$B = \theta \begin{bmatrix} \mu_1 & \mu_0 & & & \\ \mu_1 p & \mu_0 & 0 & & \\ & \mu_1 p & \mu_0 & 0 & \\ & & \mu_1 p & \mu_0 & 0 \\ & & & \ddots & \ddots & \ddots \\ & & & & \ddots & \ddots & \ddots \end{bmatrix},$$

where $\theta = (\mu_0 + \mu_1 + \mu_2)^{-1}$.

In order to apply Theorem 16 we need to consider the following matrix

$$\begin{aligned}
 P &= P(\eta) = A + \eta B + \eta^2 C \\
 &= \theta \begin{bmatrix}
 \eta\mu_1 & \eta\mu_0 + \eta^2\mu_2 & & & \\
 \mu_1(1-p) + \eta\mu_1p & \eta\mu_0 & \eta^2\mu_2 & & \\
 & \mu_1(1-p) + \eta\mu_1p & \eta\mu_0 & \eta^2\mu_2 & \\
 & & \ddots & \ddots & \ddots \\
 & & & \ddots & \ddots & \ddots
 \end{bmatrix}
 \end{aligned}$$

where

$$\begin{aligned}
 \gamma_0 &= \theta\eta\mu_1, \gamma = \theta\eta\mu_0 \\
 \lambda_0 &= \theta(\eta\mu_0 + \eta^2\mu_2), \lambda = \theta\eta^2\mu_2 \\
 \mu_0 &= \mu = \theta\mu_1(1-p + \eta p).
 \end{aligned}$$

The convergence radius for Q , constructed from P by deleting boundaries, is given by,

$$\bar{\alpha} = \frac{\mu_0 + \mu_1 + \mu_2}{\eta\mu_0 + 2\eta\sqrt{\mu_1\mu_2(1-p + \eta p)}}. \quad (3.2.6)$$

The convergence radius of P is given by $\alpha = 1/\eta$ where η is a zero of the characteristic polynomial (3.1.10) satisfying (3.1.8). Letting $\beta = \frac{1}{\eta}$ and substituting $\theta = (\mu_0 + \mu_1 + \mu_2)^{-1}$ in (3.1.10) we have

$$\begin{aligned}
 0 &= \mu_0^2\mu_1(1-p) + \eta[2\mu_0\mu_1\mu_2(1-p) - \mu_0^2\mu_1(1-p) - \mu_0(\mu_0 + \mu_1 + \mu_2)^2] \\
 &\quad + \eta^2[\mu_0\mu_1\mu_2(2p-1) + \mu_1\mu_2^2(1-p) + \mu_1^2\mu_2 + (\mu_1 - \mu_0)\mu_2(\mu_0 + \mu_1 + \mu_2)] \\
 &\quad + \eta^3\mu_2^2[\mu_1p - 2(\mu_0 + \mu_1 + \mu_2)]
 \end{aligned}$$

or

$$0 = d + c\eta + b\eta^2 + a\eta^3 \quad (3.2.7)$$

where

$$\begin{aligned}
 a &= \mu_2^2[\mu_1 p - 2(\mu_0 + \mu_1 + \mu_2)] \\
 b &= \mu_0 \mu_1 \mu_2 (2p - 1) + \mu_1 \mu_2^2 (1 - p) + \mu_1^2 \mu_2 + (\mu_1 - \mu_0) \mu_2 (\mu_0 + \mu_1 + \mu_2) \\
 c &= 2\mu_0 \mu_1 \mu_2 (1 - p) - \mu_0^2 \mu_1 (1 - p) - \mu_0 (\mu_0 + \mu_1 + \mu_2)^2 \\
 d &= \mu_0^2 \mu_1 (1 - p)
 \end{aligned}$$

which needs to be explored further. When $\eta = 0$ equation (3.2.7) reduces to

$$\mu_0^2 \mu_1 (1 - p) > 0.$$

When $\eta = 1$ equation (3.2.7) reduces to

$$\begin{aligned}
 &(\mu_0 + \mu_1 + \mu_2)(2\mu_1 \mu_2 - (\mu_0 \mu_1 + 2\mu_0 \mu_2 + \mu_0^2 + 2\mu_2^2)) \\
 &\left\{ \begin{array}{ll} < 0 & \text{if } \mu_0 \mu_1 + 2\mu_0 \mu_2 + \mu_0^2 + 2\mu_2^2 > 2\mu_1 \mu_2 \\ > 0 & \text{if } \mu_0 \mu_1 + 2\mu_0 \mu_2 + \mu_0^2 + 2\mu_2^2 < 2\mu_1 \mu_2 \\ = 0 & \text{if } \mu_0 \mu_1 + 2\mu_0 \mu_2 + \mu_0^2 + 2\mu_2^2 = 2\mu_1 \mu_2 \end{array} \right.
 \end{aligned}$$

In order for there to be at least one root of equation (3.2.7) in the interval $(0, 1)$ the parameters μ_0, μ_1, μ_2 and p must satisfy the stability conditions (3.2.5) and the condition

$$\mu_0 \mu_1 + 2\mu_0 \mu_2 + \mu_0^2 + 2\mu_2^2 \leq 2\mu_1 \mu_2. \quad (3.2.8)$$

We need to determine whether the set of parameters satisfying (3.2.8) and (3.2.5), respectively, are disjoint, coincide, overlap, or if one is contained in the other. If we let

$$\begin{aligned}
 f(x, y, z) &= xy + 2xz + x^2 + 2z^2 - 2yz \\
 g(x, y, z) &= x(1 - p) - y \\
 h(x, y, z) &= x(1 - p) - z
 \end{aligned}$$

then this reduces to determining the region in $\mathbb{R}^3$ such that

$$f(x, y, z) \geq 0, g(x, y, z) < 0 \text{ and } h(x, y, z) < 0$$

for values of p ranging between 0 and 1.

Once $\eta \in (0, 1)$ has been determined we need to determine if inequality (3.1.8) holds. This is the same as determining if $\alpha \leq \bar{\alpha}$ where

$$\alpha = 1/\eta \text{ and } \bar{\alpha} = \frac{\mu_0 + \mu_1 + \mu_2}{\mu_0\eta + 2\eta\sqrt{\mu_1\mu_2(1-p+\eta p)}},$$

where $\bar{\alpha}$ is given by equation (3.1.7). $\alpha \leq \bar{\alpha}$ is true due to the equivalence of the following statements.

$$\begin{aligned} \eta &< 1 + \frac{(\mu_1 - \mu_2)^2}{4\mu_1\mu_2p}, \\ 4\mu_1\mu_2p(-1 + \eta) &< (\mu_1 - \mu_2)^2, \\ 4\mu_1\mu_2p(-1 + \eta) &< (\mu_1 + \mu_2)^2 - 4\mu_1\mu_2, \\ 4\mu_1\mu_2(1 - p + \eta p) &< (\mu_1 + \mu_2)^2, \\ 2\sqrt{\mu_1\mu_2(1 - p + \eta p)} &< \mu_1 + \mu_2, \\ \mu_0 + 2\sqrt{\mu_1\mu_2(1 - p + \eta p)} &< \mu_0 + \mu_1 + \mu_2, \\ \alpha = \frac{1}{\eta} &< \frac{\mu_0 + \mu_1 + \mu_2}{\mu_0\eta + 2\eta\sqrt{\mu_1\mu_2(1 - p + \eta p)}} = \bar{\alpha}. \end{aligned}$$

Since $\eta \in (0, 1)$ and $0 < \frac{(\mu_1 - \mu_2)^2}{4\mu_1\mu_2p}$, $\eta < 1 + \frac{(\mu_1 - \mu_2)^2}{4\mu_1\mu_2p}$.

Note: The above equivalences hold for any values of μ_0, μ_1, μ_2, p and η .

By Theorem 16, P is $1/\eta$ -positive. Consequently, the $1/\eta$ -invariant measure $\mathbf{x}$ and vector, $\mathbf{y}$ for P exist. By Corollary 17, $\mathbf{xy} < \infty$. Equation (3.1.14) may be used to determine the $1/\eta$ -invariant measure $\mathbf{x}$ for P where

$$R(1/\eta) = \min \left(\frac{\mu_1 + \mu_2 \pm \sqrt{(\mu_1 - \mu_2)^2 - 4\mu_1\mu_2p(\eta - 1)}}{2\mu_1(1 - p + \eta p)/\eta} \right)$$

and

$$R_{01}(1/\eta) = \left(\frac{\mu_0}{\eta\mu_2} + \eta \right) R(1/\eta).$$

The G measure is given by

$$\begin{aligned} G(1/\eta) &= \frac{\mu}{\lambda} R(1/\eta) = \frac{\mu_1(1-p+\eta p)}{\eta^2\mu_2} R(1/\eta) \\ &= \frac{\mu_1 + \mu_2 \pm \sqrt{(\mu_1 - \mu_2)^2 - 4\mu_1\mu_2 p(\eta - 1)}}{2\eta\mu_2} \end{aligned}$$

and

$$G_{10}(1/\eta) = \frac{\mu_0}{\mu} G(1/\eta) = G(1/\eta).$$

The $1/\eta$ -invariant measure is given by equation (3.1.14) or

$$\mathbf{x} = (1, R_{01}(1/\eta), R_{01}(1/\eta)R(1/\eta), R_{01}(1/\eta)R(1/\eta)^2, \dots)$$

and the $1/\eta$ -invariant vector by equation (3.1.21) or

$$\mathbf{y} = (1, G(1/\eta), G(1/\eta)^2, \dots)^T.$$

In order for $P(1/\eta)$ to be $1/\eta$ -positive we need that $\mathbf{xy}^T < \infty$. This is equivalent to

$$\begin{aligned} \mathbf{xy}^T &= 1 + R_{01}G + R_{01}RG^2 + R_{01}R^2G^3 + R_{01}R^3G^4 + \dots + R_{01}R^nG^{n+1} + \dots \\ &= 1 + R_{01}R(\mu/\lambda) + R_{01}(\mu/\lambda)^2R^3 + R_{01}(\mu/\lambda)^3R^5 + R_{01}(\mu/\lambda)^4R^7 + \dots \\ &\quad + R_{01}(\mu/\lambda)^{n+1}R^{2n+1} + \dots \\ &= 1 + R_{01}R(\mu/\lambda) + R_{01} \sum_{n=1}^{\infty} (\mu/\lambda)^{n+1} R^{2n+1} \\ &= 1 + R_{01}R(\mu/\lambda) + R_{01}R(\mu/\lambda) \sum_{n=1}^{\infty} (\mu/\lambda R^2)^n \\ &= 1 + R_{01}R(\mu/\lambda) + R_{01}R(\mu/\lambda) \sum_{n=1}^{\infty} (RG)^n < \infty \end{aligned}$$

or

$$|R(1/\eta)G(1/\eta)| < 1.$$

Condition 3c. remains to be verified. This is equivalent to determining whether or not equation (3.1.25) holds or not. For this particular example

$$\begin{aligned} a_0 = 0 = \bar{a}_0, \quad a_1 = 0 = \bar{a}_1, \quad a_2 = \theta\mu_1(1-p) = \bar{a}_2 \\ c_0 = \theta\mu_2 = \bar{c}_0, \quad c_1 = 0 = \bar{c}_1, \quad c_2 = 0 = \bar{c}_2 \end{aligned}$$

From here we see that

$$\begin{aligned} \frac{1}{\eta}a_0 = \frac{1}{\eta}\bar{a}_0 = 0 &\neq \eta\bar{c}_0 = \eta\theta\mu_2 = \eta c_0, \\ \frac{1}{\eta}a_1 = \frac{1}{\eta}\bar{a}_1 = 0 &= \eta\bar{c}_1 = \eta c_1, \end{aligned}$$

and

$$\frac{1}{\eta}a_2 = \frac{1}{\eta}\bar{a}_2 = \frac{1}{\eta}\theta\mu_1(1-p) \neq 0 = \eta\bar{c}_2 = \eta c_2.$$

Therefore,

$$\frac{1}{\eta}\mathbf{x}\bar{\mathbf{A}}\mathbf{y}^T \neq \eta\mathbf{x}\bar{\mathbf{C}}\mathbf{y}^T.$$

Note: This holds for any value of μ_0, μ_1, μ_2, p and η .

According to Theorem 15 we have $\mathbf{x}R = \eta\mathbf{x}$ and $R\mathbf{z}^T = \eta\mathbf{z}^T$ where $\mathbf{z}$ is given as in Theorem 15. Now the result in Theorem 18 holds where the definition of $R(A)$ is as in Chapter 2.1, i.e.

$$\pi_m \sim c\eta^m \mathbf{x}.$$

In Table 3.1, for different sets of parameters, the root η of equation (3.2.7) lying in the interval $(0, 1)$, the corresponding $\bar{\alpha}$ and α values are given and in Table 3.2 the $R(\alpha)$ and $G(\alpha)$ measures for P are given.

The numerical results for Tables 3.1 and 3.2 were done using Maple. Maple was used to solve for η from equation (3.2.7) for various parameter values. Maple was also used to find $\bar{\alpha}$ and R and G measures using the expressions in this section, for various parameter values.

Interchanging the level and phase variables, $X(t) = (X_1(t), X_2(t))$ becomes the Markov process considered. When applying Theorem 16 to this model the resulting root of the characteristic equation is $\eta = 1$. This value does not satisfy the conditions in Theorem 16. So P is not α -positive and asymptotics of π_{mn} along the $X_1(t)$ direction cannot be found using this method. However, from figure 3.3 this model is a candidate for the compensation method. So there may be a possibility that π_{mn} is jointly geometric.

μ_0	μ_1	μ_2	p	η	α	$\bar{\alpha}$
1	2	3	0.5	0.03355382575	29.80286086	39.54610458
1	1	3	0.5	0.02234627095	44.75019578	64.35715696
2	2	3	0.5	0.04479580895	22.32351694	28.20233240
2	3	2	0.5	0.06690345103	14.94691207	18.75697369
2	3	2	0.1	0.1288284752	7.762259069	7.999860090
2	3	2	0.2	0.1126700014	8.875476947	9.642670084
2	3	2	0.3	0.09695172285	10.31441186	11.67719499
2	3	2	0.4	0.08169247917	12.24102892	14.53138967
2	3	2	0.6	0.05258921623	19.01530526	25.50788787
2	3	2	0.7	0.03874883578	25.80722698	37.62017693
2	3	2	0.8	0.02537689930	39.40591749	64.15813298
2	3	2	0.9	0.01246448368	80.22795213	154.5486699
1	3	1	0.1	0.1359870871	7.353639388	8.528786820
1	3	1	0.2	0.1174920307	8.511215561	10.27040979
1	3	1	0.3	0.09999478979	10.00052105	12.62779251
1	3	1	0.4	0.08342585653	11.98669144	15.95294311
1	3	1	0.5	0.06771775316	14.76717631	20.91035238
1	3	1	0.6	0.05280646221	18.93707622	28.90268233
1	3	1	0.7	0.03863216102	25.88516856	43.41637317
1	3	1	0.8	0.02513950568	39.77802956	75.76133288
1	3	1	0.9	0.01227763376	81.44891919	189.0307956
2	4	2	0.5	0.06940053800	14.40911020	18.78487564
2	7	2	0.5	0.06548819579	15.26992747	22.50988174
2	8	2	0.5	0.06294536172	15.88679408	24.34081943
2	9	2	0.5	0.06030507988	16.58235097	26.35895394
2	10	2	0.5	0.05768797449	17.33463532	28.53631500
2	12	2	0.5	0.8445652090	1.184041196	1.660425113
2	20	2	0.5	0.6550163623	1.526679420	2.712775168
2	22	2	0.5	0.6360478228	1.572208825	2.920057928
2	25	2	0.5	0.6148358766	1.626450307	3.206981225
1	1	1	0.5	0.05930928767	16.86076565	20.59919288

Table 3.1: η , α and $\bar{\alpha}$ values for Wang's modified Model discussed in Section 3.2.2 for various values of μ_0 , μ_1 , μ_2 and p .

μ_0	μ_1	μ_2	p	$R(\alpha)$	$R_{01}(\alpha)$	$G(\alpha)$	$R(\alpha)G(\alpha)$
1	2	3	0.5	0.02354854078	0.008639657227	7.205958475	0.1696898070
1	1	3	0.5	0.01877573198	0.006678144921	6.406687240	0.1202902425
2	2	3	0.5	0.03150855042	0.022417151287	5.468444680	0.1723027649
2	3	2	0.5	0.03151199271	0.03362025377	5.633330220	0.1775174608
2	3	2	0.1	0.07625163210	0.08607501360	6.291174290	0.4797123074
2	3	2	0.2	0.06179377735	0.06875608233	6.005833150	0.3711231165
2	3	2	0.3	0.05011092875	0.05496926963	5.830302930	0.2921618947
2	3	2	0.4	0.04017822592	0.04346048480	5.713469380	0.2295570635
2	3	2	0.6	0.02383220481	0.025085521787	5.578234380	0.1329416242
2	3	2	0.7	0.01695580485	0.017612822557	5.541215050	0.09395576102
2	3	2	0.8	0.01075303074	0.011025909327	5.517756440	0.05933260462
2	3	2	0.9	0.005126493090	0.005190392179	5.504758460	0.02822010621
1	3	1	0.1	0.04355569466	0.04947870670	6.455447215	0.2811714878
1	3	1	0.2	0.03629959142	0.04056450413	6.496342805	0.2358145895
1	3	1	0.3	0.02989055463	0.03287945436	6.546699620	0.1956844826
1	3	1	0.4	0.02418772197	0.02620560339	6.603475910	0.1597230393
1	3	1	0.5	0.01908249523	0.02037471893	6.664657035	0.1271782861
1	3	1	0.6	0.01448867052	0.015253765957	6.728859720	0.09749223146
1	3	1	0.7	0.01033639104	0.01073570816	6.795108430	0.07023689789
1	3	1	0.8	0.006568169635	0.006733290173	6.862701150	0.04507538531
1	3	1	0.9	0.003136127838	0.003174632067	6.931124815	0.02173689348
2	4	2	0.5	0.02683127422	0.02869337909	5.957389860	0.1598443610
2	7	2	0.5	0.01621690946	0.017278925607	7.050643110	0.1143396409
2	8	2	0.5	0.01389279959	0.014767286897	7.454237750	0.1035602312
2	9	2	0.5	0.01200038712	0.01272407142	7.872285250	0.09447047052
2	10	2	0.5	0.01044869984	0.01105146417	8.302124610	0.08674640808
2	12	2	0.5	0.1386446112	0.2557390262	1.075602710	0.1491265195
2	20	2	0.5	0.06429456195	0.1064085520	1.240058690	0.07972903027
2	22	2	0.5	0.05680865215	0.09294167167	1.263553480	0.07178077012
2	25	2	0.5	0.04839074276	0.07814310750	1.291969594	0.06251936828
1	1	1	0.5	0.03518132602	0.03726790541	5.297361035	0.1863681856

Table 3.2: R and G measure values for Wang's modified model discussed in Section 3.2.2 for various values of μ_0, μ_1, μ_2 and p .

3.2.3 Priority Queue

The queueing system considered consists of a single server whose service times are exponentially distributed. The queueing system has a preemptive priority discipline. The resume strategy is applied when service recommences for a preempted lower priority customer. There are two classes of customers. Class i customers arrive according to a Poisson process with rate λ_i and are served at an exponential rate μ_i where $i = 1, 2$. Class 1 customers receive priority over class 2 customers. $X_1(t)$ and $X_2(t)$ are the number of class 1 and class 2 customers in the system at time t , respectively. $X(t) = (X_2(t), X_1(t))$ is a Markov chain with state space $S = \{(i, j); i, j \geq 0, i, j = 0, 1, 2, \dots\}$. This may be seen by considering the transition rate diagram in figure 3.4.

Figure 3.4: Transition diagram for Priority queue.

The equilibrium equations are

$$(\lambda_1 + \lambda_2 + \mu_1 + \mu_2)\pi_{ij} = \mu_2\pi_{ij} + \lambda_1\pi_{i,j-1} + \mu_1\pi_{i,j+1} + \lambda_2\pi_{i-1,j}$$

$$(\lambda_1 + \lambda_2 + \mu_1 + \mu_2)\pi_{i0} = \mu_2\pi_{i+1,0} + \mu_1\pi_{i,0} + \lambda_2\pi_{i-1,0} + \mu_1\pi_{i1}$$

$$(\lambda_1 + \lambda_2 + \mu_1 + \mu_2)\pi_{0j} = \mu_1\pi_{0,j+1} + \mu_2\pi_{0,j} + \lambda_1\pi_{0,j-1}$$

$$(\lambda_1 + \lambda_2 + \mu_1 + \mu_2)\pi_{00} = (\mu_1 + \mu_2)\pi_{00} + \mu_2\pi_{10} + \mu_1\pi_{01}.$$

The transition rate matrix has a quasi-birth-death structure as in (3.1.1) where

$$B_0 = \theta \begin{bmatrix} \mu_1 + \mu_2 & \lambda_1 & & & \\ & \mu_1 & \mu_2 & \lambda_1 & \\ & & \mu_1 & \mu_2 & \lambda_1 \\ & & & \mu_1 & \mu_2 & \lambda_1 \\ & & & & \ddots & \ddots & \ddots \end{bmatrix},$$

$$A_0 = \theta \begin{bmatrix} \lambda_2 & & & & \\ & \lambda_2 & & & \\ & & \lambda_2 & & \\ & & & \ddots & \\ & & & & \ddots \end{bmatrix}, C_1 = \theta \begin{bmatrix} \mu_2 & & & \\ & 0 & & \\ & & 0 & \\ & & & \ddots & \\ & & & & \ddots \end{bmatrix},$$

$$A = \theta \begin{bmatrix} \lambda_2 & & & & \\ & \lambda_2 & & & \\ & & \lambda_2 & & \\ & & & \ddots & \\ & & & & \ddots \end{bmatrix}, C = \theta \begin{bmatrix} \mu_2 & & & \\ & 0 & & \\ & & 0 & \\ & & & \ddots & \\ & & & & \ddots \end{bmatrix},$$

and

$$B = \theta \begin{bmatrix} \mu_1 & \lambda_1 & & & \\ \mu_1 & \mu_2 & \lambda_1 & & \\ & \mu_1 & \mu_2 & \lambda_1 & \\ & & \mu_1 & \mu_2 & \lambda_1 \\ & & & \ddots & \ddots & \ddots \end{bmatrix},$$

where $\theta = (\lambda_1 + \lambda_2 + \mu_1 + \mu_2)^{-1}$. In order to apply Theorem 16, we need to consider the matrix

$$\begin{aligned} P &= P(\eta) = A + \eta B + \eta^2 C \\ &= \theta \begin{bmatrix} \lambda_2 + \eta\mu_1 + \eta^2\mu_2 & \eta\lambda_1 & & & \\ \eta\mu_1 & \lambda_2 + \eta\mu_2 & \eta\lambda_1 & & \\ & \eta\mu_1 & \lambda_2 + \eta\mu_2 & \eta\lambda_1 & \\ & & \eta\mu_1 & \lambda_2 + \eta\mu_2 & \eta\lambda_1 \\ & & & \ddots & \ddots & \ddots \\ & & & & \ddots & \ddots & \ddots \end{bmatrix} \end{aligned}$$

where

$$\begin{aligned} \gamma_0 &= (\lambda_2 + \eta\mu_1 + \eta^2\mu_2)\theta, \gamma = (\lambda_2 + \eta\mu_2)\theta \\ \mu_0 &= \mu = \eta\mu_2\theta \\ \lambda_0 &= \lambda = \eta\lambda_1\theta. \end{aligned}$$

The radius of convergence, $\bar{\alpha}$, for Q , constructed from P by deleting the boundaries, is given by

$$\bar{\alpha} = \frac{\lambda_1 + \lambda_2 + \mu_1 + \mu_2}{\lambda_2 + \eta\mu_2 + 2\eta\sqrt{\lambda_1\mu_2}}.$$

The convergence radius of P is given by $\alpha = 1/\eta$ where η is a zero of the characteristic polynomial (3.1.10) satisfying inequality (3.1.8). Letting $\beta = \frac{1}{\eta}$, substituting $\theta =$

$\frac{1}{\lambda_1 + \lambda_2 + \mu_1 + \mu_2}$ and simplifying equation (3.1.10) we arrive at

$$0 = (\eta - 1)(\eta^2 \mu_2^2 + \eta(\mu_1 \mu_2 - \mu_2^2 - \mu_2 \lambda_1 - \lambda_2 \mu_2) + \lambda_2(\mu_2 - \mu_1)).$$

The roots of

$$0 = \eta^2 \mu_2^2 + \eta(\mu_1 \mu_2 - \mu_2^2 - \mu_2 \lambda_1 - \lambda_2 \mu_2) + \lambda_2(\mu_2 - \mu_1)$$

are required. Using the quadratic equation these are

$$\eta = \frac{\lambda_1 + \lambda_2 + \mu_2 - \mu_1 \pm \sqrt{(\lambda_1 + \lambda_2)^2 + 2(\lambda_1 - \lambda_2)(\mu_2 - \mu_1) + (\mu_2 - \mu_1)^2}}{2\mu_2}.$$

Therefore, the roots of the characteristic equation are

$$\eta = 1, \frac{\lambda_1 + \lambda_2 + \mu_2 - \mu_1 \pm \sqrt{(\lambda_1 + \lambda_2)^2 + 2(\lambda_1 - \lambda_2)(\mu_2 - \mu_1) + (\mu_2 - \mu_1)^2}}{2\mu_2}. \quad (3.2.9)$$

In order for P to be $1/\eta$ -recurrent, $\alpha \leq \bar{\alpha}$. This is equivalent to determining if

$$\alpha = \frac{1}{\eta} \leq \frac{\lambda_1 + \lambda_2 + \mu_1 + \mu_2}{\lambda_2 + \eta\mu_2 + 2\eta\sqrt{\lambda_1\mu_2}} = \bar{\alpha}$$

$$\frac{\lambda_2}{\lambda_1 + \lambda_2 + \mu_1 - 2\sqrt{\lambda_1\mu_2}} \leq \eta$$

is satisfied.

The R measure, given by equations (3.1.15) and (3.1.16), is

$$R(1/\eta) = \frac{\lambda_1 + \lambda_2 + \mu_1 - \lambda_2/\eta}{2\mu_2} - \frac{\sqrt{(\lambda_1 + \lambda_2 + \mu_1 - \lambda_2/\eta)^2 - 4\lambda_1\mu_1}}{2\mu_2} \quad (3.2.10)$$

and

$$R_{01}(1/\eta) = \frac{\lambda_0}{\lambda} R(1/\eta) = R(1/\eta)$$

where η is the root of the characteristic equation lying in the interval $(0, 1)$. The G measure, given by equations (3.1.22) and (3.1.23), is

$$G(1/\eta) = \frac{\mu}{\lambda} R(1/\eta) = \frac{\mu_1}{\lambda_1} R(1/\eta)$$

and

$$G_{10}(1/\eta) = \frac{\mu_0}{\mu} G(1/\eta) = G(1/\eta).$$

The $1/\eta$ -invariant measure and vector may be determined using equations (3.1.14) and (3.1.21), respectively.

Another way to determine if P is $1/\eta$ -positive is to ensure the condition $\mathbf{xy}^T < \infty$ is satisfied. This reduces to

$$\begin{aligned} \mathbf{xy}^T &= (1, R(1/\eta), R(1/\eta)^2, \dots)(1, \mu_1/\lambda_1 R(1/\eta), (\mu_1/\lambda_1 R(1/\eta))^2, \dots)^T \\ &= \sum_{n=0}^{\infty} \left(\frac{\mu_1}{\lambda_1} \right)^n R(1/\eta)^{2n} \\ &= \sum_{n=0}^{\infty} \left(\frac{\mu_1}{\lambda_1} R(1/\eta)^2 \right)^n < \infty \end{aligned}$$

if and only if

$$\left| \frac{\mu_1}{\lambda_1} R(1/\eta)^2 \right| < 1. \quad (3.2.11)$$

There will be a range of η values for which P is $1/\eta$ -positive, depending on the parameter values, $\lambda_1, \lambda_2, \mu_1$ and μ_2 . This range occurs when η from (3.2.9) lying in the interval $(0, 1)$ is substituted into the expression for the R measure for P given in (3.2.10). This is then substituted into (3.2.11) and gives a region in which the parameter values must lie. In the special case when $\mu_1 = \mu_2$ the region can be found quite easily. This will be shown in the next section.

The final condition to be verified is condition 3(c). Namely, that $\frac{1}{\eta} \mathbf{x} A \mathbf{y}^T \neq \eta \mathbf{x} C \mathbf{y}^T$.

$$\frac{1}{\eta} \mathbf{x} A \mathbf{y}^T = \frac{\theta \lambda_2}{\eta} \sum_{i=0}^{\infty} x_i y_i \neq \theta \eta \mu x_0 y_0 = \eta \mathbf{x} C \mathbf{y}^T$$

as required.

Special Case: Equal Service Rates, $\mu_1 = \mu_2 = \mu$

A special case of the above model is when the service rates are equal. That is $\mu_1 = \mu_2 = \mu$. In this case the entries of P in equation (3.1.3) are given by

$$\begin{aligned}\gamma_0 &= (\lambda_2 + \eta\mu + \eta^2\mu)\theta, \gamma = \theta(\lambda_2 + \eta\mu) \\ \mu_0 &= \mu = \eta\mu\theta \\ \lambda_0 &= \lambda = \eta\lambda_1\theta.\end{aligned}$$

The radius of convergence, $\bar{\alpha}$, for Q , constructed from P by deleting the boundaries, is given by

$$\bar{\alpha} = \frac{\lambda_1 + \lambda_2 + 2\mu}{\lambda_2 + \eta\mu + 2\eta\sqrt{\lambda_1\mu}}.$$

The characteristic equation reduces to

$$\begin{aligned}0 &= (\eta - 1)(\eta^2\mu^2 - \eta\mu(\lambda_1 + \lambda_2)) \\ 0 &= \eta(\eta - 1)(\eta\mu - (\lambda_1 + \lambda_2))\mu.\end{aligned}$$

Therefore,

$$\eta = \frac{\lambda_1 + \lambda_2}{\mu} = \rho < 1.$$

The R measure is given by

$$R(1/\eta) = \frac{\eta(\lambda_1 + \lambda_2 + \mu) - \lambda_2 - \sqrt{(\eta(\lambda_1 + \lambda_2 + \mu) - \lambda_2)^2 - 4\lambda_1\mu\eta^2}}{2\eta\mu}$$

$$R(1/\eta) = \begin{cases} \frac{\rho_1}{\rho}, & \rho_1 < \rho^2 \\ \rho, & \rho_1 > \rho^2 \end{cases} \quad (3.2.12)$$

and

$$R_{01}(1/\eta) = \frac{\lambda_0}{\lambda} R(1/\eta) = R(1/\eta).$$

And the G measure is given by

$$G(1/\eta) = \frac{\mu}{\lambda} R(1/\eta) = \frac{\mu}{\lambda_1} R(1/\eta)$$

and

$$G_{10}(1/\eta) = \frac{\mu_0}{\mu} G(1/\eta) = G(1/\eta).$$

The $1/\rho$ -invariant measure and vector may be determined by using equations (3.1.14) and (3.1.21), respectively.

It can be checked that if $\rho_1 < \rho^2$ then $R(1/\eta) = \rho_1/\rho$, $\mathbf{x} = (1, \rho_1/\rho, (\rho_1/\rho)^2, \dots)$

and

$$\begin{aligned} \frac{1}{\rho} \mathbf{x} P &= \frac{\theta}{\rho} (1, \rho_1/\rho, (\rho_1/\rho)^2, \dots) \begin{bmatrix} \lambda_2 + \rho\mu + \rho^2\mu & \rho\lambda_1 & & & \\ & \rho\mu & \lambda_2 + \rho\mu & \rho\lambda_1 & \\ & & \rho\mu & \lambda_2 + \rho\mu & \rho\lambda_1 \\ & & & \ddots & \ddots & \ddots \\ & & & & \ddots & \ddots & \ddots \end{bmatrix} \\ &= (1, \rho_1/\rho, (\rho_1/\rho)^2, \dots) = \mathbf{x}. \end{aligned}$$

Also, $\mathbf{y} = (1, 1/\rho, 1/\rho^2, 1/\rho^3, \dots)$ and

$$\frac{1}{\rho} P \mathbf{y}^T = \frac{\theta}{\rho} \begin{bmatrix} \lambda_2 + \rho\mu + \rho^2\mu & \rho\lambda_1 & & & \\ & \rho\mu & \lambda_2 + \rho\mu & \rho\lambda_1 & \\ & & \rho\mu & \lambda_2 + \rho\mu & \rho\lambda_1 \\ & & & \ddots & \ddots & \ddots \\ & & & & \ddots & \ddots & \ddots \end{bmatrix} \begin{bmatrix} 1 \\ 1/\rho \\ 1/\rho^2 \\ \vdots \\ \vdots \end{bmatrix} = \begin{bmatrix} 1 \\ 1/\rho \\ 1/\rho^2 \\ \vdots \\ \vdots \end{bmatrix} = \mathbf{y}^T.$$

But, if $\rho_1 > \rho^2$ then $R(1/\eta) = \rho$, $\mathbf{x} = (1, \rho, \rho^2, \rho^3, \dots)$ and

$$\begin{aligned} \frac{1}{\rho} \mathbf{x}P &= \frac{\theta}{\rho} (1, \rho, \rho^2, \rho^3, \dots) \begin{bmatrix} \lambda_2 + \rho\mu + \rho^2\mu & \rho\lambda_1 & & & \\ & \rho\mu & \lambda_2 + \rho\mu & \rho\lambda_1 & \\ & & \rho\mu & \lambda_2 + \rho\mu & \rho\lambda_1 \\ & & & \ddots & \ddots & \ddots \\ & & & & \ddots & \ddots & \ddots \end{bmatrix} \\ &= \frac{\theta}{\rho} \begin{bmatrix} \lambda_2 + \rho\mu + \rho^2\mu + \rho^2\mu \\ \lambda_1\mu + \rho\lambda_2 + \rho^2\mu + \rho^3\mu \\ \rho(\lambda_1\mu + \rho\lambda_2 + \rho^2\mu + \rho^3\mu) \\ \vdots \\ \vdots \end{bmatrix} \neq (1, \rho, \rho^2, \rho^3, \dots) = \mathbf{x}. \end{aligned}$$

Therefore, when $\rho_1 > \rho^2$ then $R(1/\eta) = \rho$ and $\mathbf{x} = (1, \rho, \rho^2, \rho^3, \dots)$ is not an $1/\rho$ -invariant measure for P . However, when $\rho_1 < \rho^2$ then $R(1/\eta) = \rho_1/\rho$ which gives $\mathbf{x} = (1, \rho_1/\rho, (\rho_1/\rho)^2, (\rho_1/\rho)^3, \dots)$, an $1/\rho$ -invariant measure and $\mathbf{y} = (1, 1/\rho, 1/\rho^2, \dots)$, an $1/\rho$ -invariant vector for P . Thus, $\rho_1 < \rho^2$ and so $R(1/\eta) = \rho_1/\rho$ is the only possibility.

To determine if $P(1/\eta)$ is $1/\eta$ -positive the condition $\mathbf{xy}^T < \infty$ must be verified.

This reduces to

$$\begin{aligned} \mathbf{xy}^T &= (1, R(1/\eta), R(1/\eta)^2, \dots)(1, (\mu/\lambda_1)R(1/\eta), ((\mu/\lambda_1)R(1/\eta))^2, \dots)^T \\ &= \sum_{n=0}^{\infty} \left(\frac{\mu}{\lambda_1}\right)^n R(1/\eta)^{2n} = \sum_{n=0}^{\infty} \left(\frac{\mu}{\lambda_1} R(1/\eta)^2\right)^n = \sum_{n=0}^{\infty} \left(\frac{1}{\rho_1} \left(\frac{\rho_1}{\rho}\right)^2\right)^n \\ &= \sum_{n=0}^{\infty} \left(\frac{\rho_1}{\rho^2}\right)^n < \infty \end{aligned}$$

if and only if

$$\left| \frac{\rho_1}{\rho^2} \right| < 1 \text{ or } \rho_1 < \rho^2,$$

which holds from above.

The final condition to be verified is Condition 3(c). Namely, that $\frac{1}{\eta}\mathbf{x}A\mathbf{y}^T \neq \eta\mathbf{x}C\mathbf{y}^T$.

$$\frac{1}{\eta}\mathbf{x}A\mathbf{y}^T = \frac{\theta\lambda_2}{\rho} \sum_{i=0}^{\infty} x_i y_i = \frac{\theta\lambda_2}{\rho} \left(\frac{\rho^2}{\rho^2 - \rho_1} \right) \neq \theta\rho\mu = \theta\rho\mu x_0 y_0 = \eta\mathbf{x}C\mathbf{y}^T$$

as required.

3.2.4 Symmetric Shortest Queue

Many people have studied the symmetric shortest queue, SSQ, model. Among these are Kingman [36], Flatto and McKean [20], Cohen and Boxma, Adan [1], Zhao and Grassman [65]. Kingman [36], Flatto and McKean [20], used generating functions to analyze the model. Cohen and Boxma [16] also used generating functions and showed that they may be reduced to that of a Reimann-Hilbert boundary value problem. Adan [1] used the compensation method. Zhao and Grassman [65] provided a stable numerical algorithm.

The model may be described as follows, a system consists of two identical servers. Customers arrive according to a Poisson process with rate 2ρ , where $0 < \rho < 1$. Upon arrival a job joins the shortest queue and if the queues are equal in length then the customer joins either queue with probability $1/2$. The servers are exponentially distributed with service rate 1. The service times are independent. Let $Y_1(t)$ and $Y_2(t)$ be the lengths of the two queues at time t . $Y(t) = (Y_1(t), Y_2(t))$ is a Markov chain with state space $S = \{(i, j); i, j = 0, 1, 2, \dots\}$ but its infinitesimal generator is level dependent. To eliminate this dependence we use the representation considered by Adan [1]. Let $X_1(t) = \min(Y_1(t), Y_2(t)) = m$ and $X_2(t) = Y_2(t) - Y_1(t) = n$. The rates in the region $n \geq 0$ can be obtained by reflection in the m -axis. By symmetry $\pi_{mn} = \pi_{m, -n}$. Hence, the analysis can be restricted to the region $n \geq 0$. Now

$X(t) = (X_1(t), X_2(t))$ is a Markov chain with state space $S = \{(i, j); i, j = 0, 1, 2, \dots\}$ and a level independent transition matrix. The transition rate diagram may be found in figure 3.5.

Figure 3.5: Transition diagram for the Symmetric Shortest Queue.

From here the equilibrium equations are given by

$$(2\rho + 2)\pi_{ij} = \pi_{i,j+1} + \pi_{i+1,j-1} + 2\rho\pi_{i-1,j+1}$$

$$(2\rho + 2)\pi_{i0} = (1 + \rho)\pi_{i0} + 2\rho\pi_{i-1,1} + \pi_{i1}, i > 0$$

$$(2\rho + 2)\pi_{0j} = \pi_{0j} + \pi_{0,j+1} + \pi_{1,j-1}, j > 1$$

$$(2\rho + 2)\pi_{00} = (2 + \rho)\pi_{00} + \pi_{01}$$

$$(2\rho + 2)\pi_{i1} = \rho\pi_{i0} + 2\rho\pi_{i-1,2} + \pi_{i+1,0}, i > 0.$$

The transition matrix has a quasi-birth-death structure as in (3.1.1) where

$$\begin{aligned}
 B_0 &= \theta \begin{bmatrix} 2+\rho & \rho & & & \\ & 1 & 1 & 0 & \\ & & 1 & 1 & 0 \\ & & & 1 & 1 & 0 \\ & & & & \ddots & \ddots & \ddots \\ & & & & & \ddots & \ddots & \ddots \end{bmatrix}, \\
 A_0 &= \theta \begin{bmatrix} 0 & & & & \\ 2\rho & 0 & & & \\ & 2\rho & 0 & & \\ & & \ddots & \ddots & \\ & & & \ddots & \ddots \end{bmatrix}, C_1 = \theta \begin{bmatrix} 0 & 1 & & & \\ & 0 & 1 & & \\ & & 0 & 1 & \\ & & & \ddots & \ddots & \ddots \\ & & & & \ddots & \ddots & \ddots \end{bmatrix}, \\
 A &= \theta \begin{bmatrix} 0 & & & & \\ 2\rho & 0 & & & \\ & 2\rho & 0 & & \\ & & \ddots & \ddots & \ddots \\ & & & \ddots & \ddots & \ddots \end{bmatrix}, C = \theta \begin{bmatrix} 0 & 1 & & & \\ & 0 & 1 & & \\ & & 0 & 1 & \\ & & & \ddots & \ddots & \ddots \\ & & & & \ddots & \ddots & \ddots \end{bmatrix},
 \end{aligned}$$

and

$$B = \theta \begin{bmatrix} 1+\rho & \rho & & & \\ & 1 & 0 & 0 & \\ & & 1 & 0 & 0 \\ & & & \ddots & \ddots & \ddots \\ & & & & \ddots & \ddots & \ddots \end{bmatrix}.$$

where $\theta = (2\rho + 2)^{-1}$.

In order to use Theorem 16 we need to consider the matrix

$$\begin{aligned}
 P &= P(\eta) = A + \eta B + \eta^2 C \\
 &= \begin{bmatrix} \gamma_0 & \lambda_0 & 0 & 0 & 0 & \cdots \\ \mu_0 & \gamma & \lambda & 0 & 0 & \cdots \\ 0 & \mu & \gamma & \lambda & 0 & \cdots \\ 0 & 0 & \mu & \gamma & \lambda & \cdots \\ \vdots & \vdots & & \ddots & \ddots & \ddots \end{bmatrix},
 \end{aligned}$$

where

$$\begin{aligned}
 \gamma_0 &= \frac{\eta}{2}, & \gamma &= 0 \\
 \lambda_0 &= \frac{\eta(\rho+\eta)}{2(\rho+1)}, & \lambda &= \frac{\eta^2}{2(\rho+1)} \\
 \mu_0 &= \frac{(2\rho+\eta)}{2(\rho+1)}, & \mu &= \frac{(2\rho+\eta)}{2(\rho+1)}.
 \end{aligned}$$

The convergence radius $\bar{\alpha}$ of Q , constructed from P by deleting boundaries, is given by

$$\bar{\alpha} = \frac{\rho+1}{\eta\sqrt{2\rho+\eta}}. \quad (3.2.13)$$

The convergence radius of P is given by $\alpha = 1/\eta$, where η is a root of the following equation

$$(\eta - \rho^2)(\eta - 1)(\eta + 2\rho) = 0 \quad (3.2.14)$$

which was obtained by letting $\beta = 1/\eta$ in equation (3.1.10) and simplifying. The roots of equation (3.2.14) are $\eta = \rho^2, 1, -2\rho$, the only one of which that is in the interval $(0, 1)$ is ρ^2 , where $0 < \rho < 1$. In order to determine if P is α -recurrent we need to determine if $\alpha \leq \bar{\alpha}$ or equivalently, if $\frac{1}{\rho^2} \leq \frac{\rho+1}{\rho^2\sqrt{2\rho+\rho^2}}$. This is true since, $\sqrt{2\rho+\rho^2} \leq \sqrt{1+2\rho+\rho^2} = \rho+1$. This inequality is actually strict, $\alpha < \bar{\alpha}$, therefore, P is $1/\rho^2$ -positive and by Theorem 16 the $1/\rho^2$ -invariant measure $\mathbf{x}$ and vector $\mathbf{y}$ exist. Equations (3.1.14) and (3.1.21) may be used to determine $\mathbf{x}$ and $\mathbf{y}$, respectively.

The R and G measures for P are given by,

$$G\left(\frac{1}{\rho^2}\right) = \frac{1}{\rho}, \quad G_{10}\left(\frac{1}{\rho^2}\right) = \frac{\mu_0}{\mu} G\left(\frac{1}{\rho^2}\right) = \frac{1}{\rho},$$

and

$$R\left(\frac{1}{\rho^2}\right) = \frac{\rho^2}{2+\rho}, \quad R_{01}\left(\frac{1}{\rho^2}\right) = \frac{\lambda_0}{\lambda} R\left(\frac{1}{\rho^2}\right) = \left(\frac{\rho+1}{\rho}\right) \left(\frac{\rho^2}{2+\rho}\right).$$

By equation (3.1.14) the α -invariant measure $\mathbf{x}$ of P is given by,

$$\begin{aligned} x_0 &= z_0 = \frac{2+\rho-\rho^2}{2+2\rho-\rho^2} > 0, \\ x_n &= z_0 \frac{\lambda_0}{\lambda} R^n\left(\frac{1}{\rho^2}\right) = z_0 \left(\frac{\rho+1}{\rho}\right) \left(\frac{\rho^2}{2+\rho}\right)^n, \quad n \geq 1. \end{aligned}$$

Therefore,

$$\mathbf{x} = \frac{2+\rho-\rho^2}{2+2\rho-\rho^2} \left(\frac{\rho+1}{\rho}\right) \left(\frac{\rho}{\rho+1}, \frac{\rho^2}{2+\rho}, \left(\frac{\rho^2}{2+\rho}\right)^2, \dots, \left(\frac{\rho^2}{2+\rho}\right)^n, \dots\right).$$

By equation (3.1.21) the α -invariant vector is given by

$$\begin{aligned} y_0 &= w_0 = 1 \\ y_n &= \frac{1}{\rho} \left(\frac{1}{\rho}\right)^{n-1} w_0 = \frac{1}{\rho^n}, \quad n \geq 1. \end{aligned}$$

Another way of determining the α -positivity of P is to consider the convergence of $\mathbf{xy}^T$. Given $\mathbf{x}$ and $\mathbf{y}$ from above

$$\begin{aligned} \mathbf{xy}^T &= \left(\frac{2+\rho-\rho^2}{2+2\rho-\rho^2}\right) \left(\frac{\rho+1}{\rho}\right) \left(\frac{\rho}{\rho+1}, \left(\frac{\rho}{\rho+2}\right)^2, \dots, \left(\frac{\rho}{\rho+2}\right)^n, \dots\right) \\ &\quad \left(1, \frac{1}{\rho}, \left(\frac{1}{\rho}\right)^2, \dots, \left(\frac{1}{\rho}\right)^n, \dots\right)^T \\ &= \left(\frac{2+\rho-\rho^2}{2+2\rho-\rho^2}\right) \left(\frac{\rho+1}{\rho}\right) \left(\frac{\rho}{\rho+1} + \sum_{n=1}^{\infty} \left(\frac{\rho}{\rho+2}\right)^n\right) < \infty \end{aligned}$$

since $\frac{\rho}{\rho+2} < 1$. Therefore, P is α -positive.

Next we need to determine if $\frac{1}{\eta}\mathbf{x}A\mathbf{y}^T \neq \eta\mathbf{x}C\mathbf{y}^T$, which is equivalent to

$$\begin{aligned} \frac{1}{\eta}\mathbf{x}A\mathbf{y}^T &= \frac{1}{\eta}(x_0, x_1, \dots)A(y_0, y_1, \dots)^T \\ &= \frac{2\theta\rho}{\eta}(x_1, x_2, \dots)(y_0, y_1, \dots)^T \\ &= \frac{2\rho\theta}{\eta} \sum_{i=0}^{\infty} x_{i+1}y_i \\ &\neq \eta(x_0, x_1, \dots)C(y_0, y_1, \dots)^T \\ &= \theta\eta(0, x_0, \dots)(y_0, y_1, \dots)^T \\ &= \theta\eta \sum_{i=0}^{\infty} x_i y_{i+1} = \eta\mathbf{x}C\mathbf{y}^T. \end{aligned}$$

We see that $\frac{1}{\eta}\mathbf{x}A\mathbf{y}^T \neq \eta\mathbf{x}C\mathbf{y}^T$ is satisfied. Finally, we need to determine if $\pi_1\mathbf{y} < \infty$.

We know $\pi_1\mathbf{e}^T < 1 = \mathbf{x}\mathbf{e}^T$ and that $\mathbf{x}\mathbf{y}^T < \infty$. Thus we have that

$$\pi_m \sim c\eta^n \mathbf{x} = c \left(\frac{1}{\rho^2}\right)^m \frac{2+\rho-\rho^2}{2+2\rho-\rho^2} \left(\frac{\rho+1}{\rho}\right) \left(\frac{\rho}{\rho+1}, \frac{\rho^2}{2+\rho}, \left(\frac{\rho^2}{2+\rho}\right)^2, \dots, \left(\frac{\rho^2}{2+\rho}\right)^n, \dots\right).$$

Interchanging level and phase variable, $X(t) = (X_2(t), X_1(t))$ becomes the Markov process considered. When applying Theorem 16 to this model the resulting root of the characteristic equation is $\eta = 1$ which does not lie within the prescribed range given in Theorem 16. Thus, the conditions for α -positivity of P are not satisfied. However, this model has been studied by Adan in [1] using the compensation method. The compensation method gives an explicit solution for the stationary distribution. Analysis of the expression for the stationary distribution shows that it does have a geometric tail along the X_2 direction. From here we see that α -positivity is a sufficient condition for a geometric tail along a given direction but it is not a necessary condition.

Chapter 4

Determination of η and x for Generalized P

In this chapter we will be generalizing the method introduced in Chapter 3, Section 3.1 to the case when the infinite repeating blocks, A, B and C of P' have a more general repeating structure. In this situation the *characteristic polynomial* is no longer a polynomial of degree eight but is rather an infinite series. In fact, we now have two infinite series that need to be solved for two unknowns. One is either the G measure or R measure for P and the other is the eigenvalue η for the R -measure of P' . Section 4.1 will be concerned with these two equations and Section 4.2 will consider examples.

4.1 Generalizations

In this section we will first be looking at the situation when A, B and C have a $M/G/1$ type scalar repeating structure, Section 4.1.1, and then, when A, B and C have a $GI/M/1$ type scalar repeating structure, Section 4.1.2.

4.1.1 Irreducible $M/G/1$ Type

This section is similar to, and the results are parallel to those in Section 3.1. However, in this section the infinite blocks of P' in equation (3.1.1) have a scalar, repeating $M/G/1$ type structure. Let the repeating blocks A, B and C each be of $M/G/1$ type, i.e.

$$A = \begin{bmatrix} \overline{a}_1 & \overline{a}_2 & \overline{a}_3 & \cdots & \cdots & \cdots & \cdots \\ \overline{a}_0 & a_1 & a_2 & \cdots & \cdots & \cdots & \cdots \\ & a_0 & a_1 & a_2 & \cdots & \cdots & \cdots \\ & & a_1 & a_1 & a_2 & \cdots & \cdots \\ & & & \ddots & \ddots & \ddots & \\ & & & & \ddots & \ddots & \ddots \end{bmatrix}$$

where $\overline{a}_i$ and a_i for $i \geq 0$ are scalars, and similarly for B and C . As a result the matrix P in equation (3.1.3) also has a $M/G/1$ type structure. For $0 < \eta < 1$ consider the matrix $P = P(\eta) = A + \eta B + \eta^2 C$ which is substochastic, due to our choice of η . Expanding equation (3.1.3) we have,

$$P = \begin{bmatrix} D_1 & D_2 & D_3 & \cdots & \cdots & \cdots & \cdots \\ D_0 & C_1 & C_2 & \cdots & \cdots & \cdots & \cdots \\ & C_0 & C_1 & C_2 & \cdots & \cdots & \cdots \\ & & C_0 & C_1 & C_2 & \cdots & \cdots \\ & & & \ddots & \ddots & \ddots & \\ & & & & \ddots & \ddots & \ddots \end{bmatrix} \quad (4.1.1)$$

where

$$\begin{aligned} D_i &= \overline{a}_i + \eta \overline{b}_i + \eta^2 \overline{c}_i, \quad i \geq 0 \\ C_i &= a_i + \eta b_i + \eta^2 c_i, \quad i \geq 0 \end{aligned}$$

are scalars. Also, denote by Q the submatrix of P without the boundary values,

$$Q = \begin{bmatrix} C_1 & C_2 & \cdots & \cdots & \cdots & \cdots \\ C_0 & C_1 & C_2 & \cdots & \cdots & \cdots \\ & C_0 & C_1 & C_2 & \cdots & \cdots \\ & & \ddots & \ddots & \ddots & \\ & & & \ddots & \ddots & \ddots \end{bmatrix}. \quad (4.1.2)$$

We will assume P and Q are irreducible.

As in Section 3.1 the radius of convergence $\bar{\alpha}$ for Q will be of interest to us. The result of Kijima in Theorem 10 may be used to find it. We need to first determine the convergence norm which is defined as $\nu = \frac{1}{\bar{\alpha}}$ where $\bar{\alpha}$ is the radius of convergence for Q . If we are able to find ν for Q , we know $\bar{\alpha}$ for Q . We use the method outlined by Kijima in Lemma 7 to find ν . We need to determine the Perron-Frobenius eigenvalue, $\chi(z)$, for the generating function, $C^*(z)$, of Q . The generating function for Q is $C^*(z) = \sum_{i=0}^{\infty} C_i z^i$. Now we need to solve,

$$\chi(\theta) = \nu\theta \text{ and } \chi'(\theta) = \nu$$

for ν or,

$$\chi(\theta) = \sum_{i=0}^{\infty} C_i \theta^i = \nu\theta \quad (4.1.3)$$

$$\chi'(\theta) = \sum_{i=1}^{\infty} i C_i \theta^{i-1} = \nu. \quad (4.1.4)$$

Substituting equation (4.1.4) into (4.1.3) we need to solve the infinite series,

$$C_0 + \sum_{i=1}^{\infty} (1-i) C_i \theta^i = 0 \quad (4.1.5)$$

for θ , then substitute this value into (4.1.4) to obtain ν . Once ν is determined, $\bar{\alpha} = 1/\nu$.

Consider βP , the discounted transition matrix at rate β , for $0 < \beta \leq \alpha$ where α is the radius of convergence for P . Let $U_0(\beta)$ be the censored matrix of βP with censoring set equal to level zero and let $u_0(\beta)$ be the maximal eigenvalue of $U_0(\beta)$ given in equation (2.2.20), i.e.

$$U_0(\beta) = \beta D_1 + \sum_{k=1}^{\infty} \beta D_{k+1} G(\beta)^{k-1} N_0(\beta) \beta D_0$$

where

$$G_{10}(\beta) = \beta N_0(\beta) D_0.$$

$G_{10}(\beta)$ may also be given by

$$G_{10}(\beta) = \frac{D_0}{C_0} G(\beta).$$

Since $U_0(\beta)$ is a scalar function, $u_0(\beta) = U_0(\beta)$. This gives

$$U_0(\beta) = u_0(\beta) = \beta D_1 + \sum_{k=1}^{\infty} \beta D_{k+1} \frac{D_0}{C_0} G^k$$

where $G(\beta)$ is the minimal, non-negative solution of

$$G(\beta) = \sum_{k=0}^{\infty} \beta C_k G(\beta)^k. \quad (4.1.6)$$

Let $\beta^* = \frac{1}{\eta}$ satisfy

$$\frac{1}{\eta} < \bar{\alpha} \quad (4.1.7)$$

and substitute $\beta = \beta^*$ into $u_0(\beta) = 1$. Then we have η equal to a zero of

$$u_0(1/\eta) - 1 = 0. \quad (4.1.8)$$

Therefore, η must satisfy the equations (4.1.6) and the *characteristic equation*

$$\frac{1}{\eta} D_1 + \frac{1}{\eta} \frac{D_0}{C_0} \sum_{k=1}^{\infty} D_{k+1} G(\beta)^k = 1. \quad (4.1.9)$$

Theorem 20 Suppose η is in the interval $(0, 1)$ satisfying equations (4.1.9) and (4.1.6). Then the radius of convergence for P given in (4.1.1) is $\alpha = \frac{1}{\eta}$ and P is $1/\eta$ -recurrent.

Proof: This result follows from Theorem 11(ii).

In order for P to be $1/\eta$ -positive, the conditions in Theorem 12 must hold. This is summarized in the following corollary.

Corollary 21 If

$$\sum_{k=1}^{\infty} k D_k G^{k-1}(\alpha) < \infty, \quad \sum_{k=1}^{\infty} k C_k G^{k-1}(\alpha) < \infty \quad (4.1.10)$$

and inequality 4.1.7 then P is α -positive, otherwise P is α -null.

Once P is $1/\eta$ -positive and provided the $1/\eta$ -invariant measure and vector $\mathbf{x}$ and $\mathbf{y}$, respectively, exist, then by Theorem 7, $\mathbf{xy} < \infty$. This is stated in the following corollary.

Corollary 22 If $\mathbf{x}$ and $\mathbf{y}$ are, respectively, an $1/\eta$ -invariant measure and vector of P then $\mathbf{xy} < \infty$.

By [39, Theorem 4] $\mathbf{x}$ is given by,

$$\begin{aligned} x_0 &= x_0 \\ x_k &= x_0 R_{0k}(\alpha) + \sum_{i=1}^{k-1} x_i R_{k-i}(\alpha) \end{aligned}$$

where x_0 is the solution to $x_0(I - U_0(\alpha)) = 0$ and

$$R_{0k}(\beta) = \frac{1}{C_0} \sum_{i=1}^{\infty} D_{k+i} G(\beta)^i$$

and

$$R_k(\beta) = \frac{1}{C_0} \sum_{i=1}^{\infty} C_{k+i} G(\beta)^i$$

for $k = 1, 2, \dots$

In order to determine the $1/\eta$ -invariant vector the factorization given in equation (2.2.24) may be used

$$I - \frac{1}{\eta}P(\eta) = (I - R_U(\alpha))(I - U_D(\alpha))(I - G_L(\alpha)),$$

where $I - R_U(\alpha)$, $I - U_D(\alpha)$ and $I - G_L(\alpha)$ are given in Theorem 13. We want to solve

$$0 = \left[I - \frac{1}{\eta}P(\eta) \right] \mathbf{y}^T = (I - R_U(\alpha))(I - U_D(\alpha))(I - G_L(\alpha))\mathbf{y}^T.$$

This is done by letting $\mathbf{z}^T = (I - G_L(\alpha))\mathbf{y}^T$ and solving $\mathbf{0}^T = (I - R_U(\alpha))(I - U_D(\alpha))\mathbf{z}^T$. Solving $\mathbf{0}^T = (I - R_U(\alpha))(I - U_D(\alpha))\mathbf{z}^T$ gives,

$$\begin{aligned} \mathbf{0}^T &= (I - R_U(\alpha))(I - U_D(\alpha))\mathbf{z}^T \\ &= \begin{bmatrix} I & -R_{01} & -R_{02} & -R_{03} & \cdots \\ & I & -R_1 & -R_2 & \cdots \\ & & I & -R_1 & \cdots \\ & & & \ddots & \ddots \\ & & & & \ddots & \ddots \end{bmatrix} \begin{bmatrix} I - U_0 & & & & \\ & I - U & & & \\ & & I - U & & \\ & & & \ddots & \\ & & & & \ddots \end{bmatrix} \begin{bmatrix} z_0 \\ z_1 \\ z_2 \\ \vdots \end{bmatrix} \\ &= \begin{bmatrix} I - U_0 & -R_{01}(I - U) & -R_{02}(I - U) & \cdots \\ & I - U & -R_1(I - U) & -R_2(I - U) & \cdots \\ & & I - U & -R_1(I - U) & \cdots \\ & & & \ddots & \ddots \\ & & & & \ddots & \ddots \end{bmatrix} \begin{bmatrix} z_0 \\ z_1 \\ z_2 \\ \vdots \end{bmatrix} \end{aligned}$$

$$\begin{aligned}
0 &= (I - U_0)z_0 + \sum_{i=1}^{\infty} -R_{0i}(I - U)z_i \\
0 &= (I - U)z_1 + \sum_{i=1}^{\infty} -R_i(I - U)z_{i+1} \\
0 &= (I - U)z_j + \sum_{i=1}^{\infty} -R_i(I - U)z_{i+j}, \quad j \geq 1.
\end{aligned}$$

There exists an $z_0 \neq 0$ such that $(I - U_0)z_0 = 0$, therefore, $\mathbf{z} = (z_0, 0, 0, \dots)$ is a solution to $\mathbf{0}^T = (I - R_U)(I - U_D)\mathbf{z}^T$. Now we must solve for $\mathbf{y}$ in terms of $\mathbf{z}$ using $\mathbf{z}^T = (I - G_L)\mathbf{y}^T$.

$$\begin{aligned}
\mathbf{z}^T &= (I - G_L)\mathbf{y}^T \\
\begin{bmatrix} z_0 \\ z_1 \\ z_2 \\ \vdots \end{bmatrix} &= \begin{bmatrix} I & & & & \\ -G_{10} & I & & & \\ & -G & I & & \\ & & -G & I & \\ & & & \ddots & \ddots \\ & & & & \ddots & \ddots \end{bmatrix} \begin{bmatrix} y_0 \\ y_1 \\ y_2 \\ \vdots \end{bmatrix} \\
z_0 &= y_0 \\
0 &= -G_{10}y_0 + y_1 \\
0 &= -Gy_j + y_{j+1}, \quad j \geq 1
\end{aligned}$$

or

$$\begin{aligned}
y_0 &= z_0 \\
y_{j+1} &= Gy_j = G^j G_{10} y_0, \quad j \geq 0.
\end{aligned}$$

Provided Theorem 20, Corollaries 21 and 22 hold the result [48, Theorem 4.1] holds, i.e.

$$\pi_n \sim \eta^{n-1} \frac{\mathbf{x}\mathbf{y}^T}{\mathbf{x} \left(\frac{1}{\eta} P(\eta) \right)' \mathbf{y}^T} \pi_1 \quad (4.1.11)$$

where the derivative of $\frac{1}{\eta}P(\eta)$ is taken componentwise with respect to $1/\eta$.

4.1.2 Irreducible $GI/M/1$ Type

In this section we consider the situation when the blocks A, B and C of P' have a scalar $GI/M/1$ type structure. Technically, this section is similar to Section 4.1.1. Let the repeating blocks A, B and C each be of $GI/M/1$ type, i.e.

$$A = \begin{bmatrix} \overline{a_1} & \overline{a_0} & & & \\ \overline{a_2} & a_1 & a_0 & & \\ \overline{a_3} & a_2 & a_1 & a_0 & \\ \vdots & & \ddots & \ddots & \ddots \\ \vdots & & & \ddots & \ddots & \ddots \end{bmatrix}$$

where $\overline{a_i}$ and a_i for $i \geq 0$ are scalars. Similarly for B and C . As a result the matrix P in equation (3.1.3) also has a $GI/M/1$ type structure. Namely,

$$P = \begin{bmatrix} D_1 & D_0 & & & \\ D_2 & C_1 & C_0 & & \\ D_3 & C_2 & C_1 & C_0 & \\ \vdots & & \ddots & \ddots & \ddots \\ \vdots & & & \ddots & \ddots & \ddots \end{bmatrix} \quad (4.1.12)$$

where

$$\begin{aligned} D_i &= \overline{a_i} + \eta \overline{b_i} + \eta^2 \overline{c_i}, \quad i \geq 0, \\ C_i &= a_i + \eta b_i + \eta^2 c_i, \quad i \geq 0. \end{aligned}$$

are scalars. The submatrix of P without the boundary values is given by

$$Q = \begin{bmatrix} C_1 & C_0 & & & \\ C_2 & C_1 & C_0 & & \\ C_3 & C_2 & C_1 & C_0 & \\ \vdots & & \ddots & \ddots & \ddots \\ \vdots & & & \ddots & \ddots \end{bmatrix}. \quad (4.1.13)$$

Once again we will assume P and Q are irreducible.

Lemma 7 and Theorem 10 by Kijima, [35], was used to determine the radius of convergence for Q but for when Q has an $M/G/1$ type structure. In [35, Remark 2.3], Kijima gives an explanation on how to determine the radius of convergence for a substochastic matrix Q with a $GI/M/1$ type structure, as in (4.1.13).

The essence of Kijima's remark is that he wants to try and construct an $M/G/1$ type model from the $GI/M/1$ type model, apply his $M/G/1$ type results given in Lemma 7 and Theorem 10 and then infer results about the $GI/M/1$ type models. For example, let $\hat{Q} = Q^T/\phi$, where ϕ is chosen so that $\hat{Q}$ is substochastic. This is possible since $C = \sum_{k=0}^{\infty} C_k < \infty$. Let

$$\hat{C}^*(z) = \frac{1}{\phi} \sum_{k=0}^{\infty} C_k^T z^k = \frac{1}{\phi} C^*(z)^T \quad (4.1.14)$$

and let $\chi(z)$, $\mathbf{u}(z)$ and $\mathbf{v}(z)$ be the Perron-Frobenius eigenvalue, left and right eigenvectors of $C^*(z)$, respectively. By (4.1.14) the Perron-Frobenius eigenvalue for $\hat{C}^*(z)$ is given by $\hat{\chi}(z) = \chi(z)/\phi$. The associated left and right eigenvectors are given by $\mathbf{v}(z)$ and $\mathbf{u}(z)$, respectively. The convergence norm of $\hat{Q}$ is determined by considering $\hat{\nu}$ in Lemma 7. It follows that $\hat{\nu} = \nu/\phi$, since $\hat{\chi}(z) = \chi(z)/\phi$. Finally, Q is $\bar{\alpha}$ -transient

with $\bar{\alpha} = \frac{1}{\nu}$. Since the entries of Q are scalars, taking the transpose of Q we have

$$Q^T = \begin{bmatrix} C_1 & C_2 & C_3 & \cdots \\ C_0 & C_1 & C_2 & \cdots \\ & C_0 & C_1 & \cdots \\ & & \ddots & \ddots \\ & & & \ddots & \ddots \end{bmatrix}.$$

Since Q is substochastic, $Q\mathbf{e} \leq \mathbf{e}$, $C < 1$ and $Q^T\mathbf{e} \leq \mathbf{e}$ also. Therefore, we may take $\phi = 1$ giving $\hat{Q} = Q^T$. Now, $\hat{\nu} = \nu$ and $\bar{\alpha} = \frac{1}{\nu} = \frac{1}{\hat{\nu}}$. Thus, all that is required is to find the convergence norm for Q^T . This involves solving (4.1.5) for θ and substituting it into (4.1.4) to obtain ν .

As in Sections 3.1 and 4.1.1 consider βP , the discounted transition matrix at rate β , for $0 < \beta \leq \alpha$ where α is the radius of convergence for P .

$$U_0(\beta) = \beta D_1 + \sum_{i=1}^{\infty} \beta D_0 N_0(\beta) R(\beta)^{i-1} \beta D_{i+1} = u_0(\beta)$$

where

$$R_{01}(\beta) = \beta D_0 N_0(\beta)$$

or

$$R_{01}(\beta) = \beta \frac{D_0}{C_0} R(\beta).$$

This gives

$$U_0(\beta) = u_0(\beta) = \beta D_1 + \beta \frac{D_0}{C_0} \sum_{i=1}^{\infty} R(\beta)^i D_{i+1}$$

where $R(\beta)$ is the minimal, non-negative solution of

$$R(\beta) = \beta \sum_{i=0}^{\infty} C_i R(\beta)^i. \quad (4.1.15)$$

Again, letting $\beta^* = \frac{1}{\eta}$ we have the inequality in (4.1.7). Therefore, η must satisfy the equations (4.1.15) and (4.1.8) or

$$\frac{1}{\eta}D_1 + \frac{1}{\eta} \frac{D_0}{C_0} \sum_{i=1}^{\infty} R\left(\frac{1}{\eta}\right)^i D_{i+1} = 1. \quad (4.1.16)$$

Theorem 23 *Suppose η is in the interval $(0, 1)$ satisfying equations (4.1.16) and (4.1.15). Then the radius of convergence of P given in (4.1.12) is $\alpha = \frac{1}{\eta}$ and P is $1/\eta$ -recurrent.*

Proof: This result follows from Theorem 6(ii), [39].

In order for P to be $1/\eta$ -positive, the conditions in [39, Theorem 7] must hold. This is summarized in the following corollary.

Corollary 24 *When P is α -recurrent and $C = \sum_{i=0}^{\infty} C_i$ is strictly substochastic, necessary and sufficient conditions for P to be $1/\eta$ -positive are,*

1. $\alpha < \bar{\alpha}$, if $R(\alpha) < 1$.
2. $\alpha < \bar{\alpha}$, $\sum_{k=1}^{\infty} kR(\alpha)^{k-1}C_k < \infty$ and $\sum_{k=1}^{\infty} kR(\alpha)^{k-1}D_k < \infty$, if $R(\alpha) > 1$.
3. $\alpha < \bar{\alpha}$, $\sum_{k=1}^{\infty} kC_k < \infty$ and $\sum_{k=1}^{\infty} kD_k < \infty$, if $R(\alpha) = 1$.

Once P is $1/\eta$ -positive and provided the $1/\eta$ -invariant measure and vector $\mathbf{x}$ and $\mathbf{y}$, respectively, exist, then by Theorem 7, $\mathbf{xy} < \infty$. This is stated in the following corollary.

Corollary 25 *If $\mathbf{x}$ and $\mathbf{y}$ are, respectively, an $1/\eta$ -invariant measure and vector of P then $\mathbf{xy}^T < \infty$.*

By [39, Theorem 4] $\mathbf{x}$ is given by,

$$\begin{aligned} x_0 &= w_0 \\ x_k &= w_0 R_{01}(\alpha) R^{k-1}(\alpha), k \geq 1 \end{aligned}$$

where w_0 is the solution to $w_0(I - U_0(\alpha)) = 0$.

In order to determine the $1/\eta$ -invariant vector the factorization given in equation (2.2.24) may be used

$$I - \frac{1}{\eta}P(\eta) = (I - R_U(\alpha))(I - U_D(\alpha))(I - G_L(\alpha)),$$

where $I - R_U(\alpha)$, $I - U_D(\alpha)$ and $I - G_L(\alpha)$ may be found below. We want to solve

$$0 = \left[I - \frac{1}{\eta}P(\eta) \right] \mathbf{y}^T = (I - R_U(\alpha))(I - U_D(\alpha))(I - G_L(\alpha))\mathbf{y}^T.$$

This is done by letting $\mathbf{z}^T = (I - G_L(\alpha))\mathbf{y}^T$ and solving $\mathbf{0}^T = (I - R_U(\alpha))(I - U_D(\alpha))\mathbf{z}^T$. Solving $\mathbf{0}^T = (I - R_U(\alpha))(I - U_D(\alpha))\mathbf{z}^T$ gives,

$$\begin{aligned} \mathbf{0}^T &= (I - R_U(\alpha))(I - U_D(\alpha))\mathbf{z}^T \\ &= \begin{bmatrix} I & -R_{01}(\alpha) & & & \\ & I & -R(\alpha) & & \\ & & I & -R(\alpha) & \\ & & & \ddots & \ddots \\ & & & & \ddots & \ddots \end{bmatrix} \begin{bmatrix} I - U_0(\alpha) & & & & \\ & I - U(\alpha) & & & \\ & & I - U(\alpha) & & \\ & & & \ddots & \\ & & & & \ddots \end{bmatrix} \begin{bmatrix} z_0 \\ z_1 \\ z_2 \\ \vdots \end{bmatrix} \\ &= \begin{bmatrix} I - U_0(\alpha) & -R_{01}(\alpha)(I - U(\alpha)) & & & \\ & I - U(\alpha) & -R(\alpha)(I - U(\alpha)) & & \\ & & I - U(\alpha) & -R(\alpha)(I - U(\alpha)) & \\ & & & \ddots & \ddots \\ & & & & \ddots & \ddots \end{bmatrix} \begin{bmatrix} z_0 \\ z_1 \\ z_2 \\ \vdots \end{bmatrix} \end{aligned}$$

or, equivalently,

$$\begin{aligned} 0 &= (I - U_0(\alpha))z_0 - R_{01}(\alpha)(I - U(\alpha))z_1 \\ 0 &= (I - U(\alpha))z_1 - R(\alpha)(I - U(\alpha))z_2 \\ &\vdots \\ 0 &= (I - U(\alpha))z_{n-1} - R(\alpha)(I - U(\alpha))z_n \quad n \geq 1. \end{aligned}$$

There exists an $z_0 \neq 0$ such that $(I - U_0(\alpha))z_0 = 0$, therefore, $\mathbf{z} = (z_0, 0, 0, \dots)$ is a solution to $\mathbf{0}^T = (I - R_U(\alpha))(I - U_D(\alpha))\mathbf{z}^T$. Now we must solve for $\mathbf{y}$ in terms of $\mathbf{z}$ using $\mathbf{z}^T = (I - G_L(\alpha))\mathbf{y}^T$.

$$\begin{aligned} \mathbf{z}^T &= (I - G_L(\alpha))\mathbf{y}^T \\ \begin{bmatrix} z_0 \\ z_1 \\ z_2 \\ \vdots \end{bmatrix} &= \begin{bmatrix} I & & & & \\ -G_{10}(\alpha) & I & & & \\ -G_{20}(\alpha) & -G_1(\alpha) & I & & \\ -G_{30}(\alpha) & -G_2(\alpha) & -G_1(\alpha) & I & \\ \vdots & \vdots & \ddots & \ddots & \ddots \\ \vdots & \vdots & & \ddots & \ddots \end{bmatrix} \begin{bmatrix} y_0 \\ y_1 \\ y_2 \\ \vdots \end{bmatrix} \\ y_0 &= z_0 \\ y_1 &= G_{10}(\alpha)z_0 \\ y_n &= -G_{n0}(\alpha)z_0 + \sum_{i=1}^{n-1} G_i(\alpha)y_{n-i}, \quad n \geq 1. \end{aligned}$$

The result in (4.1.11) holds provided Theorem 23 and Corollaries 24 and 25 hold.

4.2 Examples

In this section, we will have a look at some examples of the use of the generalized method in the previous section. The first example is just the case when the infinite blocks A, B and C of P' have a repeating, tridiagonal structure as in Chapter 3. In other words, Section 3.1 is a special case of Section 4.1.1 and Section 4.1.2. The second example is an example of the $M/G/1$ type case.

4.2.1 Tridiagonal structure

When A , B and C have a tridiagonal structure it is a special case of when their structure is $M/G/1$ or $GI/M/1$ type. This is the situation that is described in Chapter 3. We will show that equations (4.1.6) and (4.1.9) give the characteristic polynomial (3.1.10) in Chapter 3 which is required to determine the radius of convergence for P . In this situation,

$$\begin{aligned} D_2 &= \bar{a}_2 + \eta \bar{b}_2 + \eta^2 \bar{c}_2 \\ D_1 &= \bar{a}_1 + \eta \bar{b}_1 + \eta^2 \bar{c}_1 \\ D_0 &= \bar{a}_0 + \eta \bar{b}_0 + \eta^2 \bar{c}_0 \\ C_2 &= a_2 + \eta b_2 + \eta^2 c_2 \\ C_1 &= a_1 + \eta b_1 + \eta^2 c_1 \\ C_0 &= a_0 + \eta b_0 + \eta^2 c_0. \end{aligned}$$

From equation (4.1.6) the G measure of βP is the minimal, non-negative solution of

$$G = \beta(C_0 + C_1 G + C_2 G^2)$$

or

$$0 = \beta C_0 + (\beta C_1 - I)G + \beta C_2 G^2.$$

Solving for G we get

$$G = \frac{(1 - \beta C_1) - \sqrt{(\beta C_1 - 1)^2 - 4\beta^2 C_2 C_0}}{2\beta C_2}.$$

Equation (4.1.9) becomes

$$\begin{aligned} 1 &= \beta D_1 + \beta \frac{D_2 D_0}{C_0} G \\ 1 &= \beta D_1 + \frac{D_2 D_0}{2C_2 C_0} ((1 - \beta C_1) - \sqrt{(\beta C_1 - 1)^2 - 4\beta^2 C_2 C_0}) \\ 1 &= U_0(\beta) \end{aligned}$$

as in equation (3.1.9).

4.2.2 Geometric Jumps

Adan *et. al.*, [3], consider two-dimensional Markov chains with geometric jumps. These are Markov chains that do not necessarily only have transitions to nearest neighbours. In one direction there may be transitions that are greater than one unit. Furthermore, these transitions along this direction are geometrically distributed. In Adan *et. al.*, [3], they generalize the compensation method to consider such a situation. In this section we will consider one of the examples mentioned in [3].

An example of a two-dimensional random walk with geometric jumps to the north west may be found in Figure 4.1. There are two stations IO and CP. At the CP station a job has with probability $1 - p$ an exponentially distributed service length with parameter μ_2 and with probability p a service length equal to 0. If a job is served with positive service length at the CP station then a geometrically distributed number of jobs leaves queue I instantaneously and joins queue II. The jobs at queue II wait there until they are processed at the IO station. After a job from queue II receives service with an exponential distribution with parameter μ_1 at the IO station the job leaves the system with probability q or joins queue I with probability $1 - q$. If queue II is empty, the IO station takes a job from an infinite source IS, [3].

This system can be modelled by a two-dimensional Markov process

$$X(t) = (X_1(t), X_2(t))$$

with state space $S = \{(i, j); i, j = 0, 1, 2, \dots\}$ where $X_1(t)$ is the number of jobs in queue I and $X_2(t)$ is the number of jobs in queue II. The transition rate diagram may be found in Figure 4.2.

Figure 4.1: Diagram for Geometric Jump model.

The equilibrium equations are

$$(\mu_1 + \mu_2)\pi_{mn} = \pi_{m,n+1}q\mu_1 + \pi_{m-1,n+1}(1-q)\mu_1 + \sum_{i=0}^{n-1} (1-p)p^i \pi_{m+i+1,n-i-1}\mu_2,$$

for $m > 0, n > 0$

$$(\mu_1 + \mu_2)\pi_{0n} = \pi_{0,n+1}q\mu_1 + \sum_{i=0}^{n-1} p^i \pi_{i+1,n-i-1}\mu_2,$$

for $n > 0$

$$(\mu_1 + \mu_2)\pi_{m0} = \mu_2\pi_{m0} + \pi_{m-1,0}\mu_1 + \pi_{m1}q\mu_1 + \pi_{m-1,1}(1-q)\mu_1,$$

for $m > 0$

$$(\mu_1 + \mu_2)\pi_{00} = \mu_2\pi_{00} + q\mu_1\pi_{01}.$$

Figure 4.2: Transition rate diagram for geometric jump model where $p_i = (1-p)p^i$ for $i = 0, 1, 2, \dots$

The transition matrix has a $GI/M/1$ type structure

$$P' = \begin{bmatrix} B_0 & A_1 & & & \\ B_{-1} & A_0 & A_1 & & \\ B_{-2} & A_{-1} & A_0 & A_1 & \\ B_{-3} & A_{-2} & A_{-1} & A_0 & A_1 \\ \vdots & \vdots & \vdots & \ddots & \ddots \\ \vdots & \vdots & \vdots & & \ddots & \ddots \end{bmatrix}$$

where the infinite blocks have a $M/G/1$ type structure with scalar entries. The

boundary blocks are given by

$$B_0 = \begin{bmatrix} \mu_2 & & & & \\ q\mu_1 & \mu_2 & & & \\ & q\mu_1 & \mu_2 & & \\ & & q\mu_1 & \mu_2 & \\ & & & \ddots & \ddots \\ & & & & \ddots & \ddots \end{bmatrix}$$

and

$$B_{-m} = \begin{pmatrix} 0 & 1 & \dots & \dots & m-1 & m & m+1 & \dots & \dots \\ 0 & 0 & \dots & \dots & 0 & p^{m-1}\mu_2 & 0 & \dots & \dots \\ & 0 & 0 & \dots & \dots & 0 & p^{m-1}\mu_2 & 0 & \dots \\ & & \ddots & \ddots & & & \ddots & \ddots & \ddots \\ & & & \ddots & \ddots & & & \ddots & \ddots \end{pmatrix},$$

for $m = 1, 2, 3, \dots$ and the repeating blocks are given by

$$A_1 = \begin{bmatrix} \mu_1 & & & & \\ (1-q)\mu_1 & 0 & & & \\ & (1-q)\mu_1 & 0 & & \\ & & (1-q)\mu_1 & 0 & \\ & & & \ddots & \ddots \\ & & & & \ddots & \ddots \end{bmatrix},$$

$$A_0 = \begin{bmatrix} 0 & & & & \\ q\mu_1 & 0 & & & \\ & q\mu_1 & 0 & & \\ & & q\mu_1 & 0 & \\ & & & \ddots & \ddots \\ & & & & \ddots & \ddots \end{bmatrix}, \quad A_{-1} = \begin{bmatrix} 0 & p_0\mu_2 & & & \\ & 0 & p_0\mu_2 & & \\ & & 0 & p_0\mu_2 & \\ & & & 0 & p_0\mu_2 \\ & & & & \ddots & \ddots \\ & & & & & \ddots & \ddots \end{bmatrix},$$

and

$$A_{-m} = \begin{pmatrix} 0 & 1 & \cdots & \cdots & m-1 & m & m+1 & \cdots & \cdots \\ 0 & 0 & \cdots & \cdots & 0 & p_{m-1}\mu_2 & 0 & \cdots & \cdots \\ & 0 & 0 & \cdots & \cdots & 0 & p_{m-1}\mu_2 & 0 & \cdots \\ & & \ddots & \ddots & & & \ddots & \ddots & \ddots \\ & & & \ddots & \ddots & & & \ddots & \ddots \end{pmatrix}$$

for $m = 1, 2, 3, \dots$, where we have assumed $\mu_1 + \mu_2 = 1$. The counterpart of equation (3.1.2) for this example is

$$P = P(\eta) = \sum_{i=0}^{\infty} \eta^i A_{1-i}$$

$$= \begin{bmatrix} \mu_1 & \eta^2 p_0 \mu_2 & \eta^3 p_1 \mu_2 & \eta^4 p_2 \mu_2 & \eta^5 p_3 \mu_2 & \cdots \\ (1-q)\mu_1 + \eta q \mu_1 & 0 & \eta^2 p_0 \mu_2 & \eta^3 p_1 \mu_2 & \eta^4 p_2 \mu_2 & \cdots \\ & (1-q)\mu_1 + \eta q \mu_1 & 0 & \eta^2 p_0 \mu_2 & \eta^3 p_1 \mu_2 & \cdots \\ & & (1-q)\mu_1 + \eta q \mu_1 & 0 & \eta^2 p_0 \mu_2 & \cdots \\ & & & (1-q)\mu_1 + \eta q \mu_1 & 0 & \ddots \\ & & & & \ddots & \ddots \end{bmatrix}$$

where

$$C_0 = D_0 = ((1-q)\mu_1 + \eta q \mu_1)$$

$$C_1 = 0, \quad D_1 = \mu_1$$

$$C_2 = D_2 = \eta^2 p_0 \mu_2$$

$$C_3 = D_3 = \eta^3 p_1 \mu_2$$

$$C_j = D_j = \eta^j p_{j-2} \mu_2, \quad j \geq 2.$$

In order to determine the radius of convergence, $\bar{\alpha}$, for Q, P without boundaries, equation (4.1.5) must be solved for θ which is then substituted into equation (4.1.4)

to give ν or $\bar{\alpha} = 1/\nu$. Equation (4.1.5) may be given by,

$$\begin{aligned}
 0 &= \frac{((1-q)\mu_1 + \eta q \mu_1)}{\mu_1 + \mu_2} + \sum_{i=2}^{\infty} (1-i) \frac{\eta^i p_{i-2} \mu_2}{\mu_1 + \mu_2} \theta^i \\
 &= (1-q + \eta q) \mu_1 - \sum_{i=0}^{\infty} (1+i) \eta^{i+2} p_i \mu_2 \theta^{i+2} \\
 &= (1-q + \eta q) \mu_1 - \eta^2 \theta^2 (1-p) \mu_2 \sum_{i=0}^{\infty} (1+i) \eta^i p^i \theta^i \\
 &= (1-q + \eta q) \mu_1 - \eta^2 \theta^2 (1-p) \mu_2 \sum_{i=0}^{\infty} (1+i) (\eta p \theta)^i \\
 \frac{(1-q + \eta q) \mu_1}{\eta^2 \theta^2 (1-p) \mu_2} &= \sum_{i=0}^{\infty} (\eta p \theta)^i + \sum_{i=0}^{\infty} i (\eta p \theta)^i \\
 &= \frac{1}{1 - \eta p \theta} + \frac{\eta p \theta}{(1 - \eta p \theta)^2},
 \end{aligned}$$

for $|\eta p \theta| < 1$, and equation (4.1.4) by,

$$\begin{aligned}
 \nu &= \sum_{i=2}^{\infty} i \frac{\eta^i p_{i-2} \mu_2}{\mu_1 + \mu_2} \theta^{i-1} \\
 \nu &= \sum_{i=0}^{\infty} (i+2) \eta^{i+2} p_i \mu_2 \theta^{i+1} \\
 \nu &= \eta^2 (1-p) \mu_2 \theta \sum_{i=0}^{\infty} (i+2) (\eta p \theta)^i \\
 \frac{\nu}{\eta^2 (1-p) \mu_2 \theta} &= \sum_{i=0}^{\infty} (i+2) (\eta p \theta)^i \\
 &= \sum_{i=0}^{\infty} i (\eta p \theta)^i + 2 \sum_{i=0}^{\infty} (\eta p \theta)^i \\
 &= \eta p \theta \sum_{i=0}^{\infty} i (\eta p \theta)^{i-1} + \frac{2}{1 - \eta p \theta} \\
 \frac{\nu}{\eta^2 (1-p) \mu_2 \theta} &= \frac{\eta p \theta}{(1 - \eta p \theta)^2} + \frac{2}{1 - \eta p \theta},
 \end{aligned}$$

for $|\eta p \theta| < 1$. The radius of convergence for P is given by $\alpha = 1/\eta$ where η satisfies

equations (4.1.6) and (4.1.9) and inequality (4.1.7). Equation (4.1.6) is given by,

$$\begin{aligned}
 G(\beta) &= \beta \sum_{k=0}^{\infty} C_k G^k(\beta) \\
 &= \beta C_0 + \beta \sum_{k=2}^{\infty} C_k G^k(\beta) \\
 &= \beta \frac{(1-q+\eta q)\mu_1}{\mu_1 + \mu_2} + \beta \sum_{k=2}^{\infty} \frac{\eta^k p_{k-2} \mu_2}{\mu_1 + \mu_2} G^k(\beta) \\
 &= \beta \frac{(1-q+\eta q)\mu_1}{\mu_1 + \mu_2} + \beta \sum_{k=0}^{\infty} \eta^{k+2} p_k \frac{\mu_2}{\mu - 1 + \mu_2} G^{k+2}(\beta) \\
 &= \left(\frac{\mu_2}{\mu_1 + \mu_2} \right) \left(\frac{1-q+\eta q}{\eta} \right) + \frac{\eta^2(1-p)\mu_2}{\eta(\mu_1 + \mu_2)} G^2(\beta) \sum_{k=0}^{\infty} \eta^k p^k G^k(\beta)
 \end{aligned}$$

or

$$G(\beta) = \frac{\mu_1(1-q+\eta q)}{\eta} + \eta(1-p)\mu_2 G^2(\beta) \sum_{k=0}^{\infty} (\eta p G(\beta))^k, \quad (4.2.1)$$

and equation (4.1.9) by,

$$\begin{aligned}
 1 &= \frac{1}{\eta} D_1 + \frac{1}{\eta} \frac{D_0}{C_0} \sum_{k=1}^{\infty} D_{k+1} G^k(\beta) \\
 &= \frac{\mu_1}{\eta(\mu_1 + \mu_2)} + \frac{1}{\eta} \sum_{k=1}^{\infty} \eta^{k+1} p_{k-1} \frac{\mu_2}{\mu_1 + \mu_2} G^k(\beta) \\
 &= \frac{\mu_1}{\eta(\mu_1 + \mu_2)} + \frac{\mu_1}{\mu_1 + \mu_2} \sum_{k=0}^{\infty} \eta^{k+1} p_k G^{k+1}(\beta) \\
 &= \frac{\mu_1}{\eta(\mu_1 + \mu_2)} + \frac{\eta \mu_2(1-p)}{\mu_1 + \mu_2} G(\beta) \sum_{k=0}^{\infty} \eta^k p^k G^k(\beta)
 \end{aligned}$$

or

$$1 = \frac{\mu_1}{\eta} + \eta \mu_2(1-p) G(\beta) \sum_{k=0}^{\infty} (\eta p G(\beta))^k. \quad (4.2.2)$$

Multiplying equation (4.2.2) by $G(\beta)$ and subtracting equation (4.2.1) from it we get

$$\frac{G(1/\eta) - 1 + q - \eta q}{\eta} = 0$$

or

$$G(1/\eta) = 1 - q + \eta q. \quad (4.2.3)$$

Note: This expression for the G measure of βP is specifically for the G measure that arises for η that satisfies equations (4.1.6) and (4.1.9), i.e. when $\alpha = 1/\eta$.

Given the relationship between the G measure for βP and η in equation (4.2.3), it may be substituted into either equation (4.2.2) or (4.2.1) giving a series in one unknown variable. A numerical procedure may be needed to determine a range for the η values given various parameter values.

Chapter 5

Determination of η and x for Reducible, Scalar, Tridiagonal P

Consider $A(z) = \sum_{k=0}^{\infty} z^k A_k$ and $A = A(1)$, for

$$P' = \begin{bmatrix} D_1 & D_0 & & & \\ D_2 & A_1 & A_0 & & \\ D_3 & A_2 & A_1 & A_0 & \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}$$

where all entries are finite, square blocks and P' is irreducible. In [49] Neuts discusses the case when A is reducible. More specifically, he considers the case when A is upper triangular and stochastic. Since the R measure for P is the non-negative minimal solution of $R = A(R)$, R is also upper triangular and as a result the eigenvalues of R are the entries, R_{jj} , along the diagonal of R which satisfy the equation

$$R_{jj} = \sum_{k=0}^{\infty} (A_k)_{jj} R_{jj}^k$$

In this chapter, we consider a similar situation. Again P' is irreducible but with a quasi-birth-death structure with infinite blocks, the sum of whose repeating blocks is

reducible.

Chapter 3 is first concerned with determining the character of P , i.e. whether P is α -positive or not, but under the assumption that P is irreducible. This chapter will again be interested in the character of P but under the assumption that it is reducible. In Chapter 3 we went on to determine the tail behaviour of π_{mn} along the level direction, i.e. as $m \rightarrow \infty$. Through examples in this chapter we will illustrate that for the reducible case α -positivity is no longer sufficient for a geometric tail. We will be considering the case when P has an upper tridiagonal structure.

5.1 Reducible Case

Consider a quasi-birth-death process with a countable number of phases in each level, namely a discrete time Markov chain $\{X(n); n \geq 0\}$ on a two dimensional state space $S = \{(i, j); i, j = 0, 1, 2, \dots\}$ with transition probability matrix

$$P' = \begin{bmatrix} B_0 & A_0 & & & \\ C_1 & B & A & & \\ & C & B & A & \\ & & C & B & A \\ & & & \ddots & \ddots & \ddots \end{bmatrix}$$

after partitioning the state space into levels $\mathcal{L}_i = \{(i, j); j = 0, 1, 2, \dots\}$ for $i = 0, 1, 2, \dots$. Assume P' is irreducible. Let the repeating blocks A, B , and C have entries along the diagonal and superdiagonal and be level independent, that is they

have the following upper triangular structure,

$$A = \begin{bmatrix} a_1 & a_0 & & & \\ & a_1 & a_0 & & \\ & & a_1 & a_0 & \\ & & & a_1 & a_0 \\ & & & & \ddots & \ddots \end{bmatrix}, B = \begin{bmatrix} b_1 & b_0 & & & \\ & b_1 & b_0 & & \\ & & b_1 & b_0 & \\ & & & b_1 & b_0 \\ & & & & \ddots & \ddots \end{bmatrix},$$

and

$$C = \begin{bmatrix} c_1 & c_0 & & & \\ & c_1 & c_0 & & \\ & & c_1 & c_0 & \\ & & & c_1 & c_0 \\ & & & & \ddots & \ddots \end{bmatrix}.$$

We would like to find the eigenvalue or radius of convergence for the R measure of P' . For $0 < \eta < 1$ consider the matrix,

$$P = P(\eta) = A + \eta B + \eta^2 C \quad (5.1.1)$$

or

$$P = \begin{bmatrix} \gamma & \lambda & & & \\ & \gamma & \lambda & & \\ & & \gamma & \lambda & \\ & & & \ddots & \ddots \end{bmatrix} \quad (5.1.2)$$

where

$$\gamma = a_1 + \eta b_1 + \eta^2 c_1 \neq 0, \quad \lambda = a_0 + \eta b_0 + \eta^2 c_0 \neq 0,$$

which is substochastic since $0 < \eta < 1$ and *reducible*.

Theorem 15 gives conditions under which R for P' has an eigenvalue of η . In essence, these conditions require us to show, given $0 < \eta < 1$ and vectors $\mathbf{x}$ and $\mathbf{y}$ such that

Condition 5 *a. all states i of $P(\eta)$ are $1/\eta$ -positive and,*

b. $\eta^{-1}\mathbf{x}\mathbf{A}\mathbf{y}^T \neq \eta\mathbf{x}\mathbf{C}\mathbf{y}^T$,

where $1/\eta$ is the radius of convergence for $P(\eta)$.

Since the matrix P is not irreducible the radius of convergence for each series

$$P_{ij}(z) = \sum_{n=0}^{\infty} z^n p_{ij}^n$$

is not necessarily the same. In the following lemma we show that the radius of convergence for each series $P_{ij}(z)$ is in fact the same.

Lemma 9 *The radius of convergence for P given by equation (5.1.2) is equal to $1/\gamma$.*

Proof: To find the radius of convergence of P in equation (5.1.2) we need to find the infimum of all the radii of convergence for each power series, $P_{ij}(z) = \sum_{n=0}^{\infty} (P^n)_{ij} z^n$, for all $i, j \geq 0$. We need to consider powers of P . Taking powers of P we have

$$P^2 = \begin{bmatrix} \gamma^2 & 2\gamma\lambda & \lambda^2 & & \\ & \gamma^2 & 2\gamma\lambda & \lambda^2 & \\ & & \gamma^2 & 2\gamma\lambda & \lambda^2 \\ & & & \ddots & \ddots & \ddots \end{bmatrix},$$

$$P^3 = \begin{bmatrix} \gamma^3 & 3\gamma^2\lambda & 3\gamma\lambda^2 & \lambda^3 & & \\ & \gamma^3 & 3\gamma^2\lambda & 3\gamma\lambda^2 & \lambda^3 & \\ & & \gamma^3 & 3\gamma^2\lambda & 3\gamma\lambda^2 & \lambda^3 \\ & & & \ddots & \ddots & \ddots & \ddots \end{bmatrix},$$

$$P^4 = \begin{bmatrix} \gamma^4 & 4\gamma^3 & 6\gamma^2\lambda^2 & 4\gamma\lambda^3 & \lambda^4 & & \\ & \gamma^4 & 4\gamma^3 & 6\gamma^2\lambda^2 & 4\gamma\lambda^3 & \lambda^4 & \\ & & \gamma^4 & 4\gamma^3 & 6\gamma^2\lambda^2 & 4\gamma\lambda^3 & \lambda^4 \\ & & & \ddots & \ddots & \ddots & \ddots & \ddots \end{bmatrix},$$

and continuing to take powers of P we have that for any integer $n > 0$,

$$P^n = \begin{bmatrix} x_0(n) & x_1(n) & x_2(n) & \cdots & x_n(n) \\ & x_0(n) & x_1(n) & x_2(n) & \cdots & x_n(n) \\ & & \ddots & \ddots & \ddots & & \ddots \\ & & & \ddots & \ddots & \ddots & \ddots & \ddots \end{bmatrix}$$

where,

$$x_k(n) = \binom{n}{k} \gamma^{n-k} \lambda^k \quad (5.1.3)$$

for $k = 0, 1, 2, \dots, n$. Notice that $P_{ij}^n = P_{0,j-i}^n$ for $j \geq i$. So

$$P_{ij}(z) = \sum_{n=1}^{\infty} P_{ij}^n z^n = \sum_{n=j-i}^{\infty} P_{0,j-i}^n z^n$$

which means we only need to consider the polynomials

$$P_{0k}(z) = \sum_{n=k}^{\infty} P_{0k}^n z^n = \sum_{n=k}^{\infty} \binom{n}{k} \gamma^{n-k} \lambda^k z^n = \left(\frac{\lambda}{\gamma}\right)^k \sum_{n=k}^{\infty} \frac{n!}{k!(n-k)!} (\gamma z)^n.$$

To find the radius of convergence of P we need to find the infimum of all the radii of convergence for each power series, $P_{0k}(z)$. For any $k \geq 0$ the radius of convergence for $P_{0k}(z)$ may be found using the ratio test. The ratio test states:

- (a) If $\lim_{n \rightarrow \infty} \left| \frac{a_{n+1}}{a_n} \right| = L < 1$, then the series $\sum_{n=1}^{\infty} a_n$ is absolutely convergent (and therefore convergent).
- (b) If $\lim_{n \rightarrow \infty} \left| \frac{a_{n+1}}{a_n} \right| = L > 1$ or $\lim_{n \rightarrow \infty} \left| \frac{a_{n+1}}{a_n} \right| = \infty$ then the series $\sum_{n=1}^{\infty} a_n$ is divergent.

[55, The Ratio Test (10.32)]. Applying the ratio test to $P_{0k}(z)$ where $a_n = \frac{n!}{k!(n-k)!}(\gamma z)^n$ we have

$$\begin{aligned}
 & \lim_{n \rightarrow \infty} \left| \frac{a_{n+1}}{a_n} \right| \\
 &= \lim_{n \rightarrow \infty} \left| \frac{\frac{(n+1)!}{k!(n+1-k)!}(\gamma z)^{n+1}}{\frac{n!}{k!(n-k)!}(\gamma z)^n} \right| \\
 &= |\gamma z| \lim_{n \rightarrow \infty} \left| \frac{n+1}{n+1-k} \right|, \\
 &= |\gamma z|
 \end{aligned}$$

which implies $P_{0k}(z)$ is convergent for z satisfying

$$|z| < \frac{1}{\gamma}.$$

So the radius of convergence for $P_{0k}(z)$, for any $k \geq 0$, is $1/\gamma$. Since all the radii of convergence are equal, it is sufficient to find it for one k , that is for one power series $P_{0k}(z)$. ■

Note: According to Cooke, [17], γ is the eigenvalue of P .

We will first show that Condition 5a. is satisfied.

Theorem 26 *Every state i of P given in (5.1.2) is $1/\eta$ -positive with $\eta = \gamma$.*

Before proving Theorem 26, a few properties of interest will be shown.

Properties 1 (a) $u_0(\beta) < 1$ if and only if $N_0(\beta) < \infty$

(b) If $\beta < \alpha$ then $N_0(\beta) < \infty$ and if $\beta > \alpha$ then $N_0(\beta) = \infty$.

(c) $u_0(\beta)$ is nondecreasing.

Proof:

(a) Suppose $u_0(\beta) < 1$. Since $u_0(\beta)$ is the maximal eigenvalue of $U_0(\beta)$, there exists a nonzero eigenvector $\mathbf{y}$ such that $\mathbf{y}U_0(\beta) = u_0(\beta)\mathbf{y}$. Consider

$$\mathbf{y}N_0(\beta) = \mathbf{y} \sum_{n=0}^{\infty} U_0^n(\beta) = \sum_{n=0}^{\infty} \mathbf{y}U_0^n(\beta) = \sum_{n=0}^{\infty} u_0^n(\beta)\mathbf{y} = \frac{1}{1 - u_0(\beta)}\mathbf{y} = n_0(\beta)\mathbf{y} < \infty.$$

Therefore, $N_0(\beta) < \infty$. The fact that $N_0(\beta) = \sum_{n=0}^{\infty} U_0^n(\beta)$ and the relation between $u_0(\beta)$ and $n_0(\beta)$ in equation (2.2.20) and the assumption were used.

Conversely, let $N_0(\beta) < \infty$. Since $\sum_{n=0}^{\infty} U_0^n(\beta) = N_0(\beta) < \infty$, $N_0(\beta)$ may be written as, $N_0(\beta) = (I - U_0(\beta))^{-1}$. Let $\mathbf{x}$ be the eigenvector of $N_0(\beta)$ corresponding to the maximum eigenvalue $n_0(\beta)$ of $N_0(\beta)$, i.e.

$$n_0(\beta)\mathbf{x} = \mathbf{x}N_0(\beta) = \mathbf{x}(I - U_0(\beta))^{-1}$$

or

$$\mathbf{x}(I - U_0(\beta)) = \frac{1}{n_0(\beta)}\mathbf{x}$$

or

$$\left(1 - \frac{1}{n_0(\beta)}\right)\mathbf{x} = \mathbf{x}U_0(\beta).$$

Since $N_0(\beta) < \infty$, $n_0(\beta) < \infty$. Therefore, $u_0(\beta) = 1 - \frac{1}{n_0(\beta)} < 1$.

(b) The definition of the radius of convergence for P is

$$\alpha = \sup \left\{ z; \sum_{n=0}^{\infty} z^n P^n < \infty \right\}$$

If $\beta < \alpha$ then $\sum_{n=0}^{\infty} \beta^n P^n < \infty$, which implies $N_0(\beta) < \infty$ since $N_0(\beta)$ is the $(0, 0)^{th}$ block of $\widehat{\beta P} = \sum_{n=0}^{\infty} \beta^n P^n < \infty$

Since α is the supremum such that $\sum_{n=0}^{\infty} \alpha^n P^n < \infty$, for any $\beta > \alpha$, $\widehat{\beta P} = \sum_{n=0}^{\infty} \beta^n P^n = \infty$. In our particular situation, since the radii of convergence are equal for all $P_{ij}(z) = \sum_{n=0}^{\infty} z^n p_{ij}^n$ for $j \geq i$, $\widehat{\beta P} = \infty$ implies $N_0(\beta) = \infty$.

(c) Take $0 < \beta_1 \leq \beta_2$. Then,

$$\widehat{\beta_1 P} = \sum_{n=0}^{\infty} \beta_1^n P^n \leq \sum_{n=0}^{\infty} \beta_2^n P^n = \widehat{\beta_2 P}$$

which means each element of $\widehat{\beta_1 P}$ is less than or equal to the corresponding element in $\widehat{\beta_2 P}$, i.e.

$$(\widehat{\beta_1 P})_{ij} \leq (\widehat{\beta_2 P})_{ij}, \quad \text{for } i, j \geq 0.$$

Since $N_0(\beta)$ is the $(0, 0)^{th}$ block of $\widehat{\beta P}$, $N_0(\beta_1) \leq N_0(\beta_2)$, as a result $\det N_0(\beta_1) \leq \det N_0(\beta_2)$. So the maximal eigenvalue of $N_0(\beta_1)$ is less than or equal to the maximal eigenvalue of $N_0(\beta_2)$, i.e. $n_0(\beta_1) \leq n_0(\beta_2)$. Using relation $u_0(\beta_1) \leq u_0(\beta_2)$. ■

Now we shall prove Theorem 26

Proof: In order to determine whether all states i of P are $1/\eta$ -positive we first need to determine whether or not they are $1/\eta$ -recurrent. In order to do this we use Theorem 6. In our situation since p_{ii}^n is the same for all i Lemma 4 gives more than the classification for state i but the classification for all states $i \geq 0$. Thus, we have the classification of the matrix P . Consider one state, say $i = 0$. So,

$$P_{00}(\beta) = \sum_{n=0}^{\infty} \beta^n p_{00}^n.$$

Notice that $P_{00}(\beta)$ equals the $(0,0)^{th}$ entry of the fundamental matrix of βP , i.e. $\widehat{\beta P} = \sum_{n=0}^{\infty} (\beta P)^n$. Or, $P_{00}(\beta) = N_0(\beta)$. Since the fundamental matrix is invariant under censoring, Lemma 6, we may first censor βP to level zero to obtain $U_0(\beta)$ and then take the fundamental matrix of $U_0(\beta)$ which in turn will be equal to $N_0(\beta)$. That is,

$$\sum_{n=0}^{\infty} \beta^n p_{00}^n = P_{00}(\beta) = N_0(\beta) = \sum_{n=0}^{\infty} U_0^n(\beta).$$

In our situation, $p_{00}^n = \gamma^n$ and $U_0(\beta) = \beta\gamma$.

In order to determine the classification of P we need to determine whether

$$P_{00}(\beta) = \sum_{n=0}^{\infty} \beta^n p_{00}^n = \infty \text{ or } < \infty.$$

According to Property 1(a), $N_0(\beta) = \infty$ if and only if $u_0(\beta) = 1$. And according to equation (2.2.20), $u_0(\beta) = 1$ if and only if $n_0(\beta) = \infty$. Thus determining for what values of β , $n_0(\beta) = \infty$, equivalently when $u_0(\beta) = 1$, will give us values of β when $N_0(\beta) = \infty$ and in turn when P is β -recurrent. If $1 - u_0(\beta) = 0$ can be solved for β and if there is one solution, call this solution β^* . If there are multiple solutions, take the smallest and call it β^* . The reason for this is if there are two distinct values, β_1 and β_2 such that $u_0(\beta_i) = 1$ and $\beta_1 < \beta_2$ then it is not true that in the region $0 \leq \beta < \beta_2$, $u_0(\beta) < 1$, for all β . Actually, $u_0(\beta) = 1$ for all $\beta_1 \leq \beta \leq \beta_2$, since $u_0(\beta)$ is monotone nondecreasing. However, it is not the case in the region $0 \leq \beta < \beta_1$.

Notice that for any β , $0 \leq \beta < \beta^*$, $u_0(\beta) < 1$ implies $n_0(\beta) < \infty$ which implies $N_0(\beta) < \infty$. Also, β^* is the smallest β such that $u_0(\beta^*) = 1$ implies $n_0(\beta^*) = \infty$ which implies $N_0(\beta^*) = \infty$. So, β^* may be thought of as the supremum of all β values such that $N_0(\beta) = \sum_{n=0}^{\infty} U_0^n(\beta) < \infty$ which is the definition of the radius of convergence for $N_0(\beta)$ and in turn for P since each $P_{ij}(z)$ has the same radius of convergence. So, $\beta^* = \alpha$, $N_0(\alpha) = \infty$ and thus P is α -recurrent. In our particular example $u_0(\beta) = U_0(\beta) = \gamma\beta$ which implies $\beta^* = 1/\gamma = \alpha$ and $N_0(\alpha) = \sum_{n=0}^{\infty} \alpha^n \gamma^n = \infty$.

In order to determine whether or not P is α -positive or α -null we need to appeal to the definition in Karlin and Taylor, [33, p.85]. In our situation, since P is reducible not all states are part of the same class. However, p_{ii}^n is equal for all i . So determining

$$\pi_i = \lim_{n \rightarrow \infty} (\alpha P)_{ii}^n = \lim_{n \rightarrow \infty} \alpha^n p_{ii}^n$$

for one state i will give the classification for each state and hence the entire state space. In our situation $p_{ii}^n = \gamma^n$ and $\alpha = 1/\gamma$ giving $\pi_i = 1 > 0$. Thus, by Lemma 5 P is α -positive. ■

Note: If we show that $\pi_i = \lim_{n \rightarrow \infty} \alpha^n p_{ii}^n > 0$ then the series $\sum_{n=0}^{\infty} \alpha^n p_{ii}^n$ cannot possibly converge by a result from elementary calculus. Therefore, $\sum_{n=0}^{\infty} \alpha^n p_{ii}^n = \infty$ implies i is α -recurrent and by the definition of α -positive, i is α -positive. So, just showing that $\pi_i > 0$ leads to a more efficient proof of Theorem 26.

We can directly try to calculate the left $1/\gamma$ -invariant measure $\mathbf{x}$ and right $1/\gamma$ -invariant vector $\mathbf{y}$ for P in (5.1.2). Multiplying out, $\mathbf{x}P = \gamma\mathbf{x}$ we get

$$\begin{aligned} \gamma x_0 &= \gamma x_0 \\ \lambda x_0 + \gamma x_1 &= \gamma x_1 \\ \lambda x_1 + \gamma x_2 &= \gamma x_2 \\ &\vdots \\ \lambda x_{n-1} + \gamma x_n &= \gamma x_n \\ &\vdots \end{aligned}$$

From the first equation we have that x_0 is arbitrary but from the second equation we have that $x_0 = 0$ and x_1 is arbitrary. The third equation gives that $x_1 = 0$ and x_2 is arbitrary. By induction we have that $\mathbf{x} = \mathbf{0}$. So, by our definition there exists no $1/\gamma$ -invariant measure.

In order to determine the $1/\gamma$ -invariant vector $\mathbf{y}$ we multiply out $P\mathbf{y}^T = \mathbf{y}^T$ to give

$$\begin{aligned} \gamma y_0 + \lambda y_1 &= \gamma y_0 \\ \gamma y_1 + \lambda y_2 &= \gamma y_1 \\ &\vdots \\ \gamma y_n + \lambda y_{n+1} &= \gamma y_n \\ &\vdots \end{aligned}$$

From the first equation we may take y_0 to be arbitrary, then $y_1 = 0$. From the remaining equations we conclude that $y_n = 0$ for $n \geq 1$. Therefore, the $1/\gamma$ -invariant vector is $\mathbf{y} = (1, 0, 0, \dots)^T$.

The case when P is lower triangular

$$P = \begin{bmatrix} \gamma & & & & \\ \mu & \gamma & & & \\ & \mu & \gamma & & \\ & & \mu & \gamma & \\ & & & \ddots & \ddots \end{bmatrix} \quad (5.1.4)$$

may also be considered. Since the radius of convergence of a matrix and its transpose are the same, [54, p.205] we may consider the transpose of P in (5.1.4), P^T , then Theorem 26 and its proof hold for P^T . Therefore, P in (5.1.4) has a radius of convergence equal to $1/\eta = 1/\gamma$ and each i of P is $1/\eta$ -positive. The $1/\eta$ -invariant measure and vector for P in (5.1.4) are interchanged with those for P in equation (5.1.2), [54, p.205]. That is, $\mathbf{x} = (1, 0, \dots)$ and $\mathbf{y} = \mathbf{0}$ for P in (5.1.4). In both cases, if P is reducible, the α -positivity will not lead to a geometric tail claimed in Theorem 18 for the irreducible case.

5.2 Examples

In this section two examples are considered. The first is an $M/M/1$ queue with a gated random order of service, studied by Resing and Rietman, [52], and the second an $M/M/1$ priority queue considered by Miller, [46]. The result in Section 5.1 is used to show that the procedure used to determine the radius of convergence for P in (5.1.1) also gives the eigenvalue for the R measure for the model considered.

5.2.1 Gated Random Order of Service

The model considered by Resing and Rietman in [52], as the title suggests, is an $M/M/1$ queue with gated random order of service. In this service discipline customers first gather in an unordered waiting room before being put into random order in the service queue once the service queue becomes empty. Customers arrive according to a Poisson process with rate λ_1 at a single server system. The service times of the customers have an exponential distribution with parameter μ . Assume $\rho = \lambda_1/\mu < 1$. Customers in the waiting room must remain there until the service queue becomes empty at which time they are instantaneously transferred to the service queue and put in random order. In a sense, customers in the waiting room remain behind a gate until the service queue is empty before being randomly placed in the service queue. Hence, the name of the service discipline, gated random order of service. Whatever order the customers are placed in the service queue determines the order in which they are served. If there are no customers in the waiting room the moment the service queue becomes empty, the server waits until a customer arrives to the waiting room then transfers them directly to the service queue where they are immediately served. From here we see that the service queue cannot be empty unless the waiting room is empty. Let $X_1(t)$ be the number of customers in the waiting room at time t and $X_2(t)$

From Figure 5.1 the equilibrium equations are given by

$$\begin{aligned}(\lambda_1 + \mu)\pi_{00} &= \mu\pi_{01} + \mu\pi_{00}, \\(\lambda_1 + \mu)\pi_{01} &= \lambda_1\pi_{00} + \mu\pi_{11} + \mu\pi_{02}, \\(\lambda_1 + \mu)\pi_{0n} &= \mu\pi_{n1} + \mu\pi_{0,n+1}, \quad n \geq 2, \\(\lambda_1 + \mu)\pi_{kn} &= \lambda_1\pi_{k-1,n} + \mu\pi_{k,n+1}, \quad k, n \geq 1.\end{aligned}$$

The transition matrix has the following form

$$P' = \begin{bmatrix} B_0 & A_0 & & & \\ C_1 & B & A & & \\ \vdots & & \ddots & \ddots & \\ C_n & & & B & A \\ \vdots & & & & \ddots & \ddots \end{bmatrix}$$

where

$$B_0 = \frac{1}{\lambda_1 + \mu} \begin{matrix} & \begin{matrix} 0 & 1 & 2 & 3 & \dots & \dots \end{matrix} \\ \begin{matrix} 0 \\ 1 \\ 2 \\ \vdots \\ \vdots \end{matrix} & \begin{pmatrix} \mu & \lambda_1 & & & \\ \mu & 0 & 0 & & \\ & \mu & 0 & 0 & \\ & & \ddots & \ddots & \ddots \\ & & & \ddots & \ddots & \ddots \end{pmatrix} \end{matrix},$$

$$A_0 = \frac{1}{\lambda_1 + \mu} \begin{matrix} & \begin{matrix} 1 & 2 & 3 & 4 & \dots & \dots \end{matrix} \\ \begin{matrix} 0 \\ 1 \\ 2 \\ \vdots \\ \vdots \end{matrix} & \begin{pmatrix} 0 & & & & \\ \lambda_1 & 0 & & & \\ & \lambda_1 & 0 & & \\ & & \ddots & \ddots & \\ & & & \ddots & \ddots \end{pmatrix} \end{matrix},$$

$$C_1 = \frac{1}{\lambda_1 + \mu} \begin{matrix} & & & 0 & 1 & 2 & 3 & \dots & \dots \\ & & & & & & & & \\ 1 & \left(\begin{array}{cccccccc} 0 & \mu & & & & & & \\ 0 & 0 & 0 & & & & & \\ 0 & & 0 & 0 & & & & \\ \vdots & & & \ddots & \ddots & \ddots & & \\ \vdots & & & & \ddots & \ddots & \ddots & \end{array} \right) \end{matrix},$$

$$C_2 = \frac{1}{\lambda_1 + \mu} \begin{matrix} & & & 0 & 1 & 2 & 3 & \dots & \dots & \dots \\ & & & & & & & & & \\ 1 & \left(\begin{array}{cccccccc} 0 & 0 & \mu & & & & & \\ 0 & 0 & 0 & 0 & & & & \\ 0 & & 0 & 0 & 0 & & & \\ \vdots & & & \ddots & \ddots & \ddots & & \\ \vdots & & & & \ddots & \ddots & \ddots & \end{array} \right) \end{matrix},$$

$$C_n = \frac{1}{\lambda_1 + \mu} \begin{matrix} & & & 0 & 1 & 2 & \dots & n-1 & n & \dots & \dots \\ & & & & & & & & & & \\ 1 & \left(\begin{array}{cccccccc} 0 & 0 & 0 & \dots & 0 & \mu & \dots & \dots \\ 0 & 0 & 0 & 0 & \dots & 0 & 0 & \dots \\ \vdots & & \ddots & \ddots & \ddots & & \ddots & \ddots \\ \vdots & & & \ddots & \ddots & \ddots & \ddots & \ddots \end{array} \right) \end{matrix}, \text{ for } n \geq 1,$$

$$B = \frac{1}{\lambda_1 + \mu} \begin{matrix} & & & 1 & 2 & 3 & \dots & \dots \\ & & & & & & & \\ 1 & \left(\begin{array}{cccc} 0 & & & \\ \mu & 0 & & \\ \mu & 0 & & \\ \vdots & & \ddots & \ddots \\ \vdots & & & \ddots & \ddots \end{array} \right) \end{matrix},$$

and

$$A = \frac{1}{\lambda_1 + \mu} \begin{matrix} & \begin{matrix} 1 & 2 & 3 & \dots & \dots \end{matrix} \\ \begin{matrix} 1 \\ 2 \\ 3 \\ \vdots \\ \vdots \end{matrix} & \begin{pmatrix} \lambda_1 & & & & \\ 0 & \lambda_1 & & & \\ & 0 & \lambda_1 & & \\ & & \ddots & \ddots & \\ & & & \ddots & \ddots \end{pmatrix} \end{matrix}.$$

Since we are interested in the R measure for P' the boundary blocks of P' do not play any role. We need to consider the matrix P in equation (5.1.1). Expanding equation (5.1.1) we see that P has the form in (5.1.4), where

$$\gamma = \frac{\rho}{\rho + 1}, \quad \mu = \frac{\eta}{\rho + 1}$$

or

$$\begin{aligned} P(\eta) &= A + \eta B + \eta^2 C \\ &= \frac{1}{\rho + 1} \begin{bmatrix} \rho & & & & \\ \eta & \rho & & & \\ & \eta & \rho & & \\ & & \ddots & \ddots & \\ & & & \ddots & \ddots \end{bmatrix}. \end{aligned}$$

Since P satisfies the conditions of Theorem 26, every state i of P is $1/\eta$ -positive where

$$\eta = \gamma = \frac{\rho}{\rho + 1}.$$

Resing and Rietman, [52], used the compensation method to determine the stationary distribution for P' . When using the compensation method a starting pair (α, β) is determined. The starting pair determined in [52] was (α_1, β_1) where $\beta_1 = 0$ and $\alpha_1 = \rho/(\rho + 1)$ which is equal to η .

It can be shown that along the level direction the joint stationary distribution π_{kn} does not have a geometric tail. The joint stationary distribution is given by

$$\pi_{kn} = \sum_{m=1}^{\infty} c_m \alpha_m^k \beta_m^{n-1}$$

where

$$\alpha_m = \frac{\rho(1-\rho^m)}{1-\rho^{m+1}}, \quad \beta_m = \frac{\rho(1-\rho^{m-1})}{1-\rho^m}$$

and

$$c_m = \frac{(1-\rho)^3 \rho^m}{(1-\rho^m)(1-\rho^{m+1})}, \quad m \geq 1.$$

For any $m \geq 1$ the following statements are equivalent,

$$\begin{aligned} 0 &\leq (\rho-1)^2 \\ 0 &\leq \rho^2 - 2\rho + 1 \\ 2\rho &\leq \rho^2 + 1 \\ 2\rho^{m+1} &\leq \rho^m(\rho^2 + 1) \\ -2\rho^{m+1} &\geq -\rho^m(\rho^2 + 1) \\ 1 - 2\rho^{m+1} + \rho^{2m+2} &\geq 1 - \rho^{m+2} - \rho^m + \rho^{2m+2} \\ (1 - \rho^{m+1})(1 - \rho^{m+1}) &\geq (1 - \rho^m)(1 - \rho^{m+2}) \\ \alpha_{m+1} = \frac{\rho(1-\rho^{m+1})}{1-\rho^{m+2}} &\geq \frac{\rho(1-\rho^m)}{1-\rho^{m+1}} = \alpha_m \end{aligned}$$

implying that α_m is an increasing sequence. Since $\rho < 1$, $\frac{1-\rho^m}{1-\rho^{m+1}} < 1$, which implies $\alpha_m < \rho$ for all m . Therefore,

$$\lim_{m \rightarrow \infty} \alpha_m = \lim_{m \rightarrow \infty} \frac{\rho(1-\rho^m)}{1-\rho^{m+1}} = \rho.$$

Since $\beta_m = \alpha_{m-1}$, $\beta_m \uparrow \rho$. Also, since $\alpha_m \uparrow \rho$ and $\beta_m \uparrow \rho$, for all $m \geq 1$, $\alpha_m/\rho < 1$ and $\beta_m/\rho < 1$. This implies

$$\left(\frac{\alpha_m}{\rho}\right)^k < 1 \text{ and } \left(\frac{\beta_m}{\rho}\right)^{n-1} < 1$$

and as a result

$$\left(\frac{\alpha_m}{\rho}\right)^k \downarrow 0, \text{ as } k \rightarrow \infty \text{ and } \left(\frac{\beta_m}{\rho}\right)^{n-1} \downarrow 0, \text{ as } n \rightarrow \infty.$$

Next we show that $\sum_{m=1}^{\infty} c_m < \infty$. Since $0 < \rho < 1$, and

$$(1 - \rho^m)(1 - \rho^{m+1}) \geq (1 - \rho^m)^2 \geq (1 - \rho^2)^2,$$

$c_m \leq (1 - \rho)/(1 + \rho)^2 \rho^m$. Therefore,

$$\begin{aligned} 0 &\leq \sum_{m=1}^{\infty} c_m < \frac{(1 - \rho)}{(1 + \rho)^2} \sum_{m=1}^{\infty} \rho^m \\ &= \frac{(1 - \rho)}{(1 + \rho)^2} \left(\sum_{m=0}^{\infty} \rho^m - 1 \right) \\ &= \frac{\rho}{(1 + \rho)^2} < \infty. \end{aligned}$$

Therefore, for all $\epsilon > 0$ there exists N_0 such that

$$\sum_{m=N_0}^{\infty} c_m < \frac{\epsilon}{2}.$$

We want to show that π_{kn} does not have a geometric tail along the level direction or as $k \rightarrow \infty$

$$\frac{\pi_{kn}}{\rho^k \beta_m^{n-1}} \asymp c,$$

where $c \neq 0$ is a constant. For any $k, n \geq 1$

$$\frac{\pi_{kn}}{\rho^k \beta_m^{n-1}} = \sum_{m=1}^{\infty} c_m \left(\frac{\alpha_m}{\rho}\right)^k \leq \sum_{m=1}^{\infty} c_m < \infty.$$

Since $(\alpha_m/\rho)^k \downarrow 0$ as $k \rightarrow \infty$, there exists k such that

$$\left(\frac{\alpha_m}{\rho}\right)^k < \frac{\epsilon/2}{\sum_{m=1}^{N_0-1} c_m}.$$

Now we have

$$\begin{aligned}
\frac{\pi_{kn}}{\rho^k \beta_m^{n-1}} &= \sum_{m=1}^{\infty} c_m \left(\frac{\alpha_m}{\rho} \right)^k \\
&= \sum_{m=1}^{N_0-1} c_m \left(\frac{\alpha_m}{\rho} \right)^k + \sum_{m=N_0}^{\infty} c_m \left(\frac{\alpha_m}{\rho} \right)^k \\
&< \frac{\epsilon}{2} + \sum_{m=N_0}^{\infty} c_m < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon.
\end{aligned}$$

Therefore,

$$\frac{\pi_{kn}}{\rho^k \beta_m^{n-1}} \xrightarrow{k \rightarrow \infty} 0$$

or π_{kn} does not have a geometric tail along the level direction. From here we see that π_{kn} decays faster than geometrically along the level direction. As to how much faster, this cannot be determined from the above analysis. As a result, the α -positivity condition will not lead to geometric tail asymptotics in the case of a reducible P .

5.2.2 Priority Queue

This is the same priority queueing model as considered in Section 3.2.3. where $X_2(t)$, the number of Class 2 customers, is taken as the level variable. Instead, here we take $X_1(t)$, the number of Class 1 customers, as the level variable. The equilibrium equations are given by

$$\begin{aligned}
(\lambda_1 + \lambda_2 + \mu_1 + \mu_2)\pi_{ij} &= \mu_2\pi_{ij} + \lambda_1\pi_{i-1,j} + \mu_1\pi_{i+1,j} + \lambda_2\pi_{i,j-1} \\
(\lambda_1 + \lambda_2 + \mu_1 + \mu_2)\pi_{i0} &= \mu_2\pi_{i0} + \mu_1\pi_{i+1,0} + \lambda_1\pi_{i-1,0} \\
(\lambda_1 + \lambda_2 + \mu_1 + \mu_2)\pi_{0j} &= \mu_1\pi_{0j} + \mu_2\pi_{0,j+1} + \lambda_2\pi_{0,j-1} + \mu_1\pi_{1j} \\
(\lambda_1 + \lambda_2 + \mu_1 + \mu_2)\pi_{00} &= (\mu_1 + \mu_2)\pi_{00} + \mu_1\pi_{10} + \mu_2\pi_{01}
\end{aligned}$$

according to the transition diagram in Figure 3.4. The infinite blocks of the transition matrix are given by

$$B_0 = \theta \begin{bmatrix} \mu_1 + \mu_2 & \lambda_2 & & & \\ & \mu_2 & \mu_1 & \lambda_2 & \\ & & \mu_2 & \mu_1 & \lambda_2 \\ & & & \mu_2 & \mu_1 & \lambda_2 \\ & & & & \ddots & \ddots & \ddots \end{bmatrix},$$

$$\begin{aligned} A_0 &= \theta \begin{bmatrix} \lambda_1 & & & & \\ & \lambda_1 & & & \\ & & \lambda_1 & & \\ & & & \ddots & \\ & & & & \ddots \end{bmatrix}, & C_1 &= \theta \begin{bmatrix} \mu_1 & & & & \\ & \mu_1 & & & \\ & & \mu_1 & & \\ & & & \ddots & \\ & & & & \ddots \end{bmatrix}, \\ A &= \theta \begin{bmatrix} \lambda_1 & & & & \\ & \lambda_1 & & & \\ & & \lambda_1 & & \\ & & & \ddots & \\ & & & & \ddots \end{bmatrix}, & C &= \theta \begin{bmatrix} \mu_1 & & & & \\ & \mu_1 & & & \\ & & \mu_1 & & \\ & & & \ddots & \\ & & & & \ddots \end{bmatrix}, \end{aligned}$$

and

$$B = \theta \begin{bmatrix} \mu_2 & \lambda_2 & & & \\ & \mu_2 & \lambda_2 & & \\ & & \mu_2 & \lambda_2 & \\ & & & \mu_2 & \lambda_2 \\ & & & & \ddots & \ddots \end{bmatrix},$$

where $\theta = (\lambda_1 + \lambda_2 + \mu_1 + \mu_2)^{-1}$.

The matrix P given in equation (5.1.1) needs to be considered. Expanding equation (5.1.1) we see that P has the form given in (5.1.2) where

$$\gamma = \theta(\lambda_1 + \eta\mu_2 + \eta^2\mu_1), \quad \lambda = \eta\theta\lambda_2$$

or

$$P = \theta \begin{bmatrix} \lambda_1 + \eta\mu_2 + \eta^2\mu_1 & \eta\lambda_2 & & & \\ 0 & \lambda_1 + \eta\mu_2 + \eta^2\mu_1 & \eta\lambda_2 & & \\ \vdots & 0 & \lambda_1 + \eta\mu_2 + \eta^2\mu_1 & \eta\lambda_2 & \\ \vdots & \vdots & \ddots & \ddots & \ddots \\ \vdots & \vdots & & \ddots & \ddots \end{bmatrix}.$$

Since P satisfies the conditions of Theorem 26, every state i of P is $1/\eta$ -positive where

$$\eta = \gamma = \theta(\lambda_1 + \eta\mu_2 + \eta^2\mu_1).$$

Solving for η reduces to solving the quadratic equation

$$\mu_1\eta^2 - (\lambda_1 + \lambda_2 + \mu_1)\eta + \lambda_1 = 0. \quad (5.2.1)$$

Using the quadratic formula we have

$$\eta = \frac{(\lambda_1 + \lambda_2 + \mu_1) \pm \sqrt{(\lambda_1 + \lambda_2 + \mu_1)^2 - 4\mu_1\lambda_1}}{2\mu_1}.$$

Since we want η to lie within the interval $(0, 1)$,

$$\eta = \frac{(\lambda_1 + \lambda_2 + \mu_1) - \sqrt{(\lambda_1 + \lambda_2 + \mu_1)^2 - 4\mu_1\lambda_1}}{2\mu_1}.$$

Miller, [46], actually determined the R measure for P' and found it to be upper triangular with the following pattern

$$R = \begin{bmatrix} r_0 & r_1 & r_2 & \cdots & \\ & r_0 & r_1 & r_2 & \cdots \\ & & r_0 & r_1 & r_2 & \cdots \\ & & & \ddots & & \\ & & & & \ddots & \end{bmatrix},$$

where

$$r_0 = \frac{(\lambda_1 + \lambda_2 + \mu_1) - \sqrt{(\lambda_1 + \lambda_2 + \mu_1)^2 - 4\mu_1\lambda_1}}{2\mu_1},$$

which is equal to η . By direct calculations the $1/r_0$ -invariant vector is $\mathbf{y} = \frac{r_0}{\sum_{i=0}^{\infty} r_i} \mathbf{e}$ and it turns out that the $1/r_0$ -invariant measure does not exist, i.e. $\mathbf{x} = \mathbf{0}$. However, it is unlikely that π_{mn} has a geometric tail along the level direction.

Chapter 6

Application: Two-Stage Inventory-Queue Model

In this chapter, a two-stage inventory-queue model is studied. Not only is such a model itself basic and important, but also it may serve as a building block in approximating many complex manufacturing systems and supply chains. We are interested in the performance of the system, particularly the starvation probability at Station 2. By wisely selecting state variables, a continuous-time two-dimensional Markov chain is defined. We show that the tail asymptotics of the stationary joint probability distribution of the Markov chain is geometric along certain directions. A modification of this result is used to analyze the asymptotics of the starvation probability. The method presented in Chapter 3 for tail analysis is used here.

6.1 Introduction

A contemporary manufacturing system or supply chain can be represented by a network of processes and storages, interconnected according to the structure of the system. A process can be a machine or a group of machines, a factory, or a distribution center among other things. The storage follows the process and stations the outputs from the process. It can be a material buffer, warehouse, or simply a counter to track the output quantity. In general, a node in this network can have one of a number of possible configurations: (1) one process, (2) one storage, (3) one process and one storage in front of the process serving as an input buffer, (4) one process and one storage following the process an output buffer, and (5) one process and two storages in front of the process serving as an input buffer and following the process serving as an output buffer. We consider configuration (4) and call such a two-component node an inventory-queue.

In an inventory-queue, demands arrive at the output storage following a specific, (stochastic), arrival process. A demand will be satisfied by a unit from the inventory if it is available or when it becomes available. At the same time, a job is created for the process which will be released, (perhaps together with other jobs), to the process according to the inventory control policy used at the storage. The released job(s) joins a job queue in front of the process just like in a regular queue. However, different from the standard queues, a job can only be processed when the required materials/components as well as the process are available. Thus to completely define an inventory-queue, we also need to specify the material supply in addition to the service process, the number of servers, the demand arrival process, and the inventory control policy. It is clear that this inventory-queue modeling formulation can more closely represent manufacturing and supply chain systems than either the standard

queueing model formulation or the standard inventory model formulation. It can be used to study simultaneously the queueing phenomenon as well as the inventory performance under a specific control policy. While similar model structures have been proposed and analyzed for production and inventory systems for some time, see, for example, in Buzacott and Shanthikumar's book [14], (primarily from an inventory control perspective), the use of "inventory-queue" and the formalism first appeared in Ettli *et. al* [18] and later developed in Liu [42] and Liu, Liu, and Yao [43].

In this chapter, we will study a two-stage inventory-queue model with station 1 in the upstream and station 2 in the downstream. Each station is a single-server inventory-queue. In this two-stage serial network, the external demands arrive at the output storage, or station 2, of station 2 one at a time. If there are stocks on hand, an arriving demand is satisfied and the inventory level is reduced by 1. Otherwise, the demand is back-ordered. We assume that a base-stock control policy is used at both stations, which works as follows: whenever the stock level is below a target level, i.e. the base-stock level, an order or job is created so that the stock level will return to the base-stock level. From a queueing perspective, this assumption means that the external arrival process feeds both stations exactly at the same time, with one additional job being released to each of the two job queues at the time of an arrival. The material supply at station 1 is assumed to be unlimited. Thus, with the base-stock policy described above, the job queue at station 1 behaves exactly the same as that of a standard single server queue. The input material supply to station 2 is the output of station 1 and hence depends on the state of station 1. Assuming that the time to move a unit from station 1 to station 2 is negligible, a job can start at station 2 when the server becomes free if station 1 is not empty. If station 1 is empty when the server at station 2 becomes free, the server becomes idle even if there are jobs waiting. We call this phenomenon "starvation".

Let R_1 and R_2 denote the base-stock levels for stations 1 and 2, respectively; and I_1 and I_2 denote the on-hand inventory levels, B_1 and B_2 the backorder levels at stations 1 and 2, respectively. For the two-stage model, the inventory control policy needs to be further specified to echelon base-stock policy or Kanban base-stock policy or others. First, we define *echelon inventory* as follows: The echelon inventory of station 2 equals the inventory level, i.e. the stock on hand minus the back-orders at station 2; The echelon inventory of station 1 equals to the sum of the inventory levels at station 1 and the echelon inventory of the downstream station, i.e. $(I_1 - B_1) + (I_2 - B_2)$. The echelon base-stock policy for the two-stage system works as follows: The job queues of the two stations are increased by 1 simultaneously, as described earlier when there is a demand arrival. Station 2 processes jobs as long as $I_1 + B_1 < R_1$ and the input materials are available. Station 1 works as long as its echelon inventory level is smaller than $R_1 + R_2$. We note that because I and B cannot be positive at the same time, the above mechanism is equivalent to station 2, works as long as there are input materials and its job queue is not empty, i.e. $I_2 < R_2$ while station 1 works as long as its job queue is not empty, or $I_1 + I_2 < R_1 + R_2$. We further note that under the echelon base-stock control, it is possible for $I_1 > R_1$; in fact, there is no upper bound on I_1 technically in a stochastic environment. Kanban policy is a different base-stock policy in that station 1 stops as soon as $I_1 = R_1$, while the rest of the above mechanism for echelon policy remain applicable. With the Kanban control, there is another important phenomenon called “blocking”, i.e. station 1 may stop when the base-stock level R_1 is reached in station 1 even when the server is free and the job queue is not empty. Blocking does not exist in the two-stage system under the echelon base-stock control.

Let us assume that customers arrive to station 2 according to a Poisson process with rate λ_1 , and the service/processing times at Stations 1 and 2 are both exponential

with rates μ_1 and μ_2 , respectively. We have now defined the model for the simplest network of inventory-queues. While this model was defined and studied in Liu, Liu, and Yao [43] for a supply chain system, much more needs to be said about the behavior and properties of this system. For example, the performance of the approximation proposed in [43] depends, to a great extent, on the tail properties of this system. The starvation probability may also be needed for the development of a new decomposition approach so that the time a job spends at a specific stage or in the whole system can be studied. In this chapter, we construct a two-dimensional Markov chain and use the matrix-analytic method to analyze this two-stage inventory queue. The main contributions of this chapter include:

- 1) A proof that the tail asymptotics of the stationary joint distribution of the two-dimensional Markov chain is geometric along the level direction.
- 2) An approximation formula for the starvation probability at Station 2, as a function of the base-stock level R_1 .

For two-dimensional models, computations are demanding and the methodologies used are often model specific. For example, the compensation method introduced by Adan, [1], can be used to compute the joint probability distribution of a specific class of QBD processes. But it does not work for our two-stage inventory model. In this situation the tail analysis of the invariant probability distributions along either the level or phase direction becomes a good alternative, although still challenging.

6.2 Mathematical model

In this section we present the mathematical models of the two-stage inventory-queue under the echelon control policy and the Kanban control policy. First, the notation

for the model is given below, where $i = 1, 2$:

N_i = number of jobs in the job queue, (single server queue), for Station i ;

I_i = inventory level for Station i ;

R_i = base-stock level at station i ;

μ_i = service rate for Station i ;

λ_1 = arrival rate into system;

U = number of jobs in queue 2 with a matching component in station 1;

H = number of semi finished products in the system.

6.2.1 The echelon base-stock policy

Recall that under the echelon control policy, Station 1 continues production until there are no more jobs in the job queue. We have the following relationships, for the echelon base-stock system, among the variables defined above:

$$U = \min(N_2, H), \quad (6.2.1)$$

$$I_1 = \begin{cases} [R_1 - N_1 + N_2 - 1]^+ & \text{if } N_2 > 0 \\ [R_1 - N_1]^+ & \text{if } N_2 = 0 \end{cases}, \quad (6.2.2)$$

$$I_2 = [R_2 - N_2]^+ = \max(0, R_2 - N_2), \quad (6.2.3)$$

$$H = \begin{cases} I_1 & N_2 = 0 \\ I_1 + 1 & N_2 > 0 \text{ and not starved,} \\ 0 & N_2 > 0 \text{ and starved} \end{cases} \quad (6.2.4)$$

$$\bar{U} = N_2 - U. \quad (6.2.5)$$

For the stability conditions under the echelon base-stock policy we have,

$$\rho_1 = \frac{\lambda_1}{\mu_1} < 1 \text{ and } \rho_2 = \frac{\lambda_1}{\mu_2} < 1.$$

Figure 6.1: Transition diagram for echelon controlled two-stage inventory queue.

If we choose $N_1(t)$ and $U(t)$ as state variables, the values of the variables N_1 and U at time t , then $(N_1(t), U(t))$ is a continuous-time two-dimensional Markov chain. In this case, N_1 and U are the stationary state variables. Let $\pi_{i,j} = P(N_1 = i, U = j)$.

We can easily write out the system equations as,

$$\begin{aligned}
 \lambda_1 \pi_{0,0} &= \mu_2 \pi_{0,1} + \mu_1 \pi_{1,0}, \\
 (\lambda_1 + \mu_2) \pi_{0,j} &= \mu_2 \pi_{0,j+1} + \mu_1 \pi_{1,j}, \quad j \geq 1, \\
 (\lambda_1 + \mu_1) \pi_{i,0} &= \mu_1 \pi_{i+1,0} + \mu_2 \pi_{i,1}, \quad 0 < i \leq R_1 - 1, \\
 (\lambda_1 + \mu_2 + \mu_1) \pi_{i,j} &= \lambda_1 \pi_{i-1,j-1} + \mu_1 \pi_{i+1,j} + \mu_2 \pi_{i,j+1}, \quad 0 \leq i \leq R_1 - 1, j \geq 1 \\
 (\lambda_1 + \mu_1) \pi_{i,0} &= \mu_2 \pi_{i,1}, \\
 (\lambda_1 + \mu_2 + \mu_1) \pi_{R_1,j} &= \lambda_1 \pi_{R_1-1,j-1} + \mu_1 \pi_{R_1+1,j-1} + \mu_2 \pi_{R_1,j+1}, \quad j \geq 1, \\
 (\lambda_1 + \mu_1) \pi_{i,0} &= \lambda_1 \pi_{i-1,0} + \mu_2 \pi_{i,1}, \quad i \geq R_1 + 1, \\
 (\lambda_1 + \mu_2 + \mu_1) \pi_{i,j} &= \lambda_1 \pi_{i-1,j} + \mu_1 \pi_{i+1,j-1} + \mu_2 \pi_{i,j+1}, \quad i \geq R_1 + 1, j \geq 1.
 \end{aligned}$$

Finding an explicit solution for the above system or developing a numerical procedure for computing the joint probabilities π_{ij} seems very challenging. Instead, we will provide simple formulas for tail asymptotics. From the above equations, we can also write out in detail the infinitesimal generator of the Markov chain which, when partitioned according to level and phase in a block-tri-diagonal form, is given by

$$Q' = \begin{matrix} \text{level} & 0 & 1 & 2 & \cdots & \cdots & R_1 - 1 & R_1 & R_1 + 1 & R_1 + 2 & \cdots & \cdots & \cdots \\ \begin{matrix} 0 \\ 1 \\ 2 \\ \vdots \\ \vdots \\ R_1 - 1 \\ R_1 \\ R_1 + 1 \\ R_1 + 2 \\ \vdots \\ \vdots \\ \vdots \end{matrix} & \begin{pmatrix} \bar{B}_0 & \bar{A}_0 & & & & & & & & & & \\ \bar{C}_0 & \bar{B} & \bar{A}_0 & & & & & & & & & \\ & \bar{C}_0 & \bar{B} & \bar{A}_0 & & & & & & & & \\ & & \ddots & \ddots & \ddots & & & & & & & \\ & & & \ddots & \ddots & \ddots & & & & & & \\ & & & & \ddots & \ddots & \ddots & & & & & \\ & & & & & \bar{C}_0 & \bar{B} & \bar{A}_0 & & & & \\ & & & & & & \bar{C}_0 & \bar{B} & \bar{A} & & & \\ & & & & & & & \bar{C} & \bar{B} & \bar{A} & & \\ & & & & & & & & \bar{C} & \bar{B} & \bar{A} & \\ & & & & & & & & & \ddots & \ddots & \ddots \\ & & & & & & & & & & \ddots & \ddots \\ & & & & & & & & & & & \ddots & \ddots \\ & & & & & & & & & & & & \ddots & \ddots \end{pmatrix} \end{pmatrix},
 \end{matrix}$$

where

$$\bar{A}_0 = \begin{pmatrix} 0 & \lambda_1 & & & \\ & 0 & \lambda_1 & & \\ & & 0 & \lambda_1 & \\ & & & \ddots & \ddots \\ & & & & \ddots & \ddots \end{pmatrix},$$

$$\bar{B}_0 = \begin{pmatrix} -\lambda_1 & & & & \\ \mu_2 & -(\lambda_1 + \mu_2) & & & \\ & \mu_2 & -(\lambda_1 + \mu_2) & & \\ & & \ddots & \ddots & \\ & & & \ddots & \ddots \end{pmatrix},$$

$$\bar{C}_0 = \mu_1 I, \bar{A} = \lambda_1 I,$$

$$\bar{B} = \begin{pmatrix} -(\lambda_1 + \mu_1) & & & & \\ \mu_2 & -(\lambda_1 + \mu_1 + \mu_2) & & & \\ & \mu_2 & -(\lambda_1 + \mu_1 + \mu_2) & & \\ & & \ddots & \ddots & \\ & & & \ddots & \ddots \end{pmatrix},$$

and

$$\bar{C} = \begin{pmatrix} 0 & \mu_1 & & & \\ & 0 & \mu_1 & & \\ & & 0 & \mu_1 & \\ & & & \ddots & \ddots \\ & & & & \ddots & \ddots \end{pmatrix}.$$

Since the method for the tail analysis developed in Chapter 3 can be more conveniently explained by applying it to a transition matrix, either stochastic or substochastic, we convert the above infinitesimal generator Q' to a transition matrix P' by

$$P' = I + \theta Q', \quad (6.2.6)$$

where

$$\theta = \frac{1}{\lambda_1 + \mu_1 + \mu_2}.$$

We partition P' into blocks in the same fashion and denote the corresponding blocks as A_0 , B_0 , C_0 , A , B and C , which are given respectively by $A_0 = \theta \bar{A}_0$,

$$B_0 = \theta \begin{pmatrix} \mu_1 + \mu_2 & & & & \\ & \mu_2 & \mu_1 & & \\ & & \mu_2 & \mu_1 & \\ & & & \ddots & \ddots \\ & & & & \ddots & \ddots \end{pmatrix},$$

$$C_0 = \theta \bar{C}_0, A = \theta \bar{A},$$

$$B = \theta \begin{pmatrix} \mu_2 & & & & \\ \mu_2 & 0 & & & \\ & \mu_2 & 0 & & \\ & & \ddots & \ddots & \\ & & & \ddots & \ddots \end{pmatrix},$$

and $C = \theta \bar{C}$.

Remark 6 We can also construct continuous time two-dimensional Markov chains with other pairs of state variables, such as $(N_1, N_2), (N_2, U), (N_2, \bar{U})$. However, it is most convenient to compute the starvation probability with (N_1, U) . Similarly for the Kanban control policy, there are different ways to construct a Markov chain. We choose one that is most convenient for writing out the starvation probability.

6.2.2 The Kanban base-stock policy

For the Kanban control policy, Station 1 keeps on serving jobs in its job queue until its inventory level I_1 reaches its base-stock level of R_1 . Relations (6.2.4) and (6.2.5) are also true for the Kanban system in addition to

$$N_2 = \begin{cases} N_1 + 1 - (R_1 - I_1), & \max(0, I_1 - N_1) \leq I_1 \text{ and system is not starved} \\ N_1 - R_1, & I_1 = 0 \text{ and system is starved} \end{cases} \quad (6.2.7)$$

For stability in the Kanban control policy, we require $\rho_2 < 1 - \pi^*$ where,

$$\pi^* = \begin{cases} 1/(R_1 + 2), & \mu_1 = \mu_2 \\ (1 - \mu_2/\mu_1)(\mu_2/\mu_1)^{R_1+1}[1 - (\mu_2/\mu_1)^{R_1+2}]^{-1}, & \mu_1 \neq \mu_2 \end{cases}, \quad (6.2.8)$$

[42]. By choosing N_2 and N_1 as state variables, (N_2, N_1) is a continuous time two-dimensional Markov chain. From the relationship between N_2 and N_1 under the Kanban control policy, we see that (N_2, N_1) is a QBD process with finite blocks and infinitesimal generator,

$$Q = \begin{bmatrix} D_0 & E_0 & & \\ F_1 & D & E & \\ & F & D & E \\ & & \ddots & \ddots & \ddots \end{bmatrix},$$

where

$$D_0 = \begin{bmatrix} -\lambda & & & & \\ \mu_1 & -(\lambda + \mu_1) & & & \\ & \mu_1 & -(\lambda + \mu_1) & & \\ & & \ddots & \ddots & \\ & & & \mu_1 & -(\lambda + \mu_1) \end{bmatrix}_{(R_1+1) \times (R_1+1)},$$

$$D = \begin{bmatrix} -(\lambda + \mu_2) & & & & \\ \mu_1 & -(\lambda + \mu_1 + \mu_2) & & & \\ & \mu_1 & -(\lambda + \mu_1 + \mu_2) & & \\ & & \ddots & \ddots & \\ & & & \mu_1 & -(\lambda + \mu_1 + \mu_2) \\ & & & & \mu_1 & -(\lambda + \mu_1) \end{bmatrix}_{(R_1+2) \times (R_1+2)},$$

$$E_0 = \begin{bmatrix} 0 & \lambda & & & \\ & 0 & \lambda & & \\ & & \ddots & \ddots & \\ & & & 0 & \lambda \end{bmatrix}_{(R_1+1) \times (R_1+2)},$$

$$E = \begin{bmatrix} \lambda & & & \\ & \lambda & & \\ & & \ddots & \\ & & & \lambda \end{bmatrix}_{(R_1+2) \times (R_1+2)}$$

$$F_1 = \begin{bmatrix} \mu_2 & & & \\ & \mu_2 & & \\ & & \ddots & \\ & & & \mu_2 \\ 0 & 0 & \dots & 0 \end{bmatrix}_{(R_1+2) \times (R_1+1)}, \quad F = \begin{bmatrix} 0 & \mu_2 & & & \\ & 0 & \mu_2 & & \\ & & \ddots & \ddots & \\ & & & 0 & \mu_2 \\ & & & & 0 \end{bmatrix}_{(R_1+2) \times (R_1+2)}.$$

6.3 Tail asymptotics of the echelon base-stock model

We now justify Theorem 16 for the echelon controlled model in detail. For this model, each infinite block, A_0, B_0, C_1, A, B and C , has a scalar tridiagonal structure where P in equation (3.1.2) is given by,

$$P(\eta) = \begin{bmatrix} \gamma_0 & \lambda & & & \\ \mu & \gamma & \lambda & & \\ & \mu & \gamma & \lambda & \\ & & \ddots & \ddots & \ddots \end{bmatrix},$$

where $\gamma_0 = \theta(\lambda_1 + \eta\mu_2)$, $\lambda = \theta\eta^2\mu_1$, $\gamma = \theta\lambda_1$ and $\mu = \theta\eta\mu_2$. The radius of convergence, $\bar{\alpha}(\eta)$, of Q may be easily found using the method outlined in Lemma 7,

$$\bar{\alpha} = \bar{\alpha}(\eta) = \frac{\lambda_1 + \mu_1 + \mu_2}{\lambda_1 + 2\eta\sqrt{\eta\mu_1\mu_2}}. \quad (6.3.1)$$

By Lemma 6.6 in [34] the censored matrix $U_0(1/\eta)$ of P to level zero is given by,

$$U_0(1/\eta) = \frac{1}{\eta}\theta(\lambda + \eta\mu_2) + \frac{1}{2} \left(1 - \frac{\theta\lambda}{\eta} - \sqrt{\left(\frac{\theta\lambda}{\eta} - 1\right)^2 - 4\theta^2\eta\mu_1\mu_2} \right),$$

which is equal to its maximal eigenvalue $u_0(1/\eta)$. The radius of convergence, α , of P is given by $\alpha = 1/\eta$ where η is a root of the characteristic equation (3.1.10), satisfying

inequality (3.1.8). Letting $\beta = 1/\eta$, substituting $\theta = (\lambda_1 + \mu_1 + \mu_2)^{-1}$ and simplifying equation (3.1.10) we get

$$0 = \eta^2 \mu_1 + \lambda_1 - \eta(\lambda_1 + \mu_1),$$

for the model described in Section 6.2. The only η value that satisfies the characteristic equation and inequality (3.1.8) is λ_1/μ_1 .

We need to determine if P is $1/\eta$ -positive which is equivalent to determining if $\alpha < \bar{\alpha}$. Due to the equivalences of the following statements,

$$\begin{aligned} 0 &< (\lambda_1 - \mu_2)^2 \\ \mu_1 + 2\sqrt{\lambda_1 + \mu_2} &< \lambda_1 + \mu_1 + \mu_2 \\ 1 &< \frac{\lambda_1 + \mu_1 + \mu_2}{\mu_1 + 2\sqrt{\lambda_1 \mu_2}} \\ \alpha = \frac{\mu_1}{\lambda_1} &< \frac{\mu_1}{\lambda_1} \left(\frac{\lambda_1 + \mu_1 + \mu_2}{\mu_1 + 2\sqrt{\lambda_1 \mu_1}} \right) = \bar{\alpha}, \end{aligned}$$

P is $1/\eta$ -positive.

To verify Theorem 16, we also need $\mathbf{x}$ and $\mathbf{y}$, the α -invariant measure and vector, respectively, of $P(\alpha)$. Notice that $\alpha = \mu_1/\lambda$ and

$$P(\alpha) = \rho_1 \theta \begin{pmatrix} \mu_1 + \mu_2 & \lambda_1 & & & \\ & \mu_2 & \mu_1 & \lambda_1 & \\ & & \mu_2 & \mu_1 & \lambda_1 \\ & & & \ddots & \ddots & \ddots \\ & & & & \ddots & \ddots & \ddots \end{pmatrix}, \quad (6.3.2)$$

which is $\alpha = \rho_1$ multiplied by a positive recurrent stochastic matrix. Therefore, $\mathbf{x}$ is simply equal to the invariant probability measure of this positive recurrent matrix, which is given by

$$\mathbf{x} = (1 - \rho_2)(1, \rho_2, \rho_2^2, \dots),$$

and $\mathbf{y}$ is simply $\mathbf{e}$, the column vector of ones.

Another way to determine the $1/\eta$ -positivity of P is by considering the convergence of $\mathbf{x}\mathbf{y}^T$. Since $\mathbf{x}$ is the invariant probability measure of $P(1/\rho_1)$, $\mathbf{x}\mathbf{e}^T = \mathbf{x}\mathbf{y}^T = 1 < \infty$. Therefore, $P(1/\rho_1)$ is $1/\eta$ -positive.

When the queue at Station 1 is viewed on its own, it behaves like an $M/M/1$ queue. The marginal probability distribution for such a queue would be $\tilde{\pi}_i = (1 - \rho_1)\rho_1^i$. The asymptotic behaviour of the joint probability distribution along the level direction is $\pi_i \sim c\rho_1^i \mathbf{x}$ where $\mathbf{x}$ is taken to be a probability measure, i.e. $\mathbf{x}\mathbf{e}^T = 1$. So the asymptotic marginal distribution for Station 1 is $\pi_i \mathbf{e}^T \sim c\rho_1^i$. Comparing this to the actual distribution for an $M/M/1$ queue we may take the constant c to be $c = 1 - \rho_1$.

6.4 Starvation probability

Recall that station 2 is *starved* if it has a positive job queue but remains inactive because its upstream is out of stock. There are two scenarios under which starvation may occur. In the first scenario, the job queue at Station 2 is greater than 1 with one job being currently in service and station 1 is out of stock. If the next event is the completion of service at station 2, Station 2 will not be able to start the service of the next job in the queue. In the second scenario, the job queue at station 2 is zero, station 1 is out of stock, and station 2 is idle. If the next event is a demand arrival, Station 2 enters the starvation state.

It follows from the definition of the Markov chain in Section 2 that the starvation probability at Station 2 is expressed by

$$p_s^K = \pi_{1,R_1+1} + \pi_{2,R_1+2} + \pi_{3,R_1+3} + \dots, \quad (6.4.1)$$

under the Kanban control, and

$$p_s^E = \pi_{R_1+1,0} + \pi_{R_1+2,0} + \pi_{R_1+3,0} + \cdots, \quad (6.4.2)$$

under the echelon control.

As p_s^K can be computed directly, this section focuses on the asymptotic properties of p_s^E . From (6.4.2) we see that p_s^E is equal to the first entry of the sum of the vectors π_{R_1+k} , $k \geq 1$. We want to approximate this sum as $R_1 \rightarrow \infty$. However, using the approximation obtained in the previous section, $\pi_i \sim c\eta^i \mathbf{x}$ is only valid for $R_1 < i \rightarrow \infty$. In other words, we actually have π_{R_1+k} for $\pi_{R_1+k} \sim c\eta^{R_1+k} \mathbf{x}$ as $k \rightarrow \infty$. The reason for this is that the approximation is based on the non boundary blocks of P' and those blocks that are far away from the boundary, that is level R_1 . If we write P' as

$$P' = \begin{pmatrix} P_0 & T_0 \\ T_1 & P_1 \end{pmatrix},$$

where

$$P_0 = \begin{matrix} & \text{level} & 0 & 1 & 2 & \cdots & \cdots & R_1 - 1 & R_1 \\ \begin{matrix} 0 \\ 1 \\ 2 \\ \vdots \\ \vdots \\ R_1 - 1 \\ R_1 \end{matrix} & \begin{pmatrix} B_0 & A_0 & & & & & \\ C_0 & B & A_0 & & & & \\ & C_0 & B & A_0 & & & \\ & & \ddots & \ddots & \ddots & & \\ & & & \ddots & \ddots & \ddots & \\ & & & & C_0 & B & A_0 \\ & & & & & C_0 & B \end{pmatrix} \end{matrix},$$

$$P_1 = \begin{matrix} & \text{level} & R_1 + 1 & R_1 + 2 & \dots & \dots & \dots \\ \begin{matrix} R_1 + 1 \\ R_1 + 2 \\ \vdots \\ \vdots \\ \vdots \end{matrix} & \left(\begin{array}{cccccc} B & A & & & & \\ C & B & A & & & \\ & \ddots & \ddots & \ddots & & \\ & & \ddots & \ddots & \ddots & \\ & & & \ddots & \ddots & \ddots \end{array} \right) \end{matrix},$$

and

$$T_0 = \begin{matrix} & \text{level} & R_1 + 1 & \dots & \dots & \dots \\ \begin{matrix} 0 \\ 1 \\ \vdots \\ R_1 \end{matrix} & \left(\begin{array}{cccc} 0 & \dots & \dots & \dots \\ \vdots & & & \\ \vdots & & & \\ A & 0 & \dots & \dots \end{array} \right) \end{matrix}, \quad T_1 = \begin{matrix} & \text{level} & 0 & 1 & \dots & \dots & R \\ \begin{matrix} R_1 + 1 \\ R_1 + 2 \\ \vdots \\ \vdots \end{matrix} & \left(\begin{array}{cccccc} 0 & \dots & \dots & \dots & C \\ 0 & \dots & \dots & \dots & 0 \\ \vdots & & & & \vdots \\ \vdots & & & & \vdots \end{array} \right) \end{matrix},$$

the asymptotic expression for π_{R_1+k} comes from the non-boundary blocks of P_1 that are not influenced by the boundary of P_1 . The starvation probability, on the other hand, involves values of the invariant measure that are close to the boundary of P_1 and hence will be influenced by the boundary values.

When considering the tail of the invariant probability measure

$$\pi = \begin{bmatrix} \pi^0 & \pi^1 \end{bmatrix},$$

where

$$\begin{aligned} \pi^0 &= \begin{bmatrix} \pi_0 & \pi_1 & \pi_2 & \dots & \pi_{R_1-1} & \pi_{R_1} \end{bmatrix}, \\ \pi^1 &= \begin{bmatrix} \pi_{R_1+1} & \pi_{R_1+2} & \dots & \pi_i & \dots \end{bmatrix}, \end{aligned}$$

the last component of π^0 , i.e. π_{R_1} , may be viewed as the boundary entry upon which the entries for π^1 are determined.

When considering the tail of π_i for the non-boundary case, we need to take into account the asymptotic behaviour of the π_{R_1} from the boundary case as described above.

In order to consider the asymptotic behaviour of π_{R_1} the method presented in the previous section is used and applied to the infinite matrix obtained by extending the structure present in P_0 to infinity. Matrix $P_0(\eta)$ for this model is given by

$$\begin{aligned}
 P_0(\eta) &= A_0 + \eta B + \eta^2 C_0 \\
 &= \theta \begin{bmatrix} \eta\mu_2 + \eta^2\mu_1 & \lambda_1 & & & \\ & \eta\mu_2 & \eta^2\mu_1 & \lambda_1 & \\ & & \eta\mu_2 & \eta^2\mu_1 & \lambda_1 \\ & & & \eta\mu_2 & \eta^2\mu_1 & \lambda_1 \\ & & & & \ddots & \ddots & \ddots \end{bmatrix}.
 \end{aligned}$$

The radius of convergence for the non boundary part of $P_0(\eta)$ is given by

$$\bar{\alpha}(\eta) = \frac{\lambda_1 + \mu_1 + \mu_2}{\eta^2\mu_1 + 2\sqrt{\eta\mu_2\lambda_1}}.$$

The characteristic equation may be simplified to

$$\eta^2\mu_1\theta - \eta(1 - \theta\mu_2) + \lambda_1\theta = 0,$$

with roots $\eta_1 = \rho_1$ or 1, of which ρ_1 is the reciprocal of the radius of convergence for $P(\eta)$. Simple calculations show that

$$0 < \alpha = \frac{1}{\eta_1} \leq \bar{\alpha}(\eta_1)$$

is satisfied. Therefore, $P_0(\eta_1)$ is $1/\eta_1$ -positive. $P_0(\rho_1)$ is given by

$$P_0(\rho_1) = \theta \rho_1 \begin{bmatrix} \lambda_1 + \mu_2 & \mu_1 & & & \\ & \mu_2 & \lambda_1 & \mu_1 & \\ & & \mu_2 & \lambda_1 & \mu_1 \\ & & & \mu_2 & \lambda_1 & \mu_1 \\ & & & & \mu_2 & \lambda_1 & \mu_1 \\ & & & & & \ddots & \ddots & \ddots \end{bmatrix}.$$

By inspection we see that

$$P_0(\rho_1)\mathbf{e}^T = \rho_1\mathbf{e}^T.$$

So we may take $\mathbf{y}_1 = \mathbf{e}$ as the $1/\rho_1$ -invariant vector of $P_0(\rho_1)$, and a little work gives

$$\mathbf{x}_1 = \frac{\mu_2 - \mu_1}{\mu_2} (1, \mu_1/\mu_2, (\mu_1/\mu_2)^2, (\mu_1/\mu_2)^3, \dots)$$

for $\mu_1 < \mu_2$. Another way of determining whether $P_0(\rho_1)$ is $1/\rho_1$ -positive is to consider the convergence of $\mathbf{x}_1\mathbf{y}_1^T$. When $\mu_1 < \mu_2$ then $\mathbf{x}_1\mathbf{y}_1^T = \mathbf{x}_1\mathbf{e}^T = 1 < \infty$. Furthermore, $\pi_1\mathbf{z} \leq \pi_1\mathbf{y} = \pi_1\mathbf{e} < \infty$. Thus, $P_0(\rho_1)$ is $1/\eta_1$ -positive. Now we have, for $\mu_1 < \mu_2$

$$\pi_{R_1} \sim c_1\eta_1^{R_1}\mathbf{x}_1 \sim (1 - \rho_1)\rho_1^{R_1} \left(\frac{\mu_2 - \mu_1}{\mu_2} \right) (1, \mu_1/\mu_2, (\mu_1/\mu_2)^2, (\mu_1/\mu_2)^3, \dots),$$

for $R_1 \rightarrow \infty$.

We now consider the tail of π_{R_1+k} , $k = 0, 1, 2, \dots$. We argue that when $k = 0$, π_{R_1} has the tail behaviour deduced from extending P_0 to infinity and hence as k tends to 0 the asymptotic behaviour of π_{R_1+k} tends to that of π_{R_1} . Similarly, as k approaches infinity the asymptotic behaviour of π_{R_1+k} can be deduced by considering the infinite block P_1 . By this reasoning, we construct the following relation

$$\begin{aligned} \pi_{R_1+k} &\sim c_1\eta_1^{R_1+k}a_k\mathbf{x}_1 + c\eta^{R_1+k}b_k\mathbf{x} \\ &\sim (1 - \rho_1)\rho_1^{R_1+k}(a_k\mathbf{x}_1 + b_k\mathbf{x}), \end{aligned}$$

or

$$\pi_{R_1+k} \sim (1-\rho_1)\rho_1^{R_1+k} \left(a_k \left(1 - \frac{\mu_1}{\mu_2} \right) \begin{bmatrix} 1 \\ \mu_1/\mu_2 \\ (\mu_1/\mu_2)^2 \\ \vdots \end{bmatrix}^T + b_k(1-\rho_2) \begin{bmatrix} 1 \\ \rho_2 \\ \rho_2^2 \\ \vdots \end{bmatrix}^T \right), \quad (6.4.3)$$

where $a_k \rightarrow 1$ and $b_k \rightarrow 0$ as $k \rightarrow 0$ and $a_k \rightarrow 0$ and $b_k \rightarrow 1$ as $k \rightarrow \infty$. Then,

$$\begin{aligned} \pi_{R_1+k} &\sim c_1 \eta_1^{R_1} \mathbf{x}_1, \quad \text{as } k \rightarrow 0 \\ \text{and } \pi_{R_1+k} &\sim c\eta^{R_1+k} \mathbf{x}, \quad \text{as } k \rightarrow \infty \end{aligned}$$

as required. Notice that when $\mu_1/\mu_2 \approx \lambda_1/\mu_2$, $\mathbf{x}_1 \approx \mathbf{x}$ and the asymptotic result for π_i obtained from the non boundary matrix P_1 is quite good. In this case,

$$p_s^E \sim (1-\rho_2)\rho_1^{R_1+1} = p_T^E. \quad (6.4.4)$$

Numerical evidence provided in the following section supports this, and we thus have the following properties.

Property 1: p_s^E decreases exponentially when R_1 increases.

Property 2: p_s^E decreases linearly as the utility of the second station increases.

6.5 Numerical studies

In the previous section, we argued that the asymptotic result given in Theorem 1 is quite good for approximating the starvation probability when $\mu_1/\mu_2 \approx \lambda_1/\mu_2$. Numerical results reported in Table 6.7 support this argument. When $\lambda_1 < \mu_1 < \mu_2$, choosing the sequences a_k and b_k appropriately can improve upon the asymptotic result deduced from considering the non boundary matrix P_1 alone. A number of

possibilities for a_k and b_k are considered:

$$\begin{aligned}
 a_k &= \rho_2^k, & b_k &= 1 - \left(\frac{\mu_1}{\mu_2}\right)^k; \\
 a_k &= \frac{1}{k+1}, & b_k &= 1 - \left(\frac{\mu_1}{\mu_2}\right)^k; \\
 a_k &= \rho_2^k, & b_k &= \frac{1}{k+1}; \\
 a_k &= \frac{1}{k^2+1}, & b_k &= \frac{k^2}{k^2+1}; \\
 a_k &= \frac{1}{k^2+1}, & b_k &= \frac{k}{k+1}; \\
 a_k &= \frac{1}{k+1}, & b_k &= \frac{k}{k+1}.
 \end{aligned}$$

Of these the last one seems to give the best results as shown in Tables 6.8 and 6.9. In Table 6.8 the approximation for the starvation probability is calculated using equation (6.4.4) and in Table 6.9 the new, improved approximation given by equation (6.5.1) is used to calculate the starvation probability. Notice that there is an improvement in the relative error of about 10%.

In this particular situation, since the tail of π_{R_1+k} is given in equation (6.4.3) the starvation probability may be approximated by,

$$\pi_s^E \sim (1 - \rho_1) \rho_1^{R_1} \sum_{k=1}^{\infty} \rho_1^k \left[\frac{1}{k+1} \left(1 - \frac{\mu_1}{\mu_2} \right) + \frac{k}{k+1} (1 - \rho_2) \right] = p_T^E(\text{new}). \quad (6.5.1)$$

Intuitively, when $\mu_1 \geq \mu_2$ starvation, in particular, $I_1 = 0$ while $U > 0$ is less likely to occur. This is illustrated in Tables 6.9-6.13. They show that the starvation probability is very small and almost negligible for $\mu_1 > \mu_2$. When we interchange the values of μ_1 and μ_2 , the starvation probability is much larger, as expected.

We now provide more numerical results for the case when $\mu_1 < \mu_2$. The power method, [56], was used to calculate the invariant probability for a truncated echelon model and programmed using C. The truncation level used was $R_1 + 500$. The customer arrival rate was kept fixed at $\lambda_1 = 1.0$. The relative error used to decide when to stop

the iterations in the power method is 1×10^{-11} . The starvation probability, p_s^E , was calculated by using equation (6.4.2), where the sum was taken over all values, $\pi_{R_1+k,0}$, for $1 \leq k \leq 500$ and the tail of the starvation probability, p_T^E , was calculated using equation (6.4.4). The relative error in Tables 6.7-6.8 was calculated using $\frac{|p_s^E - p_T^E|}{p_T^E}$. $p_T^E(\text{new})$ was calculated using equation (6.5.1).

A numerical comparison between the Kanban and echelon control systems was also done. This is illustrated in Tables 6.1-6.6. In these tables a comparison between the computed starvation probabilities for both the echelon, p_s^E , and Kanban, p_s^K , control systems is given where p_s^K is calculated using equation (6.4.1). The improvement of the echelon controlled system over the Kanban controlled system is also given and was calculated using, $\frac{|p_s^E - p_s^K|}{p_s^K}$. As expected, the starvation probability for the echelon controlled system is less than that for the Kanban controlled system. Algorithms from [37] were used to find the stationary distribution and hence the starvation probability for the Kanban controlled system. Matlab was used to program the algorithms.

6.6 Conclusions

In this chapter, we considered a two-stage inventory-queue under the echelon base-stock control. We constructed a two dimensional QBD model and demonstrated that the stationary joint distribution decays geometrically along the level direction. We also provided explicit formulas to calculate the parameters of the tail probabilities.

One of the key performance measures for this two-stage inventory-queue is the starvation probability, which can be expressed explicitly but is difficult to compute exactly. We use the tail asymptotics to construct approximations of the relevant stationary probabilities, i.e. (6.4.4) and (6.5.1), for two different cases so that the starvation probability can be computed simply. Numerical results show that the

approximations, and hence our tail asymptotics, work well.

Queueing analysis of inventory-queues is relatively new and much work needs to be done. From our perspective, there are a number of directions worthy of further investigations. The most obvious possible extension is to relax the requirement of exponential service times and Poisson arrival process of the two-stage inventory-queue. Another important possible extension is to examine whether the two-stage tail asymptotics, albeit some necessary modifications, are still true for multi-stage systems. Since the addition of even one dimension to the Markov chain will make it next to impossible to analyze, one needs to find a way to construct the tail asymptotics from those of two-stage system. In particular, it may be easier to construct approximations of the starvation probabilities of multi-stage systems using tail asymptotics of the two-stage system, since it is primarily an outcome of the system dynamics between an upstream station and a downstream station. Starvation probabilities are important from an application point of view since it is the key component of one network decomposition method. Only when equipped with an effective decomposition, we can then study more complex networks of inventory-queues.

Another interesting topic is the tail behaviour of other aspects of (two-stage) inventory queues, for example, sojourn times and inter-departure times. Apart from their application potentials, their research values also lie in the fact that new methodologies for asymptotics analysis is likely needed, making it academically challenging.

R_1	μ_1	μ_2	p_s^E	p_s^K	improvement
15	1.5	2	8.256300×10^{-05}	$1.4072622813 \times 10^{-01}$	0.999413308
20	1.5	2	1.079200×10^{-05}	$1.4027800594 \times 10^{-01}$	0.999923067
30	1.5	2	1.866000×10^{-07}	$1.4018055208 \times 10^{-01}$	0.999998669
40	1.5	2	3.227500×10^{-09}	$1.3975815879 \times 10^{-01}$	0.999999977
50	1.5	2	4.905700×10^{-11}	$1.1482388878 \times 10^{-01}$	1
15	2	1.5	2.309100×10^{-06}	$4.9617480769 \times 10^{-01}$	0.999995346
20	2	1.5	7.188300×10^{-08}	$4.9604708015 \times 10^{-01}$	0.999999855
30	2	1.5	5.618200×10^{-11}	$3.7848224754 \times 10^{-01}$	1
40	2	1.5	2.276100×10^{-14}	$1.2454439162 \times 10^{-01}$	1
50	2	1.5	2.217800×10^{-17}	$1.2454439162 \times 10^{-01}$	1
5	1.5	3	$5.3547880 \times 10^{-02}$	$6.2515603557 \times 10^{-02}$	0.143447764
10	1.5	3	$6.9869240 \times 10^{-03}$	$5.7165555726 \times 10^{-02}$	0.877777738
15	1.5	3	$9.1848370 \times 10^{-04}$	$5.6111186348 \times 10^{-02}$	0.983631006
20	1.5	3	$1.2090260 \times 10^{-04}$	$5.5912971735 \times 10^{-02}$	0.997837665
25	1.5	3	$1.5919600 \times 10^{-05}$	$5.5878248392 \times 10^{-02}$	0.999715102
30	1.5	3	$2.0963470 \times 10^{-06}$	$5.5870776946 \times 10^{-02}$	0.999962479
5	3	1.5	2.477300×10^{-04}	$7.0006160518 \times 10^{-01}$	0.999646131
10	3	1.5	1.007000×10^{-06}	$7.0011378332 \times 10^{-01}$	0.999998562
15	3	1.5	4.119200×10^{-09}	$6.9738630664 \times 10^{-01}$	0.999999994
20	3	1.5	5.667100×10^{-12}	$1.8757657297 \times 10^{-01}$	1
30	3	1.5	9.594200×10^{-17}	$1.8757657294 \times 10^{-01}$	1
10	1.5	4	8.250609×10^{-03}	$3.1346255548 \times 10^{-02}$	0.736791242
15	1.5	4	1.086280×10^{-03}	$3.0738230784 \times 10^{-02}$	0.964660295
20	1.5	4	1.430458×10^{-04}	$3.0623935746 \times 10^{-02}$	0.995328954
25	1.5	4	1.883725×10^{-05}	$3.0603689918 \times 10^{-02}$	0.999384478
30	1.5	4	2.480625×10^{-06}	$3.0599516342 \times 10^{-02}$	0.999918933

Table 6.1: Comparison of calculated p_s^E and p_s^K for various values of μ_1 and μ_2 .

R_1	μ_1	μ_2	p_s^E	p_s^K	improvement
10	4	1.5	4.117911×10^{-10}	$8.1103133271 \times 10^{-01}$	0.9999999999
15	4	1.5	4.932050×10^{-14}	$4.1405344953 \times 10^{-01}$	1
20	4	1.5	7.248259×10^{-18}	$2.2643332408 \times 10^{-01}$	1
25	4	1.5	2.311962×10^{-21}	$2.2643332408 \times 10^{-01}$	1
30	4	1.5	1.869160×10^{-24}	$2.2643332408 \times 10^{-01}$	1
5	1.5	6	7.168849×10^{-02}	$1.5237037800 \times 10^{-02}$	3.704883649
10	1.5	6	2.098800×10^{-04}	$1.3793220128 \times 10^{-02}$	0.984783829
15	1.5	6	1.242859×10^{-03}	$1.3513866109 \times 10^{-02}$	0.908030834
20	1.5	6	1.636686×10^{-04}	$1.3460133072 \times 10^{-02}$	0.987840492
25	1.5	6	2.155306×10^{-05}	$1.3450759628 \times 10^{-02}$	0.998397633
30	1.5	6	2.838264×10^{-06}	$1.3448542547 \times 10^{-02}$	0.999788954
35	1.5	6	3.737631×10^{-07}	$1.3444174195 \times 10^{-02}$	0.999972199
5	6	1.5	2.983100×10^{-06}	$9.3032254657 \times 10^{-01}$	0.999996793
10	6	1.5	3.940246×10^{-13}	$8.9131601174 \times 10^{-01}$	1
5	6	1.5	2.194398×10^{-18}	$2.7233700515 \times 10^{-01}$	1
20	6	1.5	7.230839×10^{-23}	$2.7233700515 \times 10^{-01}$	1
25	6	1.5	8.934038×10^{-27}	$2.7233700515 \times 10^{-01}$	1
30	6	1.5	3.665100×10^{-26}	$2.7233700514 \times 10^{-01}$	1
35	6	1.5	1.477526×10^{-34}	$2.7233700515 \times 10^{-01}$	1
10	1.5	7.5	9.890608×10^{-03}	$8.8314295370 \times 10^{-03}$	0.119932844
15	1.5	7.5	1.302465×10^{-03}	$8.6488023380 \times 10^{-03}$	0.84940516
20	1.5	7.5	1.715181×10^{-04}	$8.6144351040 \times 10^{-03}$	0.980089455
25	1.5	7.5	2.258674×10^{-05}	$8.6083988130 \times 10^{-03}$	0.997376197
30	1.5	7.5	2.974386×10^{-06}	$8.6072632790 \times 10^{-03}$	0.999654433
35	1.5	7.5	3.916887×10^{-07}	$8.6061132200 \times 10^{-03}$	0.999954487

Table 6.2: Comparison of calculated p_s^E and p_s^K for various values of μ_1 and μ_2 .

R_1	μ_1	μ_2	p_s^E	p_s^K	improvement
10	7.5	1.5	7.582375×10^{-15}	$2.9326755786 \times 10^{-01}$	1
15	7.5	1.5	1.051989×10^{-20}	$2.9326755786 \times 10^{-01}$	1
20	7.5	1.5	2.741449×10^{-25}	$2.9326755785 \times 10^{-01}$	1
25	7.5	1.5	1.155244×10^{-29}	$2.9326755783 \times 10^{-01}$	1
30	7.5	1.5	4.868191×10^{-34}	$2.9326755782 \times 10^{-01}$	1
35	7.5	1.5	2.051452×10^{-38}	$2.9326755782 \times 10^{-01}$	1
10	1.5	9	1.018503×10^{-02}	$6.1440294110 \times 10^{-03}$	0.657711791
15	1.5	9	1.341239×10^{-03}	$6.0155745300 \times 10^{-03}$	0.777038919
20	1.5	9	1.766240×10^{-04}	$5.9913903690 \times 10^{-03}$	0.970520365
30	1.5	9	3.062931×10^{-06}	$5.9861107420 \times 10^{-03}$	0.999488327
35	1.5	9	4.033489×10^{-07}	$5.9834937220 \times 10^{-03}$	0.99993259
10	9	1.5	2.840748×10^{-16}	$3.0822776395 \times 10^{-01}$	1
15	9	1.5	2.475855×10^{-22}	$3.0822776395 \times 10^{-01}$	1
20	9	1.5	3.908692×10^{-27}	$3.0822776394 \times 10^{-01}$	1
30	9	1.5	1.121002×10^{-36}	$3.0822776394 \times 10^{-01}$	1
35	9	1.5	1.898426×10^{-41}	$3.0822776392 \times 10^{-01}$	1
5	2	3	8.170370×10^{-03}	$9.9279048928 \times 10^{-02}$	0.917702979
10	2	3	2.422336×10^{-04}	$9.7713954748 \times 10^{-02}$	0.997520993
15	2	3	7.422271×10^{-06}	$9.7629510396 \times 10^{-02}$	0.999923975
20	2	3	2.298675×10^{-07}	$9.7598990223 \times 10^{-02}$	0.999997645
25	2	3	7.150840×10^{-09}	$9.6902563405 \times 10^{-02}$	0.999999926
30	2	3	2.229399×10^{-10}	$8.0668771442 \times 10^{-02}$	0.999999997

Table 6.3: Comparison of calculated p_s^E and p_s^K for various values of μ_1 and μ_2 .

R_1	μ_1	μ_2	p_s^E	p_s^K	improvement
5	3	2	1.459600×10^{-04}	$3.4375623836 \times 10^{-01}$	0.999575397
10	3	2	3.042176×10^{-07}	$3.4356000677 \times 10^{-01}$	0.999999115
15	3	2	6.230170×10^{-10}	$3.4182978013 \times 10^{-01}$	0.999999998
20	3	2	4.764900×10^{-12}	$1.4082780722 \times 10^{-01}$	1
25	3	2	3.273270×10^{-15}	$1.4082780719 \times 10^{-01}$	1
30	3	2	8.325916×10^{-18}	$1.4082780719 \times 10^{-01}$	1
10	2	4	3.166562×10^{-04}	$5.5190136475 \times 10^{-02}$	0.994262449
25	2	4	9.623677×10^{-09}	$5.4693231034 \times 10^{-02}$	0.999999824
30	2	4	3.007235×10^{-10}	$4.9135018199 \times 10^{-02}$	0.999999994
10	4	2	3.467641×10^{-09}	$4.0878419708 \times 10^{-01}$	0.999999992
25	4	2	1.150554×10^{-19}	$1.7340622217 \times 10^{-01}$	1
30	4	2	7.371267×10^{-23}	$1.7340622217 \times 10^{-01}$	1
5	2.75	7	1.793589×10^{-03}	$2.9823429268 \times 10^{-02}$	0.939859733
10	2.75	7	1.133461×10^{-05}	$2.9745446386 \times 10^{-02}$	0.999618946
15	2.75	7	7.200924×10^{-08}	$2.9740086187 \times 10^{-02}$	0.999997579
20	2.75	7	4.577969×10^{-10}	$2.8544566032 \times 10^{-02}$	0.999999984
5	7	2.75	3.60293×10^{-07}	$2.9154817189 \times 10^{-01}$	0.999998764
10	7	2.75	3.04029×10^{-12}	$2.4473307347 \times 10^{-01}$	1
15	7	2.75	3.31734×10^{-17}	$1.6763484887 \times 10^{-01}$	1
20	7	2.75	7.13884×10^{-22}	$1.6763484887 \times 10^{-01}$	1
5	3	4	8.490100×10^{-05}	$9.1788075031 \times 10^{-02}$	0.999075032
10	3	4	3.352400×10^{-07}	$9.1690006315 \times 10^{-02}$	0.999996344
15	3	4	1.354500×10^{-09}	$9.1523756732 \times 10^{-02}$	0.999999985
20	3	4	4.149500×10^{-12}	$5.9634591363 \times 10^{-02}$	1
30	3	4	6.969600×10^{-17}	$5.9634591349 \times 10^{-02}$	1

Table 6.4: Comparison of calculated p_s^E and p_s^K for various values of μ_1 and μ_2 .

R_1	μ_1	μ_2	p_s^E	p_s^K	improvement
5	4	3	5.16077×10^{-05}	$1.9275079556 \times 10^{-01}$	0.999732257
10	4	3	2.70783×10^{-08}	$1.9269095452 \times 10^{-01}$	0.999999859
15	4	3	1.65237×10^{-11}	$1.1142471718 \times 10^{-01}$	1
20	4	3	1.09032×10^{-14}	$1.1142471716 \times 10^{-01}$	1
30	4	3	6.33380×10^{-21}	$1.1142471716 \times 10^{-01}$	1
5	3	5	7.581200×10^{-05}	$6.1224138707 \times 10^{-02}$	0.99876173
10	3	5	3.049900×10^{-07}	$6.1149002674 \times 10^{-02}$	0.999995012
15	3	5	1.244700×10^{-09}	$6.0950049524 \times 10^{-02}$	0.99999998
30	3	5	6.950200×10^{-17}	$4.3198102092 \times 10^{-02}$	1
5	5	3	8.49408×10^{-06}	$2.1914162984 \times 10^{-01}$	0.999961239
10	5	3	1.00623×10^{-09}	$2.1790696713 \times 10^{-01}$	0.999999995
15	5	3	1.43025×10^{-13}	$1.2900731874 \times 10^{-01}$	1
30	5	3	1.35641×10^{-24}	$1.2900731873 \times 10^{-01}$	1
5	3.25	8	6.611863×10^{-04}	$2.8388474758 \times 10^{-02}$	0.97670934
15	3.25	8	4.983165×10^{-09}	$2.8237594396 \times 10^{-02}$	0.999999824
20	3.25	8	1.374094×10^{-11}	$2.2725117322 \times 10^{-02}$	0.999999999
25	3.25	8	3.789640×10^{-14}	$2.2725117320 \times 10^{-02}$	1
30	3.25	8	1.045156×10^{-16}	$2.2725117320 \times 10^{-02}$	1
5	8	3.25	1.966003×10^{-07}	$2.3530990630 \times 10^{-01}$	0.999999165
15	8	3.25	6.261006×10^{-18}	$1.4917178731 \times 10^{-01}$	1
20	8	3.25	9.605036×10^{-23}	$1.4917178731 \times 10^{-01}$	1
25	8	3.25	2.868544×10^{-27}	$1.4917178731 \times 10^{-01}$	1
30	8	3.25	8.754103×10^{-32}	$1.4917178731E \times 10^{-01}$	1
5	5.5	10	2.539973×10^{-05}	$3.1797944464 \times 10^{-02}$	0.999201215
15	5.5	10	9.563921×10^{-13}	$2.6834837619 \times 10^{-02}$	1
20	5.5	10	1.896649×10^{-16}	$2.6834837618 \times 10^{-02}$	1

Table 6.5: Comparison of calculated p_s^E and p_s^K for various values of μ_1 and μ_2 .

R_1	μ_1	μ_2	p_s^E	p_s^K	improvement
25	5.5	10	3.768543×10^{-20}	$2.6834837618 \times 10^{-02}$	1
30	5.5	10	7.487901×10^{-24}	$2.6834837618 \times 10^{-02}$	1
5	10	5.5	1.344410×10^{-07}	$1.1822357927 \times 10^{-01}$	0.999998863
15	10	5.5	1.913748×10^{-18}	$9.0707409377 \times 10^{-02}$	1
20	10	5.5	1.466311×10^{-23}	$9.0707409311 \times 10^{-02}$	1
25	10	5.5	1.466311×10^{-28}	$9.0707409378 \times 10^{-02}$	1
30	10	5.5	1.466311×10^{-33}	$9.0707409377 \times 10^{-02}$	1
10	6	10	1.750806×10^{-09}	$3.3507175568 \times 10^{-02}$	0.999999948
15	6	10	2.213075×10^{-13}	$2.8875503275 \times 10^{-02}$	1
20	6	10	2.838641×10^{-17}	$2.8875503275 \times 10^{-02}$	1
25	6	10	3.650515×10^{-21}	$2.8875503275 \times 10^{-02}$	1
35	6	10	6.037285×10^{-29}	$2.8875503274 \times 10^{-02}$	1
10	10	6	6.51228×10^{-13}	$8.1590563268 \times 10^{-02}$	1
20	10	6	2.85288×10^{-23}	$8.1590563267 \times 10^{-02}$	1
25	10	6	2.85288×10^{-28}	$8.1590563267 \times 10^{-02}$	1
30	10	6	2.85288×10^{-33}	$8.1590563267 \times 10^{-02}$	1
35	10	6	2.85288×10^{-38}	$8.1590563267 \times 10^{-02}$	1
5	6.5	10	8.426061×10^{-06}	$3.6344924143 \times 10^{-02}$	0.999768164
10	6.5	10	6.759406×10^{-10}	$3.5674952383 \times 10^{-02}$	0.999999981
15	6.5	10	5.672658×10^{-14}	$3.0789729802 \times 10^{-02}$	1
20	6.5	10	4.867354×10^{-18}	$3.0789729802 \times 10^{-02}$	1
25	6.5	10	4.194944×10^{-22}	$3.0789729802 \times 10^{-02}$	1
5	10	6.5	1.984415×10^{-07}	$9.2688423482 \times 10^{-02}$	0.999997859
10	10	6.5	8.922940×10^{-13}	$7.3833420788 \times 10^{-02}$	1
15	10	6.5	5.068593×10^{-18}	$7.3833420787 \times 10^{-02}$	1
20	10	6.5	4.448777×10^{-23}	$7.3833420780 \times 10^{-02}$	1
25	10	6.5	4.448777×10^{-28}	$7.3833420787 \times 10^{-02}$	1

Table 6.6: Comparison of calculated p_s^E and p_s^K for various values of μ_1 and μ_2 .

R_1	μ_1	μ_2	p_s^E	p_T^E	relative error
5	1.05	7	6.389689×10^{-01}	6.396132×10^{-01}	0.001007323
10	1.05	7	5.006479×10^{-01}	5.011537×10^{-01}	0.001009224
15	1.05	7	3.922708×10^{-01}	3.926670×10^{-01}	0.001009045
20	1.05	7	3.073544×10^{-01}	3.076649×10^{-01}	0.001009163
25	1.05	7	2.408202×10^{-01}	2.410635×10^{-01}	0.001009224
30	1.05	7	1.886889×10^{-01}	1.888795×10^{-01}	0.001009373
5	1.05	9	6.628971×10^{-01}	6.633026×10^{-01}	0.000611297
10	1.05	9	5.193970×10^{-01}	5.197149×10^{-01}	0.000611727
15	1.05	9	4.069611×10^{-01}	4.072102×10^{-01}	0.000611826
20	1.05	9	3.188647×10^{-01}	3.190599×10^{-01}	0.000611734
25	1.05	9	2.498388×10^{-01}	2.499918×10^{-01}	0.000611878
30	1.05	9	1.957553×10^{-01}	1.958751×10^{-01}	0.000611557
5	1.125	7	4.216574×10^{-01}	4.228030×10^{-01}	0.002709572
10	1.125	7	2.339884×10^{-01}	2.346256×10^{-01}	0.002715971
15	1.125	7	1.298469×10^{-01}	1.302006×10^{-01}	0.002716269
20	1.125	7	7.205583×10^{-02}	7.225206×10^{-02}	0.00271592
5	1.25	7	2.233310×10^{-01}	2.246949×10^{-01}	0.006069819
10	1.25	7	7.318018×10^{-02}	7.362801×10^{-02}	0.006082343
15	1.25	7	2.397968×10^{-02}	2.412643×10^{-02}	0.0060824
20	1.25	7	7.857661×10^{-03}	7.905747×10^{-03}	0.006082469
25	1.25	7	2.574798×10^{-03}	2.590555×10^{-03}	0.006082606
30	1.25	7	8.437099×10^{-04}	8.488732×10^{-04}	0.006082499

Table 6.7: Approximation of p_s^E when $\mu_1/\mu_2 \approx \lambda/\mu_2$

R_1	μ_1	μ_2	P_s^E	P_t^E	rel error 1
5	2	3	8.170370×10^{-03}	1.041667×10^{-02}	2.156447×10^{-01}
10	2	3	2.422336×10^{-04}	3.255208×10^{-04}	2.558583×10^{-01}
15	2	3	7.422271×10^{-06}	1.017253×10^{-05}	2.703614×10^{-01}
20	2	3	2.298675×10^{-07}	3.178914×10^{-07}	2.768993×10^{-01}
25	2	3	7.150840×10^{-09}	9.934107×10^{-09}	2.801728×10^{-01}
30	2	3	2.229399×10^{-10}	3.104409×10^{-10}	2.818604×10^{-01}
5	3	7	1.037854×10^{-03}	1.175779×10^{-03}	1.173052×10^{-01}
10	3	7	4.229391×10^{-06}	4.838597×10^{-06}	1.259055×10^{-01}
15	3	7	1.737631×10^{-08}	1.991192×10^{-08}	1.273413×10^{-01}
20	3	7	7.148635×10^{-11}	8.194206×10^{-11}	1.275988×10^{-01}
5	4	9	1.867583×10^{-04}	2.170139×10^{-04}	1.394178×10^{-01}
10	4	9	1.799619×10^{-07}	2.119276×10^{-07}	1.508331×10^{-01}
15	4	9	1.753022×10^{-10}	2.069606×10^{-10}	1.529682×10^{-01}
20	4	9	1.711308×10^{-13}	2.021099×10^{-13}	1.532785×10^{-01}
25	4	9	1.671194×10^{-16}	1.973730×10^{-16}	1.532814×10^{-01}
30	4	9	1.632025×10^{-19}	1.927471×10^{-19}	1.532817×10^{-01}
5	6	9	1.328199×10^{-05}	1.905197×10^{-05}	3.028548×10^{-01}
10	6	9	1.580209×10^{-09}	2.450100×10^{-09}	3.550431×10^{-01}
15	6	9	1.970706×10^{-13}	3.150848×10^{-13}	3.745474×10^{-01}
20	6	9	2.515113×10^{-17}	4.052017×10^{-17}	3.792936×10^{-01}
25	6	9	3.234457×10^{-21}	5.210927×10^{-21}	3.792934×10^{-01}
30	6	9	4.159538×10^{-25}	6.701295×10^{-25}	3.792934×10^{-01}
5	8	10	2.046006×10^{-06}	3.433228×10^{-06}	4.040576×10^{-01}
10	8	10	5.379610×10^{-11}	1.047738×10^{-10}	4.865501×10^{-01}
15	8	10	1.529628×10^{-15}	3.197442×10^{-15}	5.216088×10^{-01}

Table 6.8: Approximation, p_t^E , of p_s^E using result from Theorem 18

R_1	μ_1	μ_2	P_s^E	$P_t^E(new)$	rel error 2
5	2	3	8.170370×10^{-03}	8.404717×10^{-03}	2.788279×10^{-02}
10	2	3	2.422336×10^{-04}	2.626470×10^{-04}	7.772181×10^{-02}
15	2	3	7.422271×10^{-06}	8.207731×10^{-06}	9.569758×10^{-02}
20	2	3	2.298675×10^{-07}	2.564916×10^{-07}	1.038011×10^{-01}
25	2	3	7.150840×10^{-09}	8.015363×10^{-09}	1.078582×10^{-01}
30	2	3	2.229399×10^{-10}	2.504801×10^{-10}	1.099497×10^{-01}
5	3	7	1.037854×10^{-03}	1.006157×10^{-03}	3.150304×10^{-02}
10	3	7	4.229391×10^{-06}	4.140563×10^{-06}	2.145312×10^{-02}
15	3	7	1.737631×10^{-08}	1.703936×10^{-08}	1.977480×10^{-02}
20	3	7	7.148635×10^{-11}	7.012080×10^{-11}	1.947425×10^{-02}
5	4	9	1.867583×10^{-04}	1.802150×10^{-04}	3.630830×10^{-02}
10	4	9	1.799619×10^{-07}	1.759912×10^{-07}	2.256192×10^{-02}
15	4	9	1.753022×10^{-10}	1.718664×10^{-10}	1.999111×10^{-02}
20	4	9	1.711308×10^{-13}	1.678383×10^{-13}	1.961710×10^{-02}
25	4	9	1.671194×10^{-16}	1.639046×10^{-16}	1.961385×10^{-02}
30	4	9	1.632025×10^{-19}	1.600631×10^{-19}	1.961351×10^{-02}
5	6	9	1.328199×10^{-05}	1.345966×10^{-05}	1.320018×10^{-02}
10	6	9	1.580209×10^{-09}	1.730924×10^{-09}	8.707199×10^{-02}
15	6	9	1.970706×10^{-13}	2.225982×10^{-13}	1.146802×10^{-01}
20	6	9	2.515113×10^{-17}	2.862631×10^{-17}	1.213981×10^{-01}
25	6	9	3.234457×10^{-21}	3.681368×10^{-21}	1.213981×10^{-01}
30	6	9	4.159538×10^{-25}	4.734269×10^{-25}	1.213980×10^{-01}
5	8	10	2.046006×10^{-06}	2.157476×10^{-06}	5.166686×10^{-02}
10	8	10	5.379610×10^{-11}	6.584094×10^{-11}	1.829385×10^{-01}
15	8	10	1.529628×10^{-15}	2.009306×10^{-15}	2.387282×10^{-01}

Table 6.9: Improved approximation for p_s^E using $a_k = 1/(k+1)$ and $b_k = k/(k+1)$ in (6.4.3).

R_1	μ_1	μ_2	p_s^E	μ_1	μ_2	p_s^E
10	7	1.05	2.127326×10^{-16}	1.05	7	5.006479×10^{-01}
15	7	1.05	4.049343×10^{-22}	1.05	7	3.922708×10^{-01}
20	7	1.05	2.174865×10^{-26}	1.05	7	3.073544×10^{-01}
25	7	1.05	1.294024×10^{-30}	1.05	7	2.408202×10^{-01}
30	7	1.05	7.699313×10^{-35}	1.05	7	1.886889×10^{-01}
10	9	1.05	2.202930×10^{-18}	1.05	9	5.193970×10^{-01}
15	9	1.05	4.701990×10^{-24}	1.05	9	4.069611×10^{-01}
20	9	1.05	7.962861×10^{-29}	1.05	9	3.188647×10^{-01}
25	9	1.05	1.348518×10^{-33}	1.05	9	2.498388×10^{-01}
30	9	1.05	2.283726×10^{-38}	1.05	9	1.957553×10^{-01}
10	7	1.125	8.755412×10^{-16}	1.125	7	2.339884×10^{-01}
15	7	1.125	1.151259×10^{-21}	1.125	7	1.298469×10^{-01}
15	2	1.5	2.309100×10^{-06}	1.5	2	8.256300×10^{-05}
20	2	1.5	7.188300×10^{-08}	1.5	2	1.079200×10^{-05}
30	2	1.5	5.618200×10^{-11}	1.5	2	1.866000×10^{-07}
40	2	1.5	2.276100×10^{-14}	1.5	2	3.227500×10^{-09}
50	2	1.5	2.217800×10^{-17}	1.5	2	4.905700×10^{-11}
5	3	1.5	2.477300×10^{-04}	1.5	3	$5.3547880 \times 10^{-02}$
10	3	1.5	1.007000×10^{-06}	1.5	3	$6.9869240 \times 10^{-03}$
15	3	1.5	4.119200×10^{-09}	1.5	3	$9.1848370 \times 10^{-04}$
20	3	1.5	5.667100×10^{-12}	1.5	3	$1.2090260 \times 10^{-04}$
30	3	1.5	9.594200×10^{-17}	1.5	3	$2.0963470 \times 10^{-06}$
10	4	1.5	4.117911×10^{-10}	1.5	4	8.250609×10^{-03}
15	4	1.5	4.932050×10^{-14}	1.5	4	1.086280×10^{-03}
20	4	1.5	7.248259×10^{-18}	1.5	4	1.430458×10^{-04}
25	4	1.5	2.311962×10^{-21}	1.5	4	1.883725×10^{-05}
30	4	1.5	1.869160×10^{-24}	1.5	4	2.480625×10^{-06}

Table 6.10: Comparison of calculated p_s^E for when $\mu_1 \geq \mu_2$ and when $\mu_1 \leq \mu_2$.

R_1	μ_1	μ_2	p_s^E	μ_1	μ_2	p_s^E
5	6	1.5	2.983100×10^{-06}	1.5	6	7.168849×10^{-02}
10	6	1.5	3.940246×10^{-13}	1.5	6	2.098800×10^{-04}
15	6	1.5	2.194398×10^{-18}	1.5	6	1.242859×10^{-03}
20	6	1.5	7.230839×10^{-23}	1.5	6	1.636686×10^{-04}
25	6	1.5	8.934038×10^{-27}	1.5	6	2.155306×10^{-05}
30	6	1.5	3.665100×10^{-26}	1.5	6	2.838264×10^{-06}
35	6	1.5	1.477526×10^{-34}	1.5	6	3.737631×10^{-07}
10	7.5	1.5	7.582375×10^{-15}	1.5	7.5	9.890608×10^{-03}
15	7.5	1.5	1.051989×10^{-20}	1.5	7.5	1.302465×10^{-03}
20	7.5	1.5	2.741449×10^{-25}	1.5	7.5	1.715181×10^{-04}
25	7.5	1.5	1.155244×10^{-29}	1.5	7.5	2.258674×10^{-05}
30	7.5	1.5	4.868191×10^{-34}	1.5	7.5	2.974386×10^{-06}
35	7.5	1.5	2.051452×10^{-38}	1.5	7.5	3.916887×10^{-07}
10	9	1.5	2.840748×10^{-16}	1.5	9	1.018503×10^{-02}
15	9	1.5	2.475855×10^{-22}	1.5	9	1.341239×10^{-03}
20	9	1.5	3.908692×10^{-27}	1.5	9	1.766240×10^{-04}
30	9	1.5	1.121002×10^{-36}	1.5	9	3.062931×10^{-06}
35	9	1.5	1.898426×10^{-41}	1.5	9	4.033489×10^{-07}
5	3	2	1.459600×10^{-04}	2	3	8.170370×10^{-03}
10	3	2	3.042176×10^{-07}	2	3	2.422336×10^{-04}
15	3	2	6.230170×10^{-10}	2	3	7.422271×10^{-06}
20	3	2	4.764900×10^{-12}	2	3	2.298675×10^{-07}
25	3	2	3.273270×10^{-15}	2	3	7.150840×10^{-09}
30	3	2	8.325916×10^{-18}	2	3	2.229399×10^{-10}
10	4	2	3.467641×10^{-09}	2	4	3.166562×10^{-04}
25	4	2	1.150554×10^{-19}	2	4	9.623677×10^{-09}
30	4	2	7.371267×10^{-23}	2	4	3.007235×10^{-10}

Table 6.11: Comparison of calculated p_s^E for when $\mu_1 \geq \mu_2$ and when $\mu_1 \leq \mu_2$.

R_1	μ_1	μ_2	p_s^E	μ_1	μ_2	p_s^E
5	7	2.75	3.60293×10^{-07}	2.75	7	1.793589×10^{-03}
10	7	2.75	3.04029×10^{-12}	2.75	7	1.133461×10^{-05}
15	7	2.75	3.31734×10^{-17}	2.75	7	7.200924×10^{-08}
20	7	2.75	7.13884×10^{-22}	2.75	7	4.577969×10^{-10}
5	4	3	5.16077×10^{-05}	3	4	8.490100×10^{-05}
10	4	3	2.70783×10^{-08}	3	4	3.352400×10^{-07}
15	4	3	1.65237×10^{-11}	3	4	1.354500×10^{-09}
20	4	3	1.09032×10^{-14}	3	4	4.149500×10^{-12}
30	4	3	6.33380×10^{-21}	3	4	6.969600×10^{-17}
5	5	3	8.49408×10^{-06}	3	5	7.581200×10^{-05}
10	5	3	1.00623×10^{-09}	3	5	3.049900×10^{-07}
15	5	3	1.43025×10^{-13}	3	5	1.244700×10^{-09}
30	5	3	1.35641×10^{-24}	3	5	6.950200×10^{-17}
5	8	3.25	1.966003×10^{-07}	3.25	8	6.611863×10^{-04}
15	8	3.25	6.261006×10^{-18}	3.25	8	4.983165×10^{-09}
20	8	3.25	9.605036×10^{-23}	3.25	8	1.374094×10^{-11}
25	8	3.25	2.868544×10^{-27}	3.25	8	3.789640×10^{-14}
30	8	3.25	8.754103×10^{-32}	3.25	8	1.045156×10^{-16}
5	10	5.5	1.344410×10^{-07}	5.5	10	2.539973×10^{-05}
15	10	5.5	1.913748×10^{-18}	5.5	10	9.563921×10^{-13}
20	10	5.5	1.466311×10^{-23}	5.5	10	1.896649×10^{-16}
25	10	5.5	1.466311×10^{-28}	5.5	10	3.768543×10^{-20}
30	10	5.5	1.466311×10^{-33}	5.5	10	7.487901×10^{-24}

Table 6.12: Comparison of calculated p_s^E for when $\mu_1 \geq \mu_2$ and when $\mu_1 \leq \mu_2$.

R_1	μ_1	μ_2	p_s^E	μ_1	μ_2	p_s^E
10	10	6	6.51228×10^{-13}	6	10	1.750806×10^{-09}
15	10	6	3.31813×10^{-18}	6	10	2.213075×10^{-13}
20	10	6	2.85288×10^{-23}	6	10	2.838641×10^{-17}
25	10	6	2.85288×10^{-28}	6	10	3.650515×10^{-21}
35	10	6	2.85288×10^{-38}	6	10	6.037285×10^{-29}
5	10	6.5	1.984415×10^{-07}	6.5	10	8.426061×10^{-06}
10	10	6.5	8.922940×10^{-13}	6.5	10	6.759406×10^{-10}
15	10	6.5	5.068593×10^{-18}	6.5	10	5.672658×10^{-14}
20	10	6.5	4.448777×10^{-23}	6.5	10	4.867354×10^{-18}
25	10	6.5	4.448777×10^{-28}	6.5	10	4.194944×10^{-22}

Table 6.13: Comparison of calculated p_s^E for when $\mu_1 \geq \mu_2$ and when $\mu_1 \leq \mu_2$.

Chapter 7

Further Research

Chapter 3 dealt with creating a method for determining the radius of convergence for P in equation (3.1.3), or the eigenvalue for the R measure of P' in equation (3.1.1) and consequently, determining the asymptotic behaviour of the stationary distribution along one direction. Chapter 3 was confined to the case when P has a tridiagonal scalar structure. Considering when P has a tridiagonal finite block structure could prove to be worthwhile. One challenge here may be to determine the characteristic equation, as it would involve finding the maximal eigenvalue of the finite matrix $U_0(\beta)$. However, determining the R measure for P will prove to be quite challenging. In the scalar case the R measure for P is a scalar that can be found by solving a quadratic equation. In the block situation this is not the case. Appealing to numerical methods to find R may be necessary.

The method has only been presented for the phase independent situation. Another way of generalizing the method would be to consider phase dependence. The conditions for P to be $1/\eta$ -positive remain the same, i.e. η must be a zero of some characteristic equation and η must satisfy a particular inequality. It is verifying these conditions that may provide technical difficulties. For example in the tridiagonal

case, due to the level independent repeating structure of P the R measure for P can be found by solving a finite number of quadratic equation, thus the characteristic equation (3.1.11) has a simple form. Without the phase independence the R measure will not be a single value but rather a sequence of values. Each member of the sequence a solution of a quadratic equation dependent on the phase. As a result, the characteristic equation will have a more complicated form. A numerical procedure may be needed to solve it for the required root, η .

In the above two generalizations, the inside structure of P changes but the conditions for P to be $1/\eta$ -positive remain the same. However, the verification of these conditions present new challenges in each case. Consequently, we see that the internal structure of P does not affect the conditions to be verified.

The symmetric shortest queue is an example that has been studied intensively, so many results regarding it are available. The discussion at the end of section 3.2.4 brings up the issue that α -positivity, while sufficient to determine a geometric tail behaviour of stationary distribution along one direction, it is not necessary. An interesting question that arises is, what conditions are necessary to ensure a geometric tail, both for irreducible and reducible P . The importance of this question is reinforced for the reducible P case in Chapter 5 through the Gated Random order of service example in section 5.2.1. In this example we were able to show that P is α -positive but that the stationary distribution is not geometric along the level direction.

Throughout this dissertation we have only considered α -positive matrices P . For cases when P is α -null or α -transient it may be interesting to see if this, along with other conditions, tells us anything about the tail behaviour of the stationary distribution along a single direction.

There are different ways of considering the asymptotic behaviour of the tail of the stationary probability distribution. It may be considered jointly as in Flatto and

Mckean, [19]. For example

$$\pi_{mn} \sim c\alpha^m\beta^n, \text{ as } m, n \rightarrow \infty,$$

or along one direction as in this dissertation, Takahashi *et. al*, [57],

$$\pi_m \sim c\eta^m \mathbf{x},$$

or the marginal distribution as in Miller, [46],

$$\sum_{i=0}^{\infty} \pi_{ij} \sim c(1-\rho)\rho^j.$$

While it would seem at first sight that there is some sort of relationship between the three forms of asymptotic behaviour, there actually is no obvious relationship. Consequently, determining this relationship and which form is best to study is still uncertain.

We have only considered light tails that are geometric. It would be valuable to consider light tails that are not necessarily geometric as well tails that are heavy.

Bibliography

- [1] I. Adan, *A Compensation Approach for Queueing Problems*, Ph.D. Thesis, Eindhoven University of Technology, (1991).
- [2] I.J.B.F. Adan, O.J. Boxma and J.A.C. Resing, Queueing models with multiple waiting lines, *Queueing Systems*, **37**, (1/3), (2001), 65-98.
- [3] I.J.B.F. Adan, J.W.C.H. Visschers and J. Wessels, Two-dimensional Markov chains with geometric jumps, Technical Reprot, Eindhoven University of Technology, (1999).
- [4] I.J.B.F. Adan, J. Wessels and W.H.M. Zijm, A compensation Approach for 2-D Markov processes, *Advances in Applied Probability*, **25**, (1993), 783-817.
- [5] J. William Anderson, *Continuous-Time Markov Chains - An applications oriented approach*, Springer-Verlag, New York Inc., (1991).
- [6] E. Arjas and T.P. Speed, Symmetric Wiener-Hopf factorizations in Markov Additive Processes, *Z. Wahrscheinlich. verw. Geb.*, **26**, (1973), 105-118.
- [7] N.G. Bean, Jian-Min Li and P.G. Taylor, Some Asymptotic Properties of Two-stage Tandem Networks of $PH/PH/1$ Queues, *Advances in Matrix-analytic Methods for Stochastic Models*, Attahiru S. Alfa and Srinivas R. Chakravathy, Eds., Notable Publications, Inc., (1998), 171-192.

- [8] N.G. Bean and B.F. Nielsen, Decay Rates of Discrete Phase-Type distributions with infinitely many phases, *In Matrix-Analytic Method: Theory and Applications*, G. Latouche and P. Taylor, Eds., World Scientific, (2002), 17-38.
- [9] Patrick Billingsley, *Probability and Measure*, Third Edition, John Wiley and Sons, Inc. New York, (1995).
- [10] J.P.C. Blanc, The Power-Series Algorithm Applied to the Shortest-Queue Model, *Operations Research*, **40**, (1), (1992), 157-167.
- [11] W. John Braun, Wei Li and Yiqiang Q. Zhao, Infinite Block-Structured Transition Matrices and Their Properties, *Adv. Appl. Prob.*, **30**, (1998), 365-384.
- [12] Leo Breiman, *Probability*, Addison-Wesley Publishing Company, Inc. (1968).
- [13] J.W.M. Bertrand and J.C. Wortmann and J. Wijngaard, *Production Control a Structural and Design Oriented Approach*, Elsevier, New York, (1990).
- [14] J.A. Buzacott and J. G. Shanthikumar, *Stochastic models of manufacturing systems*, Prentice Hall, Englewood Cliffs, New Jersey, (1993).
- [15] X. Chao and M. Miyazawa, Truncation Properties of Finite-Buffer Queues, *Probability in the Engineering and Informational Sciences*, **14**, (4), (2000), 409-423.
- [16] J.W Cohen and O.J Boxma, *Boundary Value Problems in Queueing System Analysis*, North Holland, Amsterdam, (1983).
- [17] Richard G. Cooke, *Infinite Matrices and Sequence Spaces*, Dover Publications, Inc. New York, (1955).

- [18] M. Ettli, G.E. Feigin, G.Y. Lin and Yao, D.D., A supply network model with base-stock control and service requirements, *Operations Research*, 48, (2000), 216-232.
- [19] L. Flatto and S. Hahn, Two Parallel Queues Created by Arrivals with Two Demands I, *SIAM J. Appl. Math.*, 44, (5), (1984), 1041-1053.
- [20] L. Flatto and H.P. McKean, Two Queues in Parallel, *Communications on Pure and Applied Mathematics*, 30, (1977), 255-263.
- [21] R.D. Foley and D.R. McDonald, Join the Shortest Queue: Stability and Exact Asymptotics, (2000), submitted to *Annals of Applied Probability*.
- [22] David Freedman, *Approximating Countable Markov Chains*, Holden-Day, Inc., San Francisco, (1971).
- [23] David Freedman, *Markov Chains*, Holden-Day, Inc. San Francisco, (1971).
- [24] Kuo Fujimoto and Yukio Takahashi, Tail behaviour of the steady-state distribution in two stage tandem queues: Numerical experiment and conjecture, *Journal of the Operations Research Society of Japan*, 39, (4), (1996).
- [25] W.K. Grassmann, The Factorization of Queueing Equations and their Interpretation, *J. Opl. Res. Soc.*, 36, (11), (1985), 1041-1050.
- [26] W.K. Grassmann and Daniel P. Heyman, Equilibrium distribution of block structured Markov chains with repeating rows, *J. Appl. Prob.*, 27, (1990), 557-586.
- [27] Gene H. Golub and Charles F. VanLoan, *Matrix Computation*, The John Hopkins University Press, Baltimore, (1996).

- [28] Lani Haque, Liming Liu and Yiqiang Q. Zhao, Tail Asymptotics and Starvation Probability of a Two-Stage Inventory-Queue Model, (2003), submitted for publication.
- [29] Lani Haque and Yiqiang Q. Zhao, Geometric tail of a QBD process with countable level and phase, (2003), submitted for publication.
- [30] Daniel P. Heyman, A Decomposition Theorem for Infinite Stochastic Matrices, *J. Appl. Prob.*, **32**, (1995), 893-901.
- [31] D.P. Heyman and M.J. Sobel, *Stochastic Models in Operations Research, Volume 1: Stochastic processes and Operating Characteristics*, Mc-Graw-Hill Inc., Toronto, (1982).
- [32] P.G. Hoel, S.C. Port and C.J. Stone, *Introduction to Stochastic Processes*, Houghton Mifflin Company, Boston, (1972).
- [33] S. Karlin and H.M. Taylor, *A first Course in Stochastic Processes*, Academic Press Inc., New York, (1975).
- [34] J.G. Kemeny, J.L. Snell and A.W. Knapp, *Denumerable Markov Chains*, Springer-Verlag, New York, (1976).
- [35] Masaaki Kijima, Quasi-Stationary Distributions of Single-Server Phase Type Queues, *Mathematics of Operations Research*, **18**, (2), (1993).
- [36] J.F.C. Kingman, Two Similar Queues in Parallel, *Annals of Mathematical Statistics*, **32**, (1961), 1314-1323.
- [37] G. Latouche, Algorithms for infinite Markov chains with repeating columns, *In Linear Algebra, Queueing Models and Markov Chains*, C.D. Meyer and R.J. Plemmons, Eds., Springer-Verlag, (1993), 231-265.

- [38] G. Latouche and V. Ramaswami, *Introduction to Matrix Analytic Methods in Stochastic Modeling*, ASA, SIAM, (1999).
- [39] Quan-Lin Li and Yiqiang Q. Zhao, β -Invariant Measure for Transition Matrices of $GI/M/1$ Type, *Stochastic Models*, **19**, (2), (2002), 201-233.
- [40] Quan-Lin Li and Yiqiang Q. Zhao, A Constructive Method for Finding β Invariant Measures for Transition Matrices of $M/G/1$ Type, *In Matrix-Analytic Method: Theory and Applications*, G. Latouche and P. Taylor, Eds., World Scientific, (2002), 237-263.
- [41] Dannielle Liu and Y. Quennel Zhao, The Censored Markov Chain and the Best Augmentation, *J. Appl. Prob.*, **33**, (1996), 623-629.
- [42] X. M. Liu, *Performance analysis and optimization of supply networks*, Ph.D. Thesis, The Hong Kong University of Science and Technology, (1999).
- [43] L. Liu, X.M. Liu, and D. D. Yao, Analysis and Optimization of a multi-stage inventory-queue system, **2002**, submitted for publication.
- [44] Jerrold E. Marsden and Michael J. Hoffman, *Elementary Classical Analysis*, W.H. Freeman and Company, New York, (1993).
- [45] D.R. McDonald, Asymptotics of first passage times for random walk in an or-thant, *The Annals of Applied Probability*, **9**, (1), (1999), 110-145.
- [46] Douglas R. Miller, Computation of Steady-State Probabilites for $M/M/1$ Priority Queues, *Operations Research*, **29**, (5), (1981), 945-958.
- [47] Masakiyo Miyazawa, The Markov Renewal Approach for the Stationary Dis-tribution in the $M/G/1$ Type Queue with Countably many Backgroud states, (2002), submitted for publication.

- [48] Masakiyo Miyazawa and Y.Q. Zhao, The stationary tail asymptotics in the $GI/G/1$ type queue with countably many background states, (2003), submitted for publication.
- [49] Marcel F. Neuts, *Matrix-Geometric Solutions in Stochastic Models An Algorithmic Approach*, The Johns Hopkins University Press, Baltimore, (1981).
- [50] J. Prokop, Numerical Analysis of a Two-Stage Inventory-Queue Model, (1999), unpublished manuscript.
- [51] V. Ramaswami and P.G. Taylor, Some Properties of the Rate Operators in Level Dependent Quasi-Birth-and-Death processes with a Countable Number of Phases, *Communications of Statistics-Stochastic Models*, **12**, (1), (1996), 143-164.
- [52] J. Resing and R. Rietman, The $M/M/1$ queue with gated random order of service, Unpublished Manuscript.
- [53] Sheldon Ross, *Stochastic Processes*, John Wiley and Sons, Inc., Toronto, (1983).
- [54] E. Seneta, *Non-Negative Matrices and Markov Chains*, Springer-Verlag, New York, (1981).
- [55] James Stewart, *Single Variable Calculus Early Transcendentals*, Second Edition, Brooks/Cole Publishing Company, (1991).
- [56] William J. Stewart, *Introduction to the Numerical Solution of Markov Chains*, Princeton, (1994).
- [57] Yukio Takahashi, Kou Fujimoto and Naoki Makimoto, Geometric Decay of the Stead-State Probabilities in a Quasi-Birth-and-Death Process with a Countable Number of Phases, *Stochastic Models*, **17**, (1), (2001), 1-24.

- [58] R.L. Tweedie, Operator-Geometric Stationary distributions for Markov chains with Applications to queueing models, *Adv. Appl. Prob.*, **14**, (1982), 368-391.
- [59] R.N. Van Hess and W. Monhemius, *An Introduction to Production and Inventory Control*, Barnes and Noble, New York, (1972).
- [60] G.J. van Houtum, *New Approaches for Multi-Dimensional Queueing Systems*, Ph.D. Thesis, Eindhoven University of Technology, (1995).
- [61] G.J. van Houtum, W.H.M. Zijm, I.J.B.F Adan and J. Wessels, Bounds for performance characteristics; a systematic approach via cost structures, *Stochastic Models*, **14**, (1 & 2), (1998), 205-224.
- [62] Y.T. Wang, Analysis of an Intrinsic Overload Control for a Class of Queueing Systmes, *Proceedings of the 7th IFIP International Symposium on Computer Performance Modelling, Measurement and Evaluation*, Toronto, Canada, (1981), 235-243.
- [63] Yiqiang Q. Zhao, Wei Li and W. John Braun, Censoring, Factorizations, Spectral Analysis for Transition matrices with Block-Repeating Entries, *Methodology and Computing in Applied Probability*, **5**, (2003), 35-58.
- [64] Y.Q. Zhao, Censoring Technique in Studying Block-Structured Markov Chains, *Advances in Algorithmic Methods for Stochastic Models*; G. Latouche and P. Taylor, Eds., Notable Publications Inc., (2000), 417-433.
- [65] Y. Zhao and W.K Grassmann, A numerically stable algorithm for two server queue models, *Queueing Systems*, **8**, (1991), 59-80.